Private Mortgage Securitization and Loan Origination Quality - New Evidence from Loan Losses

Abdullah Yavas
Robert E. Wangard Chair
School of Business
University of Wisconsin - Madison
Madison, WI 53706
yavas@wisc.edu

and

Shuang Zhu
Associate Professor
Department of Finance
Kansas State University
Manhattan, KS 66506
shuangzhu@ksu.edu

December 11, 2018

Private Mortgage Securitization and Loan Origination Quality - New Evidence from Loan Losses

Abstract

Due to data constraints, earlier studies of the impact of securitization on loan quality have used default probability as a proxy for loan quality. In this paper, we utilize a unique data set that allows us to use loan losses, which incorporate both probability of default and loss given default, to proxy for mortgage quality. Our analysis of prime loans shows that higher expected loan losses are associated with higher probability of securitization. Lenders sell prime loans with lower observable quality and keep higher observable quality loans on their books. For subprime loans, we observe opposite results that lenders sell better quality loans and keep lower quality loans on their book. We then use the cutoff FICO score of 620 to infer the lender's screening effort with respect to unobservable loan quality. We find that securitized prime loans exhibit no significant difference in default losses for 620- versus 620+ loans. However, securitized subprime loans with a 620- score incur significantly lower loan losses than securitized subprime loans with a 620+ score. By using loan losses as the proxy of loan quality, separating the analysis into prime and subprime samples, and distinguishing between observable and unobservable risk characteristics, this study sheds additional light on the potential channels that the securitization affects loan quality.

Keywords: Securitization, Default, Loan Loss, Adverse Selection
JEL Codes: G01, G21

1 Introduction

In the years leading up to the crisis, private securitization of residential mortgages had experienced dramatic expansions. The origination of mortgages and issuance of mortgage-backed securities used to be dominated by loans to prime borrowers conforming to underwriting standards of the Government Sponsored Agencies. However, from 2001 to 2006, non-agency originations increased from $680 billion to $1.480 trillion and non-agency issuance increased from $240 billion to $1.033 trillion. By 2006, both originations and issuance by non-agency sector exceeded those of agency sector (Inside Mortgage Finance, 2007). Moreover, while private label mortgages were about 15% of all outstanding mortgages in 2009, they made up more than half of the foreclosure starts (Piskorski, Seru and Vig, 2010).

The role of private securitization in the recent subprime mortgage crisis has been studied extensively. One central question around securitization is whether the "origination-to-distribution" model reduces lenders' screening effort at origination and thus leads to lower loan quality. On the one hand, a lender's motivation to screen mortgages might be reduced if they are able to pass the credit risk to investors dispersed around the world. On the other hand, reputation concerns, regulatory oversight, and contractual provisions such as warranties and re-purchase terms might help mitigate the moral hazard problem. The effect of securitization on mortgage quality thus remains an empirical question, and the empirical findings have been mixed so far.

Mortgage performance involves both default risk (probability of default) and loss given default (LGD). The proper measure of loan performance, unconditional default loss, should include both aspects. However, due to data constraints, empirical research mainly uses the default event (probability of default) as the measure of loan quality (e.g., Ambrose et al. (2005); Elul (2015); Mian and Sufi (2009)). One exception is Zhu et al. (2018) who uses the realized loss given default (LGD) as the measure of loan quality to study the impact of securitization on loan performance of liquidated mortgages. However, conditional loan losses (loss given default or LGD) still reflect only one aspect of loan performance. Loans with higher probability of default may not experience larger losses at liquidation and vice versa. Therefore, using either probability of default or loss given default may lead to inaccurate conclusions about the true impact of securitization on loan performance. Instead, what matters for lenders, investors and policy makers is both the probability that a loan goes into default and the loss suffered in the case of a default.

This paper utilizes the *unconditional* loan losses as a proxy of loan quality and investigates the effect of securitization on loan quality. The unconditional loan losses incorporate both default probability and LGD. Specifically, this study investigates two potential channels through which securitization could impact loan quality (1) whether lenders securitize lower loan quality loans and retain higher quality loans for their portfolio, and (2) whether securitization leads to reduced screening effort by lenders.

To study whether lenders cherry pick loans with different quality to keep in their books versus those sold to the securitized pool, the main challenge is that loan quality and securitization decision could be jointly determined. To identify the potential causal relationship between securitization and loan quality, we adopt a structured approach as proposed by Ambrose et al. (1995) and Agarwal et al. (2012). We first create the lender's expected loan loss that is independent of the securitization decision, and then regress the securitization decision on the expected loan loss. To infer the potential causal relationship between securitization and the lender's screening effort, we employ the regression discontinuity design as proposed by Keys et al. (2010) and compare the performance of loans with a credit score of just above 620 and just below 620. The significance of the 620 FICO score is due to the fact that loans with a credit score above this threshold are easier to securitize. We validate in our analysis that this credit score cutoff rule is the securitization rule of thumb rather than the origination rule of thumb.

Using a nationwide data focusing on loans originated in years 2005 and 2006, the empirical results show that (1) lenders were more likely to securitize prime loans with greater expected losses while they were less likely to securitize subprime loans with greater expected losses, and (2) subprime loans with a FICO score of just below 620 performed better than subprime loans with a FICO score of just above 620 while there was no difference in the performance of prime loans with FICO scores just above and just below 620. Our findings survive various robustness tests that we conduct.

Our first result provides new evidence of adverse selection in the securitization of prime loans; lenders were more likely to sell lower quality prime loans into securitization and retain higher quality prime loans for their portfolios. Lenders do the opposite for the subprime loans where they sell higher quality subprime loans and retain lower quality ones. The absence of adverse selection for subprime loans has been documented in Agarwal, Chang and Yavas (2012) who find no statistically significant difference in the probability of default of securitized and portfolio subprime loans. Elul (2015) also who report that securitized prime loans perform worse while securitized subprime loans perform better. However, Elul (2012) argues that this result is completely explained by early defaulting loans as lenders may have originally intended to sell these loans but they failed to do so because the loans defaulted before lenders had a chance to sell them. Once this is taken into consideration, the relationship between securitization and default probability becomes insignificant for subprime loans. In our analysis, we drop the early defaulting loans from our sample and still obtain our result that subprime loans with a lower expected loss are more likely to be securitized. Thus, a comparison of our results for subprime loans to those of Agarwal et al (2012) and Elul (2012) reveal that while using probability of default as a measure of loan quality produces no significant relationship between securitization and loan quality, using expected loss as a measure of loan quality leads to higher quality subprime loans being securitized.[1]

We conduct a number of robustness tests for our adverse selection analysis. Given the lack of direct data to control for servicing practices, we control for servicing effects by studying the performance of repurchased loans.[2] These are securitized loans that buyers/investors asked the originators to purchase back because they did not meet

[1] A number of other studies have also investigated securitization and mortgage default. Mian and Sufi (2009) study aggregate trends and find that those regions in which subprime securitization expanded most rapidly were also those in which default rates subsequently increased the most. Jiang, Nelson, and Vytlacil (2013) use data on loans originated by a single lender and find that sold mortgages perform better than those held on balance sheet, although the effect goes away when they drop early payment defaults. Also using loans originated by a single lender, Ambrose et al. (2005) find that securitized loans default at lower rates than portfolio loans.

[2] Securitized loan servicers are usually third party servicers and could be more interested in maximizing their servicing fee revenues instead of minimizing loan losses (Posner and Zingales, 2009; Ambrose, Sanders and Yavas, 2016). In addition, loss mitigation practices can vary significantly across servicers (e.g., Agarwal et al., 2011; Agarwal et al. 2017) and securitized loans are less likely to be modified and more likely to be foreclosed (Piskorski, Seru, and Vig, 2010; Kruger, 2016). As a result, servicing differences can lead to different losses for securitized loans and portfolio loans even if there is no adverse selection.

underwriting criteria and found to be in violation of representations and warranties or they were delinquent, typically within 90 days of being securitized.[3] We then compare the performance of repurchased loans to that of portfolio loans. Since both groups of loans are now owned by the originator, they are most likely to be serviced by the same servicer or by the lender herself. As a result, this comparison enables us to infer the difference in the underwriting of portfolio loans versus a set of loans that were originated with the purpose of securitizing.[4] Our adverse selection results hold for repurchased loans as well; higher risk prime loans are more likely to be securitized while higher risk subprime loans are less likely to be securitized.

Our second result regarding the impact of securitization on loan screening builds on Keys et al. (2009). Keys et al. (2009) study subprime loans only and find 620+ subprime loans have a higher, not lower, probability of default than 620- subprime loans. Since they control for observable loan and borrower characteristics, they attribute this result to a reduction in screening by lenders of 620+ loans as these loans are easier to securitize. What we show is that their result holds even when we use expected loan loss, rather than probability of default, as the measure of loan quality. However, we also show that this result holds only for subprime loans. For prime loans, we find no statistically significant difference between 620+ and 620- prime loans.

