

Bubb, Ryan; Kaufman, Alex

**Working Paper**

## Further investigations into the origin of credit score cutoff rules

Working Papers, No. 11-12

**Provided in Cooperation with:**

Federal Reserve Bank of Boston

*Suggested Citation:* Bubb, Ryan; Kaufman, Alex (2011) : Further investigations into the origin of credit score cutoff rules, Working Papers, No. 11-12, Federal Reserve Bank of Boston, Boston, MA

This Version is available at:

<https://hdl.handle.net/10419/55572>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Further Investigations into the Origin of Credit Score Cutoff Rules

Ryan Bubb and Alex Kaufman

**Abstract:**

Keys, Mukherjee, and Vig (2010a) argue that the evidence presented in Bubb and Kaufman (2009) is based on an inappropriate pooling of loans sold to private-label securitizers with loans sold to the government sponsored enterprises (GSEs). In this paper we investigate the issues raised by the authors and conclude that they do not change our basic analytical approach or conclusions. We examine samples that do not pool together loans sold to these two types of purchasers—a sample of loans bought by the GSEs, a sample of loans originated in 2008–2009 after the private-label market collapsed, and a sample of jumbo loans—and find discontinuities in the number and default rate of loans at credit score cutoffs in the absence of corresponding discontinuities in the securitization rate. We also examine a key assumption underlying their estimates—that no loans are both at risk of being sold to the GSEs and at risk of being sold to private-label securitizers—and show that the data are inconsistent with that assumption. We find that 18 percent of conforming loans in our sample at some time switched between GSE and private-label ownership, demonstrating that the GSEs and private-label securitizers competed for the same loans. Additionally, we show that lender screening cutoffs grew steadily over the period 1997–2010 during which the private-label market rose and collapsed.

**JEL Classifications:** D82, G01, G18, G21, G24, G28.

---

Ryan Bubb is an assistant professor of law at the New York University School of Law. His email address is [ryan.bubb@nyu.edu](mailto:ryan.bubb@nyu.edu). Alex Kaufman is an economist at the Board of Governors of the Federal Reserve System. His email address is [alex.kaufman@frb.gov](mailto:alex.kaufman@frb.gov).

This paper presents preliminary analysis and results intended to stimulate discussion and critical comment. The views expressed herein are those of the authors and do not indicate concurrence by other members of the research staff or principals of the Board of Governors, the Federal Reserve Bank of Boston, or the Federal Reserve System.

We thank Larry Katz, Justin McCrary, Holger Spamann, Paul Willen, and numerous seminar participants for valuable comments and discussions. We are grateful to the Federal Reserve Bank of Boston and the Federal Reserve Bank of New York for supporting this research.

**This version:** August 30, 2011.

## 1. INTRODUCTION

In an important paper, Keys, Mukherjee, Seru, and Vig (2010b) (hereafter, KMSV) used variation in the default rate of loans by FICO score in a regression discontinuity design to investigate whether securitization led to lax screening by lenders. In Bubb and Kaufman (2009) (hereafter, BK) we examined the theory underlying this research design—which we refer to as the securitization rule-of-thumb theory—and found that both institutional evidence and quantitative evidence are inconsistent with the theory. In a response paper, Keys, Mukherjee, Seru, and Vig (2010a) (hereafter, KMSV2) argue that the quantitative evidence presented in BK is based on an inappropriate pooling of loans sold to private-label securitizers with loans sold to the Government-Sponsored Enterprises (GSEs).<sup>1</sup> Moreover, the authors argue that the 620 FICO rule of thumb was used only for low-doc loans, not for full-doc loans, and hence that those two types of loans should not be pooled together.

We investigated the issues raised in KMSV2 and detail our findings in this paper. Our ultimate conclusion is that these issues do not affect our basic analytical approach or change our conclusions from BK.

We begin in Section 2 by reviewing the two competing theories for the origin of lender credit score cutoff rules.

In Section 3 we consider how best to select a sample from the Lender Processing Services Applied Analytics (LPS) dataset to implement KMSV’s regression discontinuity design, and in particular the pooling issues raised by KMSV2. We begin with KMSV’s selection of a sample of only low documentation loans for their main analysis. Whether or not a lender requires full documentation of income and assets is a measure of the outcome of interest, lender screening. By selecting on the outcome, KMSV introduces posttreatment selection bias into their regression discontinuity design. Valid causal inference in a regression discontinuity design requires pooling together units with all different values of the outcome—in this case lender screening—as was done in BK. Subgroup analysis is still possible, but selection must be done on pre-treatment covariates.

We then consider whether GSE-securitized loans and private-label securitized loans should be pooled together in the conforming sample. We first show that many samples that do not pool GSE securitized loans with private-label securitized loans still exhibit default discontinuities in

---

<sup>1</sup>An abridged version of that response paper, Keys, Seru, and Vig (2011), is now forthcoming in the *Review of Financial Studies*.

the absence of securitization discontinuities. In particular we examine loans sold to the GSEs, loans originated in 2008–2009 after the private-label market shut down, and jumbo loans, which the GSEs are forbidden from securitizing. These results confirm the findings in BK that lender cutoff rules in screening are not exclusively caused by a purchasing rule-of-thumb followed by private-label securitizers. Moreover, our findings provide further evidence against the identification assumptions underlying KMSV’s research design. Lenders change their screening behavior discontinuously at these credit score cutoffs for reasons other than a change in the probability of securitization. Consequently, the jumps in default at these cutoffs are not evidence that securitization led to lax screening by lenders.

Next, we examine the assumptions underlying KMSV2’s separation of loans in the LPS data into agency and non-agency markets. The authors assume that any given loan is at risk of being sold either to the GSEs or to private-label securitizers, but that no loan is simultaneously at risk of being sold to both. We look to the data and find that 18 percent of our 2003–2007 conforming sample were at one point owned by the GSEs and at another point owned by private-label securitizers. This is a lower bound on the overlap between the two types of purchasers—many loans that were both at risk of being bought by private-label securitizers and by the GSEs were in fact held by only one type of buyer. Moreover, the institutional evidence shows that the GSEs and private-label securitizers actively competed over subprime loans in the secondary market in this period.

The fact that loans were simultaneously at risk of being sold to both GSEs and private-label securitizers implies that the discontinuity estimands in KMSV2, which are premised on separation of the two markets, are not well defined. GSE loans and private-label securitized loans, as well as portfolio loans, must be pooled together to estimate the probability of securitization at any given FICO score, as was done in BK.

In Section 4 we investigate the time-series evolution of the 620 FICO cutoff by estimating the size of the discontinuity at 620 FICO for different cohorts of loans by origination date. We find evidence that the 620 FICO cutoff existed for loans originated in the period 1997–2000, after the GSEs directed originators to adopt the 620 cutoff, and it grew steadily for later cohorts of loans through at least 2010. During this time the market for private-label securities rose, peaked in 2005–2006, and then, in 2008, collapsed to nearly nothing, where it has remained. The disconnect between the time-series pattern of the size of the 620 cutoff and the scale of the private-label

mortgage-backed securities (MBS) market is further evidence that private-label securitization was not the driving force, and certainly not the only force, behind lenders' use of the cutoff rule in screening.

In Section 5 we consider how best to measure ease of securitization. We find that there are no economically significant discontinuities in time-to-securitize or loan buy-backs at 620. We also examine KMSV2's argument that the "unconditional probability of securitization," as measured by the *number* of securitized loans, is a better measure of "ease of securitization" than is the probability of selling a loan. The number of securitized loans is the product of the total number of loans made and the securitization rate. What matters for a lender's incentive to screen is the probability that the lender sells the loan: the securitization rate. If we see the number of securitized loans jump at 620, it could be because either the lending rate jumps or because the securitization rate jumps. We find that the jump in the number of securitized loans first found in KMSV is driven by a jump in the lending rate, not the securitization rate.

To illustrate the consequences of using the number of securitized loans as a measure of ease of securitization, we show in Figure 8 panels (a) and (d) that there are jumps in the number and default rate of portfolio loans at 620. Using KMSV's "unconditional probability" approach, one would interpret the jump in number of portfolio loans as a jump in originators' skin in the game. Hence, the corresponding jump in defaults would be interpreted as evidence for the hypothesis that a *lack* of securitization reduced lenders' incentives to screen. In fact, the jump in the number of portfolio loans is due to the same increase in the overall lending rate that caused the jump in the number of private-label securitized loans.

In Section 6 we discuss the magnitudes of the estimates in KMSV2 and conclude that the changes in securitization rate at 620 that the authors estimate are too small to plausibly be the cause of the corresponding changes in default rate. For example, using data from a single mortgage originator, KMSV2 estimates a jump in the securitization rate at 620 of 1.7 percentage points, from a base of over 90 percent, and a corresponding 10 percentage point jump in the default rate, from a base of about 12 percent. Such an outsized effect of a 1.7 percentage point increase in securitization seems implausible. In contrast, these magnitudes are consistent with the origination rule-of-thumb theory. Originators change their screening behavior at 620, which in turn changes how easy it is to sell loans.

In Section 7 we discuss KMSV2’s analysis of the evidence from state anti-predatory lending laws. KMSV2 argues that these laws only had a differential effect on securitization above and below 620, not a differential effect on default directly. Hence, the authors argue, the laws provide a test of whether originators followed a credit score cutoff rule in screening for reasons other than a change in securitization. However, provisions in these laws were aimed directly at reducing defaults. For example, the Georgia Fair Lending Act required lenders to certify that the borrower has received counseling on the advisability of the loan before making certain loans. Because there are more defaults above 620 than below, the laws can be expected to have a differential effect on default above and below 620. Moreover, we show that there are discontinuities in the terms of mortgage contracts at 620 for a range of contract terms associated with default, including loan-to-value ratios, debt-to-income ratios, and whether the loan is interest only or allows for negative amortization. Thus the passage of the laws does not provide a test of the two competing theories for the origin of the credit score cutoff rules.

Section 8 concludes.

## 2. THE TWO CREDIT SCORE CUTOFF RULE THEORIES

We begin by briefly reviewing the two competing theories for the origin of credit score cutoff rules and their implications for using the cutoff rule to investigate the effect of securitization on lender screening. KMSV developed what we refer to as the *securitization rule-of-thumb theory*.<sup>2</sup> This theory posits that private-label securitizers, following the advice of the GSEs, used a rule of thumb in their purchasing decisions that made securitizers more willing to buy mortgage loans made to borrowers with FICO scores of 620 or above than loans to those with scores of 619 or below. The authors argue that this rule of thumb resulted in exogenous variation in securitization that can be used in a regression discontinuity (RD) research design to test the hypothesis that securitization led to lax screening by lenders. In particular, for loans made to borrowers at 620 the lender has a discontinuously higher probability of being able to sell the loan than for loans made at 619. They argue that any discontinuous change in the probability of default at 620 can be ascribed to the discontinuous change in the probability of securitization.

However, there are two key identifying assumptions for this research design. First, there must be a discontinuous increase in the probability of securitization at 620 FICO. In instrumental variables

---

<sup>2</sup>In the original version of BK we referred to the theory as the “securitizer-first theory.”

(IV) terms, we must have a valid first stage. Second, there must be no reason for the default rate to discontinuously change at the 620 threshold other than in response to a change in the ease of securitization. This assumption can also be framed as an IV exclusion restriction: the 620 cutoff must have an effect on defaults only through its effect on the probability of securitization. If instead there is another reason for the default rate to discontinuously change at 620, then one cannot estimate the causal effect of any change in probability of securitization at 620 by estimating the change in default rates at 620.

In BK we developed an alternative *origination rule-of-thumb theory* for the origin of credit score cutoff rules: originators adopted the credit score cutoff rules directly in response to the underwriting guidelines of the GSEs that required them to do so.<sup>3</sup> We provided institutional evidence for this theory from the history of the adoption of credit score cutoff rules in mortgage underwriting. We showed that the GSEs contractually bound lenders to adopt credit score cutoff rules long before private securitization grew to a substantial scale, and that the 620 cutoff spread through its incorporation into widely used underwriting software. Cutoff rules were a natural form for lender screening rules to take, given fixed costs in the information production function of lenders. This evidence implies a violation of the identification assumption in KMSV, implying that the jump in defaults at 620 does not provide evidence that securitization led to lax screening. Even if securitizers followed a purchasing rule of thumb at 620, any resulting incentive effect would be confounded by originators' independent use of cutoff rules in screening and the discontinuous change in GSE guidance and underwriting software output at 620.<sup>4</sup>

In BK we also looked to quantitative evidence to test the securitization rule-of-thumb theory and the origination rule-of-thumb theory. The primary dataset used in KMSV included only securitized loans, and KMSV did not show whether there was a jump in the fraction of loans securitized at 620.<sup>5</sup> Instead, KMSV argued that the number of securitized loans at any given FICO score is a

---

<sup>3</sup>In the original version of BK, we referred to this as the “lender-first theory.”

<sup>4</sup>KMSV explicitly considered the possibility that lenders followed a cutoff rule in screening for reasons other than a discontinuous change in the probability of securitization and attempted to reject it using variation induced by the passage of state anti-predatory lending laws in Georgia and New Jersey in 2002 and 2003, respectively. However, we showed in BK both theoretical and empirical reasons why these anti-predatory lending laws do not provide a way to test the theory.

<sup>5</sup>KMSV (pp. 345-347) does use the LPS dataset that we use—which contains both securitized and portfolio loans—to examine whether there is a discontinuity in GSE loans at 620, but does not use it to report any first stage showing whether the fraction of securitized loans jumps at 620. KMSV does report that the *number* of securitized loans in their sample jumps at 620, but without knowing how large the corresponding jump in the number of non-securitized loans is, this is uninformative about the probability of securitization.

good proxy for “ease of securitization” of loans at that FICO score and showed that there is a jump in the number of loans at 620.

