BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel—2023-3

Last updated: January 5, 2023

Mandatory Retention Rules and Bank Risk

by Yuteng Cheng

Banking and Payments Department
Bank of Canada
ycheng@bankofcanada.ca

DOI: https://doi.org/10.34989/swp-2023-3 | **ISSN** 1701-9397

Acknowledgements

I would like to thank Dean Corbae, Erwan Quintin, Roberto Robatto and Jean-François Houde for invaluable advice and support. I also thank Dmitry Orlov, Rishabh Kirpalani, Paolo Martellini, Timothy Riddiough, Briana Chang, Donald Hausch, Anson Zhou, Jason Choi, Marco Duarte, Yunhan Shin, Jingnan Liu and seminar participants at the University of Wisconsin–Madison for helpful comments. All errors are mine. The views expressed in this paper are those of the author and not necessarily the views of the Bank of Canada.

Abstract

This paper studies, theoretically and empirically, the unintended consequences of mandatory retention rules in securitization. The Dodd-Frank Act and the EU Securitisation Regulation both impose a 5% mandatory retention requirement to motivate screening and monitoring. I first propose a novel model showing that while retention strengthens monitoring, it may also encourage banks to shift risk. I then provide empirical evidence supporting this unintended consequence: in the US data, banks shifted toward riskier portfolios after the implementation of the retention rules embedded in Dodd-Frank. Furthermore, the model offers clear, testable predictions about policy and corresponding consequences. In the US data, stricter retention rules caused banks to monitor and shift risk simultaneously. According to the model prediction, such a simultaneous increase occurs only when the retention level is above optimal, which suggests that the current rate of 5% in the US is too high.

Topics: Financial institutions; Financial system regulation and policies; Credit risk management

JEL codes: G21, G28

Résumé

Cette étude examine, sous un angle théorique et empirique, les conséquences fortuites des règles obligatoires de rétention qui s'appliquent aux titrisations. La loi Dodd-Frank et le règlement de l'Union européenne sur la titrisation imposent tous deux une exigence de rétention du risque de 5 % pour inciter à la sélection et à la surveillance des risques. L'auteur propose d'abord un nouveau modèle montrant que si la rétention renforce la surveillance, elle peut aussi encourager les banques à déplacer le risque. Il fournit ensuite des preuves empiriques de cette conséquence imprévue : les données américaines indiquent que les banques se sont tournées vers des portefeuilles plus risqués après l'entrée en vigueur des règles de rétention enchâssées dans la loi Dodd-Frank. De plus, le modèle permet de faire des prévisions claires et vérifiables sur la politique et ses conséquences. On constate dans les données américaines que les règles de rétention plus strictes ont amené les banques à simultanément surveiller et déplacer le risque. Selon la prévision du modèle, une telle augmentation simultanée ne se produit que lorsque le niveau de rétention est supérieur au niveau optimal, ce qui donne à penser que le taux actuel de 5 % aux États-Unis est trop élevé.

Sujets : Gestion du risque de crédit; Institutions financières; Réglementation et politiques relatives au système financier

Codes JEL : G21, G28

1 Introduction

The surge in securitization activities before 2008, especially between 2005 and 2007, has received criticism for being one of the main causes of the financial crisis. With loans removed from the balance sheet, banks bypass the downside risk if a borrower ultimately defaults, which reduces their incentive to collect soft information and monitor risk (e.g., Gorton and Pennacchi (1995); Parlour and Plantin (2008); Mian and Sufi (2009); Keys et al. (2010)). Regulating securitization was the primary policy response to the crisis, the key piece in the US being the mandatory risk retention rule incorporated in Section 941 of the 2010 Dodd-Frank Wall Street Reform and Consumer Protection Act. The rule requires securitization issuers to retain *no less than* 5% of the underlying credit risk of the assets collateralizing the asset-backed securities. Former Congress member Barney Frank, who co-sponsored the act, described the requirement as the "*single most important part of the bill.*" Similar 5% retention rules were developed in Article 2(1) of the EU Securitization Regulation. Those rules intend to better align the incentives of financial intermediaries and asset-backed security (ABS) investors. Policymakers expect banks to screen and monitor their borrowers more cautiously and carefully with more skin in the game, hence reducing risk.