[3] There was a sharp increase during the 2005-2008 period in the amount of loans that originators were asked to repurchase. For instance, Fannie Mae reports that more than 2 percent of loans acquired between 2005 and 2008 resulted in bank repurchase requests, compared to less than 0.25 percent of loans acquired after 2008 (https://www.reuters.com/article/us-mortgages-repurchases/insight-fannie-mae-freddie-mac-clamping-down-on-banks-idUSBRE87D14V20120815?irpc=932).

[4] Piskorski, Seru, and Vig (2010) utilize a similar approach to study whether securitization causally induces a bias in the foreclosure decision of servicers.

This paper is organized as follows: Section 2 illustrates the main identification strategy, Section 3 discusses the empirical results, and Section 4 summarizes the paper.

2 Methodology

Loan quality varies in two dimensions associated with either observable or unobservable credit characteristics. Observable loan characteristics (also called hard information) contains information that are easily accessible to both lenders and investors. For example, borrower's credit score, loan-to-value ratio, and debt-to-income ratio are observable information that both lenders and investors can have access to. Unobservable credit information (also called soft information or lender's private information) is available to lenders, but typically not easily accessible to investors. One reason of investor's lack of access to the soft information is that soft information is not required to be reported in a standardized format and thus is likely to get lost during the securitization process (Parlour and Plantin, 2008; Albertazzi et al., 2015). Information such as employment status, price trends in the local neighborhood, and the quality of appraisal are examples of unobservable information that typically only lenders have access to. Another difference between observable and unobservable information is that observables are supposed to be priced in by investors during securitization, while the effort to collect unobservable information is typically internalized by the lender. This paper studies the potential impact of securitization on observable and unobservable loan quality separately.

We next lay out the methodologies used in our empirical analysis to connect securitization with observable and unobservable loan quality/losses. Section 2.1 introduces the structured approach in identifying whether lenders choose to keep loans with different observable quality on their books, compared to the loans they sell. Section 2.2 discusses the discontinuity design in identifying the potential change, due to ease of securitization, in lender's screening effort with regard to unobservable information.

2.1 Observable Loan Quality and Securitization Decision

Observable loan quality refers to the credit quality determined by the observable loan characteristics. Observable credit characteristics vary in many dimensions such as borrower's credit score, loan-to-value ratio, and documentation status, etc. In determining what kind of loans are sold into securitization or kept on bank's book, previous literature typically compares individual risk characteristics between portfolio loans and securitized loans (e.g., Krainer and Laderman, 2014) or regresses the probability of loan sale on the individual risk characteristics (e.g., Jiang et al. 2013). While these approaches are intuitive, two potential issues might arise. First, individual observable loan characteristics may point to different directions as for loan quality. For example, in Jiang et al. (2013) paper, sold loans have a lower loan-to-value ratio while at the same time exhibiting a higher proportion of low documentation status. Thus, often times these approach make it difficult to infer the overall riskiness level or loan quality of sold loans versus that of portfolio loans.

The second issue associated with the above mentioned approach is that securitization decision and loan characteristics could be simultaneously determined. For example, a lender might ask for full documentation if the mortgage is intended to be kept on the bank's books. On the other hand, a lender might accept low or no documentation if the loan is originated with the intention to be sold (Bubb and Kaufman, 2014). If the securitization and some loan risk characteristics are jointly determined, the reduced form regression could yield biased coefficient estimates.

To overcome the first challenge, we need to have a single comprehensive measure of loan quality that factors in the various risk dimensions in order to formally test whether lenders choose to sell loans with different quality than those they keep in their portfolios. Agarwal et al. (2012) utilize the default event as the comprehensive measure of loan quality to investigate the adverse selection in securitization. Zhu et al. (2018) utilize the conditional loan loss as a proxy of loan quality. Our approach in this paper proxies the loan quality by using the unconditional loan loss as a more comprehensive measure. The unconditional loan loss takes both default and loss given default into consideration, and reflects the overall loan quality. In addition, from an econometric perspective, as a continuous variable, loan loss is supposed to help reveal more information about loan quality than the dummy variable on default event.

To address the potential endogeneity concern, we adopt a structured approach as proposed by Agarwal et al. (2012) and Ambrose et al. (2005). The identification strategy is to create an instrumental variable that is correlated with loan loss, but independent of the securitization decision. Specifically, the whole sample is divided into an estimation sample (80% of the whole sample) and a holdout sample (20% of the whole sample). Step I uses the estimation sample to estimate the loan loss regression equation. The independent variables include only observable risk characteristics available at origination (this is the first stage regression). Information available after the loan origination, such as the loan servicing variables and changes in housing market conditions, are excluded from the regression equation, since only information at origination is available to lenders to form their expected loan quality for securitization decision. Note that the estimation equation excludes the securitization status as an independent variable. In another words, securitization status has no impact on the loan loss regression estimates. Step II applies the estimated coefficients from step I to the holdout sample to estimate the out-of-sample predicted/expected loan losses. The estimated loan loss represents lender's rational expectation of loan loss at origination based on observable risk characteristics. Step III regresses the observed securitization status on the predicted/expected loan loss (this is the second stage regression). This main regression investigates how the expected loan loss affects lender's securitization decision. In following the structured approach, the securitization outcome has no direct impact on lender's expected loan loss, and this will rule out the reverse causality issue.

The unconditional loan loss is a left censored variable, since loan loss is observable only for liquidated loans. Otherwise, the loan loss is censored at zero if the mortgage is in current status or still under default/foreclosure process. To estimate the loan loss equation, we choose the commonly used Tobit model to deal with the censored dependent variable in the first stage regression. The estimation equation is as in Equation (1) below where y_i is the dependent variable loan loss. The variable y^* is a latent variable that has a linear relation with independent variables x_i. Independent variables include loan observables at origination, state fixed effects, and closing quarter fixed effects. States have different foreclosure laws governing foreclosure procedures such as judicial versus non-judicial, foreclosure delay, and deficiency judgment. These differences in state foreclosure laws could have an impact on loan losses (Qi and Yang, 2009; Pennington-Cross, 2003; Kahn and Yavas, 1994). The state fixed effects help control the potential differences in loan losses due to varying state foreclosure laws. Lending standards change over time. All else equal, mortgages originated under lax lending environment tend to incur higher loan losses. Thus, we include the loan closing quarter fixed effects to control the impact of changing lending standard. The error term follows a normal distribution εi ~ N (0, σ2).

$$y_i = \begin{cases} y_i^* \; if \; y_i^* = x_i'\beta + \varepsilon_i > 0 \\ 0 \;\; if \; y_i^* = x_i'\beta + \varepsilon_i \leq 0 \end{cases} \tag{1}$$

The coefficient estimates β from Equation (1) are then plugged into Equations (2) and (3) to calculate lender's rationally expected loan loss for the holdout sample. Φ(·) is the normal cumulative probability function. λ(·) is the Inverse Mill's ratio. φ(·) is the normal probability density function.

$$E(y_i|x_i) = \Phi\left(\frac{x_i'\beta}{\sigma}\right) * (x_i'\beta + \sigma\lambda(x_i'\beta)) \tag{2}$$

$$\lambda(x_i'\beta) = \left(\frac{\phi(x_i'\beta)}{1-\Phi(x_i'\beta)}\right) \tag{3}$$

The second stage regression uses the Logit model to investigate whether lenders consider their expectation of loan loss in making the securitization decision. The model specification is as in Equation (4) and (5) below. Other than expected loan loss, a lender might also consider market yield information in making the securitization decision. Following the literature, we include various yield variables. Yield spread measures the difference between the original mortgage coupon rate and the 10-year Treasury bond rate at origination. Credit spread is measured as the difference between the AAA bond index and the Baa bond index. Yield curve is defined as the ratio between the 10-year risk free rate and the one-year risk free rate. Interest rate volatility (Sigma Int) is estimated as the standard deviation of the one-year risk-free rate during the fifteen months before the origination. Jumbo loan dummy is also included as a control variable.

$$Prob(Securtization = 1) = \left(\frac{e^{x'\beta}}{1+e^{x'\beta}}\right) \quad (4)$$

$$x'\beta = \beta_0 + \beta_1 * Expct\ Loss + \beta_2 * Controls \quad (5)$$

2.2 Unobservable Loan Quality and Securitization

Securitization may affect not only loan quality associated with observable information but also loan quality associated with unobservable information. Unobservable information such as employment status and family situation may be collected and utilized by lenders in making their loan origination and securitization decisions. However, these information are not easily accessible to investors. Since secondary market prices mortgages using only hard information, the screening effort in collecting soft information may not be paid off during the securitization process. Given that there are some costs in collecting soft information, this raises the question whether ease of securitization reduces lenders' screening effort on collecting unobservable information. This section lays out the methodology investigating whether securitization reduces lender's screening effort with regard to unobservable information. For observably similar loans, it is reasonable to assume that the reduced screening effort in soft information leads to lower loan quality and higher loan losses, assuming the servicing is unchanged. Therefore, our testable question is, for observably similar loans with similar treatment, whether ease of securitization leads to higher loan losses.

