Mortgage Losses:
Loss on Sale and Holding Costs

June 2018

Ben Le
Kean University, School of Accounting and Finance
leben7256@gmail.com
704-705-0180
&

Anthony Pennington-Cross
Marquette University, Department of Finance and Center for Real Estate
anthony.pennington-cross@marquette.edu
414-288-1452

Mortgage Losses:
Loss on Sale and Holding Costs

Abstract

According to the Financial Accounting Standards Board's (FASB) Current Expected Credit Loss (CECL) impairment standard, by 2021 all banks will be required to forecast losses for mortgages over the "life of the loan." This paper provides a method for banks to model such losses, focusing on the magnitude of loss for mortgages that have defaulted. Those losses have two elements: the financial loss associated with the sale of the property and costs associated with the time it takes for the default to be processed and eventually sell the property (holding costs). The results show that both the dollar loss on the sale and the time-related holding costs have substantial variations across space and over time. Most of the losses are associated with the sale of a property not the holding costs. This variation can, at least in part, be attributed to borrower and loan characteristics and economic conditions. The legal environment (borrower and lender rights) can have strong effects on the length of the holding period (the default timeline) and therefore holding costs, but there is no evidence that it has an impact on the dollar loss associated with the sale of the property.

Keywords: Mortgage Loss; Loss Given Default; Foreclosure; Foreclosure Laws; Current Expected Credit Loss

June 2018

Ben Le
Edgewood College
leben7256@gmail.com
704-705-0180

&

Anthony Pennington-Cross
Marquette University, Department of Finance and Center for Real Estate
anthony.pennington-cross@marquette.edu
414-288-1452

Introduction

Two fundamentals drive the value of a mortgage: the expected cash flows the mortgage will generate over its life, and the valuation of these cash flows in the financial market. In order to calculate projected cash flows accurately, lenders and investors must at a minimum model the probability that a loan will default, and then estimate the losses on that loan if it defaults. This paper focuses on the second issue -- the loss on a mortgage in the event that it has defaulted.

Losses arise not only out of the sale of a defaulted property, but also out of holding and carrying costs. Holding costs include legal expenses, taxes, insurance, and maintenance costs. Carrying costs include lost interest payments due to default and opportunity costs over the default timeline. These losses are incurred during the default timeline, which spans from the time the loan enters default to the end of the loan's life (i.e., when the lender or investor takes or sells the property).

An effective model of projected losses from defaults, along with risk tolerances in the financial market (and other factors), helps to determine break-even interest rates on mortgages and the amount of capital a bank is required to hold. Under the Basel II and III capital framework, banks use internally generated models of the probability of default and loss given default (LGD) to help determine their capital requirements, while regulatory agencies evaluate the validity of such models. This paper can shed further light on how best to design and evaluate internal LGD models.

On June 16th, 2016, the Financial Accounting Standards Board (FASB) issued the Current Expected Credit Loss (CECL) impairment standard, which changed how banks are required to estimate losses in the Allowance for Loan and Lease Losses (ALLL) calculation. These new rules will take effect in 2020 for all Securities Exchange Commission registrants and in 2021 for all other banks. Currently, banks are only required to estimate losses in situations where future cash flows will not be collected as originally agreed, i.e., *after* the loan is already impaired. Under the new CECL impairment standard, banks will be required to forecast future losses over the life of the loan -- for all loans, whether impaired or not. The American Bankers Association calls the CECL requirements "the most sweeping change to bank accounting ever." (https://www.aba.com/Advocacy/Issues/Pages/Issues_LoanLoss.aspx, downloaded on 5/22/2018). The research in this paper offers a method to estimate losses on defaulted mortgages and provides an empirical example using mortgages securitized by Freddie Mac.

We find that losses from the sale of a defaulted property (which make up a majority of losses) are impacted by market and borrower characteristics differently than holding and carrying costs incurred during the default timeline. For example, low and negative equity in a property drives up dollar losses from the sale of the property, but it also shortens the amount of time it takes to resolve the defaulted loan, thus reducing holding and carrying costs. Similarly, states where foreclosures are processed through the judicial system have much longer default timelines and hence higher holding costs, yet we find no measurable evidence that the longer default timeline has an impact on the loss associated with the sale of the property.

This paper contributes to the academic literature in a variety of ways. First, this paper includes both observed losses from the sale of the property and directly observed costs of holding the loan. Second, we address the potential endogeneity of these factors. Third, we provide a rich set of descriptive information about the spatial variation and time series variation of losses over the real estate cycle. Fourth and lastly, the paper extends the prior literature on the legal rights of borrowers and lenders to mortgage losses and holding costs; we use an improved identification strategy by matching similar loans across borders of states with differing legal rights.

Literature on Measuring Loss Severity

The loss given default (lgd) literature uses a variety of approaches to define losses. The simplest approach is to compare the outstanding balance on a loan at the end of a loan's life to the sale price of the defaulted property. The advantage of this approach is that it does not include any mechanical costs associated with holding the property or selling the property (Lekkas et al. 1993, Crawford and Rosenblatt 1995, and Pennington-Cross 2003). Therefore, economic and financial considerations should determine this loss rate, not servicer and lender operational efficacy. An alternative is to include the original balance of the loan in the denominator so the loss rate is a percentage of the origination loan amount instead of the amount of the loan still owed (Clauretie and Herzog 1990, and Zhang, Li, and Liu 2010). Other researchers include the costs of selling the property or the net sale proceeds (Park and Bang 2014) to calculate losses. Still others attempt to measure the lost interest payments, as well as insurance costs and real estate taxes through proxies (Calem and LaCour-Little 2004, Qi and Yang 2009 and Cordell, Geng, Goodman, and Yang 2015). This last approach comes closer to estimating the full costs associated with a default.

As shown by Qi and Yang (2009) the process of taking property can impact loss rates. In particular, loans in states with a judicial foreclosure process have higher loss rates. Cordell, Geng, Goodman, and Yang (2015) use a loss severity model from Amherst Securities Group to estimate holding cost. An and Cordell (2107) estimate the lost interest on the mortgage for the first 120 days of the default timeline and thereafter apply the Freddie Mac cost of funds as a measure of holding costs. Cutts and Merrill (2008) find that, for Freddie Mac loans originated before or at the beginning of the financial crisis, foreclosures typically take longer in states with a judicial foreclosure process. The longer timeline of a judicial foreclosure process can increase losses due to holding costs, lost interest, or depreciation of the property caused by poor maintenance.

Unlike prior literature, we directly observe (for each individual loan) the actual maintenance costs, taxes, insurance, and legal expenses associated with holding a loan. We rely on this information to estimate the holding costs. After estimating these costs, we mechanically add in carrying costs -- lost interest and opportunity costs. This allows us to measure and study the costs for the loan as a whole, not the losses experienced by Freddie Mac.[1]

We estimate three separate empirical models. The first is a loss on the sale (los_{it}) specification. The second estimates the time it takes for a loan to transition from the beginning of the default to resolution of the loan. Given this predicted default timeline, estimates of holding and carrying costs are attached.

Data and Model Specification

Summary statistics from our dataset show substantial variation, across time and space, in loss rates on sales and the length of the default timeline. Our data source is the Single Family Loan-Level Dataset from Freddie Mac (available at http://www.freddiemac.com/news/finance/sf_loanlevel_dataset.html, downloaded in February 2015). The sample starts in January 2000 and ends in December 2013. Only loans with reported losses (or gains) from a post-default sale are included in the sample.[2] It typically takes quite a lot of time for loans to move through the default process, and therefore loans with unusually short

---

[1] Therefore, we do not include any measures of recoveries made by Freddie Mac from lenders, mortgage insurance companies, or any pool level insurance.
[2] The reported loss could also be 0.

timelines will be unique and not representative of a typical loan. To reduce bias caused by any over-representation of loans with short time lines, only loans that enter default (the last payment date) by December of 2011 are included in the sample. This provides at least 24 months for a loan to complete the default timeline (which is substantially more than the average timeline of 19 months in our sample). Repurchased and modified loans are not included in the sample. Loans with key missing information such as location, purchase prices, balance or other factors are also excluded.

To calculate the loss on the sale (los_i), we use the net sale proceeds of the defaulted property (nsp_i)[3] at resolution (the end of the loan or zero balance date) and the unpaid balance (upb_i) on the loan at resolution. i indexes the loans.

$$los_i = (upb_i - nsp_i) / upb_i \quad (1)$$

Figure 1 presents the distribution of the loss on sale, showing a wide range of losses. The mean loss is 39 percent with a standard deviation of 25 percent. The peak of the distribution is for losses of 40 to 50 percent. However, some loans show a gain (negative loss) on the sale because the net sale proceeds are actually larger than the unpaid loan balance. Figure 2 shows box charts for the loss on sale for each year (the year in which the loan ends or is resolved, not the year in which the loan was originated). The figure indicates that even within each year there is a large variance in loss rates. The median loss rate increases steadily from 2000 until 2009 and declines slightly in 2010 and 2011.[4] Figure 3 reports box charts by state. Within each state there is a very wide distribution of loss rates. Across states there are also substantial differences in both the

[3] Net sale proceeds is the selling price of the property less expenses associated with the sale of the property. It does not include costs associated with holding the property.

[4] An and Cordell (2017) find that loss severity stays elevated after the introduction of new servicing rules and standards associated with the Consumer Finance Protection Bureau (January 2014) and the National Mortgage Settlement (March 2012).

median loss on sale rate and the volatility around the median. Consistent with the magnitude of the house price cycles and overall severity of the Great Recession, states with especially low loss rates are Alaska, Montana, North Dakota, South Dakota, and Wyoming. States with especially high loss rates are Arizona, California, Florida, Indiana, Michigan, Nevada, and Ohio.

The last date the borrower makes a payment is used as the start date for the default timeline.[5] The loan is fully resolved when Freddie Mac has received proceeds from selling the property and any losses are written off (the zero balance date). Figure 4 shows the default timeline distribution. The mean default timeline is approximately 19 months, the peak of the distribution is 12 to 16 months, and the shape of the distribution is approximately log-normal. Figure 5 box charts indicate that the default timeline was fairly steady from 2000 through 2007. In 2008 and 2009 both the variance and median of the timelines increased, and then in 2010 and 2011 the timelines declined from the peak.

Figure 6 shows the variation in the default timelines across states. Again, there is substantial variation both across different states and within states. States with large losses are not always states with long default timelines. For example, Arizona has very high loss rates but has the shortest default timeline and a low variance of timelines within the state. By way of contrast, Wisconsin has a long timeline as well as relatively high loss rates. This paper will try to disentangle the reasons why losses and default timelines differ.

There are multiple paths by which a loan can be resolved. In our data, the most prevalent path is for Freddie Mac to become the owner of the property through a foreclosure sale. When this occurs, the property is referred to as "real estate owned," or reo property (approximately 65

[5] Technically the loan is in default when it misses one payment. However, many of these loans cure, so we mark the start of the timeline at the last payment date. The end of the default timeline is the end of the loan’s life (the resolution date for the lender and investor).

percent of the data has an reo indicator). After Freddie Mac takes possession and ownership of the property, it contacts local representatives to sell the property in an attempt to recover losses. Loans with losses and a recorded sale that are not reo may be short sales (the borrower sells the property but does not cover all the losses), complete sales (the borrower sells the property and covers all the losses), or sales by or to third parties. There are a myriad of other avenues for resolving a loan in default, including refinancing or loan modification, but none of these options include the sale of the property. For property that does become reo, we can separately measure the time from the last payment date to the reo begin date (pre-reo timeline) as well as the time from the beginning of reo status to the resolution date when the property is sold (reo timeline).

Explanatory Variables and Model Specification

The loss on sale (los_{it}) should be strongly correlated with the amount of equity in the property. Figure 1 and table 1 highlight that the average loss is 39 percent in the data. However, some properties have 100 percent losses; and there are a non-trivial number that have gains on the sale (negative losses, which can only occur when the value of the property is greater than the outstanding balance on the loan). One explanation for this large variety in losses on sale is the amount of equity in the property. Property with lower current loan to value (ltv) ratios should have low losses and even potential gains, whereas property with high ltvs should have higher losses on the sale. There is also evidence that loans with larger unpaid balances have lower losses. This may be due to fixed costs, the different amount of effort made in the sale of lower-cost defaulted homes, the quality or "sellability" of the home, and other unobserved factors that may be reflected in the outstanding loan balance (Clauretie and Herzog 1990, Zhang, Li, and Liu 2010, Park and Bang 2014, Calem and LaCour-Little 2004, Qi and Yang 2009).

Borrower characteristics may also affect the loss. For example, homeowners who are having financial difficulty may react differently if this is their first home or if they had been very careful with their finances. Therefore, a first-time homebuyer indicator is included in our model, as well as the credit score of the borrower at origination. To proxy for overall deterioration of the labor market, the change in the county level unemployment rate from origination to the end of the loan is also included.

The type of loan may also matter, at least indirectly. For example, the channel through which the loan is originated (retail, broker, or wholesale) may reflect unobserved differences in loan quality (Jaing, Nelson and Vytlacil 2014). Consistent with these results, borrowers who extract equity through a cash-out refinance are more likely to default (Pennington-Cross and Chomsisengphet 2007). Hence, indicators for these factores are included in the specification.