To better understand the impact of this policy, this paper studies two related research questions. First, can mandatory retention have unintended consequences? Second, is the current level of retention optimal? To answer those questions, I propose a novel model in which retention strengthens monitoring (intended consequence) but may also encourage banks to shift risk (unintended consequence). More importantly, the unintended consequence is not just a theoretical possibility; I provide empirical evidence that banks shifted toward riskier portfolios after the implementation of the retention rules embedded in the Dodd-Frank Act. Furthermore, the model provides clear, testable predictions about policy and the corresponding consequences. In the data, stricter retention rules in the Dodd-Frank Act caused banks to monitor and shift risk simultaneously. According to the model prediction, such a simultaneous increase can only occur when the retention level is above optimal. Therefore, this paper suggests that the current rate of 5% in the US is too high.

In the model, upon choosing a project, the bank securitizes and sells to investors a fraction of the project, subject to the retention requirement. The bank's credit riskiness is determined by a two-dimension moral hazard friction. One dimension is *risk shifting* (Jensen and Mecking, 1976), also known as *asset substitution*, which is the opportunity for a bank to replace a high-net present value (NPV) project (a prudent project) with a low-NPV project that yields higher private returns if it succeeds (a gambling project).

The gambling project thus contains higher credit risk. While the downside risk will be absorbed by debtholders, the private returns go to shareholders if the investment strategy pays off. Hence, risk is shifted toward debtholders. The other dimension is *costly monitoring effort* after the securitization stage, which maintains the quality of the project. In practice, monitoring includes collecting payments, renegotiating, and working closely with the trustee representing investors' interests. The bank cannot commit to choosing the prudent project as well as the monitoring effort. Mandatory retention interacts with the two-dimensional moral hazard. At a low retention ratio, the bank has low interest in monitoring because they have few items left on the book, but chooses the prudent project because it is easier to securitize with its favorable market value. At a high retention ratio, the bank monitors more carefully and cautiously. Still, it holds a larger share of loans, hence the standard argument that limited liability creates an incentive to gamble becomes relevant. Consequently, the bank monitors but shifts risk. The main model focuses on bank monitoring. The mechanism is still valid when monitoring is replaced with *screening*, which is shown in Appendix C.

Given the trade-off, the model is capable of characterizing the socially optimal retention ratio. I show that the optimal ratio is interior, meaning that a certain degree of retention is the correct policy. Mandatory retention has three effects: it induces monitoring, it may also encourage risk shifting, and it reduces the gain from trade in securitization. The welfare is maximized when the prudent project is selected and monitoring is incentivized, without securitization activities being overly regulated.

The model gives clear predictions about the link between policy and the corresponding consequences. In particular, if a simultaneous increase of monitoring and risk shifting is observed, the retention ratio is above the optimal level. I exploit this result to test whether the current 5% mandatory retention ratio is too high. I use the Dodd-Frank Act as a quasi-experiment and find that this is the case in the US bank holding companies (BHC) data.

The empirical analysis features difference-in-difference estimations focusing on the behavior of BHC securitizers. The final rule of mandatory retention in Dodd-Frank was implemented at different times for different types of securitization. For residential mortgage-backed securities (RMBS), the center of the Dodd-Frank Act, the implementation date was December 2015 and one year later for other securitization categories. I use the risk-weighted assets (RWA) ratio to measure risk shifting. An increase in this ratio represents risk-increasing change in a bank's asset portfolio. I find that RMBS issuers significantly increased this ratio by 2 percentage points after the effective date of the retention rules on RMBS, signifying that banks are moving toward riskier investment strategies. One concern is that if the bank retains more tranches, the RWA ratio may automatically go up because those tranches are attached to higher risk weights. To deal with this, I use the RWA less securitization exposure to proxy portfolio risk outside tranches the bank retains in securitization. My results are robust to this specification.