To infer the causal relation between securitization and unobservable loan quality, an exogenous variation of the securitization standard would help identify the securitization effect, when keeping everything else the same. We adopt the regression discontinuity design as first proposed by Keys et al. (2010) to investigate the question. Their identification strategy is using a credit score cutoff rule as the rule of thumb for securitization. The securitization rule of thumb is that mortgages with a FICO score of 620 or above is easier to get sold than mortgages with a credit score of 619 or below. Since borrowers cannot precisely manipulate their credit score to be above or below 620, the credit cutoff creates an exogenous variation in the ease of securitization. Since borrowers' observable risk characteristics are supposed to be similar and comparable around FICO 620 as shown in Keys et al. (2010), any differences in loan performance should be attributed to the difference in unobservable or soft information.[5]

The difference between our paper and their approach is that we use loan loss as a comprehensive measure of loan quality, instead of default event. Following the commonly used approach in literature (Keys et al., 2010; DiNardo and Lee, 2004; Card et al., 2008), we calculate the means of loan losses at each FICO score and estimate Equation (6), where $\varepsilon_i \sim N(0, \sigma^2)$.

$$y_i = \alpha + \beta_1 T_i + \beta_2 f\big(FICO(i)\big) + \beta_3 T_i f\big(FICO(i)\big) + \varepsilon_i \qquad (6)$$

The dependent variable y_i is the mean loan losses at each FICO score. Variable T_i is a binary variable that equals to one if FICO ≥ 620, and equals to zero otherwise. $f(\mathrm{FICO}(i))$ are flexible polynomials. In the empirical design, we use third order, fifth order, and seventh

[5] Keys et al. (2010) have detailed discussion about the validity of using credit score 620 in the regression discontinuity design of securitization effect.

order polynomials to fit the smoothed curves from either side of the cutoff point. Given f (FICO(i)) and $T_i f$ (FICO(i)) fit the curve to the left and right of FICO score 620, β_1 should capture any jump at the cutoff line. FICO score is re-centered to 620 such that a FICO score of 620 is now equal to 0.

One potential complication for this approach is that the credit score cutoff rule could result from either the securitization rule of thumb or the origination rule of thumb. On the one hand, loans with a FICO score of lower than 620 are harder to get sold in the securitization market (Keys et al., 2010). On the other hand, Bubb and Kaufman (2014) argue that lenders adopt the cutoff rule as a response to the underwriting guidelines from Fannie Mae and Freddie Mac, hence, the credit cutoff rule arises due to the changing lending standard rather than ease of securitization.

To address this potential complication, we separately investigate the discontinuity effect for portfolio loans and securitized loans. Since loan loss could be affected by loan quality and servicer treatment, focusing on only securitized or portfolio loans helps rule out the potential differences in servicer treatment effects between portfolio loans and securitized loans, which might contribute to the difference in loan losses between the two groups. In case that portfolio loans show no change in loan losses around FICO 620 and sold loans show a jump around FICO 620, this provides evidence that the credit score cutoff rule is a securitization rule of thumb rather than an origination rule of thumb. The results in Section 3 do show that there is no jump in the loan losses for portfolio loans, while there is a jump in the loan losses for the securitized loans. This finding validates our methodology of investigating the securitization effect on loan unobservables.

3 Empirical Evidence

This section empirically investigates the effect of securitization on loan origination quality. Section 3.1 discusses data and sample. Section 3.2 presents the results of expected loan losses on lender's securitization decision. Section 3.3 studies the effect of securitization on loan performance associated with unobservable loan characteristics.

3.1 Data and Sample

The main data source of this study comes from Black Knight Financial Services, Inc (BKFS). BKFS provided us the MacDash Core Data,[6] the MacDash Property Module, and the MacDash Resolution Module. We also utilized the treasury interest rates from the US Department of Treasury, and the Corporate Bond Indexes from S&P 500.

MacDash Core Data includes residential mortgages serviced by nine out of the ten largest US mortgage servicers. This data set contains detailed mortgage-level information at origination, such as borrower's credit scores, loan-to-value ratios (LTV), and documentation status, etc. The data set also reports the subsequent monthly loan activities, such as payment, default, and foreclosure, etc. The MacDash Property Module was created by BKFS utilizing their proprietary methodology matching the MacDash Core Data with the nationwide county-level Recorder's data set. The Property Module reports the real estate transactions

[6] MacDash Core Data was previously called LPS data, which was provided by LPS Applied Analytics. LPS Applied Analytics later was acquired by Black Knight Financial Service. The LPS data has been used for academic research such as Demyanyk and Van Hemert (2011) and Agarwal et al. (2012).

associated with both mortgage originations and terminations, such as the transaction dates and the transaction prices, etc. Monthly zip code-level house price index (HPI) updated property values are also reported in the Property Module. MacDash Resolution Module tracks the transactions of foreclosed properties until liquidation. The Resolution Module was created by BKFS using their proprietary methodology by merging the MacDash Core Data with the nationwide county-level Recorder's data set.

Mortgages are heterogeneous financial products. To reduce the heterogeneity, we restrict the mortgages included in our sample to conventional, single family, first lien, new purchase, adjustable rate loans with a mortgage term of either thirty or forty years.[7] Our sample includes both portfolio loans and privately securitized loans (also called non-agency securitized loans). Loan origination time ranges from January 2005 to December 2006. We focus on mortgages originated from the beginning of year 2005 since MacDash data does not have comprehensive coverage before year 2005. We include the loans originated by the end of year 2006 since there is a structural change in the private mortgage securitization market starting from the beginning of year 2007. The structural change of the private securitization market makes it difficult to identify lenders' original intention with regard to the securitization decision of the mortgages.[8] To avoid potential data errors, mortgages are further limited to have the underlying property value between $5K and $1.5M, loan amount in between $5K and $1.5M, and the original loan-to-value ratio lower than 1.5. We also

[7] We focus on adjustable rate mortgages since those are the main type of mortgages that caused lots of trouble during the mortgage crisis.
[8] For a detailed discussion on the structural change of the private mortgage securitization market , see Zhu et al. (2018).

require the observations to have valid values for each variable. To control for survival bias, we require loans entering into the data set within four months of origination.

To investigate the effect of securitization on loan quality, the intended securitization status, whether the mortgage is intentionally to be held on bank's balance sheet or to be sold to investors, needs to be identified. We start with the final securitization status, either at liquidation or at the end of our sample time period. Since a mortgage may end up retained by the bank on their book or sold to investors for reasons other than lender's original intention, we make several adjustments to ensure that the original intended securitization status is correctly identified. First, a mortgage might fall into default too early to get securitized. These portfolio loans may not be originated with the intention to be kept on the bank's books. Thus, we exclude mortgages that default within six months of origination. Second, securitized mortgages might be repurchased by the lender and get back into the portfolio pool due to the MBS warranty clauses. Given this concern, the repurchased mortgages are excluded from the sample. Third, in order to capture the intention of securitization at origination, loans are required to be securitized within six months of loan origination. Zhu et al. (2018) shows that over 75% of securitized loans are sold with six months after origination. Although mortgages can be sold years after origination, it is unlikely that those loans are originated with the intention to be securitized.

Loan performances are tracked until either three, four, or five years after origination. If a loan is liquidated within the specified tracking time frame, loan loss rate is defined as in Equation (7). Otherwise, if for instance a mortgage is in current status or under the foreclosure procedure at the end of the specified time frame, loan loss is treated as zero. To avoid the results driven by the extreme values or some possible data errors, mortgages with the top and bottom 0.5 percentile of loan losses are excluded from the sample.

$$Loan\ Loss\ Rate = \frac{Outstanding\ Loan\ Balance - Liquidation\ Price}{Outstanding\ Loan\ Balance} \quad (7)$$

The outstanding loan balance is the unpaid loan balance at the time of default. Liquidation price is the final property sale price. Our measure of loan loss rate does not represent the total loan loss rate, and does not include items such as legal fees, servicing fees, property maintenance cost, selling expenses, and mortgage insurance payment, etc. Although these other fees or payments contribute to the total loan losses, they are not likely to be related to the initial loan quality. Since this paper use loan loss to infer the loan quality, the measure of loss in (7) serves the purpose better than the total loss.

Loan characteristics at origination include borrower's credit score (scaled by 100), low documentation dummy that equals to one for loans with no or limited documentations, loan-to-value ratio, debt-to-income ratio, owner occupied status, second lien dummy that equals to one for loans with junior liens, jumbo loan dummy that equals to one for mortgages with the purchase price higher than the OFHEO guideline for jumbo loans, and loan term (term30), which equals to one for mortgages with a 30-year loan term.

Table 1 reports the summary statistics of the loan characteristics at origination. Compared to securitized loans, portfolio loans on average have a higher credit score, a lower loan-to-value ratio, a lower proportion of second liens, a lower debt to income ratio, and a higher proportion of jumbo loans. Securitized loans seem to have a slightly higher proportion of loans with full documentation status, and a larger percentage of properties that are owner occupied.