The mortgage market experienced a series of interventions as the mortgage crisis unfolded, which are likely to affect default timelines and realized losses. In October of 2008 Fannie Mae and Freddie Mac announced that they suspended foreclosures of occupied homes (Reuters.com, 2008); the largest servicers announced a similar moratorium in February of 2009 (Wall Street Journal, 2009). In addition, some states (including California in June of 2009) and even municipalities instituted moratoriums. In September of 2010 some larger servicers announced that they were suspending foreclosures after the "robo-signing" of legal documents (in which servicers were approving foreclosures in large quantities without having reviewed appropriate documentation) was publicly revealed (Inside Mortgage Finance, 2010). The rules governing servicing have also changed over time. The Consumer Finance Protection Bureau (CFPB) issued new servicing rules that took effect in January of 2014. The National Mortgage Settlement (NMS), reached in March 2012 and affecting the five largest mortgage services, also changed servicing rules. An and Cordell (2017) discuss the long-term effect of these rule changes in more detail. Any empirical model of foreclosures, refinances, or losses must try to account for these changes in the mortgage market. Therefore, we include state and time fixed effects to control for the events and, in some specifications, interact state and time to allow for location specific changes in losses over time. This flexible approach should control for the interventions described above.

There is evidence that the way foreclosures are processed, as well as the legal rights of the lender and borrower, can affect mortgage loss outcomes. For example, in states with judicial foreclosure proceedings (as opposed to power of sale proceedings which do not involve the judicial process), the foreclosure process takes longer – though ultimately the use of judicial foreclosure proceedings has little impact on the outcome for the borrower. The effect of a right to redeem the property is even less clear in terms of default rates, though it is again associated with longer default timelines (Collins, Lam and Herbert 2011, Demiroglu, Dudley and James 2014, Gerardi, Lambie-Hanson and Willen 2013, Cordell, Geng, Goodman and Yang 2015, and Clauretie and Herzog 1990). The ability of the lender or investor to attempt to recover losses from the borrower (beyond just taking the home) has been shown to decrease foreclosure rates; this is especially true when a home is in negative equity (Ghent and Kudlyak 2011, Cha, Haughwout, Hayashi and Klaauw 2015). Harrison and Seiler (2015) also show that judicial foreclosure has little impact on mortgage interest rates. Our measures of states with judicial foreclosures, statutory rights of redemption, and recourse are taken from Cutts and Merrill (2008) and Ghent and Kudlyak (2011). How these results translate to losses on sale and holding costs is an empirical question.

Table 1 provides the summary statistics for the estimation data set. The average loss on sale is 39 percent and the average time to resolve a default is 19 months (for defaults that end in the sale of the property only). For loans that become reo, the loss on sale is very similar to non-reo loans, but the default timeline is longer, at 21 months. Most of the default timeline is spent before the property becomes reo (pre-reo). Consistent with the incentives to default, by the time a defaulted loan is resolved the ltv is on average almost 100 percent. Credit scores are less than 700 on average and county unemployment rates have risen by almost 4 percentage points since origination. The majority of the loans were refinances at origination, and about one third were originated through a retail (non-broker or wholesale) channel.

The loss on sale specification is as follows:

$$los_i = \beta_1 bal_i + \beta_2 bor_i + \beta_3 loan_i + \beta_4 legal_i + \beta_5 fe_i + \varepsilon_i \qquad (2)$$

los_i is the loss on the sale; β subscript 1 through 5 are vectors of the coefficients to be estimated; bal_i is a vector of balance related variables such as the unpaid balance (upb_i) and the current loan to value ratio (ltv_i); bor_i is a vector of borrower related variables such as credit score ($fico_i$), a first time homeowner indicator ($first_i$), and the change in the county unemployment rate ($\Delta urate_i$); $loan_i$ is a vector of loan attributes at origination such as an indicator that the loan was originated through a retail channel ($retail_i$), a cash out refinance indicator ($cashrefi_i$), and a refinance indicator when no cash is extracted ($nocashrefi_i$); $legal_i$ is a vector of variables that describe the legal processes and rights of the borrower and lender such as an indicator that local laws require a judicial procedure for foreclosure ($judicial_i$), an indicator that the borrower has the statutory right to redeem the property after the initial transfer of property ownership from the borrower to another entity (srr_i), and an indicator that the lender has the ability recover any losses from non-housing related sources such as income or other borrower assets ($recourse_i$); fe_i

is a matrix of fixed effects including fixed effects for the year the loan is originated (*origination year*$_i$), the year the last payment is made on the loan (*last payment year*$_i$), servicer fixed effects (*serv*$_i$), and a vector of state fixed effects (*state*$_i$); ε_i is a random and identically distributed error term. The standard errors are clustered at the 3-digit zip level.[6]

The default timeline specification is very similar to the loss on sale specification and is as follows:

$$months_i = \alpha_1 bal_i + \alpha_2 bor_i + \alpha_3 loan_i + \alpha_4 legal_i + \alpha_5 fe_i + \varepsilon_i \qquad (3)$$

Months represents the number of months from the last payment to the loan resolution (*month_ltz*$_i$), which is the end of the loan or zero balance and the α's represent vectors of coefficients to be estimated. Some specifications will look only at loans that become reo and break the time line into the pre-reo timeline (*month_ltr*$_i$) and the reo timeline (*month_rtz*$_i$). Each grouping of explanatory variables is the same as for the loss on sale regression, except that all time-varying variables are observed at the last payment date (the beginning of the default timeline). This is done so that the explanatory variables represent starting conditions of the default timeline instead of contemporaneous conditions.

Since there is much less literature on the default timeline and holding costs, it may be instructive to consider in more detail the potential relationships between these variables and the length of the default timeline. A variety of scenarios might shape the course of a default timeline. For instance, while losses on a sale are anticipated to be negatively associated with loan size, due to fixed costs, the market for more expensive and larger homes is less homogeneous and thinner. Therefore, larger homes should be more difficult to sell, lengthening the default timeline.

---

[6] The three digit zip code is the smallest geographic unit of observation in the data set.

Borrowers who are in better financial condition and are more financially sophisticated are more likely to understand that the taking of the home is inevitable at some point and be prepared to move before official eviction. This awareness likely speeds up the taking process. Therefore, higher credit scores may be associated with shorter default timelines. An unemployed (or under-employed) homeowner who believes they can find a new or better job next week or month will try to delay the taking of the property. Therefore, increasing unemployment rates may be associated with longer default timelines. Finally, a homeowner with modest or even a small amount of negative equity may benefit from delay, to search for alternative solutions such as a short sale or a modification. These efforts will lengthen the default timeline.

These anecdotal stories of borrower incentives, which may extend or shorten the default timeline, no longer matter after the lender has taken possession of the property. Thus, during the pre-reo time period, borrower characteristics should matter as the default process is “negotiated”; but once the property is reo, borrower characteristics should not be relevant to the timeline.

Results

Loss on Sale Results

Table 2 provides the loss on sale results for various specifications. Consistent with the prior research, specification I shows that the ltv and unpaid balance on the loan have the expected signs and provide reasonable explanatory power (R^2 over 0.25). Loans with higher current ltvs have higher losses; loans with larger unpaid balances have lower losses. Non-linear and more flexible specifications of these two variables revealed no unusual patterns and provided little or no more explanatory power, so we report the linear specification only. Specification II indicates that borrower characteristics can also impact losses. For all specifications in the table, first-time homebuyers, borrowers with higher credit scores at origination, and declining unemployment rates are associated with lower loss rates on the sale. However, compared to the balance variables (unpaid balance and ltv) the borrower characteristics provide little additional explanatory power.

Specification III indicates that the characteristics of the loan at origination can also have meaningful impacts on expected losses on the sale. Loans originated through the retail channel have lower losses. Refinance loans tend to have higher losses and cash out refinances have the highest. Purchase loans that are originated through the retail channel have much lower losses on sale.

The data we use in this study covers the housing boom, housing bust and recovery. Since there have been many interventions in the market and changing underwriting standards, origination and resolution-year fixed effects are included in specification IV. This has little impact on the point estimates or their precision for all variables except the change in the unemployment rate and credit scores.

Specification V includes state fixed effects to control for the legal environment (borrower and lender rights) and servicer fixed effects to control for unobserved differences in how servicers process defaulted loans. Taken together, the extra control variables in specifications IV and V add some explanatory power to the regression and do not materially change the results for most variables. The exceptions are credit scores and changes in unemployment rates.

Robustness

Table 3 reports on additional specification tests and Table 4 examines any biases associated with truncation of the data. Table 3 specification I includes an interaction of last payment year dummies and state dummies. This specification will control for unobserved unique shocks that vary over time, including policy changes at both the national and state level. The results are not materially affected by these additional controls.

Specification II adjusts the left hand side variable from the percentage of loan amount lost to thousands of dollars lost. The order of magnitude of the coefficients will necessarily differ because the variables are now explaining dollars and not percentages, but the direction of the effects are almost all the same. The one exception is upb_{it} (unpaid balance). However, these results are consistent with the prior results based on percentages. Larger loans are associated with smaller percentage losses but larger dollar amount losses. In specification II, upb_{it} is functioning as a scaling mechanism. It would be counterintuitive if larger loans did not have larger dollar losses. A one dollar increase in upb_{it} is associated with a 19-cent increase in loss on sale. As a result, on a percentage basis larger loans should have lower losses. In short, specification II provides almost the same story as the prior results.

Specification III only includes loans that were for purchase. There is evidence in the literature that appraisals are not random, so this may bias our result for ltv_{it}. However, the coefficient estimate for ltv_{it} is almost identical, indicating that our results are not biased by appraisal issues. However, the impact of credit scores is much smaller and statically insignificant for the purchase only loans.

Specification IV attempts to correct for any selection bias. Defaulted loans and foreclosed properties are not randomly selected. There is a long line of literature showing that less equity in the home, lower credit scores, and stressed labor markets are associated with higher default rates. This is why, in the sample of loans used to estimate losses, ltv_{it} is almost 100, credit scores are under 700, and the loans are in locations with rising unemployment rates. To help correct for this potential issue, we implement a two-stage process. In the first stage we estimate a probit model indicating whether a loan is ever foreclosed on, using information available at origination of the loan. The appendix shows the results for the first stage results, which are as expected – less equity, more debt, lower credit scores, smaller loans, and loans with one borrower are associated with a higher probability of default. The specification also allows the impact of equity to change when the ltv is above 80 percent, because these loans will be required to have private mortgage insurance and will also be underwritten by the insurance company. Origination year and state and servicer fixed effects are also included. A random sample of loans originated each year is included in the sample along with all the foreclosures used in the loss estimates. More than 519,000 loans are included in the estimation, and each loan is weighted by the inverse of the probability of being in the sample. For the second stage, the inverse mills ratio is calculated and included in the loss on sale specification V in Table 2.[7] The coefficient on the inverse mills ratio is statistically significant, indicating that selection issues are present. However, most of the key variables do not materially change. The borrower's credit score is the exception: it becomes statistically insignificant. In summary, Table 3 shows that the results are very robust across a variety of specifications.

Table 4 addresses a major concern with the estimation approach. The raw data indicates that, on average, it takes 19 to 21 months to complete the default and foreclosure process. This is why for the last 2 years of available data, only loans with more than 24 months since the last payment date were included (all prior specifications). Since there is substantial variation in the length of these timelines, table 4 extends the approach already used. For the last 2 years of available data, only loans with more than 36 or 48 months since the last payment date are included. If there is any bias associated with truncating the sample, the parameter estimates

[7] The inverse mills ratio, which is often referred to as λ, is calculated as: $\lambda = \phi(x'\beta)/ \Phi(x'\beta)$, where ϕ is the normal partial density function and Φ is the normal cumulative density function.

should change across these subsamples. The impacts of equity, loan size, and the labor market are very stable. Again, the least stable effect is for credit scores.

Default Timelines: Basic Results

Table 5 provides the basic results for four different specifications of the default timeline, which is measured as the log of the number of months from the beginning of default to resolution. Origination year, last payment year, state, and servicer fixed effects are included in the specifications. Specifications I and II use the full length of the default timeline (from the last payment date to loan resolution or zero balance). Specification I includes loans where the property never becomes reo; specifications II through IV include loans that do enter the reo state at some point during the default timeline.

In contrast to our finding that loans with higher ltvs result in higher losses upon sale, loans with higher ltvs have shorter timelines. There are intuitive reasons why this might be the case. Less equity, almost by definition, increases losses on a sale. However, a higher ltv makes it harder to cure or modify the loan. As a result, less effort is likely spent on finding an alternative to foreclosure and more effort is spent processing the foreclosure and getting the home sold, thus shortening the timeline.

Loans with larger outstanding amounts (unpaid balances) usually are associated with a longer default timeline. This is in contrast to losses on sale, where larger loans are associated with smaller losses. These contrasting findings likely relate to the amount of effort made by servicers for larger loans, the thickness of the market for selling property, and the fixed costs associated with taking and selling property.

The remaining results for foreclosure timelines are largely consistent with the loss on sale results. The default timeline is shorter for borrowers with higher credit scores and in locations with declining unemployment rates. These results indicate that the borrower may find value in delay when local labor markets are stressed, when there is equity in the home, and when the borrower has low credit scores at the beginning of the loan. In addition, refinanced loans tend to take longer to sell.