In BK we used a different dataset—the LPS dataset—that includes loans with three main types of owners: (1) portfolio loans retained by the originator; (2) loans purchased by the GSEs; and (3) private-label securitized loans. The advantage of these data over the main dataset used in KMSV is that we could investigate whether there is a jump in the probability of securitization at 620. We estimated the securitization probability at each FICO score as the fraction of loans that are either sold to the GSEs or sold to private-label securitizers (or equivalently as 1 minus the fraction of portfolio loans). We showed that in the “conforming sample” of loans for amounts below the size limits set for the GSEs, the default rate jumps at 620 but the securitization rate does not. We found the same pattern among low-documentation (“low-doc”) loans. In essence, we showed samples where there is a strong reduced form effect of the cutoff on default but no first-stage effect of the cutoff on probability of securitization. This is strong evidence against the securitization rule-of-thumb theory and the identification assumptions underlying KMSV’s RD research design.

### 3. SELECTING A SAMPLE FOR REGRESSION DISCONTINUITY DESIGN

A key issue raised in KMSV2 is the appropriate sample to select from LPS to implement their regression discontinuity design. KMSV2 argues that the results in BK are based on an inappropriate pooling of loans sold to private-label securitizers with loans sold to the GSEs. The authors assert that there are two separate markets: (1) the “non-agency” market of loans only at risk of being purchased by private-label securitizers and not by the GSEs (which they also refer to as the “subprime” market); and (2) the “agency” market of loans only at risk of being purchased by the GSEs and not by private-label securitizers (which they also refer to as the “prime” market). They assume that loans are either in one market or the other and that no loan is both at risk of being bought by private-label securitizers and by the GSEs. Moreover, they argue that the 620 FICO rule of thumb was used only by private-label securitizers and not by the GSEs, and only for low-doc loans, not for full-doc loans (KMSV2, p. 9).<sup>6</sup>

---

<sup>6</sup>There is a tension between KMSV2’s view that the 620 purchasing rule of thumb was only used by private-label securitizers and not the GSEs, and KSMV’s institutional story for the origin of the rule of thumb. KMSV argues that the rule-of-thumb originated with advice from the GSEs. To believe their story, you have to believe that (1) the GSEs’ direction to lenders to adopt the 620 rule of thumb had no direct effect on lender behavior, (2) the GSEs themselves did not follow a *purchasing* rule of thumb at 620, and (3) private-label securitizers did listen to the 620 guidance the



Based on these premises, KMSV2 separates loans from the LPS dataset used in BK into subgroups defined by their two-by-two categorization of low-doc vs. full-doc and agency vs. non-agency markets. The authors assign loans bought by private-label securitizers to the non-agency market and assign loans bought by the GSEs to the agency market. They use several different approaches to assign portfolio loans to one market or the other. They find that the only subgroup with jumps in either securitization or default is the low-doc non-agency subgroup (KMSV2, pp. 7–9). They argue that the pattern found in BK—discontinuities in the number of loans and default without a discontinuity in securitization—was an artifact of pooling agency loans with non-agency loans. The default discontinuities found in BK, they argue, are “driven solely by low-documentation non-agency loans” (KMSV2, p. 9), while the securitization discontinuity among non-agency loans was attenuated by pooling those loans with GSE loans (KMSV2, p. 14).

In this section we investigate these issues raised in KMSV2. In Section 3.1 we consider KMSV’s selection of only low-doc loans to estimate the effect of the discontinuity in securitization on lender screening behavior at 620. In 3.2 we examine samples that do not pool private-label securitized loans with GSE securitized loans and find that cutoff rules in screening exist in these samples in the absence of a corresponding discontinuity in the securitization rate. In 3.3 we analyze more formally KMSV2’s and BK’s approaches to estimating the discontinuity in securitization probability. In 3.4 we investigate the key assumption underlying KMSV2’s approach—that the agency and non-agency markets for these loans are completely separate—and find that the evidence shows that the GSEs and private-label securitizers actively competed to buy the same loans. In particular, we find that 18 percent of our 2003–2007 conforming sample were at one point owned by the GSEs and at another point owned by private-label securitizers.

**3.1. Selecting on documentation status.** We begin by considering econometric issues raised by KMSV and KMSV2’s use of samples that contain only low-doc loans to estimate the effect of discontinuities in securitization on lender screening behavior. KMSV and KMSV2 restrict their main analysis samples to low-doc loans because, the authors argue, moral hazard is a bigger risk due to these loans’ lack of hard information.<sup>7</sup> Hence, they expect that securitization has a larger effect on lender screening for low-doc loans than for full-doc loans.

---

GSEs gave to lenders about screening and adopted a rule-of-thumb in purchasing at 620. We find this chain of events implausible, even without considering the institutional and quantitative evidence against it presented in BK.

<sup>7</sup>See, e.g., KMSV, pp. 309-310.

However, the amount of documentation required by lenders is an important aspect of lender screening, which is the outcome variable of interest. Indeed, documentation status is a more direct measure of lender screening than is loan performance. Figure 1 shows that the fraction of low-doc loans jumps at the FICO scores of 620 and 660.

Selecting the sample based on the outcome when estimating a causal effect results in biased estimates—a problem referred to as posttreatment selection bias (Frangakis and Rubin, 2002). Valid causal inference in a regression discontinuity design requires pooling together units with all different values of the outcome—in this case lender screening—as was done in BK. Much as in the analysis of data from a randomized experiment, it is possible to do subgroup analysis in a regression discontinuity design to estimate heterogeneous treatment effects, but the subgroups must be defined by plausibly pretreatment covariates. Hence we consider subgroups only by loan size, origination year, and FHA / VA (Federal Housing Administration / Veterans Administration) status, and offer a breakdown by documentation status only for comparability with KMSV’s estimates.

In addition to this general econometric problem caused by selecting on the outcome, KMSV’s selection of low-doc loans raises a specific concern: the change in lender screening at 620 may shift the relative costs and benefits to loan applicants of providing full documentation vs. low documentation. To give one example of this form of selection, suppose that for a certain type of riskier-than-average loan applicant, providing full documentation of income and assets is especially costly. If these applicants are above the 620 cutoff, where lenders are less cautious, they have a good chance of receiving a loan even if they do not provide full documentation. In contrast, below the 620 cutoff, where lenders are more careful, it may become worth it to these applicants to fully document their income and assets for the lender. When considering a sample of only low-doc loans, this shift in the composition of which applicants provide low documentation may confound changes in the default rate stemming from lender screening itself.

The data provide evidence that this form of selection is indeed taking place. For example, in the 2003–2007 sample examined in Bubb and Kaufman (2011), right above 620, 30 percent of low-doc loans default, while only 26 percent of full-doc loans. In contrast, right below 620, low-doc loans are actually *less* risky than full-doc loans, with only 22 percent defaulting compared with 25 percent of full-doc loans. It appears that relatively riskier borrowers select into low-doc loans rather than full-doc loans above the cutoff, as compared with below the cutoff.

Thus, while it may indeed be the case that soft information is more important for low-doc loans, selecting on documentation status results in estimates that cannot be interpreted causally. Valid causal inference in a regression discontinuity design requires pooling together full-doc and low-doc loans.

### **3.2. Samples that do not pool private-label securitized loans with GSE securitized loans.**

We now turn to KMSV2's argument that our finding of default discontinuities in the absence of securitization discontinuities is an artifact of inappropriately pooling private-label securitized loans with GSE-securitized loans. We begin by considering samples that do not pool private-label and GSE securitized loans. Using the same sample from LPS used in the updated version of our 2009 paper, Bubb and Kaufman (2011),<sup>8</sup> we focus on three subgroups: (1) non-FHA/VA loans originated between January 2003 and December 2007 and bought by the GSEs; (2) loans originated between January 2008 and June 2009 after the private-label MBS market had shut down; and (3) jumbo loans, which the GSEs are forbidden from securitizing. We analyze the discontinuities at both 620 and 660 FICO, since the GSEs' underwriting guidelines directed lenders to use both FICO scores as cutoffs to determine how carefully to screen mortgage borrowers (see BK, pp. 12–13). The discontinuities at 660 provide additional evidence on whether the GSEs' lender underwriting guidelines had a direct effect on lender behavior at these cutoffs, as posited by the origination rule-of-thumb theory.

We define default as an indicator variable equal to 1 if the loan was ever 61 days or more delinquent within a fixed follow-up period after origination, and 0 otherwise. For each subgroup we use the maximum follow-up period that we can use consistently for the entire subgroup. For the 2003–2007 subgroup this means we use a 36-month follow-up, and for the 2008–2009 subgroup an 18-month follow-up.<sup>9</sup> We define securitization as an indicator variable equal to 1 if the loan was ever sold by the originator within the follow-up period, and 0 otherwise.

---

<sup>8</sup>Please refer to Bubb and Kaufman (2011) for details about the dataset and our sample selection.

<sup>9</sup>We downloaded our data in March of 2011, when December 2010 was the latest available month of data. This allowed 36 months of follow-up for our sample of loans originated between January 2003 and December 2007. For our January 2008 to June 2009 sample, the maximum possible was 18 months of follow-up. We chose to cut the sample at June 2009 rather than December 2009 in order to allow for a longer follow-up period in which to observe default outcomes.

3.2.1. *GSE loans.* Figure 2 shows the discontinuities in the number of loans and default rate for non-FHA/VA loans originated in 2003–2007 and bought by the GSEs.<sup>10</sup> Under the theory described in KMSV2, loans bought by the GSEs should not be affected by any 620 rule of thumb since they were not at risk of being sold to the private-label securitizers who followed the rule of thumb. However, we see large jumps in the number of loans at 620 and 660 FICO. Moreover, Figure 2 shows that there are corresponding jumps in the default rate at these two thresholds. Panel A of Table 1 provides estimates of these jumps. At 620 FICO, the default rate jumps 3.2 percentage points from a rate of 16.8 percent just below 620. At 660, the default rate jumps 1.8 percentage points from a rate of 11.2 percent just below 660.<sup>11</sup>

Figure 3 and panels B and C of Table 1 show these discontinuities for the subgroups of low-doc and full-doc loans from this GSE sample. We find jumps in the number of loans and their default rate for both the full-doc and low-doc subgroups at 620 and 660.

These results are inconsistent with KMSV2’s hypothesis that screening cutoffs exist only for low-doc loans in the “non-agency” market. Lenders’ screening behavior changed discontinuously at both 620 and 660 for loans that were ultimately sold to the GSEs.

To further investigate the securitization rule-of-thumb theory, we would also want to estimate the jump in securitization probability at 620 and 660 for these loans. Because all of the loans in this sample were at some time securitized by the GSEs, one cannot estimate the *ex ante* probability of securitization using only this sample. Instead, as we explain in detail in Section 3.4 below, estimating the securitization probability at a given FICO score in the conforming market requires pooling together the GSE loans with private-label and portfolio loans, as was done in BK. There we showed that there is no jump in the securitization probability at 620 in the conforming market.<sup>12</sup>

3.2.2. *Loans originated in 2008–2009.* Another way to investigate whether cutoffs are solely a product of a rule of thumb among private-label securitizers is to analyze a sample of loans originated in 2008–2009, after the private-label market shut down. During the peak years of 2005 and

---

<sup>10</sup>We exclude FHA- and VA-insured loans to maintain comparability with the samples used in BK and KMSV2. Our results are similar if FHA/VA loans are included. A loan is considered as bought by GSEs if it is ever observed as owned by Fannie Mae, Freddie Mac, or Ginnie Mae.

<sup>11</sup>KMSV2 also investigated whether GSE loans exhibit a discontinuity at 620 FICO and found no discontinuity in a sample from LPS originated in 2001–2006. However, we are unable to replicate this finding, even when we use a 2001–2006 sample meant to mimic KMSV2’s sample selection. We are unaware of any reason for this discrepancy in results. We asked the authors for the code used to produce their results, but they did not provide it.

<sup>12</sup>In the updated version, Bubb and Kaufman (2011), we also show that there is no jump in the securitization probability at 660.

2006, private-label securitizers issued \$1,191 billion and \$1,146 billion of MBS, respectively.<sup>13</sup> However, after the mass downgrades of private-label MBS by the rating agencies in 2007 that followed the onset of the subprime mortgage crisis, private-label MBS issuance ground to a halt. In 2007 only \$707 billion in private-label MBS were issued, followed by a mere \$58 and \$60 billion in 2008 and 2009, respectively. Of these, only \$2 billion and \$0.9 billion, respectively, were subprime MBS. This is tiny relative to the \$1,500 billion and \$1,800 billion in total mortgage loans that were originated in 2008 and 2009, respectively. If credit score cutoff rules in screening are a product of a rule of thumb among private-label securitizers, then there should be no credit score discontinuities in our sample of loans originated in 2008–2009.

The 2008–2009 sample had a very different price appreciation experience than the 2003–2007 sample, since the 2008–2009 sample was originated after the bursting of the housing bubble. Moreover, we have a much shorter follow-up period available for these loans. We use as our measure of loan performance whether the loan defaulted within 18 months of origination. We expect these differences to result in smaller jumps in default than in the 2003–2007 sample. Our ability to observe jumps in lending volume is not compromised by the shorter follow-up period. Because the 2008–2009 sample is not selected by conditioning on securitization status, we can directly look for jumps in the securitization rate at the thresholds.

Figure 4 and Panel A of Table 2 show the discontinuities at 620 and 660 FICO in the number, default rate, and securitization rate of loans originated in 2008–2009. The data reveal that the cutoff rules in screening at 620 and 660 FICO persist in this period without private-label securitization. There are large jumps in the number of loans at each of these thresholds. As expected, the jumps in default are relatively small but statistically significant. Securitization appears smooth through the cutoffs, with no statistically significant jump at 620 and a jump of only 0.7 percentage points at 660.

During this period, loans insured by the Federal Housing Administration and the Department of Veterans Affairs grew to become a substantial fraction of the mortgage market. FHA and VA loans constitute 33 percent of our 2008–2009 sample. The secondary market for FHA and VA loans is dominated by the Government National Mortgage Association (Ginnie Mae), a federal agency. Because virtually all FHA and VA loans are bought by Ginnie Mae, the securitization rate in this subgroup is close to 1 for all loans. However, unlike the earlier GSE sample, this subgroup is not

---

<sup>13</sup>These statistics are from the 2011 *Mortgage Market Statistical Annual*, published by Inside Mortgage Finance.

selected according to *ex post* securitization outcomes, but according to *ex ante* FHA or VA status. This allows us an opportunity to test whether FICO cutoffs can exist without securitization cutoffs in a market in which private-label securitizers do not operate and all loans are securitized.