The main motivation of this paper is to propose using loan loss as a comprehensive measure of loan quality rather than using just default probability. However, if loan loss and default probability are highly correlated and contain same amount of information, it may not be interesting to do the new experiment. Therefore, we first check the correlation between default probability and loss rate. Table 2 reports the Pearson correlation coefficients between default probability and loan loss rate. Loan losses are tracked 36 months, 48 months and 60 months after origination (Loss36m-36 months after origination, Loss48m-48 months after origination, and Loss60m-60 months after origination). Defaults are tracked one year and two years after origination (Default12m-12 months after origination, and Default24-24 months after origination) as those are the most commonly used measures in academic research. We also run the default estimation regression and calculate the expected default probability within one and two years after origination. The correlations range from 0.150 (between expected default probability within 12 months of origination and loss rate tracked until the end of 60 months) to 0.396 (between default within 24 months after origination and loan loss rate within 36 months of origination). The low correlation between default and loss rate indicates that information content of those two measures may be different and the loan quality measured by these different proxies may be different, and that the inferences of the securitization effect drawn from these two proxies are not necessarily to be the same.

Table 3 reports the loan loss rates for the full sample and various sub samples. As for the average loan loss, securitized loans exhibit higher loan loss rates than portfolio loans for the full sample and for each of the sub samples. For example, when loan performances are tracked until sixty months after origination, the full sample shows that securitized loans have an average loan loss rate of 5.3%, while portfolio loans have an average loan loss rate of only 2.3%. The pattern is consistent for different tracking time periods. Prime loans and loans with higher credit scores exhibit lower losses than the corresponding subprime loans and loans with lower credit scores. When comparing the relative increase in loss rates of portfolio loans versus sold loans, it seems that, for better loan quality sample (high credit score sample and prime sample), sold loans have doubled or more than doubled losses than the corresponding portfolio loans. For lower quality sample (low credit score sample and subprime sample), the relative increase in loss from portfolio loan to sold loans seem to be smaller. For example, when tracked forty-eight months after origination, for high credit score sample, the loan loss of sold loans is about 2.86 times (0.043/0.015=2.86) of the loss of portfolio loans. The loan loss of sold loans, for the low credit score sample, is only 1.28 times (0.037/0.029=1.28) of the corresponding loan loss from portfolio loans. The difference in loss between sold loans and portfolio loans might come from difference in loan observables, unobservables and/or servicing practices.

3.2 Observable Loan Quality and Securitization Decision

This section investigates the potential adverse selection issue in securitization. Specifically, we follow the structured approach as discussed in Section 2.1 to study whether lenders choose to sell loans with different observable quality than those they choose to keep in their books. If lenders are able to sell lower observable quality loans to investors, their motivation to lessen the lending standard and originate lower quality loans is likely to increase. Since investors pay for the mortgage-backed securities based on observable information, this would be especially true if the pricing of mortgage-backed securities does not properly reflect the increased risk associated with lower observable loan quality.[9] This, in turn, could lead to an overall deterioration of observable loan quality in the market.

Previous literature shows mixed empirical evidence on lenders' securitization decision. For example, focusing on individual risk characteristics, Krainer and Laderman (2014) and Jiang et al. (2013) find that lenders sell observably riskier mortgages. While Agarwal et al. (2012) adopt the structured approach by using estimated default probability as a proxy of loan quality and find no significant difference in default risk between privately securitized loans and portfolio loans.

The full sample is divided into prime and subprime subsamples according to lender's original classification based on the credit quality of the mortgage. Prime loans have better risk profile and subprime loans carry higher credit risk. Prime loans and subprime loans represent different mortgage market segments. Next, we present the empirical results on lender's securitization decision based on expected loan losses from observable risk characteristics.

As discussed in Section 2.1, securitization status could affect the outcome of loan loss, and the perceived loan quality could affect the outcome of securitization decision. Thus, the reduced form regression could lead to biased coefficient estimate. To overcome this

[9] If the secondary market pricing of mortgage-backed securities properly reflects the increased risk with lower observable quality, lenders may need to weigh in the reduced price of loan sale and the profit from loan origination to make their loan origination decision.

challenge, we construct the *expected* loan loss variable, which should reflect lender's perceived observable loan quality at origination and is estimated independent of the securitization decision. In order to do so, first, we utilize the Tobit model to estimate the loan loss rate regression equation according to Equation (1). The loss rate estimation model uses the estimation sample that consists a random 80% of the mortgages in the full sample. The dependent variable is loan loss rate. If a mortgage is liquidated within n years after origination, the loss rate is calculated by using the outstanding loan balance at default and the liquidation sale price as defined in Equation (7). Otherwise, if a mortgage remains current status, default but not yet in the foreclosure procedure, or still under the liquidation procedure, the loan loss rate is treated as censored at zero, meaning no realized loan loss yet by the end of the tracking time period. The explanatory variables include observable loan characteristics at origination. State fixed effects are included in all the regressions to control the impact of the different state foreclosure laws on loan losses. Origination quarter fixed effects are included to control the potential effect of the time-varying lending standards on loan losses.

Table 4 reports the Tobit regression results of loan loss equation. We report the coefficient estimates with the standard errors in parenthesis. We track the loan performance thirty-six, forty-eight, and sixty months after origination, and report the corresponding regression results. Regressions one to three report the results for prime mortgages with varying tracking times. Regressions four to six show the results for subprime mortgages with different loan performance tracking time frames.

The results show that, for both prime and subprime loans, higher loan loss rates are associated with risky loan features such as lower credit score, higher loan-to-value ratio, higher debt-to-income ratio, and low documentation status. Jumbo prime loans carry lower loan losses for prime loans, but does not show significant effect for subprime loans.[10] Second liens increase loan loss for prime loans. However, second liens lead to lower loan loss for subprime loans when tracking thirty-six or forty-eight months after origination.[11] The junior lien effect becomes insignificant when tracking the loan performance until sixty months after origination. Owner occupied properties incur lower loan losses in all cases except for the prime loans tracking thirty-six months after origination. Thirty-year mortgages in general incur lower loan losses than loans with longer terms. In sum, the data shows that observably riskier loans incur higher loan losses.

Second, we apply the coefficient estimates from Table 4 to the remaining 20% holdout sample, and use equations (2) and (3) to calculate the out-of-sample expected loan losses at origination. The expected loan loss incorporates only the observable information available at origination. Securitization status is not included in the first stage regression as an explanatory variable, and does not have any direct impact on the estimated loan loss. The estimated loan loss thus works as a proxy of lender's rational expectation of observable loan quality that is independent of the securitization decision.

Next, using the 20% holdout sample, we correlate the expected loss and securitization status, and run the Logit model as in Equation (4) to (5) to investigate lender's securitization decision. The dependent variable is the securitization status, which equals to one if a loan is

[10] One possible reason of jumbo loans lacking of significance for subprime sample is that there are few observations of jumbo subprime loans.

[11] Since subprime loans have lower credit quality, it might be possible that only better credit quality subprime loans are able to have a second lien.

privately securitized and equals to zero if a loan is kept on bank's own balance sheet. The independent variable of interest is the expected loan loss calculated in the previous step. Other controls include jumbo loan dummy and yield variables as discussed in Section 2.1. Table 5 reports the Logit regression results of lender's securitization decision on expected loan losses. We report the coefficient estimates with the standard errors in parenthesis. Like in Table 4, we divide the sample into prime and subprime loans, and track the loan performance for different time frames as robustness checks. Expected loan losses (Expct Loss) are estimated using the matching regression coefficients in Table 4. For example, the expected loan losses in regression one of Table 5 are estimated using the regression coefficient estimates from regression one of Table 4.

Focusing on prime loans, the results show that higher expected loan losses are associated with higher probability of securitization. The effect is statistically and economically significant at one percent level across different loan performance tracking times. For example, according to regression three, a ten percent increase in expected loan loss increases the probability of securitization about 20 percent for prime loans. This result offers evidence of adverse selection that lenders choose to sell lower observable quality prime loans to the private securitization market and keep higher observable quality loans in their own portfolio.

Because of the complex nature of mortgage-backed securities, investors typically rely on credit rating agencies to infer the riskiness level of the securities and to price the securities accordingly. However, during the pre-crisis period, rating agencies failed to rate many MBSs properly (Benmelech and Dlugosz, 2010). As an example, more than fifty percent of the structured finance issues received the highest credit rating before the subprime crisis, which is obviously hard to be justify by the later deterioration in MBS performances. In another words, it seems that, during the pre-crisis time period, rating agencies have failed to distinguish between lower observable quality prime loans and higher observable quality prime loans. Investors correspondingly were not able to price in the differences in observable loan qualities among prime loans. If prime loans with different observable qualities can be sold at similar prices, lenders have stronger incentive to sell those with lower observable quality, and keep the better quality ones on their books. Since the perceived default rate in 2005 and 2006 was low even for lower observable quality prime loans, lenders' reputation concerns were not likely to play a major role in the prime loan securitization decision. One possible outcome of this result is that it induced lenders to loosen observable lending standards at origination, and sell more low quality prime loans. This would lead to a decline in the overall quality of prime loans. Thus, lenders' loosened lending standards due to securitization could play a role in the poor performance of prime loans during the mortgage crisis. As our data set does not have the secondary market mortgage pricing information, we are not able to investigate how mortgage pricing interacts with lender's securitization decision. However, even if the secondary mortgage market pricing properly reflects the riskiness level of the prime mortgages, if lenders are able to sell lower observable quality loans to investors, their motivation to lessen the lending standard and originate lower quality loans is still likely to increase, which may lead to a deteriorating quality of prime mortgage market.