There are differences in the results across different parts of the default timeline. For example, using specifications I and II, non-reo defaults are almost 3 times more sensitive to borrower credit score. There are large differences between the pre-reo timeline and the reo timeline results (specifications III and IV). In general, after the property has become reo – i.e., after the borrower has left the home and lost any power to negotiate -- the timeline is more strongly affected by the key loan characteristics (ltv, upb, and fico) and less impacted by loan origination information, labor market conditions or other borrower characteristics.

Default Timelines: Robustness

To eliminate any bias associated with the truncation of the sample, table 6 includes only loans that have had 3 or 4 years to complete the foreclosure process. The coefficients are estimated with very similar precision and the signs of the coefficients are the same. However, the magnitude of the estimates tends to be larger, the more any potential truncation is removed from the sample. These results indicate that ignoring truncation tends to bias most coefficient estimates towards zero. As a result, any timeline cost estimates will underestimate the true cost of holding a defaulted loan.

Table 7 addresses the non-random selection of loans. The same selection equation (first step, see the appendix) is used to estimate the inverse mills ratio, which is then included in the default timeline estimates (second step) reported in table 7. The results for specifications I through III are very similar - many point estimates are almost identical. The coefficient on the inverse mills ratio is statistically significant in 3 of the 4 specifications. The point estimates for some variables modestly increase (*lnltv* and *lnupb*) and others modestly decrease (*lnfico*).

Default Timelines and Changing House Prices

Changes in a home's market value directly affect the current LTV estimate and the default timeline. The prior results indicate that declining house prices, which would increase the log of the current LTV (*lnltv*), reduce the length of the default timeline. However, house prices may exert an effect on the default timelines that is independent of the current equity position. This may occur if the ability to sell distressed property is cyclical in nature. For example, in a market where prices are declining the volume of distressed sales (foreclosure, short sales, or other forced sales) increases, which can make properties harder to sell. As a result, the increased marketing time needed to sell property extends the foreclosure timeline. This story is consistent with the findings of Cordell, Geng, Goodman and Yang (2015). In contrast, An and Cordell (2017) find that larger decreases in house prices shorten the default timeline.

Table 8 conducts an additional empirical test of this potential phenomenon. The results are presented for the full timeline (from last payment date to resolution) separately for loans that go through REO and those that do not. House price changes are measured using the 3-digit zip price index from the date of last payment through loan resolution. The first two specifications (I and II) show that when prices increase the default timeline is shorter. This is similar in spirit to the results found by An and Cordell (2017). Since changes in house prices are by definition directly correlated with the current LTV, it is not surprising that the magnitude of *lnltv* coefficient is reduced.

Specifications III and IV include dummy variables indicating how much prices have declined or increased. These results indicate that the longest default timelines occur in locations where prices have increased by more than 5 percent. This may reflect the fact that, with rising prices, there may be more value in delay, which can extend the default timeline.

Causation: Does the Default Timeline increase the Loss on Sale?

So far, our empirical approach has treated the loss on the sale as independent of the default timeline. However, one could argue that the amount of time a loan spends in a distressed state is likely to affect the value of the property and hence the loss associated with the sale. Pennington-Cross (2006) hypothesizes that one reason foreclosed property sells at a discount is because the incentive to maintain a property is very low if you know you are going to lose the home in the future through foreclosure. Campbell, Giglio and Pathak (2011) show that the timing of a forced sale due to the death of an older homeowner does not affect the price discount. They interpret this as providing indirect evidence that the discount is due to poor maintenance. They also find that lower-priced homes sell at a larger forced sale discount; they infer that these houses must be sold faster and at a discount to avoid vandalism. Goodwin and Johnson (2017) find that discount for a foreclosure sale is larger than for a short sale. Foreclosures also spend less time on the market and are more likely to sell. However, they find no evidence that days on the market affected the selling price. Harding, Miceli and Sirmans (2000) show that maintenance expenditures are lower when the current LTV is near 100 percent. Cordell, Geng, Goodman Yang (2015) discuss what they call excess depreciation, which is the extra decline in the value of the property while the owner is not making payments and likely not maintaining the property. In their empirical work, excess depreciation is inferred from the residual of an auxiliary regression. An and Cordell (2017) directly include the default timeline as an explanatory variable for total losses (including holding costs). Our focus in this section is on the impact of the default timeline on the loss on sale and whether they are jointly determined. Given concerns about distressed borrowers not maintaining property, we should expect that, as the default timeline becomes longer, losses on sale should also increase.

Table 9 tests the hypothesis that longer default timelines reduce the value of the property. Specification I includes the number of months in the default timeline as an explanatory variable for the loss on the sale. The coefficient is positive, as anticipated, but is small and statistically insignificant.[8] This provides some initial evidence that any delayed maintenance is not reducing the value of foreclosure property. As explained below, results show that expenditures made by investors can be used to bring inadequate maintenance back to standard levels, once they take back the property.

Specifications II through IV include an instrumental variables approach to solving this issue. Good instruments are capable of explaining the contemporaneous mortgage timeline but not the loss on sale. The instrument used is the average length of the foreclosure timeline in the 3-digit zip code in which the property is located. Three versions are tested -- contemporaneous 3-digit zip code based instrument, 3-digit zip code lagged one year, and 3-digit zip code lagged 2 years. The first stage robust partial R^2 indicates that these variables do a reasonable job explaining the loan specific default timelines. The long lag also makes it less likely the instrument is correlated with the error term, and long lagged region-based timelines are very unlikely to be affected by contemporaneous losses on an individual property. Therefore, the 3-year lag version of the zip code timeline is most likely to meet standards for a good instrument. However, the endogenous F-statistic is never statistically significant. Thus the default timeline

[8] An and Cordell (2017) find a positive relationship between the length of the foreclosure timeline and losses. However, this is expected because they define losses to include interest related carrying costs, which must increase with the passage of time.

(*$months_{it}$*) does not appear to be endogenous. Perhaps not surprisingly, given the F-statistic, in all specifications the default timeline has no impact on losses on the sale of the property.[9]

In summary, the results offer no evidence that default timelines are associated with losses from the sale of a foreclosed property. These results may seem to be inconsistent with the prior research. However, the research here includes only foreclosed property and does not compare these losses to non-foreclosure potential losses. The empirical work here also only includes GSE loans. The approach that the GSEs use to process foreclosed and distressed loans is likely much different than loans in the jumbo market, the alt-a market, the subprime market, and loans held in portfolio. That said, even if these results are not generalizable across other segments of the mortgage market, they are an important part of the discussion because GSEs are such a large segment of the mortgage market.

Earlier research used indirect evidence to show that maintenance and the threat of vandalism was the driving force behind price discounts. In the following section (Holding Cost and Calculating Total Cost), we directly measure maintenance costs and relate these costs to the length of the holding period. The results show that holding costs are non-trivial, varying by price segment, cohort, and time. The results support the hypothesis that the GSEs use their legal and maintenance expenditures to bring property quality back to market norms in an attempt to minimize losses.

Borrower and Lender Rights

Tables 10 and 11 present specification tests that examine borrower and lender rights. Specification I includes the typical approach to examining these issues. Dummy variables are

[9] The instrumental variables approach used here is perfectly identified, so over-identification statistics are not relevant. Additional specification tests that included the contemporaneous, one-year lag and two-year lag instruments all generated the same basic results. However, if clustered errors are not used (making it possible to test for over-identification using the Hansen J-statistic), the results indicate that over-identification is an issue.

included to indicate whether the defaulted loan is in a judicial foreclosure state, a statutory right of redemption state, or a recourse state. The prior literature would lead us to expect that judicial foreclosure processes and rights of redemption increase losses on the sale. The unobserved motivation for this is that both legal requirements increase the time it takes to recover the property, while the right of redemption also reduces the quality of the title (ownership) during the right of redemption time period.

Consider a homeowner who is not making any payments on his mortgage. The owner expects to be removed from the property eventually, but he does not know exactly when. Under these circumstances, the homeowner has no incentive to maintain the property. The primary objective is to maintain habitability (for example, keep the heat or air conditioning on) during his stay, not to invest in needed capital expenses to stop property depreciation and maintain full functionality (such as replacing the roof, repaving the driveway, or repainting the exterior). From this perspective, it is not a surprise that specification I finds that losses on sale are 3 percentage points higher in judicial foreclosure states. It is also not surprising that the default timeline is 21 to 42 percent longer (e^{β}) in these states (table 11). There is no evidence that rights of recourse and of redemption have any impact on the loss on sale, although they may shorten the default timeline.

Tables 12 and 13 enhance the identification strategy beyond simple dummy indicators. Identification is likely improved if the sample is limited to nearby locations in similar labor and housing markets with different legal rights. This helps to remove any unobserved variables that may be affecting the results. Therefore, specification II limits the sample to only loans in 3-digit zip codes that border other 3-digit zip codes with different legal rights and foreclosure laws. As a result, there is a different sample of loans, zip codes and states for each legal right being tested across the spatial discontinuity. For each state border discontinuity, a dummy variable is used to control for any unobserved economic conditions.

First, consider table 12. Once the sample is reduced to loans in the border zip codes, all coefficient estimates are statistically insignificant. To help improve identification even further, loans across borders are matched one-to-one with each other based on their unpaid balance, the current LTV estimate, and the last year a payment is made. Each loan in the judicial side of the border is matched with a similiar loan in the power of sale side of the border. The best match is determined by selecting the loan with the smallest absolute sum of percentage deviations across the three criteria. This process is repeated for the statutory right of redemption border discontinuities and recourse border discontinuities. The appendix includes a comparison of the matched loans across the three different types of borders. On average, the matched loans across the borders have very similar ltvs, balances, changes in unemployment rates, and credit scores. After these attempts to improve the identification of the legal impacts on the loss on sale, the results consistently find that judicial foreclosures, rights of redemption, and rights of recourse have no impact on losses on sale.

In sum, our results support several critical findings. Most surprisingly, longer default timelines do not increase losses on sale. Furthermore, differences in the legal rights and required procedures in foreclosure do not impact losses on sale, but rather have impacts on how long the default timeline lasts. Thus, to the extent different legal requirements and rights create higher losses on a defaulted loan, the mechanism for those losses involve holding costs (this includes maintenance costs) and carrying costs.

Table 13 repeats the same empirical approach for the default timeline. The first column reports results using the full sample; the second column reports results using the unique border samples; and the third column reports the results for the same matched loans across the border discontinuity. Each coefficient estimate reported in columns II and III is from a separate regression.

In contrast to the loss on sale results, legal rights and foreclosure processes do have an effect on the default timeline. For example, judicial foreclosures consistently take longer than power of sale foreclosures. This is true for reo and non-reo timelines. When the border sample and matched border sample is used to enhance identification, the results are only modestly affected for the overall timelines. However, for the pre-reo time period, the use of the border discontinuity increases the judicial coefficient substantially (0.061 to 0.302), so that it becomes statistically significant. The matching scheme also flips the sign of the judicial coefficient for the in-reo part of the default timeline from negative to positive. Regardless, the results consistently indicate that judicial foreclosure extended the default timeline, with the largest impact in the pre-reo time period.

The role of the statutory right of redemption is more mixed and much smaller in magnitude. Once the lender/investor has taken the property (in-reo), the timeline is longer. This is consistent with the statutory right of redemption, which clouds the title after the taking. When the redemption period ends, most lenders will then try to sell and recover any losses. Defaulted loans in states with the right of recourse tend to have shorter timelines overall. This is especially true in the pre-reo time period. However, perhaps due to the speed of the taking, the in-reo timeline is longer.

In summary, legal rights and different methods of foreclosure processing do not have an impact on loss on sale of the property, but do have potentially important impacts on the holding period for a loan. Judicial foreclosures take 28 to 33 percent longer (e^{β}) to transition from the last payment to property sale and recovery of losses. In contrast, loans in state where the lender has the right to seek recovery of losses from outside the property itself have slightly faster default timelines (6.5 to 4.3 percent shorter).

Holding Costs and Calculating Total Cost

In light of the different impacts of economic, financial, and legal factors on the loss on sale and the default timeline, it is also necessary to estimate the total cost and the total loss associated with the default. This requires using the estimated models of the loss on sale and the estimated models of the default timeline. In addition, we need a mechanism to convert the predicted default timeline into holding and carrying costs.

Cordel et al (2013) and Cutts and Merrill (2008) provide some guidance. Cutts and Merrill (2008) provide summary statistics indicating that the majority of losses are holding costs. These costs can be broken down into groupings such as maintenance, legal expenses, insurance and other costs. Cordel et al (2013) focus on taxes, insurance, and excess depreciation. The estimated costs are not collected from the loans themselves but imputed from the average state level of property taxes as reported by the Tax Foundation. However, property taxes are levied by the local municipality and can have substantial variation within a state, and even within metropolitan areas.

Our approach to calculate holding costs uses loan level and directly observed holding cost data. The holding expenses are categorized into legal, maintenance, taxes and insurance, and miscellaneous. The summary statistics are included in the Appendix. Table 14 provides the descriptive regression results designed to convert timelines into dollars of holding costs. Four different variables are included to explain holding expenses. The first variable is used to control for scale. Larger or more expensive housing should have larger holding costs. The estimated value of the property (updated using a 3-digit repeat sales price index form FHFA) is used to capture this effect. Holding costs should also increase as the holding period (the default timeline) grows longer. Holding costs are likely to vary over time and across the foreclosure crisis. Therefore, the year that the loan ends is also included. State level fixed effects are also included to capture state level foreclosure laws and borrower rights which will likely change the level of maintenance the defaulted borrower engages in and the requirements for the use of legal counsel during the holding period and through the sale.