Figure 5 and Panel B of Table 2 present the discontinuities in the subgroup of FHA / VA loans. In this subgroup we see discontinuities in the number of loans and the default rate of loans at both 620 and 660, but no corresponding changes in the probability of securitization.<sup>14</sup> As expected, the plot of the securitization rate for this sample shows that virtually all of these loans are securitized.

These results provide more evidence against the identification assumption underlying the RD research design in KMSV. For the discontinuities in default at a FICO score threshold to serve as evidence that securitization led to lax screening by lenders, it must be the case that there is no reason other than a change in probability of securitization that defaults change discontinuously at these thresholds. The data from 2008–2009 show that even when there is essentially *no* private-label securitization, and no discontinuity in GSE securitization at these thresholds, lenders screen differentially across the 620 and 660 FICO thresholds. This is strong evidence that the 620 and 660 FICO score lender screening cutoff rules operate independently of a change in private-label securitization.

**3.2.3. Jumbo loans.** We can similarly isolate a sample of loans at risk only of private-label securitization, and not GSE securitization, by examining jumbo loans. Because we were interested in possible differences between the GSEs and private-label securitizers, in BK we analyzed a sample of loans for amounts above the GSEs' conforming loan limits (“jumbo” loans). Because the law prohibits the GSEs from purchasing these loans, this approach is a clean way to isolate a purely private-label secondary market for loans. In this jumbo sample, we showed that there is a small jump in the securitization probability at 620.

In Figure 6 we present updated estimates of the jumps in the number, default rate, and securitization rate of loans at 620 and 660, using the more complete jumbo sample from Bubb and

---

<sup>14</sup>The histogram reveals one unusual feature of the data: at exactly 620 and 660, the number of loans is greater than the number of loans at surrounding FICO scores. At 620, there is a clear discontinuity in the number of loans even ignoring the data right at 620. At 660 there is no clear discontinuity if the loans at exactly 660 are ignored. One possible explanation is that mortgage borrowers with lower credit scores wait until their scores reach these thresholds—and hence they are subject to less screening—before applying for loans. At 620, there is a clear discontinuity in the number of loans even ignoring the data right at 620. In contrast, at 660 there is not a discontinuity if the loans right at 660 are ignored.

Kaufman (2011), the updated version of BK.<sup>15</sup> There are significant jumps in the number of jumbo loans as well as in their default rate at both credit score thresholds. The number of loans roughly doubles at 620—there are 1,530 loans in the sample with a FICO score of 619, compared with 3,238 loans in the sample with a FICO score of 620—and the size of the jump in number of loans at 660 is similar. We estimate default discontinuities of 3.4 and 3.2 percentage points at 620 and 660, respectively. Our regression shows a small jump in the securitization rate at 620, estimated at 2.0 percentage points, though this jump is not readily apparent from visual examination of the corresponding scatterplot. We find no jump in the securitization rate at 660.

The 660 results further confirm that the phenomenon discovered in BK—the existence of lender screening cutoffs in the absence of any change in securitization rate—is not an artifact of pooling the “agency market” with the “non-agency market.” Moreover, as we discuss in more detail in Section 6 below, the 2.0 percentage point jump in the securitization rate at 620 is too small to plausibly be the cause of the doubling in the number of loans lenders are willing to make, and the 3.4 percentage point increase in default at 620.

3.2.4. *Discussion.* The findings presented above show that for many samples of loans, including samples that do not pool together private-label securitized loans and GSE securitized loans, there are credit score cutoffs in screening for which there is no corresponding jump in the securitization rate. To put this in IV terms, there is no valid first stage: the instrument (the credit score cutoff) has no effect on the treatment (the probability of securitization). And yet, at these cutoffs we see large jumps in the default rate of loans. In IV terms, we are showing a large reduced form effect of the instrument (the credit score cutoff) on the outcome (lender screening, as measured by the default rate and the number of loans), despite the lack of a first stage effect of the instrument on the treatment. This corroborates the evidence in BK that the GSEs had a direct effect on lender behavior when they contractually required lenders to follow cutoff rules at the 620 and 660 FICO score thresholds to determine how carefully to screen mortgage borrowers and those cutoffs spread through automated underwriting systems. In IV terms, this is evidence of a failure of the exclusion restriction, which implies that the credit score cutoff rules do not provide any evidence on whether securitization led to lax screening.

---

<sup>15</sup>See Bubb and Kaufman (2011) for the full regression table with these results.

**3.3. Assumptions underlying BK’s and KMSV2’s estimators of the probability of securitization.** We now examine more formally the differences between the approaches in BK and KMSV2 to estimating the probability of securitization at a given FICO score and the assumptions underlying each approach. Consider a sample of  $N$  mortgage loans. Each loan has one ultimate owner, which is either the originator, a GSE, or a private-label securitizer, as indicated by the random variables  $O_i$ ,  $G_i$ , and  $V_i$ , respectively.  $O_i = 1$  if owned by the originator,  $= 0$  if not (and similarly for  $G_i$  and  $V_i$ ). Because each loan has a unique ultimate owner, one and only one of these ownership indicators is equal to 1 for any loan  $i$ . There is a probabilistic element to the ultimate owner of each loan when viewed from the time of origination. Let  $\sigma_i^O \equiv Pr(O_i = 1)$ , that is, the probability that the originator of loan  $i$  would ultimately have to keep loan  $i$  on its books. Similarly, let  $\sigma_i^G \equiv Pr(G_i = 1)$  and  $\sigma_i^V \equiv Pr(V_i = 1)$ . The portfolio probability is then equal to  $\sigma_i^O = 1 - \sigma_i^G - \sigma_i^V$ .

The economic hypothesis of interest is that originators will be more careful in their screening decisions with respect to a particular potential borrower the higher is the probability that they would have to keep a loan to the borrower on their books,  $\sigma_i^O$ . The higher this probability is, the more “skin in the game” the originator has. The securitization rule-of-thumb theory posits a particular source of variation in  $\sigma_i^O$  that can be used to estimate its effect on lenders’ screening decisions. In particular, it supposes that private-label securitizers follow a simple rule of thumb such that the average probability that private-label securitizers will purchase each loan with a FICO score of 620 is discontinuously higher than the average probability that securitizers will purchase each loan with a FICO score of 619. Denote the FICO score of each loan by  $FICO_i$ . Denote the average probability that private label securitizers will buy loans with  $FICO_i = x$  as

$$(1) \quad \sigma^V(x) \equiv \frac{\sum_{i:FICO_i=x} \sigma_i^V}{N_x},$$

where  $N_x \equiv \#\{i = 1, \dots, N | FICO_i = x\}$ . Using this notation, the exogenous securitizer theory thus posits that:

$$(2) \quad \sigma^V(620) > \sigma^V(619).$$

Moreover, the theory makes the still stronger assumption that this difference in the willingness of one type of purchaser to buy otherwise similar loans is reflected in the probability that any purchaser will buy the loans, that is, the securitization probability. This might not be true if there



were other securitizers that did not follow the rule of thumb. The rule-of-thumb behavior by one set of securitizers provides a market opportunity for other securitizers. These other purchasers would buy a higher proportion of the loans just below 620. KMSV's securitization rule-of-thumb theory assumes instead that such a set of securitizers does not exist or does not fully make up for the lack of demand by the rule-of-thumb securitizers for loans just below 620. Defining  $\sigma^G(x)$  and  $\sigma^O(x)$  analogously to  $\sigma^V(x)$ , this additional assumption is:

$$(3) \quad \sigma^V(620) + \sigma^G(620) > \sigma^V(619) + \sigma^G(619)$$

or equivalently:

$$(4) \quad \sigma^O(620) < \sigma^O(619)$$

3.3.1. *BK's approach to estimating the probability of securitization.* A key issue in testing the two competing cutoff rule theories with quantitative evidence is how to estimate the average portfolio probability (equivalently, 1 minus the securitization probability) at a given FICO score. The conventional estimator for such average probabilities is the sample average of  $O_i$  over the relevant subgroup, namely:

$$(5) \quad \bar{O}(x) \equiv \frac{\sum_{i:FICO_i=x} O_i}{N_x} = 1 - \frac{\sum_{i:FICO_i=x} V_i + G_i}{N_x}.$$

The sample average is an unbiased estimator for  $\sigma^O(x)$ :<sup>16</sup>

$$(6) \quad \begin{aligned} E[\bar{O}(x)] &= 1 - E\left[\frac{\sum_{i:FICO_i=x} V_i + G_i}{N_x}\right] = 1 - \frac{\sum_{i:FICO_i=x} E[V_i] + E[G_i]}{N_x} \\ &= 1 - \frac{\sum_{i:FICO_i=x} \sigma_i^V + \sigma_i^G}{N_x} = 1 - \frac{\sum_{i:FICO_i=x} 1 - \sigma_i^O}{N_x} = \sigma^O(x). \end{aligned}$$

What we are essentially estimating is the regression function  $E[O|FICO]$ . Because FICO is discrete,  $\bar{O}(x)$  is a workable estimator. In the more usual case the forcing variable is continuous,

<sup>16</sup>KMSV2 (p. 8) asserts that we “assume that lenders originate loans with no regard for whether they are intended for non-agency or agency secondary markets.” However, we make no such assumption, and our estimator of the discontinuity in the average probability of securitization of originated loans at any given FICO score is consistent whether or not lenders originate particular loans with an intent to sell to a particular set of buyers. As we discuss below, KMSV2's estimators require a strong and testable assumption—that there are no loans in their sample that are at risk of being bought by both the GSEs and private-label securitizers.

and regression-based analogs to  $\bar{O}(x)$  are needed (see Imbens and Lemieux (2008) for a sampling). This is the approach we took in BK and in Section 3.2 above.<sup>17</sup>

3.3.2. *KMSV2’s approach to estimating the probability of securitization.* In contrast, KMSV2 argues that  $\bar{O}(x)$  is not a good estimator to use to estimate the average portfolio probability at a given FICO score. The authors argue that there are two completely distinct markets for loans: (1) those originated with the intent to sell to the GSEs (the “agency market”); and (2) those originated with the intent to sell to private-label securitizers (the “non-agency market”). In terms of the simple notation introduced above, KMSV2 assumes the following:

**Assumption 1.** Separate markets: *For all  $i$ , either  $\sigma_i^G = 0$  or  $\sigma_i^V = 0$ .*

Put another way, KMSV2 assumes that there is no loan at risk of being sold to both types of purchasers (that is, with  $\sigma_i^G > 0$  and  $\sigma_i^V > 0$ ).<sup>18</sup>

Based on this assumption, KMSV2 analyzes what they call the “agency market” and the “non-agency market” separately.<sup>19</sup> They assign loans bought by the GSEs to the agency market and loans bought by private-label buyers to the non-agency market. They use two main methods for assigning portfolio loans to one market or the other. In one, they simply assign all portfolio loans to the non-agency market. In the other, they estimate the probability each portfolio loan was destined for the GSEs based on the loans’ characteristics using predicted values from a probit model estimated using only securitized loans. They then use these predicted probabilities to weight each portfolio loan observation in their agency market analysis. For the non-agency market, they use 1 minus these weights. We provide a more formal characterization of their approaches in Appendix A. The key point is that their approaches are premised on Assumption 1.

Moreover, throughout the paper KMSV2 equates “GSE” (and “agency”) with “prime.” Similarly, it equates “private-label” (and “non-agency”) with “subprime.” KMSV2 (p. 3) states: “We

<sup>17</sup>In BK, FICO is treated as an integer measure of a continuously varying latent credit risk, and standard regression discontinuity techniques are used to estimate  $\lim_{x \uparrow 620} \sigma^O(x)$  and  $\lim_{x \downarrow 620} \sigma^O(x)$ . We focus on  $\sigma^O(619)$  and  $\sigma^O(620)$  here just to simplify the exposition, but all of our analytical points can be made using the continuously varying latent credit risk approach as well.

<sup>18</sup>KMSV2 also implicitly assumes that there are no pure portfolio loans, that is, for all  $i$ ,  $\sigma_i^O < 1$ .

<sup>19</sup>KMSV2 argues that by pooling together GSE loans and private-label securitized loans, BK’s estimates of the change in probability of securitization are “mechanically dampened” to a greater degree than BK’s estimates of the change in probability of default (KMSV2, pp. 13–14). However, KMSV2’s argument is premised on, among other things, Assumption 1 (that loans are either in one market or the other). As we show below, quantitative and institutional evidence is inconsistent with Assumption 1. Moreover, as we showed above,  $\bar{o}(x)$  (using a pooled sample) is in fact an unbiased estimator of the probability of securitization of originated loans with a FICO score of  $x$ .

use the terms ‘GSE’, ‘prime’, and ‘agency’ interchangeably, as well as ‘subprime’, ‘private investors’ and ‘non-agency’ or ‘non-GSE’.”

**3.4. Testing the separate markets assumption.** We look now to evidence and find that it is inconsistent with KMSV2’s separate markets assumption.

3.4.1. *Quantitative evidence.* The LPS data provide a straightforward way to test KMSV2’s separate markets assumption. In particular, we know the entire history of ownership of each loan in the sample. It turns out that many loans that are ultimately owned by the GSEs are owned earlier by private-label securitizers. Conversely, many loans that were ultimately owned by private-label securitizers were earlier owned by the GSEs. Specifically, 18 percent of conforming loans in our 2003–2007 sample were at one point in their life held by the GSEs and at another point owned by private-label securitizers. This is a lower bound for the fraction of loans in the sample that were at risk of being sold to both the GSEs or private-label securitizers, since loans that were *ex ante* at risk of being sold to both types of buyers need not have been held by both types *ex post*.