Interestingly, for subprime loans, higher expected loan losses are associated with lower probability of securitization. Lenders choose to sell better quality subprime loans and keep lower quality ones on their balance sheets. In general, subprime loans have lower credit quality and the default rate is relatively high. If a large number of loans default soon after origination, that hurts the lender's reputation. Also, the Servicing and Pooling Agreement typically includes clauses that require lenders to repurchase loans that default soon after securitization. Since loans with extremely low quality, such as the lower quality subprime loans, are likely to default early, lenders may choose to sell relatively better observable quality subprime loans to investors. Compared with prime loans where about 30 percent (8806/(19823+8806)=30.76%) are kept in bank's balance sheet, only about 8 percent (2119/(24217+2119)=8.05%) subprime mortgages are kept as portfolio loans. Since lenders seem to securitize as many subprime loans as possible, another possible explanation is that the remaining 8 percent of subprime mortgages might be loans with very low quality that are difficult to get securitized.

Finally, Table 6 to 8 conduct various robustness checks on the effect of expected loan loss on lender's securitization decision. Table 6 divides the full sample into high and low credit score subsamples and reports the second stage Logit regression coefficient estimates and standard errors. The credit score cutoff line of 620 is chosen as mortgages with a credit score of 620 or higher are easier to get securitized than those with a credit score lower than 620 (Keys et al., 2010). The model specification is the same as in Table 5. The results are consistent with the previous findings. For higher credit score sample where loans have better quality and are easier to securitize, lenders choose to sell lower observable quality loans into the securitized pool. For lower credit score sample where loans have lower credit quality and are subject to closer examination during securitization, better observable quality loans are sold into the secondary market.

Table 7 conducts more subsample robustness checks. The model specifications are the same as Table 5. For simplicity, we only report the coefficient estimates and standard errors for the variable of interest - expected loan losses from the second stage Logit regressions. Panel A presents the robustness checks for prime loans. Panel B presents the robustness checks for subprime loans. The prime/subprime samples are further divided into jumbo and non-jumbo samples since jumbo loans can only be sold at the private securitization market or kept on the bank's books, while non-jumbo loans can also be sold to GSEs. We also run separate regressions for loans originated in year 2005 and 2006. Loan performances are tracked for thirty-six, forty-eight, and sixty months after origination. For prime loans, the results show that across different specifications, higher loan losses are consistently associated with higher probability of securitization. The results are statistically significant at one percent level. For subprime loans, expected loan losses are statistically significant only for non-jumbo loans, and the effect is significant mainly for the loans originated in 2005.

Although loan loss outcome is affected by both loan quality and servicing treatment, the expected loan loss is calculated using only information available at origination, regardless of the securitization status. Therefore, the potential servicing difference between sold loans and portfolio loans should have no direct impact on the expected loan loss. Despite of this, we conduct additional test to control the potential servicing difference between portfolio loans and securitized loans. Since the data set does not have loan servicing information, we construct the repurchased loan sample to control the treatment effect. The repurchased loan sample include both repurchased loans and portfolio loans. The repurchased loans are securitized loans that buyers/investors asked the originators to purchase back. We then compare the performance of repurchased loans to that of portfolio loans. Since both groups of loans are now owned by the originator, they are most likely to be serviced by the same servicer or by the lender herself. As a result, this comparison enables us to infer the difference in the underwriting of portfolio loans versus a set of loans that were originated with the purpose of securitizing. Our adverse selection results hold for repurchased loan sample as well: higher risk prime loans are more likely to be securitized while higher risk subprime loans are less likely to be securitized.

The various robustness checks confirm that lenders sell lower observable quality prime loans to the market, which could lead to a decrease in the overall prime market observable loan quality. Since prime mortgages constitute a large proportion of the overall mortgage market, our findings shed light on one potential channel through which prime loan quality deteriorates during the pre-crisis years. Using expected loan loss to proxy lender perceived observable loan quality, our results offer new evidence for the effect of securitization effect on observable loan quality, especially for prime loans. The results also highlight the importance of investigating prime and subprime mortgage markets separately.

3.3 Unobservable Loan Quality and Securitization

This section studies empirically the impact of securitization on unobservable loan losses that are associated with soft information. Specifically, we adopt the regression discontinuity design as described in Section 2.2 to investigate whether ease of securitization leads to reduced lender screening effort with regard to unobservables. If lenders reduce their screening effort due to ease of securitization, this implies that securitization leads to overall lower mortgage quality.

Given the debate on the cause of the "puzzling" positive relationship between FICO scores and default probability around the FICO score of 620, we plot the mean loss rate against FICO score separately for portfolio loans and securitized loans in Figure 1. A sudden change in loan losses for portfolio loans would indicate that the 620 cutoff rule is used by lenders in their screening process even in the absence of the role that 620-score plays in the probability of securitizing a loan. If that is the case, any jump around the credit cutoff line of the securitized loans needs to be interpreted with caution, as the 620 effect could be due to the changing lending standards as well as the ease of securitization. On the other hand, a lack of sudden change in the portfolio loan losses would indicate that 620 cutoff rule is not a rule of thumb in underwriting. In the absence of any changes in loan losses around the 620-score for portfolio loans, a sudden change in loan losses for securitized loans would indicate that the source of the higher losses for 620+ loans is the moral hazard created by securitization.

Figure 1 shows that there is no obvious difference in loan losses between 620+ and 620- scores for portfolio loans. This indicates that the credit cutoff may not be used by loan originators in their underwriting standards. It is obvious that there is a sudden jump in securitized loan losses around the credit score of 620; loan losses for loans with 620- scores are significantly lower than those for loans with 620+ scores. Since loans on either side of

620 are supposed to have similar observable risk characteristics, any difference in loan losses is likely to be attributed to the differences in unobservable loan quality. The evidence that 620+ securitized loans incur higher loan losses than 620- securitized loans indicates that ease of securitization leads to lower incentives for lenders to collect and use soft information for further screening.

For more detailed analysis of the 620 FICO score, we divide the sample into portfolio prime loans, portfolio subprime loans, securitized prime loans, and securitized subprime loans. Figure 2 plots the mean loan losses against the credit score for the four subsamples. The graph shows that, for both portfolio prime and portfolio subprime loans, there is no obvious change in loan losses around the credit score of 620. The securitized subprime loans show a dramatic increase in loan losses as we move from FICO scores of 620- to 620+. The change in loan losses around the 620-score is not obvious for prime securitized loans.

Table 9 reports the regression discontinuity results for the portfolio and securitized samples according to Equation (6). We report the whole, prime only, and subprime only regression results for portfolio and securitized samples separately. Loan performances are tracked until sixty months after origination. We use the 7th order polynomial in the regressions. Table 9 reports the results using loan loss rates as the dependent variable. As a robustness check, it also reports the results using dollar amount loan loss as the dependent variable. The regression results are consistent with the previous observations from Figure 1 and 2. The coefficient estimates are statistically insignificant for portfolio whole sample, portfolio prime subsample, and portfolio subprime sample. Securitized loans show a statistically significant (at 1% significance level) increase of 2.55% in loss rate, and over eight thousand dollars increase of losses, as the FICO score moves from 620- to 620+. However, the prime securitized loans do not show any significant jump in loan losses around the credit score of 620. One possible explanation is that prime loans meet certain credit standards and have better loan quality, so that the variations in unobservable loan quality are likely to be attenuated after controlling for hard information.[12] Subprime securitized loans, on the other hand, fail to satisfy certain observable credit standards and their performance are likely to be affected more by unobservable soft information. Indeed, securitized subprime loans show a statistically significant increase of 2.73 percent in loss rate as we move from a FICO score of 620- to 620+. Given that the average loan loss rate is 5.94 percent for loans with FICO scores between 615 and 619, a 2.73 percent increase in loss rate is equal to a 46 percent (2.73/5.94 = 0.46) relative increase in loan loss rate. The increase in dollar amount of losses for securitized loans is about eight thousand dollars around the 620 score.

Next, focusing on the securitized subprime sample, Table 10 presents the results of various robustness checks. We track the loan performance thirty-six months, forty-eight months, and sixty months after origination. We use the third order, the fifth order, and the seventh order polynomials to fit the data. Across different specifications, the data consistently shows that 620- loans incur a statistically and economically significant lower loan loss than the 620+ loans. Overall, the evidence provided in this section points out that securitization leads to lower unobservable loan quality for subprime loans that are easier to be securitized.

[12] Lack of enough observations for prime securitized loans with a credit score below 620 might also contribute to the insignificant results.