Specification I explains the legal costs. A one thousand dollar increase in the value of the property increases legal holding costs by $1.47. As anticipated, maintenance costs, taxes, and insurance go up even more in response to property value. A one thousand dollar increase in value is associated with an $16.25 increase in maintenance costs and $25.49 increase in taxes and insurance. The length of the holding period also drives up all holding costs. Maintenance and taxes and insurance (specifications II and III) are more sensitive to the length of the holding period. Each month increases holding costs by over 46 dollars for legal expenses, over 141 dollars for maintenance expenses, and over 327 dollars in legal expenses.

Table 15 reports the predicted losses as a percentage of the unpaid balance for a representative loan (retail, purchase, estimated fixed effect for a loan originated in 2006, estimated fixed effect for a loan last payment date 2010, and everything else at the mean value).[10] Costs are estimated for the loss on sale, the holding costs, and carrying costs. We estimate carrying costs in two ways. The first approach (labeled Carrying Cost I) follows An and Cordell (2017) and uses the lost interest for the mortgage for the first 120 days in the timeline and the 3-month constant maturity treasury yield (Freddie Mac in in receivership) for the

[10] The predicted loss is estimated using the loss on sale coefficient estimates from table 2 specification V, table 5 specification II, and table 13.

remainder of the timeline. The second approach (labeled Carrying Cost II) uses the mortgage interest rate for the whole timeline. Approach I can be viewed as a lower bound of carrying costs and approach II as an upper bound.

For the representative loan, between 42 and 48 percent of the unpaid balance is lost (depending on the carrying costs approach). 32.7 percent of the losses come from the sale, 7.9 percent come from the holding costs, and between 1.8 and 7.8 percent from carrying costs. The holding costs are very evenly spread across legal, maintenance, tax, and insurance costs. Table 15 also reports expected losses for a variety of current loan to value ratios, holding all other variables constant. As expected, overall losses increase as LTV increases. This overall change is driven by larger losses associated with the sale of the property, even though holding costs decline for maintenance, taxes and insurance costs. The holding costs decline because lower equity is associated with shorter default timelines. Holding costs are very stable as a percentage of unpaid balance for the variety of equity positions (ltvs).

Tables 16 repeats the exercise for different unpaid balances. Table 16 shows that larger loans have substantially lower overall losses. Losses associated with holding and carrying costs increase as loans grow larger, because the default timeline is longer for larger loans. In contrast, larger loans have lower losses associated with the sale. The changes in the losses on the sale dominate the changes in the holding and carrying costs, thus driving down over loss percentages. Table 17 repeats the process for a variety of credit scores.

Conclusion and Discussion

This paper examines losses on defaulted mortgages. Our approach can help lenders and their regulators design more effective loss models that are crucial for estimating economic and regulatory capital requirements. This is especially important given the scheduled implementation of the Current Expected Credit Loss (CECL) impairment standard by 2020 for SEC registered banks and all other banks by 2021. CECL requires a modeled estimate of expected loan level losses for all loans, not just impaired loans. This paper provides a description and an example of how to model losses for mortgages (selection, loss on sale, and holding and carrying costs). Descriptive statistics show that the loss on the sale and the amount of time that a loan is in default before resolution have substantial variation across space (state and metropolitan areas) and over time (during the housing run-up, collapse and recovery). However, losses associated with the sale of the property have stayed elevated after the recession.

When considering mortgage losses, it is important to consider both the loss on the sale and the cost of holding a loan through the default and foreclosure process (the default timeline). Some key factors that drive losses have substantially different effects on the losses associated with the sale than on holding costs. For example, low and negative equity aggressively drives up losses on sale but shortens the default timeline and reduces holding costs. Holding costs (legal, maintenance, and taxes and insurance expenses) vary with the length of the holding period (default timeline), the value of the property, location, and cohort factors. Carrying costs (lost interest and opportunity costs) vary in proportion to the length of the holding period and prevailing interest rates.

Our results show no evidence that the length of the holding period affects losses from sale of the property. Instead, longer holding periods drive up expenses associated with keeping a clean title on the property (paying property taxes, for example) and keeping the property reasonable well maintained. Even if borrowers in default are less likely to maintain the property, corrective maintenance expenses incurred by the lender/investor (Freddie Mac) may help explain why longer default timelines are not associated with larger losses (smaller recovery) from the sale of the property. The magnitude of these holding expenses is nontrivial but much smaller than typical losses associated with the sale of the property. Our results indicate that each month of the default timeline is associated with more than 140 dollars in maintenance expenses and more than 325 dollars in taxes and insurance expenses. Since the observed average default timeline is 18.74 months, this translates into more than 8,700 dollars in time varying expenses just for maintenance, insurance, and taxes. Legal expenses, as well as other miscellaneous and fixed components of holding expenses, can drive the cost up even more.

The legal rights of borrowers and lenders have been studied extensively in the mortgage market. However, much less attention has been paid to the impact of these rights on mortgage losses. This paper improves the typical identification strategy by limiting the sample to loans in 3-digit zip codes along state borders with different laws in place, and then matching loans across those borders so that they are as similar as possible. After applying this approach, we find no statistical evidence that judicial foreclosures, rights of redemption, or rights of recourse have any impact on losses associated with the sale of the property. However, the judicial foreclosure process does increase the default timeline by 33 percent, while rights of recourse speed the taking process by about 4.4 percent. These lengthened timelines can have nontrivial impacts on holding and carrying costs.

In summary, this paper provides strong evidence that any measure of losses for mortgages must account for the holding and carrying cost associated with the default timeline as well as the losses associated with selling the property. These two components react differently to the legal rights of the borrower and lender and a variety of other drivers of losses.

References

An, Xudong and Larry Cordell. (2017). Regime Shift and the Post-Crisis World of Mortgage Loss Severities. Working Papers of the Federal Reserve Bank of Philadelphia No. 17-08. Downloaded June 1, 2017 from https://www.philadelphiafed.org/-/media/research-and-data/publications/working-papers/2017/wp17-08.pdf.

Calem, Paul and Michael LaCour-Little. (2004). Risk-based Capital Requirements for Mortgage Loans. *Journal of Banking and Finance* 28(3): 647-672.

Campbell, John, Stefano Giglio and Parag Pathak. (2011). Forced Sales and House Prices. American Economic Review, 101: 2108. doi: 10.1257/aer.101.5.2108.

Chan, Sewin, Andrew Haughwout, Andrew Hayahsi and Wilbert vam der Klaauw. (2015). Determinants of Mortgage Default and Consumer Use: The Effects of Foreclosure Laws and Foreclosure Delays. *Journal of Money, Credit and Banking* forthcoming.

Clauretie, Terrence and Thomas Herzog. (1990). The Effect of State Foreclosure Laws on Loan Losses: Evidence from the Mortgage Insurance Industry. *Journal of Money, Credit and Banking*, 22(2): 221-233. http://www.jstor.org/stable/1992309.

Collins, J. Michael, Ken Lam and Christopher Herbert. (2011). State Foreclosure Policies and Lender Interventions; Impacts on Borrower Behavior in Default. *Journal of Policy Analysis and Management*, 30(2): 216-232. doi: 10.1002/pam.20559.

Cordell, Larry, Liang Geng, Laurie Goodman and Lidan Yang. (2015). The Cost of Foreclosure Delay. *Real Estate Economics*, 43(4): 916-956. doi: 10.1111/1540-6229.12107

Crawford, Gordon and Eric Rosenblatt. (1995). Efficient Mortgage Default Option Exercise: Evidence from Loss Severity. *Journal of Real Estate Research*, 10(5): 543-555.

Cutts, Amy C. and William A. Merrill. (2008). Interventions in Mortgage Default: Policies and Practices to Prevent Home Loss and Lower Costs. Borrowing to Live: Consumer and Mortgage Credit Revisited, eds. N. P. Retsinas and E. S. Belsky, Brookings Institution Press, Washington, DC.

Demiroglu, Cem, Evan Dudley and Christopher James. (2014). State Foreclosure Laws and the Incidence of Mortgage Default. *Journal of Law and Economics*, 57(1) 225-280.

Ghent, Andra and Marianna Kudlyak. (2011). Recourse and Residential Mortgage Default: Evidence from US States. *The Review of Financial Studies*, 24(9): 3139-3186. doi: 10.1093/rfs/hhr055.

Goodwin, Kimberly and Ken Johnson. (2017). The Short Sale Stigma. *Journal of Real Estate and Finance and Economics*, 55:416-434. doi: 10.1007/s11146-016-9571-2.

Gerardi, Kristopher, Lauren Lambie-Hanson and Paul Willen. (2013). Do Borrower Rights Improve Borrower Outcomes? Evidence from the Foreclosure Process. *Journal of Urban Economics*, 73:1-17. doi:10.1016/j.jue.2012.06.003.

Harding, John, Thomas Miceli and C. F. Sirmans. (2000). Do Owners Take Better Care of their Housing than Renters. *Real Estate Economics*, 28(4) 663–81.

Harrison, David and Michael Seiler. (2015). The Paradox of Judicial Foreclosure: Collateral Value Uncertainty and Mortgage Rates. Journal of Real Estate Finance and Economics, 50(3):377-411. doi:10.1007/s11146-014-9467-y.

Inside Mortgage Finance. (2010). Federal Agencies Dig Into Foreclosure Processing Problems, October 28, 2010, http://www.insidemortgagefinance.com/issues/imfpubs_imf/27_41/news/Federal_Agencies_Dig_Into_Foreclosure_Processing_Problems-1000015009-1.html.

Jaing, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil. (2014). Liar's Loan? Effects of Origination Channel and Information Falsification on Mortgage Delinquency. *Review of Economics and Statistics*, 96(1): 1-18. doi:10.1162/REST_a_00387.

Park, Yun and Doo Won Bang. (2014). Loss given default of residential mortgages in a low LTV regime: Role of foreclosure auction process and housing market cycles. *Journal of Banking and Finance*, 39: 192-210. http://dx.doi.org/10.1016/j.jbankfin.2013.11.015.

Pennington-Cross, Anthony. (2003). Subprime and Prime Mortgages: Loss Distributions, Working Paper 03-1. Office of Federal Housing Enterprise Oversight. http://www.fhfa.gov/PolicyProgramsResearch/Research/PaperDocuments/2003-05_WorkingPaper_03-1_N508.pdf.

Pennington-Cross, Anthony. (2006). The Value of Foreclosed Property. *Journal of Real Estate Research*, 28(2): 193-214.

Pennington-Cross, Anthony and Souphala Chomsisengphet. (2007). Subprime Refinancing: Equity Extraction and Mortgage Termination. *Real Estate Economics*, 35(2): 233–263. doi: 10.1111/j.1540-6229.2007.00189.x.

Qi, Min and Xiaolong Yang. (2009). Loss given default of high loan-to-value residential mortgages. *Journal of Banking and Finance*, 33(5): 788-799. http://dx.doi.org/10.1016/j.jbankfin.2008.09.010.

Reuters.com. (2008). Fannie Mae, Freddie Mac Suspend Some Foreclosures, November 21, 2008, http://www.reuters.com/article/2008/11/21/us-fannie-freddie-idUSTRE4AJ90520081121.

Wall Street Journal. (2009). Banks Agree to Foreclosure Moratorium, February 14 2008, http://www.wsj.com/articles/SB123454524404184109.

Lekkas, Vassilis, John Quigley, Robert Van Order. (1993). Loan Loss Severity and Optimal Mortgage Default. *Real Estate Economics*, 21(4): 353–371. doi: 10.1111/1540-6229.00616.

Zhang, Yanan, Lu Ji and Fei Liu. (2010). Local Housing Market Cycle and Loss Given Default: Evidence from Sub-Prime Residential Mortgages. *IMF Working Paper*, WP/10/167 July.