On a not-readily-observable dimension of risk, I analyze the change in the delinquency rate of banks. I show that the delinquency rate of RMBS-only securitizers decreased about 0.3 percentage points on average after the mandatory rules' implementation, meaning ex post, loans became safer. To conform with the fact that delinquency takes a longer time to happen, I also consider an alternative difference-in-difference setup in which I compare RMBS-only securitizers and banks that are mortgage sellers in an extended sample period. Both ordinary least squares (OLS) and propensity score matching deliver similar results. The higher ex ante risk of the loans and their better ex post performance are consistent with banks exerting more effort to screen and monitor borrowers after the implementation of the retention rules. In sum, there was a simultaneous increase in bank screening/monitoring and risk shifting, confirming the model prediction that the current rate of 5% is overshooting.

Lastly, my results have implications for the optimal retention form. In the final rule, banks are allowed to retain a horizontal interest that consists of the most subordinated tranches or a vertical interest in each class of ABS tranches (or a hybrid of both). In a separate discussion, I show first that the trade-off between monitoring and gambling exists under both retention forms. Second, the horizontal component generates a higher expected value for the bank when the capital requirement is not binding. This component requires the bank to retain more on the balance sheet due to more involved risk, leading to more gains in the good states and more losses in default. The net expected gains are higher thanks to limited liability. If the capital requirement is a concern, then the bank has to retain the vertical component when the equity issuance cost is too high. Finally, I show that at a fixed retention ratio, ceteris paribus, horizontal retention makes the bank more willing to monitor its borrowers after securitization; however, it also makes over-regulating and risk shifting ex ante more likely. The reason is that the larger fraction of the loan retained under horizontal retention resembles an increased retention ratio. From a welfare perspective, another trade-off arises with regard to the optimal retention form.

Related literature. This paper connects several different strands of literature. The first literature studies how securitization negatively affects banks' traditional roles. Theoretically, Pennacchi (1988), Gorton and Pennacchi (1995), Petersen and Rajan (2002), and Parlour and Plantin (2008) argue that securitization leads to a decline of the originating bank's screening and monitoring incentives. Empirically, Keys et al. (2010) and Purnanandam (2011) show that securitization led to lax screening standards of mortgages; Piskorski, Seru, and Vig (2010) and Agarwal et al. (2011) provide evidence that securitization damaged servicing of loans, in particular renegotiation of delinquent loans. While there are no policy or risk shifting implications in the above literature, this paper presents policy analysis and links it with banks' risk shifting motives.

The second literature discusses the optimal retention form, that is, which tranches banks should retain in response to the retention requirement. Fender and Mitchell (2009) and Kiff and Kisser (2014) discuss the optimality of equity and mezzanine tranches in maximizing screening efforts. Pagès (2013) finds that to implement optimal delegated monitoring by the bank, the securitization scheme should use a cash reserve account rather than retention of the residual interest. Malekan and Dionne (2014) study the optimal contract with regard to retention in the presence of moral hazard in lender screening and monitoring. In those papers, the retention requirement is fixed. Instead, this paper studies the optimal requirement, and my results hold under different retention forms. The optimal retention form is also discussed under the broader concept of moral hazard.

This paper is also related to the large literature on risk shifting uncovered by Jensen and Mecking (1976). For an early literature review, see Gorton and Winton (2003). Keeley (1990), Demsetz, Saidenberg, and Strahan (1996), and Repullo (2004) demonstrate the negative first-order effect of profitability on risk shifting. The focus of this literature is mainly capital requirements. This paper instead sheds light on risk shifting and retention requirements in securitization.

Moreover, this paper contributes to the relatively small empirical literature on the impact of retention rules. Furfine (2020) shows that after the implementation of the retention rules, loans in the commercial mortgage-backed securities (CMBS) market become safer, as measured by indexes such as interest rates, loan-to-value ratios, and income to debt-service ratios. In a similar manner, Agarwal et al. (2019) find that underwriting standards in the CMBS market are tighter after the implementation. The results in the two papers supplement my empirical findings on banks' monitoring and screening behavior, but they do not capture the risk shifting aspect.[1]

Lastly, there is a literature about signaling private information through retention, pioneered by Leland and Pyle (1977) and DeMarzo and Duffie (1999). Guo and Wu (2014) argue

[1]In a related paper, Sarkisyan and Casu (2013) show that retained interests increased bank insolvency risk before the crisis.