4 Conclusion

Due to data constraints, earlier studies of the impact of securitization on loan quality have used default probability as a proxy for default losses. In this paper, we utilize a unique data set that allows us to proxy loan quality with loan losses. Our analysis of prime loans shows that higher expected loan losses are associated with higher probability of securitization. Lenders sell prime loans with lower observable quality and keep higher observable quality loans on their books. This contradicts earlier studies that use probability of default as a proxy for default losses. We obtain opposite results for a subset of subprime loans. We then use the cutoff FICO score of 620 to infer lender's screening effort with respect to unobservable loan quality. Since loans on either side of 620 are supposed to have similar observable risk characteristics, any differences in loan losses can be attributed to unobservable risk characteristics. We find that there is no significant difference in losses between portfolio loans with 620- and 620+ credit scores. Similarly, securitized prime loans exhibit no significant difference in default losses for 620- versus 620+ loans. However, we find that securitized subprime loans with a 620- score incur significantly lower loan losses than securitized subprime loans with a 620+ score. Thus, securitization leads to lower unobservable loan quality for subprime loans, but not for prime loans. By separating the analysis into prime and subprime samples, and distinguishing between observable and unobservable risk characteristics, this study sheds additional light on the potential channels that the securitization affects loan quality.

References

Adelino, M., K. Gerardi, and P. S. Willen (2013). Why don't lenders renegotiate more home mortgages? Re-defaults, self-cures and securitization. *Journal of Monetary Economics 60*(7), 835–853.

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., Evanoff, D., 2011. The role of securitization in mortgage renegotiation. *Journal of Financial Economics* 102, 559-578.

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., Piskorski, T., Seru, A., 2017. Policy intervention in debt renegotiation: evidence from the Home Affordable Modifcation Program. *Journal of Political Economy* 125, 654-712.

Agarwal, S., Y. Chang, and A. Yavas (2012). Adverse selection in mortgage securitization. *Journal of Financial Economics 105*(3), 640–660.

Albertazzi, U., G. Eramo, L. Gambacorta, and C. Salleo (2015). Asymmetric information in securitization: An empirical assessment. *Journal of Monetary Economics 71*, 33–49.

Ambrose, B. W., J. Conklin, and J. Yoshida (2016). Credit rationing, income exaggeration, and adverse selection in the mortgage market. *The Journal of Finance 71*(6), 2637–2686.

Ambrose, B. W., M. LaCour-Little, and A. B. Sanders (2005). Does regulatory capital arbitrage, reputation, or asymmetric information drive securitization? *Journal of Financial Services Research 28*(1-3), 113–133.

Ambrose, B. W., Sanders, A. B. and Yavas, A. (2016). Servicers and Mortgage-Backed Securities Default: Theory and Evidence. *Real Estate Economics*, 44: 462–489.

Benmelech, E. and J. Dlugosz (2010). The credit rating crisis. *NBER Macroeconomics Annual 24*(1), 161–208.

Bubb, R. and A. Kaufman (2014). Securitization and moral hazard: Evidence from credit scoring cutoff rules. *Journal of Monetary Economics 63*(1), 18.

Card, D., A. Mas, and J. Rothstein (2008). Tipping and the dynamics of segregation. *The Quarterly Journal of Economics 123*(1), 177–218.

Demyanyk, Y. and O. Van Hemert (2011). Understanding the subprime mortgage crisis. *Review of financial Studies 24*(6), 1848–1880.

DiNardo, J. and D. S. Lee (2004). Economic impacts of new unionization on private sector employers: 1984–2001. *The Quarterly Journal of Economics 119* (4), 1383–1441.

Elul, R. (2015). Securitization and mortgage default. *Journal of Financial Services Research*, 1–29.

Inside Mortgage Finance (2007). The 2007 Mortgage Market Statistical Annual.

Jiang, W., A. A. Nelson, and E. Vytlacil (2013). Securitization and loan performance: Ex ante and ex post relations in the mortgage market. *The Review of Financial Studies 27*(2), 454–483.

Kahn, C.M., Yavas, A. (1994). The economic role of foreclosures. *Journal of Real Estate Finance and Economics.* 8, 35–51

Keys, B., T. Mukherjee, A. Seru, and V. Vig (2010). Did securitization lead to lax screening? Evidence from subprime loans. *Quarterly Journal of Economics 125*, 307–362.

Krainer, J. and E. Laderman (2014). Mortgage loan securitization and relative loan performance. *Journal of Financial Services Research 45*(1), 39–66.

Kruger, S. A. (2016). The effect of mortgage securitization on foreclosure and modifications,working paper.

Mian, A. and A. Sufi (2009). The consequences of mortgage credit expansion: Evidence from the US mortgage default crisis. *The Quarterly Journal of Economics 124*(4), 1449–1496.

Parlour, C. A. and G. Plantin (2008). Loan sales and relationship banking. *The Journal of Finance 63* (3), 1291–1314.

Pennington-Cross, A. (2003). Subprime & prime mortgages: Loss distributions.

Piskorski, T., Seru, A., Vig, V., 2010. Securitization and distressed loan renegotiation: evidence from the subprime mortgage crisis. *Journal of Financial Economics* 97, 369-397.

Posner, E. A. and L. Zingales (2009). A loan modification approach to the housing crisis. *American Law and Economics Review,* 11, 575-607.

Purnanandam, A. (2011). Originate-to-distribute model and the subprime mortgage crisis. *Review of Financial Studies 24* (6), 1881–1915.

Qi, M. and X. Yang (2009). Loss given default of high loan-to-value residential mortgages. *Journal of Banking & Finance 33* (5), 788–799.

Zhu, S., A. Yavas, and E. Higgins (2018). Adverse selection in mortgage securitization: Evidence from loss given default. W*orking paper*.

Table 1: Summary Statistics

	Portfolio Loan		Securitized Loan	
Variable	Mean	Std Dev	Mean	Std Dev
FICO	7.107	0.607	6.734	0.678
LTV	0.782	0.108	0.800	0.091
Second Lien	0.133	0.339	0.265	0.441
Full Doc	0.556	0.497	0.563	0.496
Low Doc	0.444	0.497	0.437	0.496
DTI	32.574	14.628	37.893	11.731
Jumbo	0.380	0.485	0.258	0.437
Owner Occupy	0.719	0.450	0.809	0.393
Term30	0.880	0.325	0.887	0.317

Table 2: Pearson Correlation Coefficients between Default Probability and Loan Loss Rates

Track Time	Loss36m	Loss48m	Loss60m
Default12m	0.245	0.186	0.163
Default24m	0.396	0.366	0.337
Expct Default12m	0.173	0.162	0.150
Expct Default24m	0.245	0.252	0.244

Table 3: Loan Loss Rates

Variable	Portfolio Loan		Securitized Loan	
	Mean	Std Dev	Mean	Std Dev
Full Sample				
Loss36m	0.007	0.055	0.023	0.101
Loss48m	0.016	0.086	0.042	0.135
Loss60m	0.023	0.103	0.053	0.152
Prime Sample				
Loss36m	0.004	0.043	0.008	0.060
Loss48m	0.010	0.067	0.020	0.095
Loss60m	0.016	0.087	0.032	0.119
Subprime Sample				
Loss36m	0.018	0.087	0.036	0.124
Loss48m	0.043	0.135	0.059	0.159
Loss60m	0.051	0.148	0.070	0.172
FICO $\geq$ 620				
Loss36m	0.007	0.053	0.023	0.101
Loss48m	0.015	0.083	0.043	0.137
Loss60m	0.022	0.100	0.056	0.155
FICO $<$ 620				
Loss36m	0.012	0.070	0.023	0.101
Loss48m	0.029	0.114	0.037	0.129
Loss60m	0.037	0.129	0.045	0.141

Table 4: Lender's Securitization Decision - First Stage Tobit Regression Full Sample

	Prime Loan			Subprime Loan		
	(1)	(2)	(3)	(4)	(5)	(6)
Variable	36m	48m	60m	36m	48m	60m
Intercept	−1.1202	−0.6917*	−0.5913*	−0.7864**	−0.7598**	−0.6994**
	(0.5967)	(0.3405)	(0.2803)	(0.1498)	(0.1142)	(0.0977)
FICO	−0.3133**	−0.2893**	−0.2520**	−0.1238**	−0.0746**	−0.0516**
	(0.0177)	(0.0118)	(0.0094)	(0.0078)	(0.0062)	(0.0057)
LTV	1.6137**	1.5684**	1.3869**	0.8265**	0.6238**	0.5360*
	(0.1226)	(0.0801)	(0.0618)	(0.0537)	(0.0430)	(0.0395)
Second Lien	0.1658**	0.1234**	0.1180**	−0.0718**	−0.0203*	−0.0023
	(0.0164)	(0.0114)	(0.0092)	(0.0109)	(0.0088)	(0.0081)
DTI	0.0016*	0.0011**	0.0010**	0.0035**	0.0038**	0.0035*
	(0.0006)	(0.0004)	(0.0003)	(0.0004)	(0.0003)	(0.0003)
Low Doc	0.1511**	0.1603**	0.1310**	0.1599**	0.1416**	0.1289*
	(0.0166)	(0.0113)	(0.0089)	(0.0089)	(0.0071)	(0.0067)
Jumbo	−0.2294**	−0.1983**	−0.1629**	0.0260*	0.0034	−0.0125
	(0.0177)	(0.0121)	(0.0097)	(0.0102)	(0.0085)	(0.0080)
Owner Occupy	0.0183	−0.0481**	−0.0642**	−0.2546**	−0.2811**	−0.2793**
	(0.0193)	(0.0125)	(0.0100)	(0.0093)	(0.0076)	(0.0071)
Term30	0.0231	−0.0721**	−0.0803**	−0.1341**	−0.1484**	−0.1288**
	(0.0273)	(0.0187)	(0.0157)	(0.0094)	(0.0077)	(0.0073)
Sigma	0.7393**	0.7057**	0.6726**	0.5984**	0.5642**	0.5546*
	(0.0147)	(0.0091)	(0.0069)	(0.0054)	(0.0040)	(0.0036)
State Effect	Y	Y	Y	Y	Y	Y
Orig Quarter	Y	Y	Y	Y	Y	Y
N	107162	107162	107162	99247	99247	99247
R2	0.2297	0.2002	0.1795	0.1954	0.2002	0.1916