In addition to asserting that the GSEs and private-label securitizers operate in separate markets, KMSV2 equates GSE purchases with the “prime” market and private-label securitization with the “subprime” market. Indeed, the authors use the phrases interchangeably. One way to examine the validity of this assumption is to examine the fractions of loans with FICO scores in a neighborhood around 620 that were bought by the GSEs. Loans to borrowers with FICO scores this low are generally considered subprime loans.<sup>20</sup> The data show that the GSEs were active purchasers of these loans. In our 2003–2007 conforming sample, of the loans with FICO scores between 600 and 619, six months after origination 56 percent were owned by the GSEs, 35 percent were owned by private-label securitizers, and 9 percent were held in portfolio. For the loans with FICO scores between 620 and 640, the corresponding numbers are 61 percent, 30 percent, and 9 percent.

For higher FICO score borrowers, the GSEs hold a higher fraction of loans. For loans in the conforming sample to borrowers with FICO scores between 750 and 800, for instance, 78 percent were owned by the GSEs, 14 percent by private-label securitizers, and 8 percent by originators.

---

<sup>20</sup>See, for example, Mian and Sufi (2009, p. 1449) and the federal banking regulators’ “Expanded Guidance for Subprime Lending Programs,” issued in 2001 (available at <http://www.occ.treas.gov/news-issuances/bulletins/2001/bulletin-2001-6a.pdf>), which includes loans with FICO scores below 660 in its definition of “subprime” loans.

But as the evidence above makes clear, it does not follow from this that the GSEs were absent from the subprime market.

3.4.2. *Institutional evidence.* Institutional evidence also reveals that the GSEs actively competed with private-label securitizers to buy subprime loans. Calomiris and Wallison (2008) write (emphasis added):

In order to curry congressional support after their accounting scandals in 2003 and 2004, Fannie Mae and Freddie Mac committed to increased financing of “affordable housing.” They became *the largest buyers of subprime and Alt-A mortgages* between 2004 and 2007, with total GSE exposure eventually exceeding \$1 trillion.

As this quotation makes clear, the GSEs were large buyers of subprime mortgages during the period in question. This is also reflected in reports from the GSEs’ federal regulator. For example, Office of Federal Housing Enterprise Oversight (2007, p. 11) reports that, in addition to being major buyers of subprime private MBS, Fannie Mae and Freddie Mac “also buy and hold subprime mortgages directly.”

Moreover, the SEC filings of subprime originators show that they sell subprime loans to both the GSEs and private-label securitizers. For example, Countrywide Financial, one of the largest originators of subprime mortgage loans, reported in its 2005 Annual Report that, in addition to pooling “Nonprime Mortgage Loans” into private-label MBS, it “pooled a portion of [its] Nonprime Mortgage Loans into securities guaranteed by Fannie Mae” (p. 95).

3.4.3. *Discussion.* The above evidence clarifies that KMSV2’s assumption that no loans are both at risk of being sold to the GSEs and at risk of being sold to private-label securitizers does not hold. Indeed, the data show that a sizable fraction of the sample was in fact sold to both. This is a lower bound on the fraction of loans that *ex ante* were simultaneously at risk of being sold to both types of purchasers. Consequently, to estimate the average probability of securitization at a FICO score, one must pool portfolio loans together with GSE loans and private-label securitized loans. The sample average,  $\bar{O}(x)$ , or its regression-based analogs used in BK, provides a straightforward approach.

#### 4. COHORT ANALYSIS

KMSV2 (pp. 14-16) examines different cohorts of privately securitized loans in the LP dataset and finds that variation across cohorts in the size of the jump in the default rate of securitized loans at 620 correlates with variation across cohorts in the size of the jump in the number of securitized loans at 620. The paper argues that this corroborates their securitization rule-of-thumb theory.

However, a time-series correlation between the size of jumps in the number of loans and in the default rate is equally consistent with the origination rule-of-thumb theory, and indeed with any theory that holds that jumps in the number of loans are driven by discontinuous lender screening intensity. When lenders reduce their screening intensity above 620, it both increases the number and lowers the credit quality of loans made above 620. Thus, the bigger the drop in screening intensity at 620 for a cohort of loans, the larger both the jump in the number of loans and the jump in default rates, *ceteris paribus*. The central question is whether this discontinuous lender screening behavior is solely a response to differences in the ease of securitization, or whether lenders independently adhere to a cutoff rule. The time-series evidence in KMSV2 cannot distinguish between the two hypotheses.

However, a different sort of time-series evidence can help distinguish the theories. KMSV2 argues that cutoff rules are driven by private securitizers buying subprime loans. Therefore, one would expect a time-series correlation between the size of jumps at the 620 cutoff for loans originated in a period and the scale of private-label securitization during that period.

Private-label MBS issuance began as a small portion of the overall mortgage market. In 2004, however, its share began to increase rapidly, reaching 39 percent of the mortgage market in 2006, before collapsing to nearly nothing after 2008. If KMSV2 is correct, we would expect to see the size of the discontinuity at 620 similarly rise and fall over this period.

To investigate this, we divide our sample into six-month cohorts by origination date. Figure 7 shows the size of the discontinuity in (log) number of loans at 620 for these cohorts from 1997 to 2010, as well as the fraction of mortgages originated during each period that are bought by private-label securitizers.<sup>21</sup> The proportion of mortgages bought by private-label securitizers follows a clear inverted U pattern.

---

<sup>21</sup>The MBS data come from the *Mortgage Market Statistical Annual*, compiled by Inside Mortgage Finance. We are able to use origination data through December 2010 here because this application of the data does not require any follow-up period for observation of outcomes.

The 620 discontinuity, however, follows a very different pattern. It is positive though small in the period following the GSEs' original 1995 credit score cutoff rule guidelines,<sup>22</sup> grows steadily throughout the boom in private-label mortgage securitization, and finally reaches its greatest size in the years 2009–2010, when the private-label market is nonexistent. This is further confirmation that the 620 discontinuity is not an artifact of private-label securitization.

KMSV2 argues in favor of a positive time-series correlation between the size of the 620 discontinuity and the size of the private market using data from LoanPerformance. There are two important limitations of these data. First, they cover only private-label securitized loans. The issue in question is whether the 620 cutoff existed before and after the rise of private-label securitizers, so data containing only private-label securitized loans is ill-suited to answer this question. Second, KMSV2 restricts the sample period to end in 2007H1, excluding the important post-bust period that allows for identification of the trend.

KMSV2 analyzes in detail a narrow slice of time: the period from 2005H2 to 2007H1. The private market was shrinking over this period, and the authors show, again using LoanPerformance data, that the size of the discontinuity is shrinking too. KMSV2 (p. 16) concludes that this pattern “further supports the characterization provided by KMSV that the ease of securitization [as measured by number of securitized loans] was directly related to the performance of loans around the FICO threshold in the subprime mortgage market.”

We reproduced this analysis using the LPS dataset and find no decrease in the size of the 620 discontinuity over this time period. Panels A and B of Table 3 show that the jumps in the number and default rates of loans at 620 FICO were largely unchanged from 2005H2–2006H1 to 2006H2–2007H1. Defaults jump 4.3 percentage points in the 2005H2–2006H1 sample, and 4.1 percentage points in the 2006H2–2007H1 sample. Moreover, when the sample is limited to low-doc, privately securitized loans to try to match KMSV2's sample from LP, we similarly find stable discontinuities in this period (results unreported for brevity).

Overall, the time-series evidence paints a strong picture of a steadily growing 620 discontinuity, unrelated to the boom and bust of the market for private-label mortgage backed securities, and is inconsistent with the securitization rule-of-thumb theory.

---

<sup>22</sup>Pooling together the 1997–2000 subgroup, the number of loans jumps by a statistically significant .06 log points at 620.

## 5. MEASURING EASE OF SECURITIZATION

KMSV2 (pp. 19–21) considers different aspects of what the authors refer to as “ease of securitization” and argues that the *number* of securitized loans at a FICO score is a better measure of ease of securitization of loans at that FICO score than is the *fraction* of loans at that FICO score that are securitized (that is, the securitization rate) because it “is likely to capture the complete impact” of securitization on lenders’ payoffs (p. 21). We show in this section that the total number of securitized loans is a problematic measure of the ease of securitization, and that the securitization rate is a good measure.

**5.1. Aspects of ease of securitization possibly not captured by the securitization rate.** KMSV2 asserts that there are two aspects of ease of securitization that are not captured by the probability of securitization: the time it takes to securitize a loan and how intensely securitizers audit purchased loans to verify their quality.

**5.1.1. Time to securitize.** If there are differences in the time it takes for an originator to sell its loans on either side of a FICO threshold, KMSV2 argues, these differences could influence lender screening behavior. The authors point out that the longer it takes to sell a loan, the greater the likelihood that it will become delinquent while on the lender’s books. And they argue that lenders have a harder time selling loans that are delinquent, and that this difference will be reflected in the total number of loans sold. KMSV2 (p. 19) finds that low-doc private-label loans to the right of the 620 threshold are securitized 1.9 months faster than loans to the left, and that during those months a full 7.1 percent of those loans enter delinquency of 31 days or more—that is, missed at least one payment. This is a much weaker definition of delinquency than the 61 day definition of default used in the rest of the paper and in KMSV. The authors also argue that “[t]he longer it takes for a loan to be securitized, the costlier it may be for the bank. This cost could come from the opportunity cost of not being able to invest elsewhere” (p. 19).

First, note that any differences in the ultimate ability to sell loans will manifest not only in the total number of securitized loans, but in the securitization rate as well. The securitization rate thus captures this dimension of lenders’ incentives. This is not, therefore, an argument against using the securitization rate as a measure of ease of securitization.

Second, note that while keeping a loan for longer on one's books may entail an opportunity cost to the lender, this opportunity cost has nothing to do with lenders' incentives to screen. A lender's funds may be in use for a loan under 620 FICO for a longer period of time, but unless lenders can shorten that wait through more intense screening (which KMSV2 asserts is unobservable to the securitizer) then that opportunity cost should not affect the lender's screening decision.

Third, we do not find substantial changes in the time it takes loans to be securitized at 620 and 660 FICO in our subgroups. Table 4 presents our results. We find no differences in time to securitize at the 620 threshold in our 2003–2007 conforming sample: our point estimate is a difference of 0.003 months. Hence, a discontinuity in time to securitize does not explain the large change in lender screening at 620 that we document in our data. The low-doc and full-doc samples have similarly small changes in time to securitize at the cutoffs. We estimate a somewhat larger change in time-to-securitize for jumbo loans at 620. There we find that loans above 620 take 0.713 fewer months for originators to sell than loans below 620. However, the change in incentives represented by this difference is still small. Only 0.63 percent of loans near the cutoff in that sample default in that additional 0.713 of a month, out of a total default rate of 31.2 percent.

Our estimates are smaller than the 1.9 month change in time to securitize for low-doc private-label securitized loans reported in KMSV2. We tried to replicate KMSV2's estimate using the 2001–2006 low-doc private-label securitized sample, presented in panel E, but we estimate in that sample only a 0.61 month change. We inquired with the authors of KMSV2 about this discrepancy, and they reported that they had made a mistake in reporting their point estimate and that the correct point estimate is 1.1 months, about half as large as the estimate reported in KMSV2 and closer to our 0.61 month estimate. To investigate the difference in our estimates, we requested from the authors of KMSV2 the code they used to generate their estimates, but they did not provide it.

Moreover, KMSV2 interprets its estimate of the change in time to securitize by calculating the fraction of loans that missed a single payment during the gap period. However, missing a single payment is a much weaker definition of default than the 60 days or more delinquency definition used in the rest of KMSV2, and many loans that miss only a single payment are ultimately paid in full.

5.1.2. *Loan buybacks by originators.* KSMV2 (p. 20) also hypothesizes that securitizers may audit loans with FICO scores below 620 more intensely, and that this auditing may force originators



to accept greater numbers of buybacks, but provides no evidence that this is the case. We investigate this by examining whether there is a change at 620 and 660 FICO in the fraction of loans that are bought back by the originator. We define buybacks as loans that are first securitized and then later appear in portfolio at some time in the first 36 months after origination.

For the 2003–2007 conforming sample, we find that there is no change in the fraction of buybacks at 620, and only a 0.1 percentage point decrease in buybacks above 660. When we select on a sample meant to mimic the sample in KMSV2, as before, we find a 2.3 percentage point *increase* in buybacks above 620, and no change in buybacks at 660. (Note that KMSV2’s hypothesis would suggest a decrease in buybacks above the threshold.) The evidence thus suggests that differential audit rates by securitizers do not explain the changes in lender screening at these thresholds.

**5.2. The number of securitized loans is not a useful measure of “ease of securitization”.** The jump in the number of securitized loans at a credit score cutoff cannot be used as a measure of the change in lenders’ incentives due to securitization because it conflates changes in the lending rate and changes in the securitization rate.

KSMV2 defines the “unconditional probability of securitization” as the probability that a potential borrower is given a securitized loan (as opposed to not being given a loan or being given an unsecuritized loan). Let  $L_i \in \{0, 1\}$  denote whether potential borrower  $i$  is given a loan and let  $S_i \in \{0, 1, \emptyset\}$  denote whether borrower  $i$ ’s loan is securitized (with  $S_i = \emptyset$  if borrower  $i$  is not given a loan). KMSV’s unconditional probability is then:

$$(7) \quad Pr(S_i = 1) = Pr(L_i = 1) * Pr(S_i = 1|L_i = 1).$$

The first factor on the right-hand side of this equation is the lending rate; the second factor is the securitization rate. KMSV2 uses the number of securitized loans as a measure of the unconditional probability of securitization.

The problem with using the unconditional probability of securitization is that the lending rate can change for reasons completely unrelated to securitization. For instance, under the origination rule-of-thumb theory the lending rate changes at 620 because lenders are following a cutoff rule in screening. This creates an increase in the total number of securitized loans, but is unrelated to the actions of any securitizer. The number of unsecuritized loans increases as well, and for the same reason.

To illustrate the problem with using the total number of securitized loans rather than the securitization rate as a measure of ease of securitization, we offer a thought experiment: suppose we were investigating the opposite hypothesis from KMSV, namely that a *lack* of securitization—that is, lenders having to keep loans in portfolio—causes lax screening. Applying KMSV’s approach to test this hypothesis, we would use the *number* of portfolio loans as a measure of the risk that lenders get stuck with loans. If we find jumps in the number of portfolio loans and their default rate at a threshold like 620, we would interpret that as evidence that skin in the game ruins lenders’ incentives.