Table 1: Summary Statistics and Description of Variables

Variable Name	Description	Mean	Standard Deviation
los_i	Loss on sale = 100*(unpaid balance - default sale price) / unpaid balance at loan resolution date.	38.70	25.24
$month\_ltz_i$	The timeline: Months from last payment date to loan resolution (end of loan and zero balance date).	18.74	9.60
ltv_i	100*(loan amount /house value) at loan resolution date. House value is estimated by updating the value from the origination date to the resolution date using the 3-digit zip code repeat price index reported by the Federal Housing Finance Agency.	99.78	30.53
upb_i	Unpaid balance in $1,000 at the loan resolution date.	165.33	85.65
$fico_i$	Fico score at origination divided by 100.	6.92	0.54
$first_i$	First time homebuyer indicator at origination.	0.10	0.30
$\Delta urate_i$	Change in county unemployment rate from origination to loan resolution as reported by the Bureau of Labor and Statistics. Positive indicates an increase in the rate.	3.99	2.76
$\%\Delta hp_i$	The percentage change in 3-digit zip code house prices from the last payment date to loan resolution date.	-5.74	9.40
$retail_i$	Loan originated through a retail channel indicator.	0.34	0.47
$cashrefi_i$	An indicator that the loan was originated as a refinance that extracted equity and took cash out.	0.35	0.48
$nocashrefi_i$	An indicator that the loan was originated as a refinance that did not take any cash out.	0.25	0.43
$cashrefi_i$	An indicator that the loan was originated as a refinance that extracted equity and took cash out.	0.35	0.48
$nocashrefi_i$	An indicator that the loan was originated as a refinance that did not take any cash out.	0.25	0.43
$judicial_i$	Judicial foreclose process indicator.	0.44	0.50
srr_i	Statutory right of redemption indicator.	0.65	0.48
$recourse_i$	Recourse indicator.	0.65	0.48
Loans that are real estate owned (reo) at some point during the default timeline			
los_i	Loss on sale = 100*(unpaid balance - default sale price) / unpaid balance.	38.25	27.32
$month\_ltz_i$	The timeline: Months from last payment date to loan resolution (end of loan and zero balance date).	20.79	9.11
$month\_ltr_i$	The pre-reo timeline: Months from last payment date to the beginning of the property becoming reo.	14.32	8.33
$month\_rtz_t$	The reo timeline: Months from beginning of reo to loan resolution (end of loan and zero balance date).	6.46	3.96

207,162 loans are included in the sample. 140,305 loans become real estate owned (reo) at some point during the default timeline. Source: All variables are collected from the Freddie Mac Single Family Loan Level Dataset, except $\Delta urate_{it}$ is collected from the Bureau of Labor and Statistics, and judicial, srr, and recourse are collected from Cutts and Merrill (2008) and Ghent and Kudlyak (2011).

Table 2: Loss on Sale (*los*$_{it}$) Results

	I: Balance		II: Borrower		III: Loan		IV: Year		V: Servicer & State	
Variable	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
ltv_{it}	0.48***	0.03	0.44***	0.03	0.45***	0.03	0.46***	0.03	0.50***	0.02
upb_{it}	-0.12***	0.01	-0.12***	0.01	-0.12***	0.01	-0.13***	0.01	-0.11***	0.00
$fico_i$			-1.66***	0.18	-0.92***	0.18	-1.19***	0.17	-0.88***	0.16
$first_i$			-6.24***	0.37	-1.38***	0.24	-1.65***	0.23	-1.55***	0.21
$\Delta urate_{it}$			1.07***	0.11	1.00***	0.11	0.30	0.15	0.70***	0.13
$retail_i$					-2.94***	0.23	-2.93***	0.23	-3.04***	0.22
$cashrefi_i$					7.82***	0.43	7.69***	0.46	7.82***	0.36
$nocashrefi_i$					6.26***	0.45	6.26***	0.42	5.81***	0.31
fixed effects:										
origination year							x		x	
last payment year							x		x	
servicer									x	
state									x	
constant	9.97***	2.19	22.73***	2.38	12.99***	2.59	2.37	8.76	9.21	8.59
R^2	0.28		0.30		0.31		0.35		0.38	
N	207,162		207,162		207,162		207,162		207,162	

*, **, and *** indicate that the coefficient is significant at the 10, 5, or 1 percent level. The standard errors are clustered at the three-digit zip code level. The Appendix contains the estimated fixed effects for specification V.

Table 3: Loss on Sale (los_{it}) Robustness

	I: State-Year Interactions		**II: Dollars of Losses (1,000's)**		**III: Purchase Only**		**IV: Selection Correction**	
Variable	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
ltv_{it}	0.50***	0.02	0.95***	0.03	0.49***	0.02	0.48***	0.02
upb_{it}	-0.11***	0.00	0.19***	0.01	-0.11***	0.00	-0.11***	0.00
$fico_i$	-0.88***	0.15	-0.94***	0.18	-0.14	0.17	0.41	0.31
$first_i$	-1.46***	0.21	-1.15***	0.26	-1.06***	0.21	-1.66***	0.21
$\Delta urate_{it}$	0.70***	0.13	0.97***	0.20	0.87***	0.15	0.77***	0.13
$retail_i$	-3.06***	0.21	-3.46***	0.23	-1.73***	0.28	-2.91***	0.22
$cashrefi_i$	7.75***	0.35	13.23***	0.54			8.41***	0.34
$nocashrefi_i$	5.73***	0.30	10.81***	0.49			6.25***	0.30
Inverse $Mills_i$							-4.09***	0.66
fixed effects:								
origination year	x		x		x		x	
last payment year	x		x		x		x	
last payment year *state	x							
state	x		x		x		x	
servicer	x		x		x		x	
constant	-21.06	5.99	-86.05	10.89	-8.34	10.53	-4.23	5.10
R^2	0.39		0.64		0.40		0.38	
N	207,162		207,162		82,574		202,640	

*, **, and *** indicate that the coefficient is significant at the 10, 5, or 1 percent level. The standard errors are clustered at the three-digit zip code level. The two-step process is used to correct for potential selection issues. The first stage probit selection results are in the appendix.

Table 4: Loss on Sale (*los_{it}*) Robustness Truncation Tests

	I: Minimum 3 Year Default Timeline		**II: Minimum 4 Year Default Timeline**	
Variable	Coeff.	SE	Coeff.	SE
ltv_{it}	0.50***	0.02	0.51***	0.02
upb_{it}	-0.12***	0.00	-0.12***	0.00
$fico_i$	-0.81***	0.17	-0.74***	0.19
$first_i$	-1.67***	0.22	-2.00***	0.25
$\Delta urate_{it}$	0.71***	0.13	0.77***	0.12
$retail_i$	-3.13***	0.23	-3.47***	0.25
$cashrefi_i$	7.96***	0.37	7.95***	0.39
$nocashrefi_i$	5.66***	0.32	5.57***	0.36
fixed effects:				
origination year	x		x	
last payment year	x		x	
state	x		x	
servicer	x		x	
constant	-10.73	8.37	-10.72	7.93
R^2	0.39		0.39	
N	171,181		123,068	

*, **, and *** indicate that the coefficient is significant at the 10, 5, or 1 percent level. The standard errors are clustered at the three-digit zip code level.

Table 5: Default Timeline Basic Results (log of months)

	I. No reo defaults: Full timeline		II. Reo defaults: Full timeline		III. Reo defaults: Pre-reo timeline		IV. Reo-defaults: Reo timeline	
Variable	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
$lnltv_{it}$	-0.129***	0.027	-0.096***	0.014	-0.052***	0.017	-0.183***	0.018
$lnupb_{it}$	0.020***	0.007	0.096***	0.004	0.076***	0.004	0.147***	0.007
$lnfico_{i}$	-0.993***	0.040	-0.323***	0.015	-0.302***	0.017	-0.339***	0.021
$first_{i}$	0.006	0.009	0.012***	0.003	0.012***	0.004	0.007	0.006
$\Delta urate_{it}$	0.007***	0.002	0.006***	0.002	0.011***	0.002	-0.005**	0.002
$retail_{i}$	-0.048***	0.005	-0.019***	0.002	-0.024***	0.003	-0.010***	0.003
$cashrefi_{i}$	0.059***	0.005	0.021***	0.004	0.028***	0.004	-0.001	0.005
$nocashrefi_{i}$	0.018***	0.006	0.016***	0.003	0.023***	0.004	0.000	0.005
fixed effects:								
origination year	x		x		x		x	
last payment year	x		x		x		x	
servicer	x		x		x		x	
state	x		x		x		x	
constant	5.065***	0.239	3.349***	0.106	2.611	0.131	2.618	0.183
R^2	0.22		0.33		0.39		0.18	
N	66,857		140,305		140,305		140,305	

Left hand side variable is the log of months from last the last payment date. *, **, and *** indicate that the coefficient is significant at the 10, 5, or 1 percent level. The standard errors are clustered at the three-digit zip code level.

Table 6: Default Timeline Results (log of months) Truncation Tests

	No reo default				Reo default			
	I: Minimum 3 Year Default Timeline		II: Minimum 4 Year Default Timeline		III: Minimum 3 Year Default Timeline		IV: Minimum 4 Year Default Timeline	
Variable	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
$lnltv_{it}$	-0.159***	0.030	-0.202***	0.035	-0.101***	0.016	-0.109***	0.019
$lnupb_{it}$	0.033***	0.008	0.040***	0.010	0.099***	0.004	0.102***	0.004
$lnfico_{i}$	-1.036***	0.045	-0.950***	0.054	-0.347***	0.017	-0.356***	0.020
$first_{i}$	0.015	0.010	0.018	0.013	0.016***	0.004	0.018***	0.004
$\Delta urate_{it}$	0.010***	0.003	0.013***	0.003	0.006***	0.002	0.008***	0.002
$retail_{i}$	-0.064***	0.006	-0.064***	0.008	-0.022***	0.003	-0.018***	0.003
$cashrefi_{i}$	0.060***	0.007	0.040***	0.008	0.021***	0.004	0.019***	0.004
$nocashrefi_{i}$	0.019**	0.008	0.003	0.010	0.018**	0.003	0.019***	0.004
fixed effects:								
origination year	x		x		x		x	
last payment year	x		x		x		x	
servicer	x		x		x		x	
state	x		x		x		x	
constant	5.210***	0.268	5.125***	0.284	3.382***	0.116	3.434***	0.124
R^2	0.23		0.27		0.33		0.32	
N	48,230		28,598		122,951		94,470	

Left hand side variable is the log of months from last the last payment date. *, **, and *** indicate that the coefficient is significant at the 10, 5, or 1 percent level. The standard errors are clustered at the three-digit zip code level.

Table 7: Default Timeline Results (log of months) Self Selection Correction

	I. No reo defaults: Full timeline		II. Reo defaults: Full timeline		III. Reo defaults: Pre-reo timeline		IV. Reo-defaults: Reo timeline	
Variable	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
$lnltv_{it}$	-0.167***	0.032	-0.113***	0.019	-0.046**	0.023	-0.255***	0.021
$lnupb_{it}$	0.026***	0.007	0.098***	0.004	0.076***	0.004	0.156***	0.006
$lnfico_{i}$	-0.860***	0.049	-0.282***	0.022	-0.325***	0.026	-0.139***	0.031
$first_{i}$	0.004	0.008	0.011***	0.003	0.013***	0.004	0.003	0.006
$\Delta urate_{it}$	0.009***	0.002	0.007***	0.002	0.010***	0.002	-0.003	0.002
$retail_{i}$	-0.047***	0.005	-0.018***	0.003	-0.024***	0.003	-0.007*	0.003
$cashrefi_{i}$	0.068***	0.006	0.023***	0.003	0.027***	0.004	0.011*	0.005
$nocashrefi_{i}$	0.023***	0.006	0.018***	0.003	0.023***	0.004	0.008*	0.004
$InvMills_{i}$	-0.061***	0.016	-0.020**	0.008	0.009	0.010	-0.094***	0.010
fixed effects:								
origination year	x		x		x		x	
last payment year	x		x		x		x	
servicer	x		x		x		x	
state	x		x		x		x	
constant	4.968***	0.253	2.712***	0.115	2.080***	0.139	1.689***	0.194
R^2	0.22		0.33		0.39		0.18	
N	65,345		137,295		137,295		137,280	

Left hand side variable is the log of months from last the last payment date. *, **, and *** indicate that the coefficient is significant at the 10, 5, or 1 percent level. The standard errors are clustered at the three-digit zip code level. The two-step process is used to correct for potential selection issues. The first stage probit selection results are in the appendix.

Table 8: Default Timeline (log of months) Changing House Prices for the Full Timeline

	I. Reo defaults		II. No reo defaults		III. Reo defaults		IV. No reo defaults	
Variable	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
$lnltv_{it}$	-0.056***	0.015	0.002*	0.025	-0.085***	0.015	-0.095***	0.026
$lnupb_{it}$	0.085***	0.004	0.007	0.006	0.089***	0.004	0.021***	0.006
$lnfico_{i}$	-0.315***	0.015	-0.924***	0.035	-0.301***	0.014	-0.838***	0.036
$first_{i}$	0.012***	0.003	0.002	0.008	0.013***	0.003	0.008	0.008
$\Delta urate_{it}$	0.006***	0.002	0.011***	0.003	0.007***	0.002	0.015***	0.003
$\%\Delta hp_{it}$	-0.006***	0.001	-0.021***	0.002				
$\%\Delta hp_{it}$<= -10%					-0.054***	0.020	-0.089***	0.027
$-10\%<\%\Delta hp_{it}<=0\%$					-0.244***	0.015	-0.562***	0.030
$0\%<\Delta hp_{it}<=5\%$					-0.215***	0.010	-0.548***	0.022
$retail_{i}$	-0.019***	0.002	-0.045***	0.005	-0.020***	0.002	-0.046***	0.005
$cashrefi_{i}$	0.024***	0.003	0.070***	0.005	0.020***	0.003	0.053***	0.005
$nocashrefi_{i}$	0.021***	0.003	0.030***	0.006	0.016***	0.003	0.020***	0.006
fixed effects:								
origination year	x		x		x		x	
last payment year	x		x		x		x	
servicer	x		x		x		x	
state	x		x		x		x	
constant	3.301***	0.105	4.623***	0.272	3.362***	0.104	4.868	0.167
R^2	0.34		0.25		0.36		0.32	
N	140,305		66,857		140,305		66,857	

Left hand side variable is the log of months from last the last payment date. *, **, and *** indicate that the coefficient is significant at the 10, 5, or 1 percent level. The standard errors are clustered at the three-digit zip code level. The excluded category for %Δhp is house price growth greater than 5 percent. %Δhp is measured from the last payment date to the resolution of the loan.