Notes: This table reports the coefficient estimates and standard errors of the Tobit model regression (first stage regression as in Equation (1)) of loan loss rates for the prime and subprime mortgages. The dependent variable is loan loss rate. Loan loss rates are tracked 36 months, 48 months, and 60 months after origination. ** p<0.01, * p<0.05.

Table 5: Lender's Securitization Decision - Second Stage Logit Regression Full Sample

	Prime Loan			Subprime Loan		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A	36m	48m	60m	36m	48m	60m
Intercept	1.9193^{**}	2.0750^{**}	2.1171^{**}	9.6156^{**}	9.5666^{**}	9.6667^{**}
	(0.6498)	(0.6499)	(0.6499)	(1.1703)	(1.1709)	(1.1711)
Expct Loss	27.7881^{**}	13.2156^{**}	9.7413^{**}	-2.7636^{**}	-2.3079^{**}	-2.1723^{**}
	(1.5678)	(0.7149)	(0.5046)	(0.6459)	(0.3873)	(0.3546)
Jumbo	0.0048	0.0083	0.0084	-0.2068^{**}	-0.1709^{*}	-0.1744^{*}
	(0.0261)	(0.0262)	(0.0262)	(0.0682)	(0.0683)	(0.0679)
Yield Spread	0.4025^{**}	0.3825^{**}	0.3762^{**}	-1.2636^{**}	-1.2598^{**}	-1.2531^{**}
	(0.0564)	(0.0562)	(0.0562)	(0.1138)	(0.1133)	(0.1131)
Credit Spread	0.2725	0.1234	0.1433	-10.0946^{**}	-10.0385^{**}	-10.0604^{**}
	(0.3443)	(0.3456)	(0.3457)	(0.6212)	(0.6215)	(0.6219)
Yield Curve	−0.4238	-0.4853^{*}	-0.5152^{*}	-7.0354^{**}	-7.0212^{**}	-7.0497^{**}
	(0.2196)	(0.2196)	(0.2194)	(0.3715)	(0.3707)	(0.3711)
Sigma Int	1.2230^{**}	1.0844^{**}	0.9435^{**}	6.7646^{**}	6.8232^{**}	6.8169^{**}
	(0.2887)	(0.2877)	(0.2873)	(0.4731)	(0.4725)	(0.4722)
N Securitized	19823	19823	19823	24217	24217	24217
N Portfolio	8806	8806	8806	2119	2119	2119
R2	0.0179	0.0177	0.0183	0.0917	0.0928	0.0929

Notes: This table reports the coefficient estimates and standard errors of the Logit model regression (second stage regression as in Equation (4) and (5)) of the securitization decision for the prime and subprime mortgages. The dependent variable is the securitization status. Expected loan losses are out-of-sample estimation using the coefficient estimates from Table 4. ** p<0.01, * p<0.05.

Table 6: Lender's Securitization Decision - Robustness Checks on High and Low FICO Score Samples

	FICO $\geq$ 620			FICO < 620		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A	36m	48m	60m	36m	48m	60m
Intercept	5.4480**	5.5370**	5.4541**	9.3575**	9.2550**	9.3544**
	(0.5754)	(0.5757)	(0.5760)	(1.9008)	(1.9021)	(1.9027)
Expct_Loss	12.5324**	7.8523**	7.0776**	−3.9408**	−3.6775**	−3.5143**
	(0.6070)	(0.3514)	(0.2955)	(1.4730)	(0.8222)	(0.7034)
Jumbo	−0.3338**	−0.3272**	−0.3205**	−0.7535**	−0.7029**	−0.6876**
	(0.0237)	(0.0238)	(0.0238)	(0.1159)	(0.1159)	(0.1160)
Yield Spread	0.3121**	0.3020**	0.3077**	−1.2800**	−1.2864**	−1.2779**
	(0.0501)	(0.0500)	(0.0501)	(0.1844)	(0.1840)	(0.1837)
Credit Spread	−2.1897**	−2.2610**	−2.2522**	−8.6208**	−8.5701**	−8.5842**
	(0.3054)	(0.3060)	(0.3060)	(0.9649)	(0.9652)	(0.9662)
Yield Curve	−2.1715**	−2.2101**	−2.1597**	−6.5651**	−6.5402**	−6.5646**
	(0.1932)	(0.1932)	(0.1933)	(0.5965)	(0.5954)	(0.5960)
Sigma Int	2.2746**	2.1570**	2.1487**	4.4007**	4.4807**	4.4792**
	(0.2514)	(0.2513)	(0.2516)	(0.7614)	(0.7615)	(0.7614)
R2	0.0304	0.0316	0.0329	0.0922	0.0942	0.0949
N Securitized	34018	34018	34018	10022	10022	10022
N Portfolio	10079	10079	10079	846	846	846

Notes: This table reports the coefficient estimates and standard errors of the Logit model regression (second stage regression as in Equation (4) and (5)) of the securitization decision for the high and low credit score mortgages. The dependent variable is the securitization status. Expected loan losses are out-of-sample estimation using the coefficient estimates from the corresponding Tobit regression (first stage regression). ** p<0.01, * p<0.05

Table 7: Lender's Securitization Decision - Sub Sample Robustness Checks

Panel A: Prime	Jumbo=1	Jumbo=0	OrigY 2005	OrigY 2006
		Prime 36m		
Expct Loss	34.0421**	25.2090**	57.8270**	26.7480**
	(3.0190)	(1.8058)	(5.5902)	(1.6303)
N Securitized	8610	11213	12803	7020
N Portfolio	3891	4915	5835	2971
		Prime 48m		
Expct Loss	15.1119**	12.3886**	17.1722**	12.6904**
	(1.3483)	(0.8383)	(1.6070)	(0.7989)
N Securitized	8610	11213	12803	7020
N Portfolio	3891	4915	5835	2971
		Prime 60m		
Expct Loss	11.3245**	9.0790**	11.2495**	9.4513**
	(0.9528)	(0.5917)	(0.9660)	(0.5924)
N Securitized	8610	11213	12803	7020
N Portfolio	3891	4915	5835	2971

Panel B: Subprime	Jumbo=1	Jumbo=0	OrigY 2005	OrigY 2006
		Subprime 36m		
Expct Loss	−0.5503	−3.4742**	−3.6114**	−1.3335
	(1.2833)	(0.7514)	(0.8190)	(1.1610)
N Securitized	2832	21385	13587	10630
N Portfolio	333	1786	1866	253
		Subprime 48m		
Expct Loss	−0.5244	−2.8210**	−2.5392**	−1.5125
	(0.8230)	(0.4372)	(0.4539)	(0.8008)
N Securitized	2832	21385	13587	10630
N Portfolio	333	1786	1866	253
		Subprime 60m		
Expct Loss	−0.3638	−2.6260**	−2.2920**	−1.6740*
	(0.7973)	(0.3942)	(0.4115)	(0.7380)
N Securitized	2832	21385	13587	10630
N Portfolio	333	1786	1866	253

Notes: This table reports the coefficient estimates and standard errors of the Logit model regression (second stage regression as in Equation (4) and (5)) of the securitization decision for various subsamples. The dependent variable is the securitization status. Expected loan losses are out-of-sample estimation using the coefficient estimates from the corresponding Tobit regression (first stage regression). ** p<0.01, * p<0.05

Table 8: Lender's Securitization Decision - Robustness Checks Using Repurchased Loan Sample