We can use the LPS dataset to actually implement this research design. We mimic KMSV’s approach by restricting our dataset to only unsecuritized (portfolio) loans, and we test whether increases in the number of unsecuritized loans are associated with jumps in the default rate. We estimate a jump in the number of unsecuritized loans at 620 of 0.342 log points, and a jump in the default rate of these loans of 6.5 percentage points. Figure 8 shows the jumps in the number and default rate of portfolio loan at 620 and 660, as well as equivalent graphs for samples of privately securitized loans and GSE securitized loans for comparison.

Using KMSV’s “unconditional probability” approach, one would interpret the jump in number of portfolio loans as a jump in originators’ skin in the game. Hence, the corresponding jump in defaults would be interpreted as evidence for the hypothesis that a *lack* of securitization led to lax underwriting. The problem with this interpretation of the jump in portfolio loans at 620 is that lenders changed their overall lending rate at this cutoff. This can be seen in the jumps in the number of private-label securitized loans and GSE securitized loans from the same origination years, also displayed in Figure 8.

KSMV2 (pp. 11–13) also sorts lenders by the size of the jump in the number of securitized loans and shows that lenders whose loans exhibit the largest jump in the number of securitized loans also exhibit the largest jump in default rates. However, this is uninformative about the hypothesis of interest. The number of loans made by a lender at a given FICO score and the default rate of those loans are both measures of the intensity of lender screening at that FICO score. When a lender is more strict, it makes fewer loans, and the loans it does make are less likely to default. KMSV2 has shown only that lenders whose screening changes the most at 620 (as measured by the size of the jump in the number of loans) also generate the largest jump in loan quality (as measured by

default rates). The question of why lenders' screening is changing discontinuously at 620 remains unanswered.

The right measure of ease of securitization to use to test the hypothesis that securitization led to lax screening by lenders is the probability that loans do not stay on the books of lenders ( $Pr(S_i = 1|L_i = 1)$ ). If a lender has a very high probability of selling a loan, say to a naive securitizer, then the lender's incentives to screen borrowers might be attenuated. If instead there is a large chance that the lender will keep the loan, then the moral hazard problem is less severe. This probability that a loan is kept is what is usually meant by "skin in the game." And as discussed in Section 3, the probability of securitization at a given FICO score is best estimated by calculating the fraction of loans at that FICO score that lenders sell. Doing so, as we have done in BK as well as in this paper, one finds that screening discontinuities at FICO score thresholds exist in the absence of discontinuities in the securitization rate.

## 6. IMPLAUSIBLE EFFECT SIZES

KMSV2 argues that there is a jump in the securitization rate at 620 for non-agency, low documentation loans, and that this change in ease of securitization is what is causing the jump in the default rate at 620. We also find a small jump in the securitization rate in our jumbo sample. Leaving aside that we find many subgroups in which there is no discontinuity in securitization but large discontinuities in the number of loans and the default rate at 620 and 660, and taking at face value the estimates from KMSV2 based on the separate markets assumption, we consider here whether the sizes of the estimated jumps in default and number of loans at 620 are plausibly explained by the estimated corresponding jumps in the probability of securitization.

We begin by reviewing the magnitudes of the estimates, starting with our results for the jumbo sample. As described in Section 3.2.3 above, there is a large change in lender behavior at 620 in the jumbo market. The number of loans roughly doubles, and the default rate increases 3.4 percentage points from 31.2 percent to 34.6 percent. By comparison, the change in the securitization rate is small—jumbo loans are securitized at a rate of 80.7 percent right below 620, compared to 82.7 percent just above 620, an increase of 2 percentage points.

Turning now to the estimates in KMSV2, the authors find a jump in the securitization rate at 620 of 5.4 percentage points, from about 72 percent just below 620 to about 77 percent at 620, in their

preferred sample of low-documentation, non-agency market loans.<sup>23</sup> They find a corresponding jump in default of 7.1 percentage points, from about 35 percent to about 42 percent.<sup>24</sup> KMSV2 does not report how large the jump in the number of loans is for this sample, but does report a tripling in the number of loans for lenders who have greater than the median jump in number of loans at 620.<sup>25</sup> KMSV2 also uses data from two individual subprime lenders. For lender X, KMSV2 finds a jump in securitization of 1.7 percentage points and a corresponding jump in default of 10 percentage points. For lender Y, KMSV2 finds jumps of 2.8 percentage points in securitization and 4 percentage points in default. Jiang, Nelson, and Vytlačil (2010) also uses data from a single originator and finds a 2.7 percentage point jump in the securitization rate at 620 but a 5.6 percentage point jump in default at 620.

Are these small changes in the securitization rate the cause of the large changes in lender behavior at 620? We find these effect sizes implausible. Going from a securitization rate of 80.7 percent to 82.7 percent for jumbo loans cannot plausibly cause lenders to double the number of loans they make at 620, with a corresponding 3.4 percentage point jump in default. Similarly, KMSV2's estimates of changes in the securitization rate are too small to plausibly be the cause of the corresponding changes in lender behavior that it documents. KMSV2's data from lender X provide the clearest illustration of the magnitude issue—the securitization rate jumps 1.7 percentage points at 620, from a base above 90 percent, whereas the default rate jumps by 10 percentage points, from a base of about 12 percent.

In contrast, these magnitudes are consistent with the origination rule-of-thumb theory. The jumps in the number of loans and default rate at 620 are a consequence of lenders' following a credit score cutoff rule in screening in response to GSE guidance and the use of automated underwriting systems that encode the cutoffs. The differential screening above and below the credit score cutoff results in a discontinuous change in the amount of asymmetric information the lender has about loans, as well as discontinuities in other characteristics of loans that affect the likelihood that the originator sells the loans. The differences in the securitization rate reflected in these samples are consistent with this theory in which they are a consequence, not a cause, of discontinuities in lender screening behavior.

---

<sup>23</sup>See Figure 1d of KMSV2.

<sup>24</sup>KMSV2 does not report the point estimate but displays it visually in Figure 2c. The point estimate is provided in Keys, Seru, and Vig (2011, p. 19), a revised version of KMSV2.

<sup>25</sup>See Figure 10a of KMSV2.

## 7. NEW JERSEY AND GEORGIA ANTI-PREDATORY LENDING LAW ANALYSIS

KMSV (pp. 21–23) attempted to test whether the 620 FICO score threshold was used by lenders for reasons unrelated to securitization by using variation induced by the passage of state anti-predatory lending laws in Georgia and New Jersey in 2002 and 2003, respectively. It argues that the laws made it harder for lenders to securitize mortgages but kept “everything else equal” (p. 21). KMSV further argues that if 620 represents a threshold used by lenders, independent of securitization, then the passage of these laws should have no effect on the discontinuities at 620. KMSV then shows that the discontinuity at 620 in the number of loans originated in Georgia and New Jersey gets smaller during the period the laws were in effect. Moreover, the authors find that there is no discontinuity in default at 620 when the laws were not in effect, and a *negative* discontinuity in default when the laws were in effect. These results are interpreted as supporting the theory that the discontinuities in default are exclusively caused by a securitization rule-of-thumb and that lenders do not change their behavior at 620 for other reasons.

In BK we showed that these anti-predatory lending laws did not keep everything else equal. In particular, they regulated a range of lending practices in order to reduce defaults directly. We also showed empirically that the laws did not lower the securitization rate in those states.

KMSV2 responded by arguing that the laws were intended to affect contract terms, not default per se. The authors showed in KMSV that interest rates and loan-to-value ratios (LTVs) for low-doc, private-label securitized loans are continuous at 620. Hence, they argue that there is no reason to expect the laws to have had a differential effect on defaults above and below 620.

There are two problems with the response in KMSV2. First, there are provisions in the laws that are aimed directly at reducing defaults. For example, the Georgia Fair Lending Act requires lenders to certify that the borrower has received counseling on the advisability of the loan and requires that the lender reasonably believes that the borrower will be able to make the payments on the loan before making certain loans.<sup>26</sup>

Second, there are discontinuities in contract terms at 620. Table 5 provides estimates of the discontinuities in contract terms using loans in LPS originated from 2003–2009.<sup>27</sup> We find statistically significant differences in the loan-to-value ratio, the debt-to-income ratio, adjustable rate

<sup>26</sup>O.C.G.A. § 7-6A-5 (6)-(7).

<sup>27</sup>Results are robust to examining different time periods.

loans, pre-payment penalties, whether the loan is interest only, and whether the loan allows for negative amortization. Such discontinuities are likely the result of originators treating loan applicants above and below the cutoff differently. Jiang, Nelson, and Vytlačil (2010) also finds discontinuities in contract terms at 620 using data from a single large originator.

## 8. CONCLUSION

KMSV2 raises a number of interesting issues about the analysis in BK. However, after investigation, we find that none of the issues change our analytical approach or conclusions. The institutional and quantitative evidence strongly rejects KMSV's securitization rule-of-thumb theory. Lenders change their screening behavior at certain credit score cutoffs for reasons unrelated to changes in the probability of securitization. Evidently when the GSEs directed lenders to adopt these cutoff rules in screening in the late 1990s, and lenders signed contracts in which they committed to follow the GSEs' underwriting guidelines and adopted automated underwriting systems that incorporate those cutoff rules, it had a direct effect on lender behavior. This means that, unfortunately, the regression discontinuity research design used in KMSV and KMSV2 is unable to answer the question of whether securitization led to lax screening. Researchers must continue to search for other sources of evidence that can be brought to bear on this important question.

## REFERENCES

- BUBB, R., AND A. KAUFMAN (2009): "Securitization and Moral Hazard: Evidence from a Lender Cutoff Rule," Federal Reserve Bank of Boston Public Policy Discussion Paper No. 09-5.
- (2011): "Securitization and Moral Hazard: Evidence from Credit Score Cutoff Rules," Federal Reserve Bank of Boston Public Policy Discussion Paper No. 11-6.
- CALOMIRIS, C., AND P. WALLISON (2008): "Blame Fannie Mae and Congress For the Credit Mess," *Wall Street Journal*.
- FRANGAKIS, C. E., AND D. B. RUBIN (2002): "Principal Stratification in Causal Inference," *Biometrics*, 58(1), 21–29.
- IMBENS, G., AND T. LEMIEUX (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142(2), 615–635.
- JIANG, W., A. NELSON, AND E. VYTLACIL (2010): "Securitization and Loan Performance: A Contrast of Ex Ante and Ex Post Relations in the Mortgage Market," Unpublished manuscript.
- KEYS, B., T. MUKHERJEE, A. SERU, AND V. VIG (2010a): "620 FICO, Take II: Securitization and Screening in the Subprime Mortgage Market," Unpublished manuscript.
- KEYS, B., T. MUKHERJEE, A. SERU, AND V. VIG (2010b): "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans," *Quarterly Journal of Economics*, 125(1).
- KEYS, B., A. SERU, AND V. VIG (2011): "Lender Screening and the Role of Securitization: Evidence from Prime and Subprime Mortgage Markets," *Review of Financial Studies*, forthcoming.

MCCRARY, J. (2008): “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test,” *Journal of Econometrics*, 142(2), 698–714.

MIAN, A., AND A. SUFI (2009): “The Consequences of Mortgage Credit Expansion: Evidence from The U.S. Mortgage Default Crisis Stability,” *The Quarterly Journal of Economics*, 124(4).

OFFICE OF FEDERAL HOUSING ENTERPRISE OVERSIGHT (2007): “2007 Performance and Accountability Report,” .

## APPENDIX A

Assume Assumption 1 holds and let  $\gamma_i$  be an indicator for the market loan  $i$  was intended for, with  $\gamma_i = 1$  indicating  $\sigma_i^G > 0$  (and  $\sigma_i^V = 0$ ) and  $\gamma_i = 0$  indicating  $\sigma_i^V > 0$  (and  $\sigma_i^G = 0$ ). The authors of KMSV2 are interested in estimating the average probability that loans with  $FICO_i = x$  that were intended for the GSEs stay on the books of the lender:

$$(8) \quad \sigma_G^O(x) \equiv 1 - \frac{\sum_{i:FICO_i=x, \gamma_i=1} \sigma_i^G}{N_{gx}},$$

where  $N_{gx} \equiv \#\{i = 1, \dots, N | FICO_i = x, \gamma_i = 1\}$  is the total number of loans with  $FICO_i = x$  originated with the intent to sell to the GSEs.

Similarly, for loans intended for the private-label market, they want to estimate the average probability that loans with  $FICO_i = x$  stay on the books of the lender:

$$(9) \quad \sigma_V^O(x) \equiv 1 - \frac{\sum_{i:FICO_i=x, \gamma_i=0} \sigma_i^V}{N_{vx}},$$

where  $N_{vx} \equiv \#\{i = 1, \dots, N | FICO_i = x, \gamma_i = 0\}$  is the total number of loans originated with the intent to sell to private-label securitizers.

$\sigma_G^O(x)$  and  $\sigma_V^O(x)$  are different estimands than  $\sigma^O(x)$ . Note that Assumption 1 is crucial for the estimands in (8) and (9) to be well defined. If some loans are at risk of being sold to either the GSEs or private-label securitizers, then we would have to further complicate the estimands by excluding those loans. Moreover, estimation of (8) and (9) requires identifying which loans were at risk of being sold to either the GSEs or private-label securitizers. In contrast, the estimator we use in BK for  $\sigma^O(x)$  is robust to all of these possibilities.

Even if Assumption 1 holds, a key challenge to KMSV2's approach remains: how to identify which market each loan was destined for. The numerator for the fraction in (8) is easy to estimate using  $\sum_{i:FICO_i=x} G_i$  (the number of loans bought by the GSEs) since  $G_i = 0$  for all loans with  $\gamma_i = 0$ . Similarly,  $\sum_{i:FICO_i=x} V_i$  can be used to estimate the numerator in the fraction in (9). But to estimate the denominators,  $N_{gx}$  and  $N_{vx}$  (the total number of loans in each market), we need to figure out which portfolio loans in the sample to assign to which market.