Table 9: Instrument Variables – Loss on Sale (los_{it})

	I: OLS		II: Contemporaneous IV		III: One Year Lag IV		IV: Two Year Lag IV	
Variable	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
$months_{it}$	0.01	0.02	0.03	0.04	-0.01	0.04	-0.06	0.05
ltv_{it}	0.50***	0.02	0.95***	0.03	0.49***	0.02	0.49***	0.02
upb_{it}	-0.11***	0.00	0.19***	0.01	-0.11***	0.00	-0.11***	0.00
$fico_i$	-0.85***	0.15	-0.94***	0.18	-0.14	0.17	-0.42	0.31
$first_i$	-1.55***	0.21	-1.15***	0.26	-1.06***	0.21	-1.53***	0.21
$\Delta urate_{it}$	0.72***	0.14	0.97***	0.20	0.87***	0.15	0.73***	0.13
$retail_i$	-3.03***	0.21	-3.46***	0.23	-1.73***	0.28	-3.01***	0.23
$cashrefi_i$	7.80***	0.35	13.23***	0.54			8.02***	0.34
$nocashrefi_i$	5.80***	0.30	10.81***	0.49			5.99***	0.30
fixed effects:								
origination year	x		x		x		x	
last payment year	x		x		x		x	
state	x		x		x		x	
servicer	x		x		x		x	
constant	-9.54	8.58	-9.83	8.59	-2.35	12.62	14.18**	4.66
R^2	0.38		0.38		0.38		0.38	
N	207,162		207,162		206,773		204,960	
Endogenous F-Stat	0.16		0.42		0.42		0.18	
First Stage Robust Partial R2	0.37		0.31		0.31		0.19	

*, **, and *** indicate that the coefficient is significant at the 10, 5, or 1 percent level. The instrument is the average number of months from the last payment date to the end of the loan's life for the 3-digit zip code the property is located in. Specification II uses the contemporaneous zip code months. Specification III uses one-year lagged months. Specification IV uses 2-year lagged months. The standard errors are clustered at the three-digit zip code level.

Table 10: Borrower and Lender Rights – Loss on Sale (*los_{it}*)

Variable	Coeff.	SE
ltv_{it}	0.46***	0.03
upb_{it}	-0.12***	0.00
$fico_i$	-1.02***	0.17
$first_i$	-1.37***	0.22
$\Delta urate_{it}$	0.37*	0.14
$retail_i$	-2.95***	0.23
$cashrefi_i$	7.67***	0.47
$nocashrefi_i$	6.29***	0.42
judicial	3.30**	1.46
srr	1.66	1.27
recourse	0.75	1.17
constant	-2.53	8.54
fixed effects:		
origination year	x	
last payment year	x	
servicer	x	
R^2	0.35	
N	207,162	

*, **, and *** indicate that the coefficient is significant at the 10, 5, or 1 percent level. The standard errors are clustered at the three-digit zip code level.

Table 11: Borrower and Lender Rights – Default Timeline (log of months)

	I. No reo defaults: complete timeline		II. Reo defaults: complete timeline	
Variable	Coeff.	SE	Coeff.	SE
$lnltv_{it}$	-0.089**	0.039	-0.091***	0.025
$lnupb_{it}$	0.051***	0.008	0.108***	0.006
$lnfico_{i}$	-0.967***	0.039	-0.303***	0.020
$first_{i}$	0.009	0.009	0.013***	0.004
$\Delta urate_{it}$	0.036***	0.003	0.009***	0.003
$retail_{i}$	-0.048***	0.005	-0.010**	0.003
$cashrefi_{i}$	0.069***	0.006	0.032***	0.004
$nocashrefi_{i}$	0.017***	0.007	0.018***	0.005
$judicial_{i}$	0.190***	0.031	0.351***	0.017
srr_{i}	-0.133***	0.023	-0.005	0.015
$recourse_{i}$	-0.071***	0.025	-0.040**	0.020
constant	4.758***	0.278	3.180***	0.145
fixed effects:				
origination year	x		x	
last payment year	x		x	
servicer	x		x	
R^2	0.54		0.27	
N	66,857		140,305	

Left hand side is the log of months from the last payment date to the end of the loans life. *, **, and *** indicate that the coefficient is significant at the 10, 5, or 1 percent level. The standard errors are clustered at the three-digit zip code level.

Table 12: Borrower and Lender Rights – Loss on Sale (*los_{it}*) – Border and Matching

	I. Complete Sample		**II. Border Samples**		**III. Matched Samples on the Borders**	
Variable	Coeff.	SE	Coeff.	SE	Coeff.	SE
$judicial_i$	3.30**	1.46	-1.73	2.23	-0.71	2.32
ssr_i	1.66	1.27	0.46	0.91	-0.04	1.41
$recourse_i$	0.75	1.17	-0.49	1.07	-1.20	1.02

Left hand side is the loss on sale. *, **, and *** indicate that the coefficient is significant at the 10, 5, or 1 percent level. The standard errors are clustered at the three-digit zip code level. Specification I uses the coefficients reported in table 7. For the remaining results, each coefficient is estimated in a different regression. In addition to the standard variables, dummy variables for each state border are included. The "Border Sample" for the judicial test only includes loans in 3-digit zip codes along state borders where the neighboring state has a different law. Following the same logic, a different sample is constructed for border zip codes for the statutory tight of redemption and recourse. The matched samples are created by matching loans based on unpaid balance, current ltv and last payment year. This matched sample is constructed within the 3-digit zip codes along each specific state border .

Table 13: Borrower and Lender Rights – Default Timeline (log of months) – Border and Matching

	I. Complete Sample		**II. Border Samples**		**III. Matched Samples on the Borders**	
Variable	Coeff.	SE	Coeff.	SE	Coeff.	SE
No Reo: Complete Timeline						
$judicial_i$	0.190***	0.031	0.233***	0.034	0.249***	0.026
srr_i	-0.133***	0.023	-0.015	0.029	0.022	0.034
$recourse_i$	-0.071***	0.025	-0.072	0.053	-0.067*	0.037
Reo: Complete Timeline						
$judicial_i$	0.351***	0.017	0.294***	0.029	0.286***	0.025
srr_i	-0.005	0.015	0.010	0.018	0.046*	0.024
$recourse_i$	-0.040**	0.020	-0.047*	0.025	-0.044**	0.021
Reo: Pre-reo Timeline						
$judicial_i$	0.061	0.061	0.302***	0.027	0.386***	0.027
srr_i	-0.325***	0.052	-0.117**	0.051	-0.052*	0.031
$recourse_i$	-0.174***	0.032	-0.103***	0.031	-0.114***	0.026
Reo: In-reo Timeline						
$judicial_i$	-0.294***	0.071	-0.260***	0.017	0.136***	0.042
srr_i	-0.384***	0.062	-0.137***	0.033	0.236***	0.032
$recourse_i$	0.297***	0.064	0.047*	0.026	0.073**	0.030

Left hand side is the loss on sale. *, **, and *** indicate that the coefficient is significant at the 10, 5, or 1 percent level. The standard errors are clustered at the three-digit zip code level. Specifications follow the set up as shown in table 8. In specification I all three legal coefficients are jointly estimated. For the remaining results, each coefficient is estimated in a different regression. In addition to the standard variables, dummy variables for each state border are included. For example, the "Border Sample" for the judicial test only includes loans in 3-digit zip codes along state borders where the neighboring state has a different law. Following the same logic, a different sample is constructed for border zip codes for the statutory tight of redemption and recourse. The matched samples are created by matching loans based on unpaid balance, current ltv and last payment year. This matched sample is constructed within the zip codes along each specific state border with a law change.

Table 14: OLS Description of Legal, Maintenance, Taxes and Insurance, and Miscellaneous Expenses

	I. Legal		II. Maintenance		III. Taxes and Insurance		IV. Miscellaneous Expenses	
Variable	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
house value$_{it}$	1.47	0.04	16.25	0.18	25.49	0.13	0.34	0.03
month_ltz$_{it}$	46.35	0.46	141.31	1.87	327.96	1.38	14.58	0.35
zero balance year								
2006	231.76	26.89	-31.99	109.60	59.71	80.93	-41.93	20.41
2007	355.50	26.41	-581.79	107.63	34.95	79.48	-112.57	20.05
2008	579.30	23.88	-1,638.82	97.31	150.30	71.86	-184.14	18.12
2009	806.27	21.76	-1,400.67	88.69	609.22	65.49	-188.49	16.52
2010	978.16	20.67	-915.26	84.23	726.03	62.20	-167.09	15.69
2011	1,109.96	20.47	768.82	83.41	825.51	61.60	-152.45	15.54
2012	1,270.87	20.70	2,220.34	84.36	1,037.38	62.29	-118.30	15.71
2013	913.45	21.31	4,900.31	86.86	745.33	64.14	408.02	16.18
state	x		x		x		x	
constant	1,486.47	114.52	-973.69	466.71	-7,609.06	344.64	205.99	86.93
R^2	0.42		0.24		0.50		0.07	
N	184,033		184,033		184,033		184,033	

Left hand side variable is dollars of expenses. House value is the estimated value of the property when the loan's life is ended in 1,000 of dollars. Month_ltz is the number of months from the last payment date until the loan is fully terminated (zero balance date). Dummy variables for the end of the loan's life are included (2005 is the excluded year) as well as state level fixed effects. Information on expenses is only available from 2005 through 2014. Estimates use Ordinal Least Squares (OLS).

Table 15: Predicted Losses (Percentage of UPB) by LTV

	Representative loan	LTV 60	80	100	120	140
Total I	42.3%	22.7%	32.5%	42.4%	52.4%	62.3%
Total II	48.3%	28.7%	38.5%	48.3%	58.2%	68.0%
Loss on sale	32.7%	12.7%	22.7%	32.8%	42.8%	52.8%
Holding Cost	7.9%	8.2%	8.0%	7.9%	7.7%	7.7%
Legal	2.4%	2.4%	2.4%	2.4%	2.4%	2.4%
Maintenance	2.3%	2.4%	2.3%	2.3%	2.3%	2.3%
Tax and Insurance	2.9%	3.1%	2.9%	2.9%	2.8%	2.7%
Miscellaneous	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%
Carrying Cost						
I: Mortgage & Treasury	1.8%	1.8%	1.8%	1.8%	1.8%	1.8%
II: Mortgage Rate	7.8%	7.8%	7.7%	7.6%	7.5%	7.5%

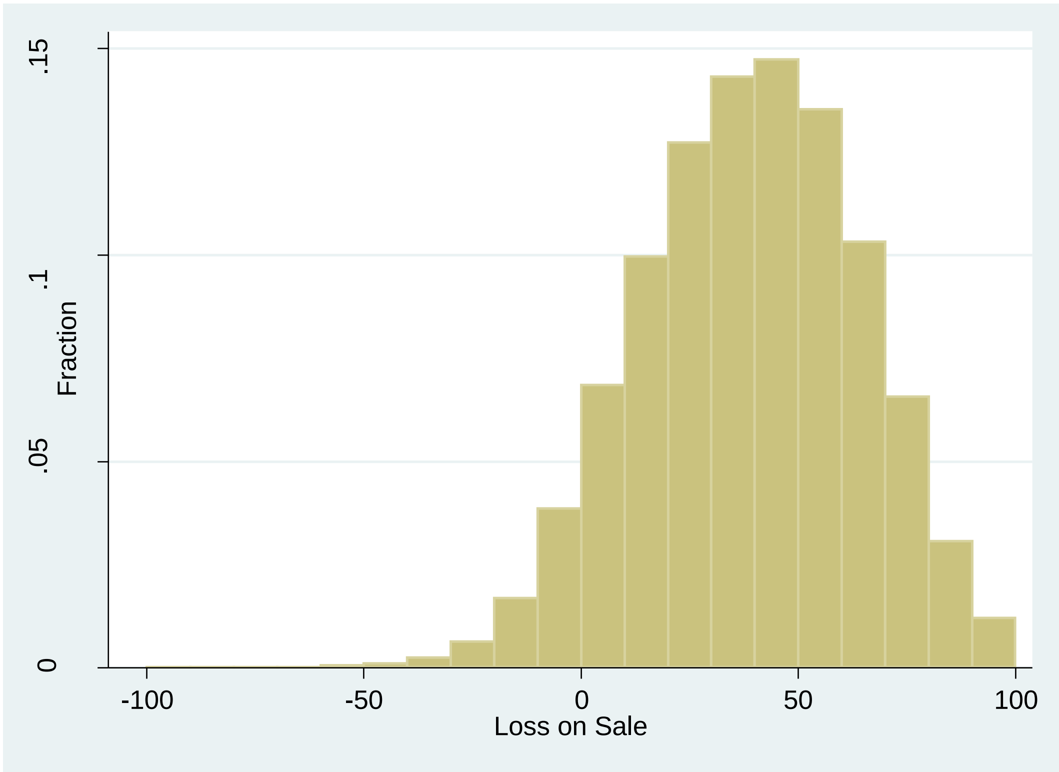
The predicted loss is estimated using the loss on sale coefficient estimates from table 2 Specification V, table 5 Specification II, and Table 11. The representative loan has the average ltv_{it}, upb_{it}, $fico_{it}$, $first_{it}$, $\Delta urate_{it}$, state fixed effect, and servicer fixed effects. The representative loan is also defined as a retail, purchase loan with the estimated fixed effect for a loan originated in 2006 with the estimated fixed effect for a loan with the last payment made in 2010 that was real estate owned during the default process.