	Prime Loan			Subprime Loan		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A	36m	48m	60m	36m	48m	60m
Intercept	4.4800	4.1999	4.1502	5.3304	4.7691	4.5221
	(2.7149)	(2.7127)	(2.7142)	(3.0799)	(3.0847)	(3.0814)
Expct Loss	13.6522**	7.7375**	6.4757**	−15.2561**	−9.9472**	−7.9856**
	(3.8550)	(2.4886)	(1.8216)	(3.1066)	(1.7293)	(1.4536)
Jumbo	−0.4949**	−0.5008**	−0.4962**	−0.0175	0.0232	−0.0376
	(0.1162)	(0.1161)	(0.1161)	(0.1782)	(0.1781)	(0.1763)
Yield Spread	0.4201	0.3652	0.3585	0.2952	0.3452	0.3375
	(0.2369)	(0.2354)	(0.2354)	(0.2713)	(0.2703)	(0.2702)
Credit Spread	1.0124	0.9970	0.9849	−1.4298	−1.2502	−1.2457
	(1.5860)	(1.5805)	(1.5802)	(1.7427)	(1.7385)	(1.7389)
Yield Curve	−3.0211**	−2.9837**	−2.9740**	−6.3286**	−5.9087**	−5.8036**
	(0.9554)	(0.9554)	(0.9544)	(1.1394)	(1.1377)	(1.1366)
Sigma Int	−5.1016**	−5.1129**	−5.1321**	4.8018**	5.1064**	5.2524**
	(1.1028)	(1.1008)	(1.1008)	(1.6487)	(1.6355)	(1.6302)
R2	0.0572	0.0565	0.0574	0.0808	0.0875	0.0852
N Repurchased	368	368	368	391	391	391
N Portfolio	8267	8267	8267	2022	2022	2022

Notes: This table reports the coefficient estimates and standard errors of the Logit model regression (second stage regression as in Equation (4) and (5)) of the securitization decision for the repurchased loan sample. Repurchased loan sample includes repurchased loans and portfolio loans. The dependent variable is the securitization status. Expected loan losses are out-of-sample estimation using the coefficient estimates from the corresponding Tobit regression (first stage regression). ** p<0.01, * p<0.05

Table 9: Regression Discontinuity - Portfolio Loan and Securitized Loan

		Loan Loss Rate (%)			Loan Loss ($)		
Sample	Variable	Estimate	P-Value	R2	Estimate	P-Value	R2
Portfolio Whole	T	0.0045	0.6205	0.5406	2339.916	0.4295	0.4411
Portfolio Prime	T	0.0110	0.4035	0.2198	8840.002	0.3190	0.1079
Portfoio Subprime	T	0.0028	0.9040	0.1276	1188.461	0.8858	0.1483
Securitized Whole	T	0.0255**	0.0000	0.9245	8003.848**	0.0000	0.9246
Securitized Prime	T	0.0066	0.3884	0.7080	1343.180	0.6049	0.7149
Secutitized Subprime	T	0.0273**	0.0034	0.7473	8008.189*	0.0108	0.7930

Notes: This table reports the estimates of the regression discontinuity as in Equation (6). The dependent variable is the loan loss rate and dollar loss amount. T equals to one if credit score is greater than or equal to 620. Otherwise, T equals to zero. Loan performances are tracked 60 months after origination. We use the 7th order polynomial in the regression. ** $p<0.01$, * $p<0.05$

Table 10: Regression Discontinuity - Robustness Checks for Securitized Subprime Loan

			Loan Loss Rate (%)			Loan Loss ($)		
Track Time	Polynomial	Variable	Estimate	P-Value	R2	Estimate	P-Value	R2
36m	3rd	T	0.0158**	0.0005	0.6655	4892.161**	0.0004	0.7265
36m	5th	T	0.0186**	0.0019	0.6711	5885.713**	0.0014	0.7291
36m	7th	T	0.0167**	0.0041	0.6712	5236.723**	0.0033	0.7292
48m	3rd	T	0.0202**	0.0025	0.7038	6137.314**	0.0077	0.7413
48m	5th	T	0.0316**	0.0004	0.7096	10058.980**	0.0010	0.7497
48m	7th	T	0.0260**	0.0026	0.7130	7602.198**	0.0094	0.7560
60m	3rd	T	0.0190**	0.0087	0.7394	5718.180*	0.0206	0.7798
60m	5th	T	0.0329**	0.0006	0.7457	10634.260**	0.0012	0.7882
60m	7th	T	0.0273**	0.0034	0.7473	8008.189*	0.0108	0.7930

Notes: This table reports the estimates of the regression discontinuity as in Equation (6) for securitized subprime loans. The dependent variable is the loan loss rate. We also use the dollar loss amount as the dependent variable for robustness check purpose. T equals to one if credit score is greater than or equal to 620. Otherwise, T equals to zero. Loan performances are tracked 36, 48 and 60 months after origination. We use the 3rd, 5th, and 7th order polynomial in the regressions. ** p<0.01, * p<0.05

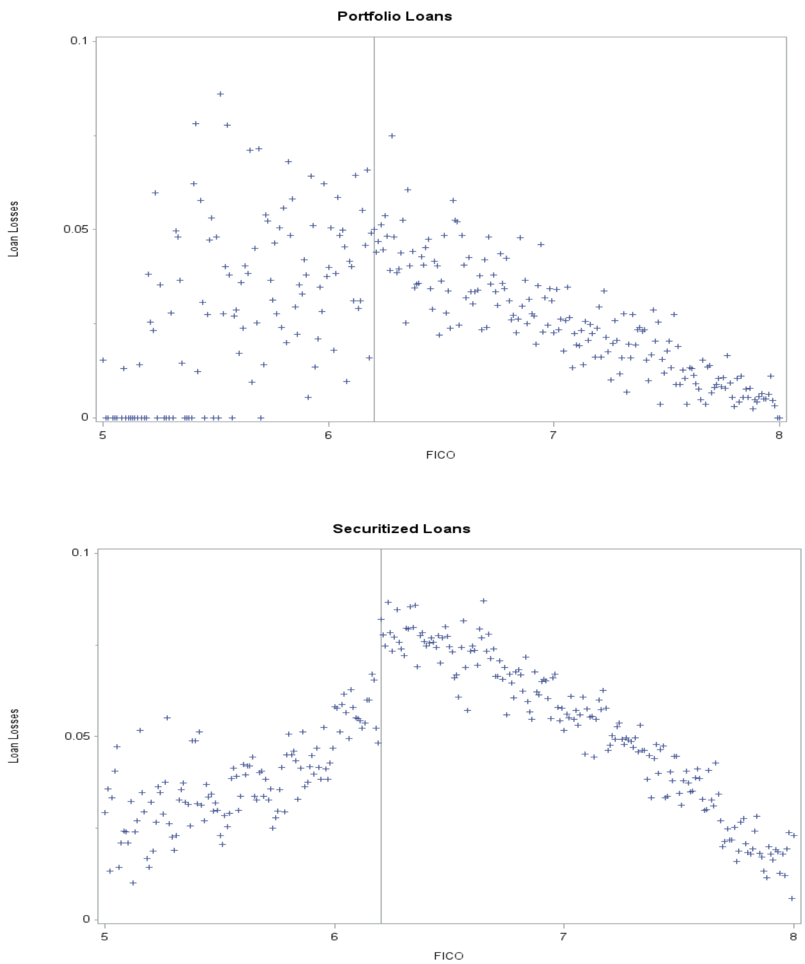

Figure 1: Loan Loss Rate by FICO Score - Portfolio versus Securitized Loans

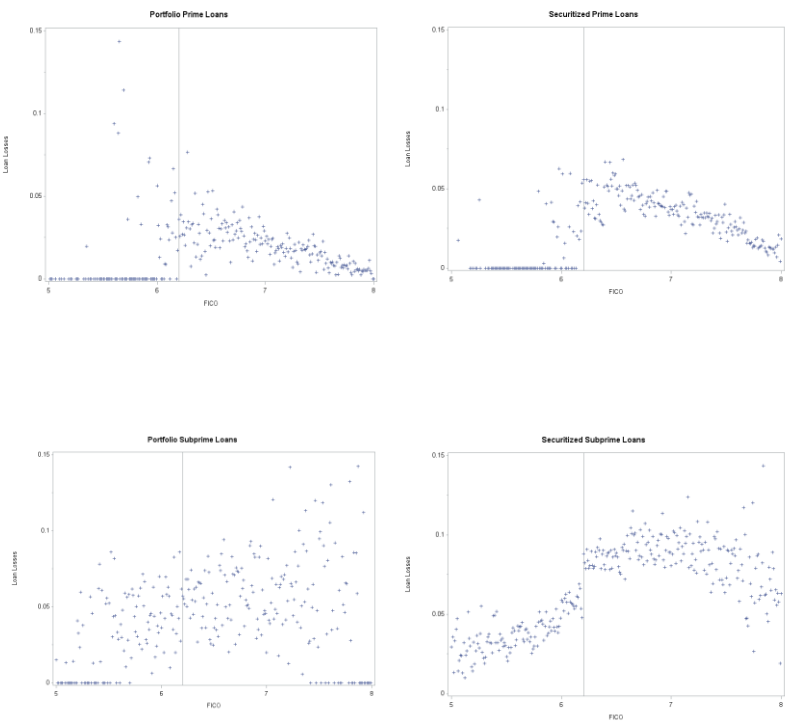

Figure 2: Loan Loss Rate by FICO Score - Portfolio versus Securitized and Prime versus Subprime Loans