We can express  $N_{gx}$  and  $N_{vx}$  as:

$$(10) \quad N_{gx} = \sum_{i:FICO_i=x} G_i + O_i \cdot 1_{\{\gamma_i=1\}}$$

$$(11) \quad N_{vx} = \sum_{i:FICO_i=x} V_i + O_i \cdot 1_{\{\gamma_i=0\}}.$$

KMSV2 use two methods for estimating  $N_{gx}$  and  $N_{vx}$ . The simpler of the two they refer to as "method 2": assume that all of the portfolio loans are intended for the private-label market. Then we have:

$$(12) \quad N_{gx} = \sum_{i:FICO_i=x} G_i$$

$$(13) \quad N_{vx} = \sum_{i:FICO_i=x} V_i + O_i.$$



Their estimators then become:

$$(14) \quad \hat{\sigma}_G^O(x) \equiv 1 - \frac{\sum_{i:FICO_i=x} G_i}{\sum_{i:FICO_i=x} G_i} = 0$$

$$(15) \quad \hat{\sigma}_V^O(x) \equiv 1 - \frac{\sum_{i:FICO_i=x} V_i}{\sum_{i:FICO_i=x} V_i + O_i}.$$

They also use a more complicated method, referred to as “method 1,” in which they use only the loans that were either bought by GSEs or by private-label securitizers to estimate a probit model predicting whether each loan was destined for the GSEs using various loan characteristics, like loan size, FICO score, etc. Denote the vector of loan characteristics by  $\mathbf{X}_i$ . They estimate a model of the form:

$$(16) \quad Pr(G_i = 1 | \mathbf{X}_i, O_i = 0) = f(\mathbf{X}_i).$$

They then use the estimate of the model in (16) to calculate predicted values for the portfolio loans in the sample. They then estimate the probability of securitization at each FICO score for the GSE market by using these predicted probabilities to weight each portfolio loan observation (and using a weight of 1 for each loan bought by the GSEs and a weight of 0 for each loan bought by private-label securitizers). Similarly, for the private-label market, they use 1 minus these predicted probabilities to weight the portfolio loans (and a weight of 0 for each loan bought by the GSEs and a weight of 1 for each loan bought by private-label securitizers).

Using this approach, their estimators for  $N_{gx}$  and  $N_{vx}$  can be written as:

$$(17) \quad \hat{N}_{gx} = \sum_{i:FICO_i=x} G_i + \hat{f}(\mathbf{X}_i)O_i$$

$$(18) \quad \hat{N}_{vx} = \sum_{i:FICO_i=x} V_i + [1 - \hat{f}(\mathbf{X}_i)]O_i,$$

where  $\hat{f}(\cdot)$  is the estimate of the model in (16).

The hope is that (17) and (18) are unbiased (and/or consistent) for  $N_{gx}$  and  $N_{vx}$ , expressed in (10) and (11), respectively. KMSV2 does not derive conditions under which that is so.

KMSV2’s method 1 estimators for  $\sigma_G^O(x)$  and  $\sigma_V^O(x)$  can be written as:<sup>28</sup>

$$(19) \quad \tilde{\sigma}_G^O(x) = 1 - \frac{\sum_{i:FICO_i=x} G_i}{\sum_{i:FICO_i=x} G_i + \hat{f}(\mathbf{X}_i)O_i}$$

$$(20) \quad \tilde{\sigma}_V^O(x) = 1 - \frac{\sum_{i:FICO_i=x} V_i}{\sum_{i:FICO_i=x} V_i + [1 - \hat{f}(\mathbf{X}_i)]O_i}.$$

<sup>28</sup>Like us, KMSV2 treats FICO as an integer measure of a continuously varying latent credit risk and so use a regression-based analogs to (19) and (18). All of our analytical points can be made using that more complicated model. We focus on the discrete case just to simplify.

APPENDIX B

TABLE 1. Discontinuities in the 2003–2007 GSE Sample

	FICO 620		FICO 660	
	log(# loans) (1)	Default (2)	log(# loans) (3)	Default (4)
PANEL A: ALL LOANS ( $N = 8,279,552$ )				
Change at cutoff	.417***	.032***	.092***	.019***
<i>s.e.</i>	(.005)	(.003)	(.003)	(.002)
Predicted below	-	.168	-	.112
PANEL B: LOW-DOC LOANS ( $N = 1,484,940$ )				
Change at cutoff	.378***	.040***	.085***	.036***
<i>s.e.</i>	(.010)	(.007)	(.008)	(.004)
Predicted below	-	.177	-	.137
PANEL C: FULL-DOC LOANS ( $N = 3,775,947$ )				
Change at cutoff	.469***	.013***	.075***	.007***
<i>s.e.</i>	(.007)	(.004)	(.005)	(.003)
Predicted below	-	.202	-	.118

*Notes:* Sample includes conventional loans originated in 2003–2007 and bought by the GSEs. Columns 1 and 3 use a local linear regression, as outlined in McCrary (2008). Default is an indicator for whether the loan ever became 61 days delinquent within 36 months of origination. Columns 2 and 4 use a 6th-order polynomial in FICO on either side of the cutoff. The number of observations in Panels B and C do not sum to the number of observations in Panel A because documentation status is missing for some loans. Heteroskedasticity-robust standard errors in parentheses. (\*\*\*) significant at 1 percent, (\*\*) significant at 5 percent, (\*) significant at 10 percent.

TABLE 2. Discontinuities in the 2008–2009 Sample

	FICO 620			FICO 660		
	log(# loans) (1)	Default (2)	Securitization (3)	log(# loans) (4)	Default (5)	Securitization (6)
PANEL A: ALL LOANS ( $N = 3,437,910$ )						
Change at cutoff	.372***	.007*	.002	.210***	.018***	.007***
<i>s.e.</i>	(.006)	(.004)	(.002)	(.005)	(.002)	(.001)
Predicted below	-	.134	.974	-	.071	.966
PANEL B: FHA/VA LOANS ( $N = 1,123,381$ )						
Change at cutoff	.338***	.009**	.001	.196***	.024***	-.000
<i>s.e.</i>	(.007)	(.004)	(.000)	(.006)	(.003)	(.000)
Predicted below	-	.134	.999	-	.072	1.000

*Notes:* Sample includes loans originated in 2008–2009. *Default* is an indicator for whether the loan ever became 61 days delinquent within 18 months of origination. *Securitization* is an indicator for whether the loan was ever sold within 18 months of origination. Columns 1 and 4 use a local linear regression, as outlined in McCrary (2008). Columns 2, 3, 5, and 6 use a 6th-order polynomial in FICO on either side of the cutoff. (\*\*\*) significant at 1 percent, (\*\*) significant at 5 percent, (\*) significant at 10 percent.

TABLE 3. Comparison of 2005H2–2006H1 and 2006H2–2007H1

	log(# loans) (1)	Default (2)	Securitization (3)
PANEL A: 2005H2–2006H1 ( $N = 3,459,648$ )			
Change at 620	.354***	.043***	-.004*
<i>s.e.</i>	(.005)	(.003)	(.002)
Predicted below	-	.260	.953
PANEL B: 2006H2–2007H1 ( $N = 3,237,504$ )			
Change at 620	.384***	.041***	-.015***
<i>s.e.</i>	(.005)	(.004)	(.003)
Predicted below	-	.368	.919

*Notes:* Subgroups based on origination date from LPS. *Default* is an indicator for whether the loan ever became 61 days delinquent within 36 months of origination. *Securitization* is an indicator for whether the loan was ever sold within 36 months of origination. Column 1 uses a local linear regression, as outlined in McCrary (2008). Columns 2 and 3 use a 6th-order polynomial in FICO on either side of the cutoff. Heteroskedasticity-robust standard errors in parentheses. (\*\*\*) significant at 1 percent, (\*\*) significant at 5 percent, (\*) significant at 10 percent.

TABLE 4. Discontinuities in Time-to-Securitize and Buybacks in the 2003–2007 Sample

	FICO 620		FICO 660	
	<i>Months to Securitize</i>	<i>Buyback Rate</i>	<i>Months to Securitize</i>	<i>Buyback Rate</i>
	(1)	(2)	(3)	(4)
PANEL A: 2003–2007 CONFORMING LOANS				
Change at cutoff	-.003	-.001	-.017*	.001*
<i>s.e.</i>	(.012)	(.001)	(.009)	(.001)
Predicted below	2.12	.063	2.00	.038
<i>N</i>	12,051,229	12,875,935	12,051,229	12,875,935
PANEL B: 2003–2007 JUMBO LOANS				
Change at cutoff	-.713***	.006	-.064	.002
<i>s.e.</i>	(.083)	(.005)	(.047)	(.002)
Predicted below	3.71	.057	2.80	.041
<i>N</i>	1,291,160	1,622,208	1,291,160	1,622,208
PANEL C: 2003–2007 LOW-DOC LOANS				
Change at cutoff	-.051	-.004*	.014	.002
<i>s.e.</i>	(.034)	(.003)	(.024)	(.001)
Predicted below	2.22	.048	2.05	.021
<i>N</i>	2,468,082	2,790,986	2,468,082	2,790,986
PANEL D: 2003–2007 FULL-DOC LOANS				
Change at cutoff	-.043***	-.007***	-.025**	-.001
<i>s.e.</i>	(.015)	(.001)	(.011)	(.001)
Predicted below	2.04	.051	1.94	.029
<i>N</i>	5,873,386	6,402,420	5,873,386	6,402,420
PANEL E: 2001–2006 LOW-DOC PRIVATE CONVENTIONAL PURCHASE LOANS				
Change at cutoff	-.608***	.015***	.009	.002
<i>s.e.</i>	(.144)	(.005)	(.079)	(.003)
Predicted below	3.392	.010	2.479	.015
<i>N</i>	349,346	360,250	349,346	360,250

*Notes:* *Months to Securitize* is defined for loans that became securitized at some point within 36 months of origination, and is equal to the number of months until securitization. *Buyback Rate* is a binary variable equal to 1 if, within 36 months of origination, a loan enters portfolio after first being securitized. Heteroskedasticity-robust standard errors in parentheses. (\*\*\*) significant at 1 percent, (\*\*) significant at 5 percent, (\*) significant at 10 percent.

TABLE 5. Discontinuities in Contract Terms

	FICO 620						
	<i>LTV</i> (1)	<i>DTI</i> (2)	<i>ARM</i> (3)	<i>PP</i> (4)	<i>IO</i> (5)	<i>NegAm</i> (6)	<i>Rate</i> (7)
Change at cutoff	-.125*	.519***	.028***	.008**	.028***	-.004***	.025***
<i>s.e.</i>	(.068)	(.070)	(.002)	(.002)	(.001)	(.001)	(.005)
Predicted below	82.4	36.5	.236	.212	.050	.036	6.62

Notes: Sample includes loans originated in 2003–2007. *LTV* stands for loan-to-value ratio. *DTI* stands for debt-to-income ratio. *ARM* stands for adjustable-rate mortgage. *PP* stands for prepayment penalty. *IO* stands for interest only loan. *NegAm* stands for a loan allowing for negative amortization. *Rate* is the interest rate at 6 months after origination. Discontinuities estimated using a 6th-order polynomial in FICO on either side of the cutoff. (\*\*\*) significant at 1 percent, (\*\*) significant at 5 percent, (\*) significant at 10 percent.

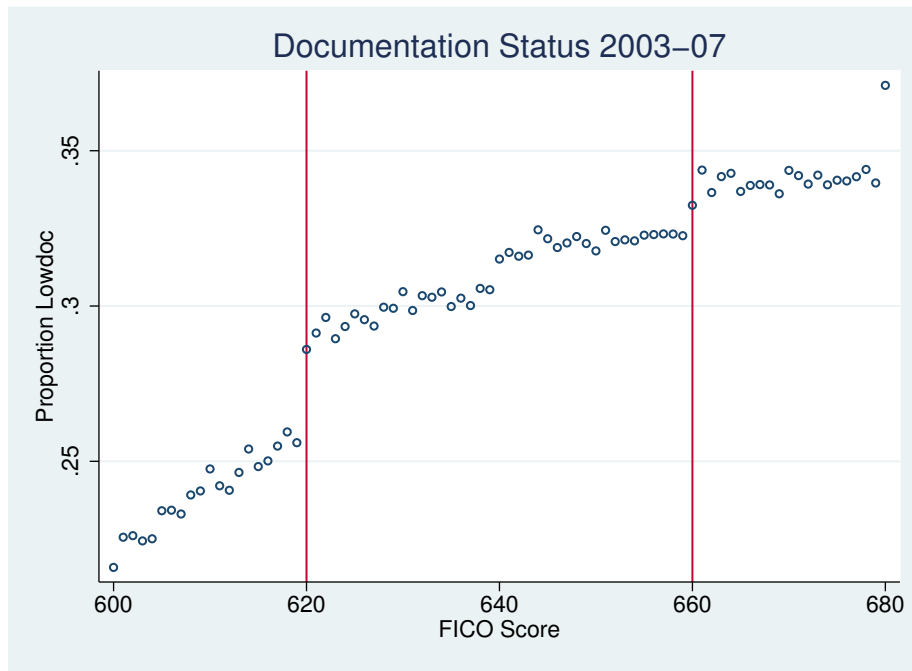
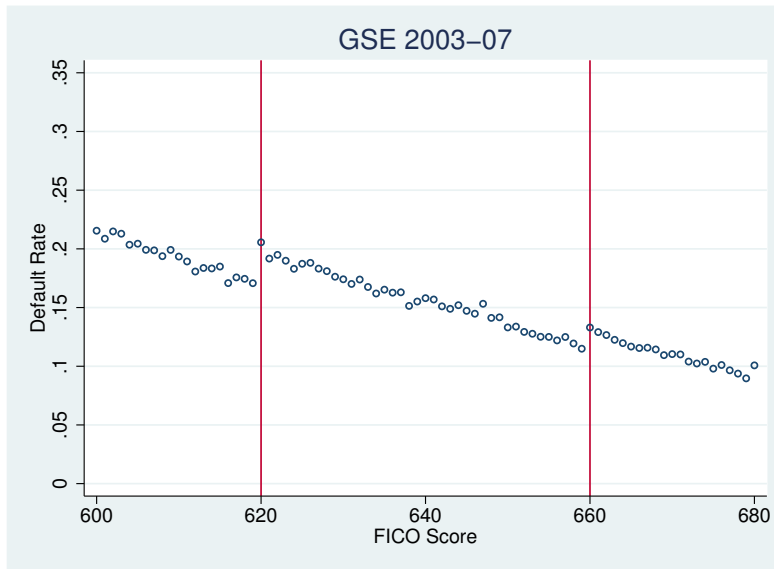


FIGURE 1. Proportion low documentation by FICO. Source: LPS data.