Table 16: Predicted Losses (Percentage of UPB) by Outstanding Loan Balance

	Outstanding Loan Balance in $1,000					
	50	100	150	200	250	300
Total I	61.3%	51.6%	44.4%	37.9%	31.7%	25.6%
Total II	66.4%	57.2%	50.2%	43.9%	37.9%	32.0%
Loss on sale	45.8%	40.1%	34.4%	28.7%	23.0%	17.3%
Holding Cost	13.6%	9.6%	8.1%	7.3%	6.8%	6.5%
Legal	7.4%	3.8%	2.6%	2.0%	1.7%	1.4%
Maintenance	3.3%	2.6%	2.4%	2.2%	2.1%	2.1%
Tax and Insurance	2.1%	2.7%	2.8%	2.9%	2.9%	2.8%
Miscellaneous	0.8%	0.4%	0.3%	0.2%	0.2%	0.2%
Carrying Cost						
I: Mortgage & Treasury	1.8%	1.8%	1.8%	1.8%	1.8%	1.8%
II: Mortgage Rate	7.0%	7.4%	7.7%	7.9%	8.0%	8.2%

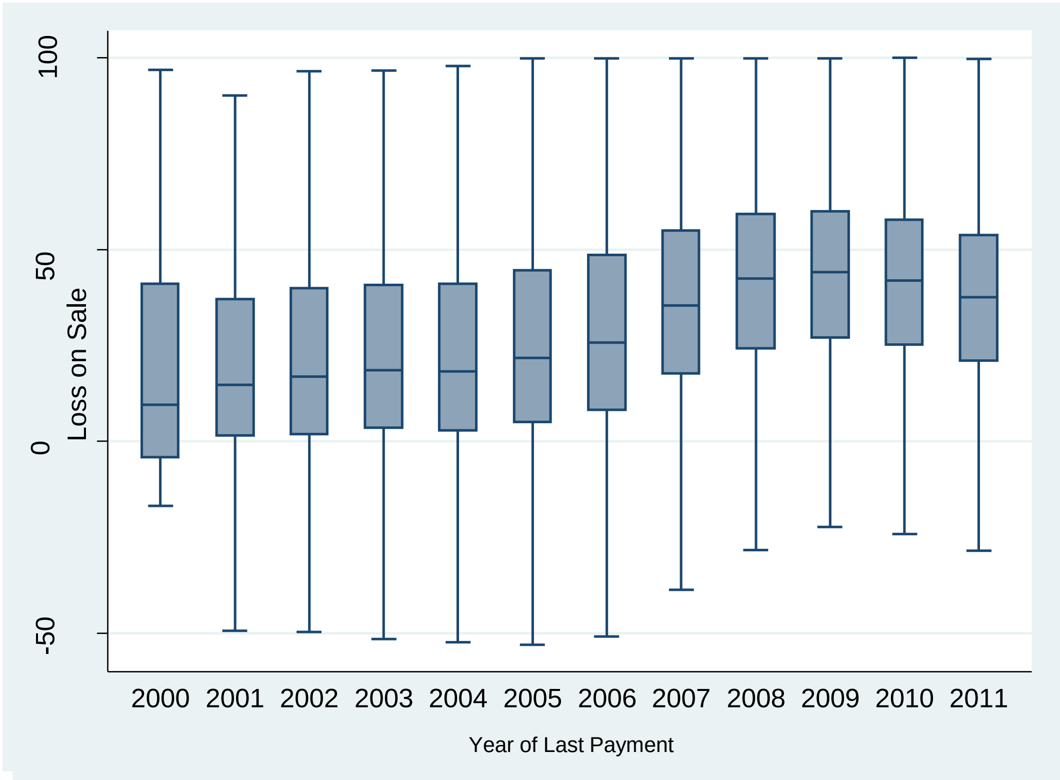
The predicted loss is estimated using the loss on sale coefficient estimates from table 2 Specification V, table 5 Specification II, and Table 11. The representative loan has the average ltv_{it}, upb_{it}, $fico_{it}$, $first_{it}$, $\Delta urate_{it}$, state fixed effect, and servicer fixed effects. The representative loan is also defined as a retail, purchase loan with the estimated fixed effect for a loan originated in 2006 with the estimated fixed effect for a loan with the last payment made in 2010 that was real estate owned during the default process.

Figure 1: Loss on Sale Distribution



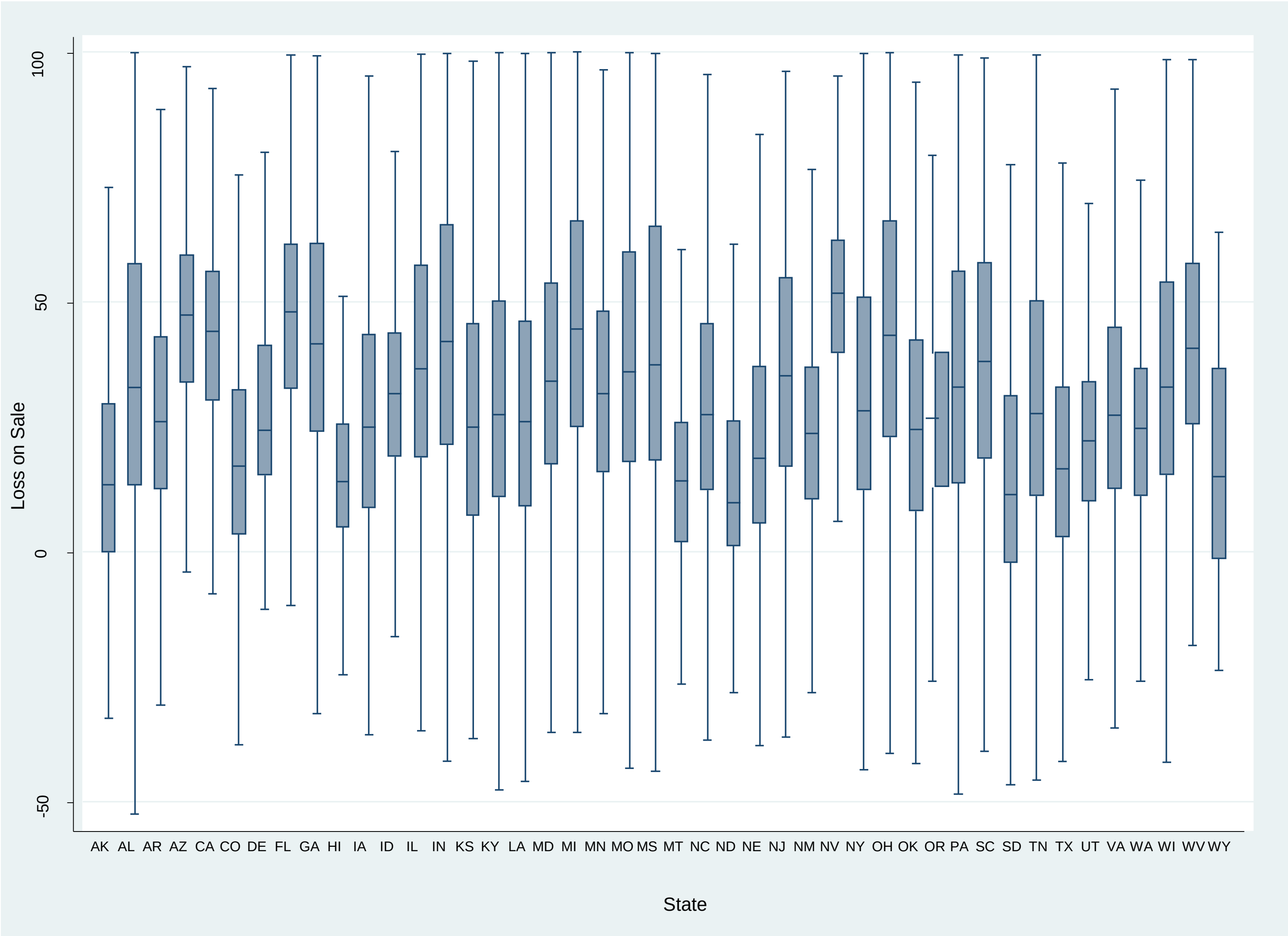
Loss on sale equals unpaid balance at the end of the loan's life less the sale price, divided by unpaid balance at the end of the loan's life. Each column represents the percentage of all loans in the bucket. The bucket is 10 percentage points wide. For example, the column just to the right of 0 on the x-axis indicates that approximately 7.5 percent of the loans had a loss percentage >= 0 and < 10.

Figure 2: Loss on sale over time – box charts



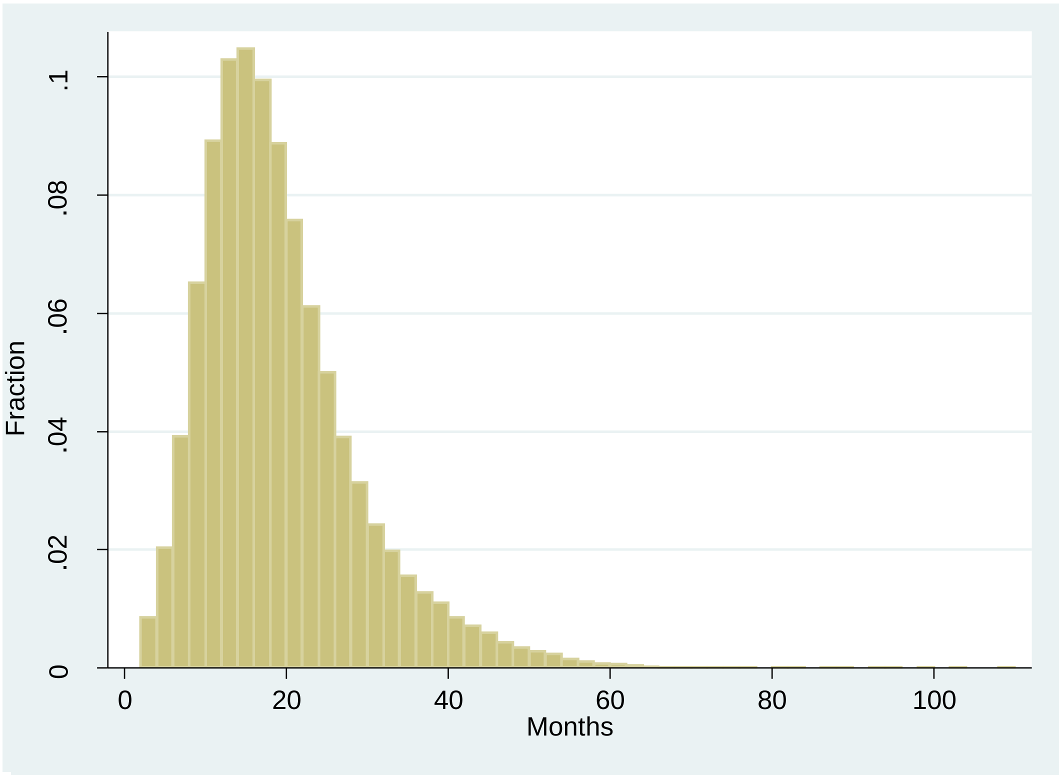
The box includes the 25th to 75th percentile of the distribution and the line in the box is the median. The whiskers or lines leading out of the box extend to the last adjacent value (next value is more than one unit away). Year is the year the resolution (zero balance or end of the loan's life) year.

Figure 3: Loss on sale by state – box charts



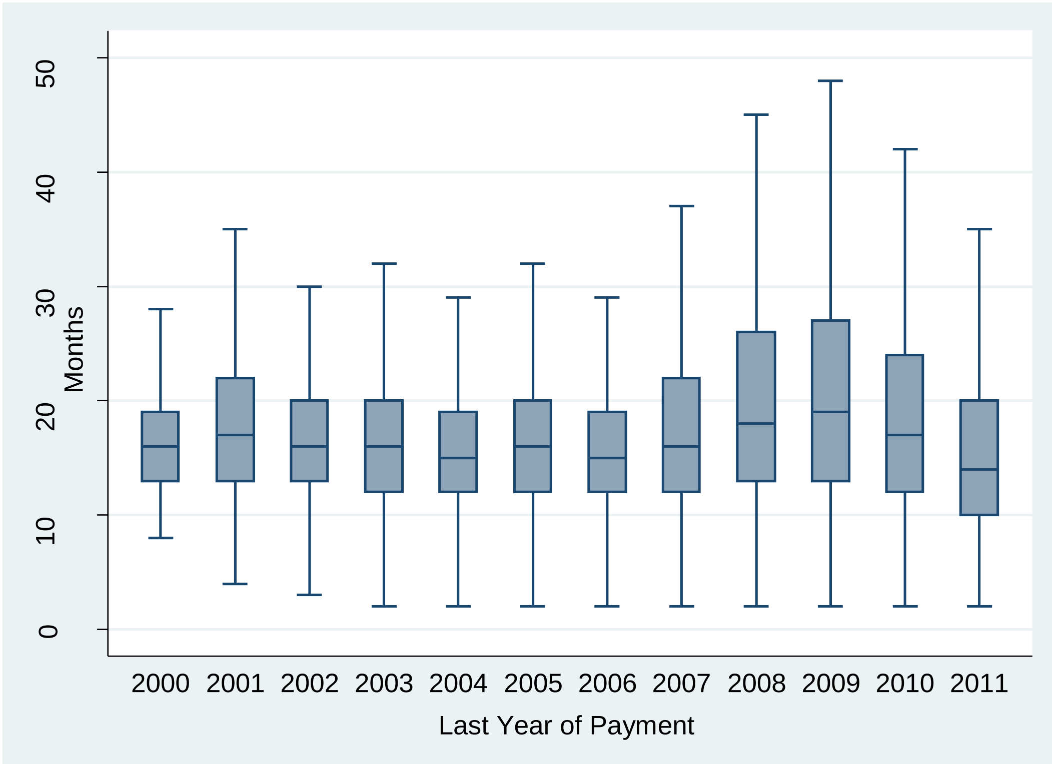
The box includes the 25th to 75th percentile of the distribution and the line in the box is the median. The whiskers or lines leading out of the box extend to the last adjacent value (next value is more than one unit away).