(a) Frequency



(b) Default

FIGURE 2. 2003–2007 GSE Sample. Source: LPS data.

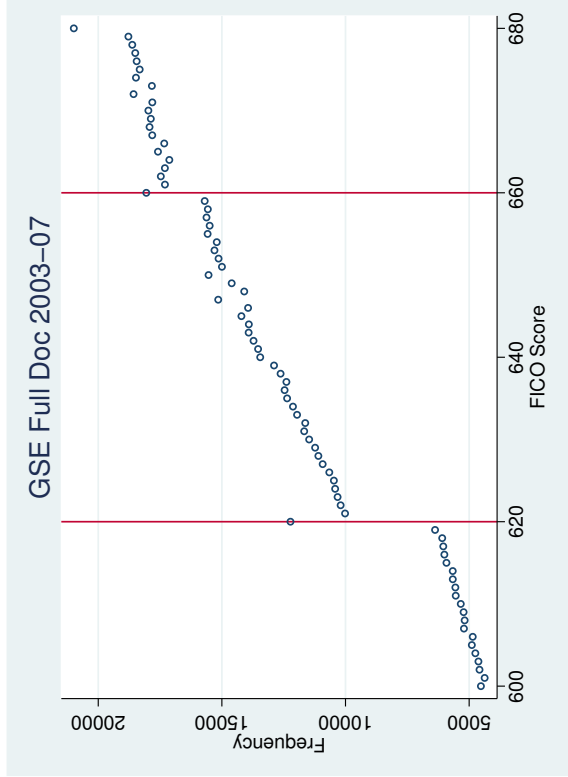
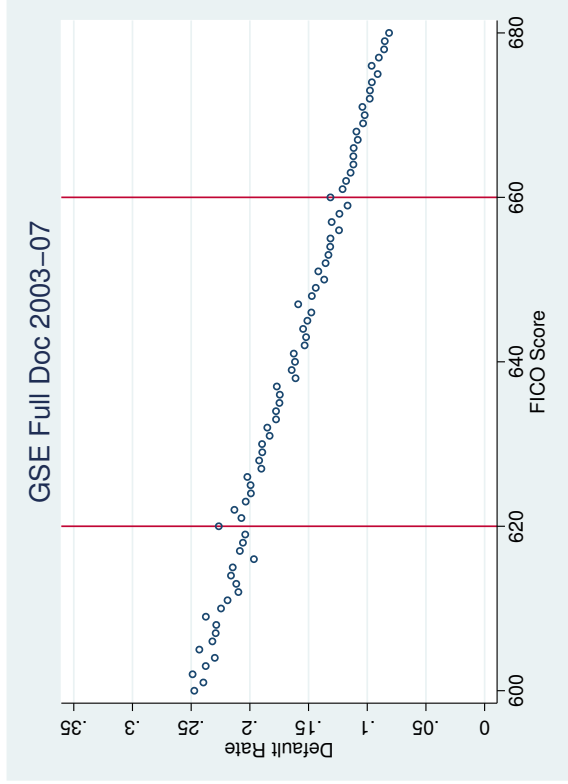
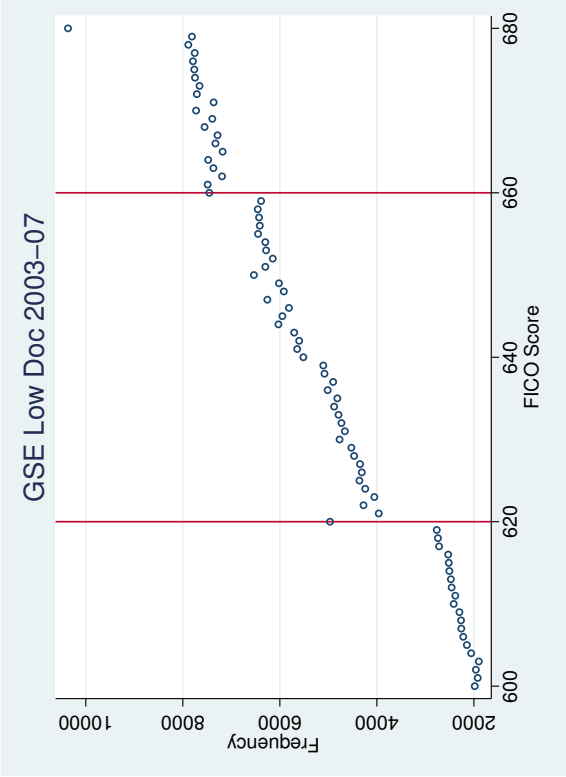
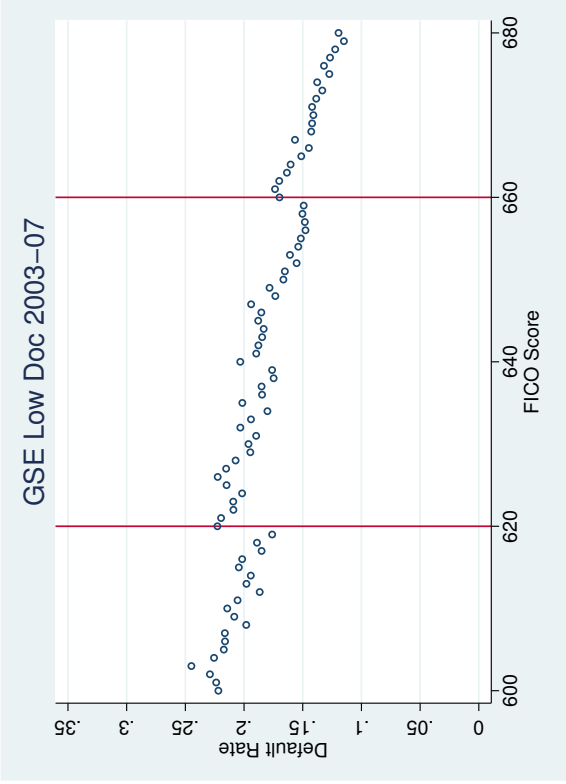
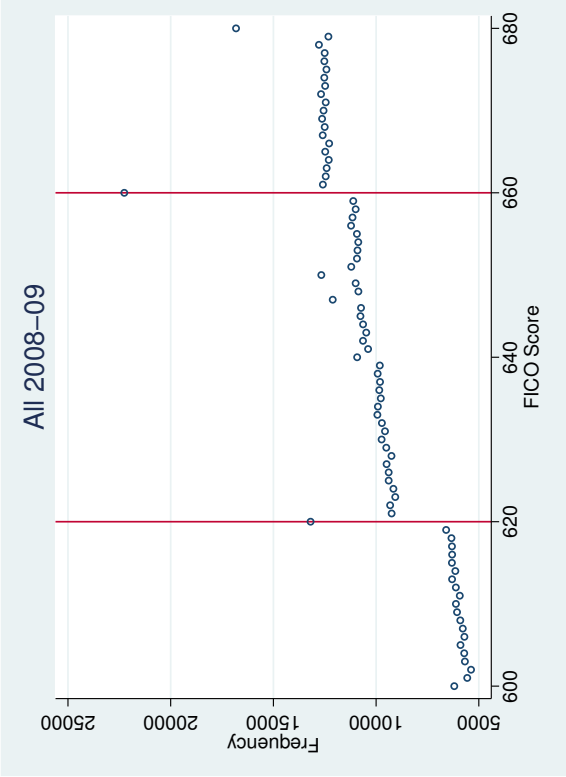
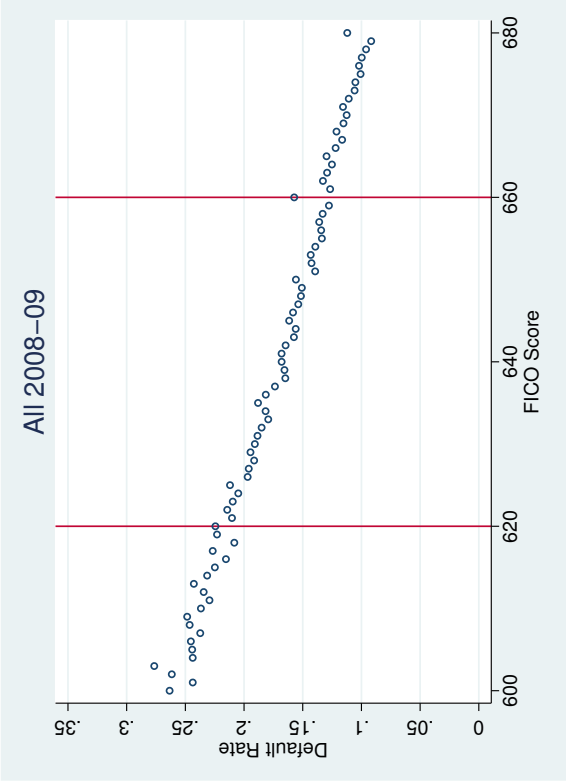
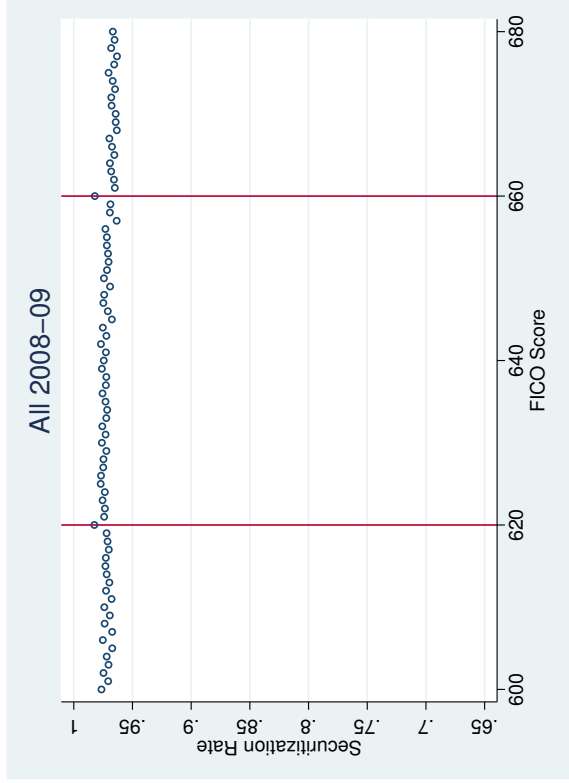


FIGURE 3. 2003–2007 GSE Sample, low-doc vs. full-doc. Source: LPS data.



(a) Frequency

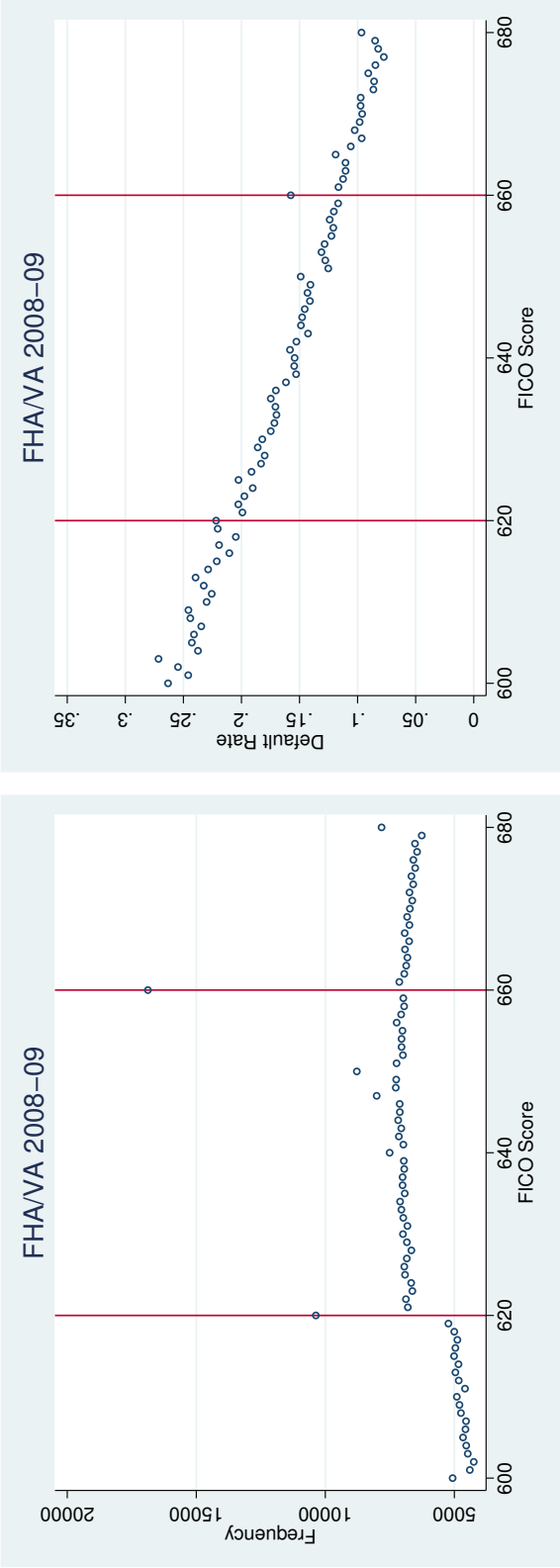
(b) Default



(c) Securitized

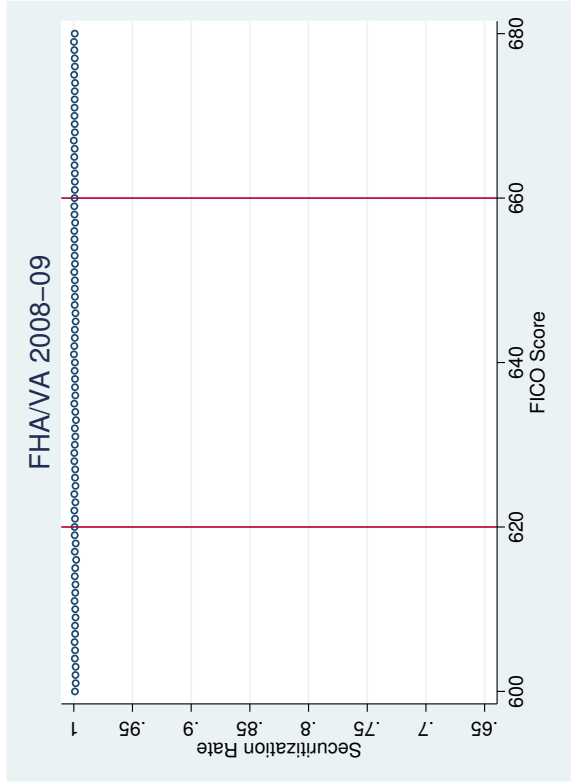
FIGURE 4. 2008–2009 Sample. Source: LPS data.





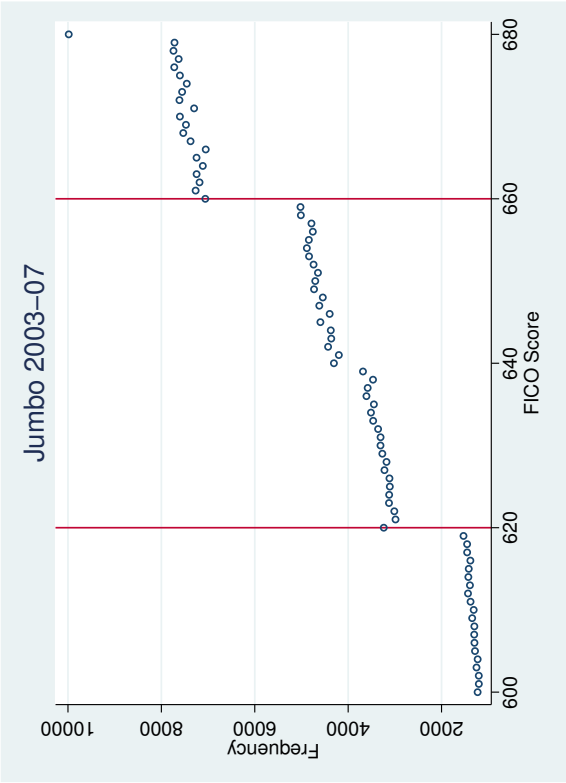
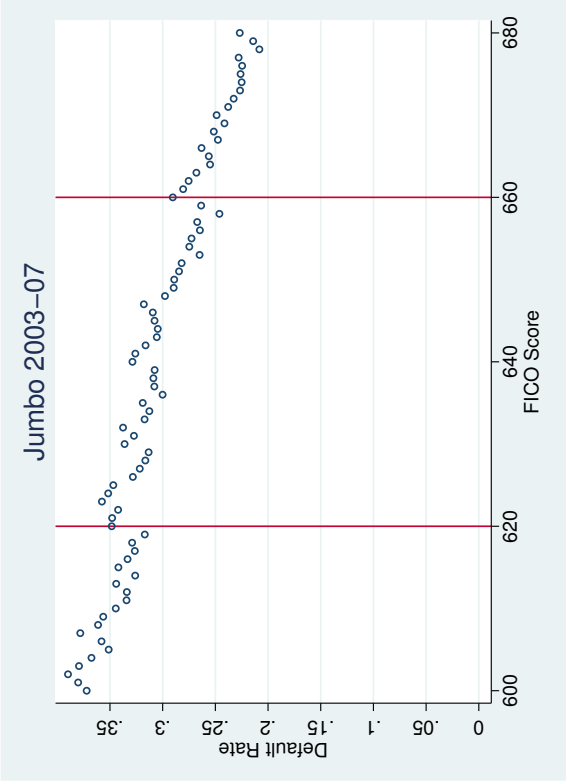
(a) Frequency

(b) Default

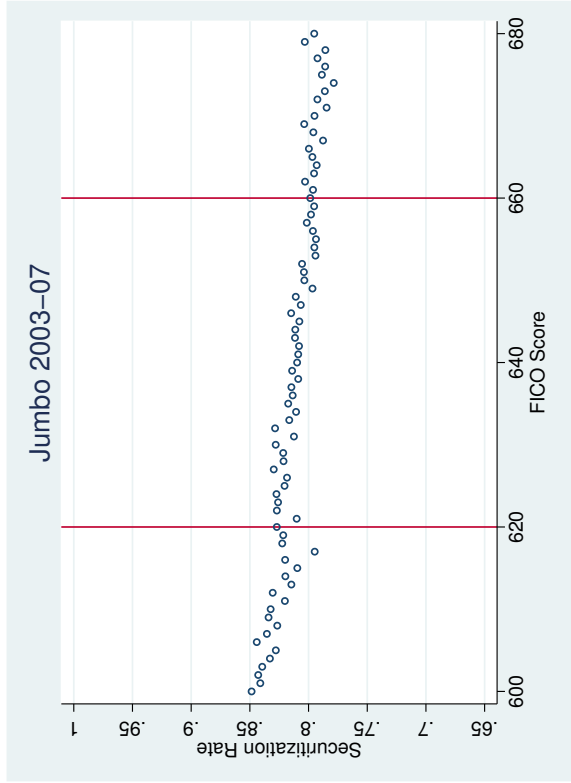


(c) Securitization

FIGURE 5. 2008–2009 Sample: FHA/VA loans. Source: LPS data.



(a) Frequency



(b) Default

(c) Securitized Rate

FIGURE 6. 2003–2007 Jumbo sample. Source: LPS data.

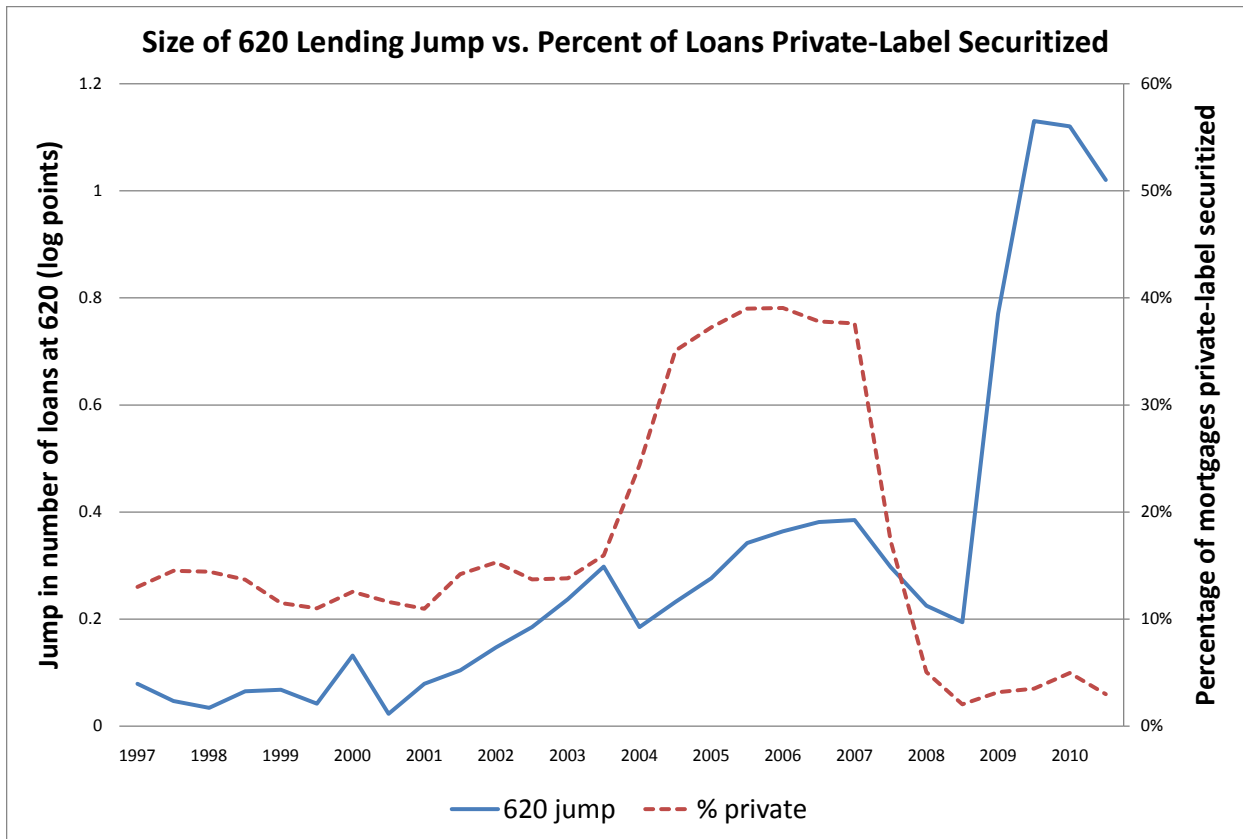
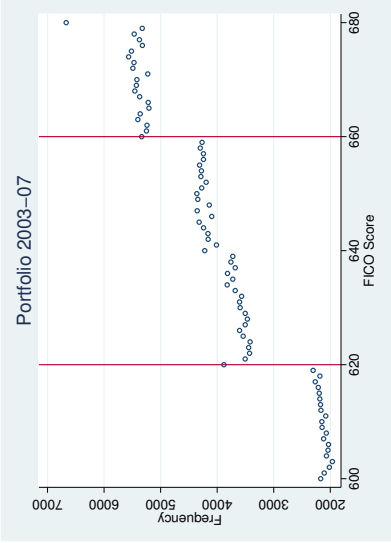
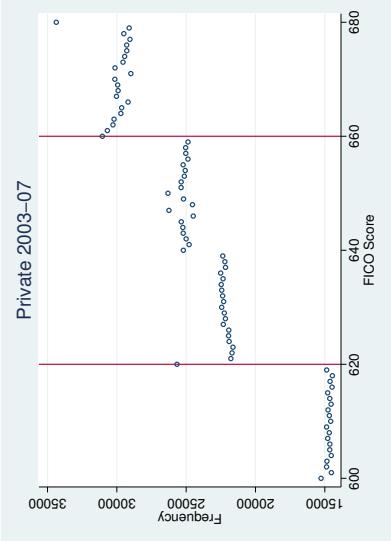
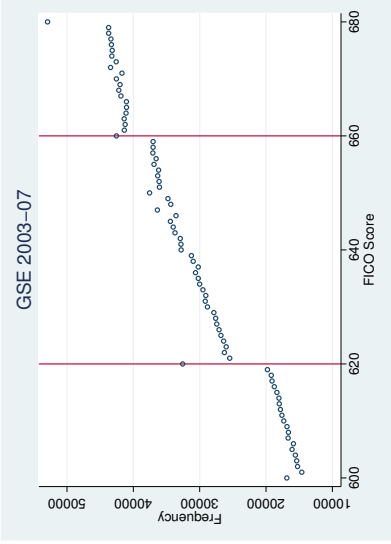


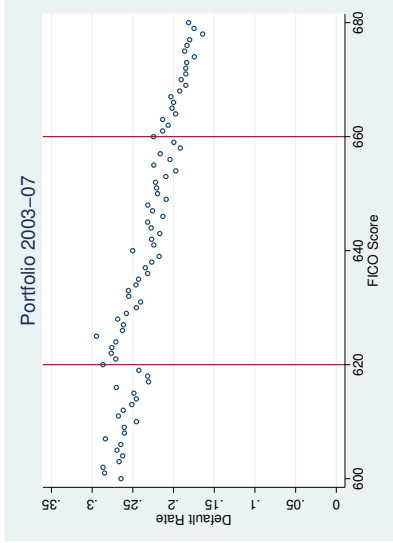
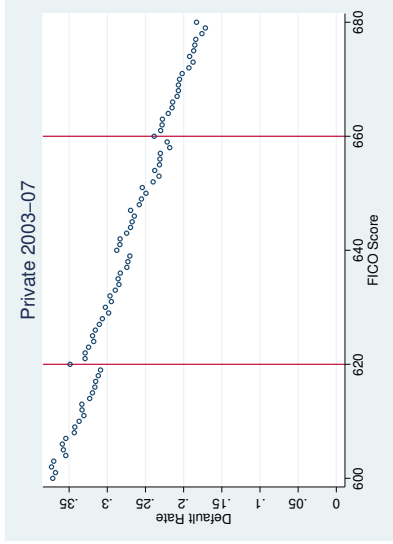
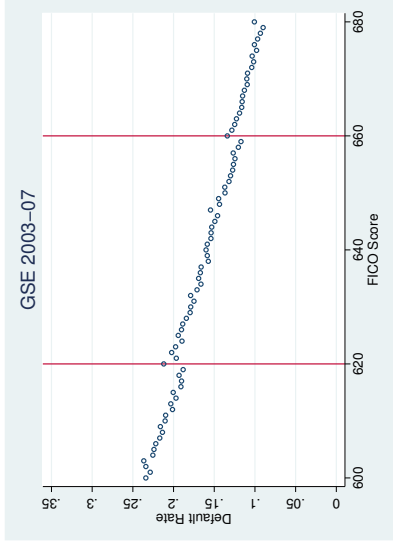
FIGURE 7. Size of 620 discontinuity vs. size of private-label MBS market for six-month cohorts based on origination date . Source: LPS data and Mortgage Market Statistical Annual.



(c) Frequency

(b) Frequency

(a) Frequency



(f) Default

(e) Default

(d) Default

FIGURE 8. Frequency and default discontinuities for portfolio, private-label securitized, and GSE securitized loans originated in 2003–2007. Source: LPS data.