Figure 4: Default timeline distribution – last payment date to resolution date



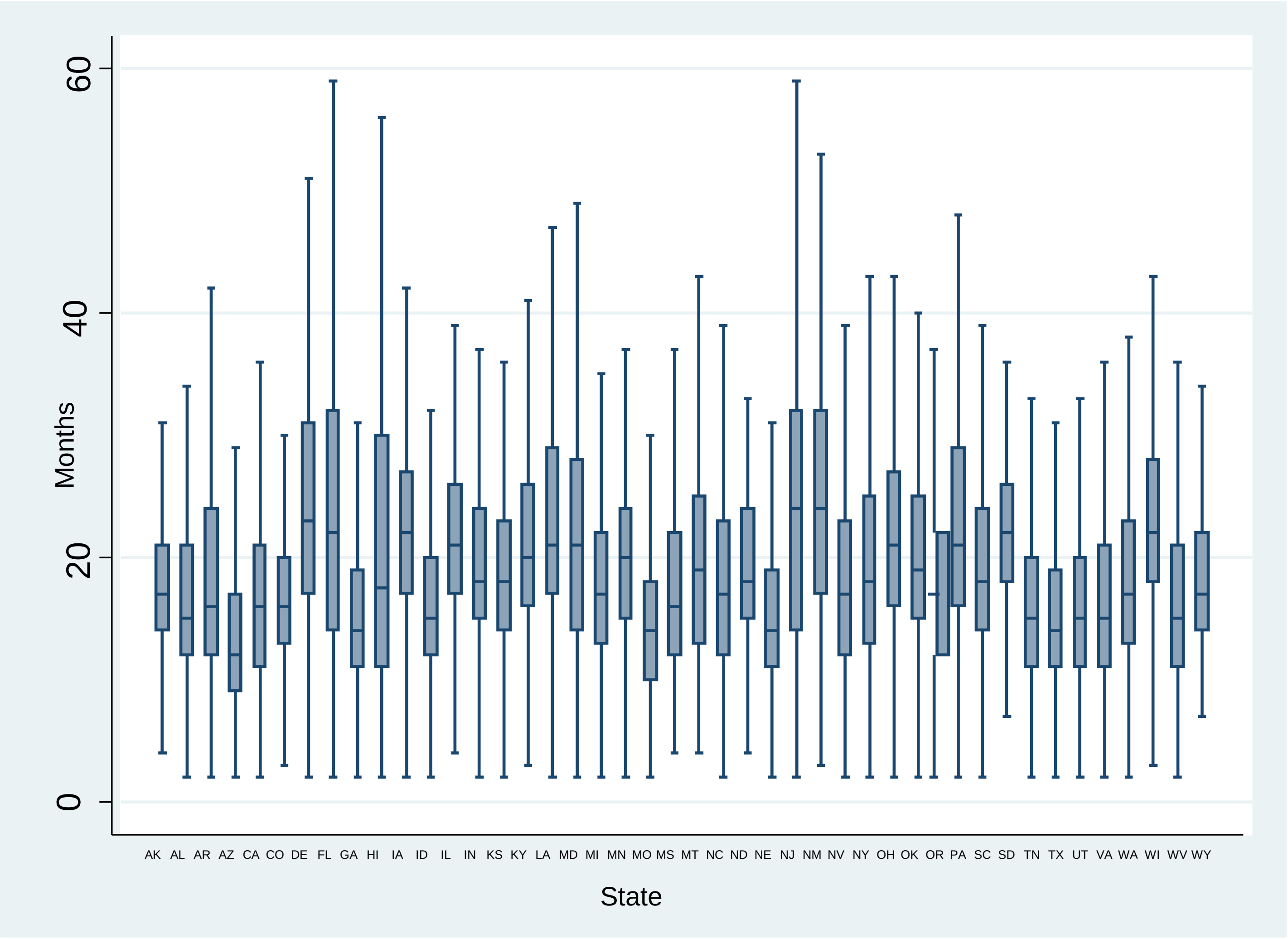
Months equals the number of months from when the last payment was made by the borrower until the loan is fully resolved by Freddie Mac. The bucket is 2 months wide. For example, the column just to the right of 0 on the x-axis indicates that 0 percent of the loans were in default for months >= 0 and < 2.

Figure 5: Default timelines over time (last payment year) – box charts



The box includes the 25th to 75th percentile of the distribution and the line in the box is the median. The whiskers or lines leading out of the box extend to the last adjacent value (next value is more than one unit away).

Figure 6: Default timelines by state – box charts



The box includes the 25th to 75th percentile of the distribution and the line in the box is the median. The whiskers or lines leading out of the box extend to the last adjacent value (next value is more than one unit away).

Appendix

Table 1: Fixed Effects for Table 2 Specification V

	Coeff	SE
Last Payment Year		
2001	0.59	7.22
2002	1.79	7.61
2003	2.72	7.91
2004	3.79	7.88
2005	7.45	7.84
2006	13.12*	7.87
2007	19.01**	7.86
2008	19.32**	7.86
2009	19.04**	7.84
2010	17.09**	7.84
2011	14.91*	7.85
Origination Year		
2001	0.84	0.73
2002	2.12***	0.76
2003	-4.91***	0.89
2004	-3.73***	0.90
2005	-2.96***	0.99
2006	-2.07**	0.98
2007	-2.22**	0.98
2008	1.22	1.08
2009	0.88	1.30
2010	-6.06***	1.62
2011	-1.85	3.91
Servicer	Coeff	SE
2	4.14***	0.69
3	2.41***	0.79
4	34.82***	4.04
5	-0.40	0.29
6	0.66**	0.26
7	-2.55***	0.60
9	7.37***	2.45
10	-1.77***	0.47
11	-3.59***	1.25
12	-0.34	1.77
13	0.49*	0.26
14	-0.25	0.60
15	-1.24	6.89

Continued … Servicer	Coeff	SE
16	0.16	1.04
17	-1.52**	0.62
19	-0.96	3.29
20	-0.52*	0.31
21	3.96	5.58
22	-3.67	2.86
23	-0.04	0.24
24	-1.08	2.59
25	-1.21	1.31
26	0.55*	1.01
27	2.99***	0.94
28	0.85**	0.39
30	5.86	8.22
31	-0.94	1.06
32	1.14**	0.49
33	-1.57	1.43
34	-2.27***	0.41
35	2.10	4.35
36	-1.47*	0.78
37	0.04	0.67
38	-1.63***	0.33
39	1.42*	0.85
40	-0.26	0.21
41	-1.12	1.03
42	-3.97**	2.02
State	Coeff	SE
AL	6.11	4.62
AR	-2.81	4.06
AZ	0.50	4.06
CA	-0.56	4.12
CO	-4.82	4.08
DE	-2.37	3.90
FL	-0.26	4.05
GA	8.95**	4.79
IA	0.68	4.21
ID	-7.24*	3.96
IL	6.54	4.09
IN	11.02***	4.26

Continued …

State	Coeff	SE
KS	-0.50	5.02
KY	3.15	4.54
LA	2.64	4.62
MD	8.88	5.77
MI	7.52*	4.16
MN	3.88	4.24
MO	7.96	5.20
MS	9.43	6.40
MT	-8.31*	4.41
NC	1.38	4.02
ND	-12.38***	4.04
NE	-3.91	4.51
NJ	7.37	5.57
NM	-1.20	5.63
NV	-10.48**	4.34
NY	5.71	4.66
OH	11.81***	4.30
OK	-0.32	4.08
OR	-4.33	3.99
PA	6.92	4.47
SC	9.21**	4.15
SD	-8.57*	4.71
TN	2.12	5.26
TX	-4.76	4.06
UT	-6.20	3.96
VA	1.19	4.13
WA	-3.64	4.03
WI	5.00	4.74
WV	9.85**	5.06
WY	-9.34**	3.92
HI	2.17	5.11

The standard errors are clustered at the three-digit zip code level.

Table 2: Probit first stage selection results

Variables	Coeff.	SE
$fico_i$	-0.404	0.004
$debt_i$ / $income_i$	0.837	0.018
ltv_i	0.027	0.000
(ltv_i>80)*ltv_i	-0.001	0.000
loan $amount_i$	-0.059	0.003
two $borrowers_i$	-0.272	0.000
fixed effects:		
origination year	x	
state	x	
servicer	x	
constant	-2.545	0.070
Psuedo R^2	0.27	
N	519,520	

*, **, and *** indicate that the coefficient is significant at the 10, 5, or 1 percent level. The standard errors are clustered at the three-digit zip code level. All variables are measured at origination. The left hand side variable is an indicator equal to 1 if the loan is ever foreclosed across all observed time periods. A random sample of loans originated in each year is included. Each loan is weighted by the inverse of the probability that it is included in the sample. Foreclosed loans are oversampled with a probability of 100%. The additional variables are the debt to income ratio ($debt_i$ / $income_i$), a spline that allows the coefficient for ltv_i to change for loans with ltv's greater than 80 ((ltv_i>80)*ltv_i), and a dummy variable indicating loans with 2 or more borrowers (two $borrowers_i$).

Table 3: Comparison of Matched Border Samples

A. Judicial Borders – 17,701 observations and 20 borders

	Power of Sale		**Judicial**	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
los_i	36.19	26.41	35.91	25.76
$month\_ltz_i$	17.39	8.31	21.41	9.77
ltv_i	84.15	16.67	83.59	15.43
upb_i	136.84	74.75	136.18	73.55
$fico_i$	6.86	0.54	6.79	0.52
$first_i$	0.10	0.30	0.10	0.30
$\Delta urate_i$	2.74	2.12	2.64	2.26
$retail_i$	0.34	0.47	0.36	0.48
$cashrefi_i$	0.31	0.46	0.27	0.45
$nocashrefi_i$	0.31	0.46	0.30	0.46

Judicial Borders: ARLA, AZNM, FLAL, FLGA, GASC, IAMO, ILMO, KSMO, KSOK, MDVA, MDWV, MIIN, MIOH, MNWI, MSLA, NCSC, NDMN, OHWV, PAWV, and TXLA

B. Statutory Right of Redemption (SRR) Borders – 18,016 observations and 21 borders

	No SRR		**SRR**	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
los_i	37.91	25.09	37.08	24.55
$month\_ltz_i$	18.38	9.20	18.92	8.67
ltv_i	97.87	29.12	96.72	28.23
upb_i	158.75	85.03	157.87	83.60
$fico_i$	6.91	0.53	6.88	0.53
$first_i$	0.10	0.29	0.13	0.33
$\Delta urate_i$	4.06	2.81	3.93	2.71
$retail_i$	0.37	0.48	0.37	0.48
$cashrefi_i$	0.32	0.46	0.29	0.45
$nocashrefi_i$	0.29	0.45	0.26	0.44

SRR Borders: ALFL, ALGA, ALMS, ARLA, AROKTX, CANV, IANE, IAWI, IDMT, IDUT, IDWA, ILIN, ILWI, KSNE, KSOK, KYIN, MIIN, MIOH, NJPA, ORNV, and ORWA

C. Recourse Borders– 11,394 observations and 11 borders

	No Recourse		**Recourse**	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
los_i	35.48	23.71	34.25	23.81
$month\_ltz_i$	18.32	8.69	18.39	8.87
ltv_i	98.56	30.86	96.86	30.45
upb_i	170.98	87.34	170.49	86.97
$fico_i$	6.89	0.54	6.87	0.56
$first_i$	0.09	0.29	0.08	0.28
$\Delta urate_i$	4.53	3.06	4.47	3.15
$retail_i$	0.33	0.47	0.36	0.48
$cashrefi_i$	0.35	0.48	0.35	0.48
$nocashrefi_i$	0.26	0.44	0.25	0.43

Recourse Borders: AZNM, CANV, IAIL, IAMO, IANE, IDWA, ILWI, NCSC, NCVA, ORNV, and TNNC

Table 4. Summary Statistics for Holding Period Costs

Variable	Mean	Standard Deviation
house value$_i$	164.64	85.95
months$_i$	21.33	9.46
legal$_i$	3,886.60	1,916.03
maintenance$_i$	6,706.20	6,839.61
taxes and insurance$_i$	5,830.80	6,255.59
miscellaneous$_i$	630.11	1,150.03

This information was only available for the 2005-2014 time period. House value is in 1,000's and all the expense types are in dollars.