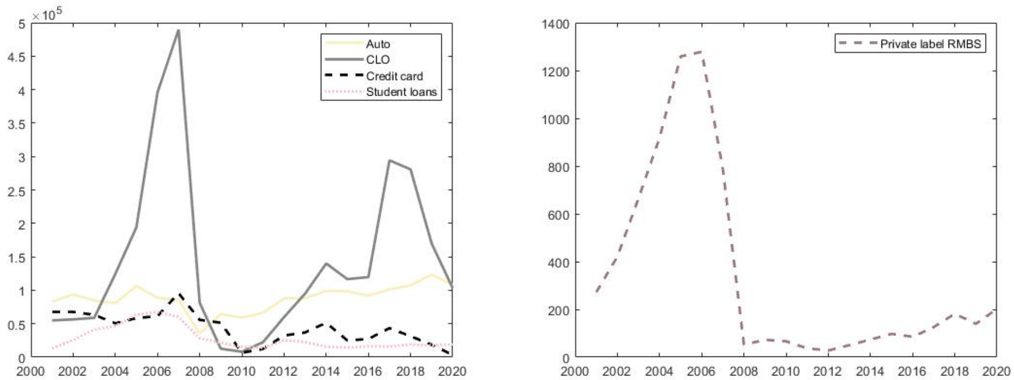

Figure 1: Issuance of securitization by class (Unit: $ billion). Source: Securities Industry and Financial Markets Association.

that mandatory retention may deteriorate the adverse selection problem because it prevents issuers from using the level of retention as a signal. On the other hand, Flynn, Ghent, and Tchistyi (2020) show that banks can in fact signal through the retention structure of vertical and horizontal interests. The retention policy is fixed in those papers. This paper examines the optimal policy but does not contain signaling. In Chemla and Hennessy (2014), screening is combined with a follow-up possible signaling process through junior tranches, and optimal retention form is also discussed. But in their paper the retention requirement is fixed and there is no risk shifting.

Layout. The paper proceeds as follows. Section 2 describes the institutional background of securitization and risk retention. Section 3 highlights relevant facts about bank risk before and after the implementation of the risk retention rule. Section 4 presents the model, and Section 5 formally tests its predictions. Section 6 and Section 7, respectively, extends the model and concludes.

2 Institutional Background

Securitization activity ground to a halt in all categories as a consequence of the recent financial crisis. As shown in Figure 1, almost all types of securitization issuance experienced a sharp increase in the 2005–2007 period, especially for collateralized loan obligations (CLOs) and private label RMBS, and a dramatic decline during the crisis. The recovery has been slow post crisis, the exception being auto loans and CLOs.

In a standard securitization process, a loan originator determines if a borrower qualifies

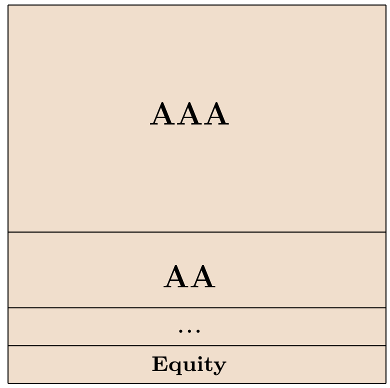

Figure 2: Subordination

for a loan and, if so, the interest rate of the loan. Having originated a loan, the originator sells it to an issuer (sometimes referred to as sponsor), who brings together the collateral assets from originators for the asset-backed security. For banks studied in this paper, they serve as issuers and originators of the portfolio of securitized assets at the same time.[2] The issuer pools assets together and sells them to an external legal entity, often referred to as a special-purpose vehicle (SPV). The structure is legally insulated from management. The SPV then issues security, dividing up the benefits (and risks) among investors on a pro-rata basis. The issuer usually keeps the servicing rights, that is, the responsibility for managing payments and working closely with the trustee who represents investors. Credit enhancements are also provided by the issuer[3] to protect investors from potential losses on the securitized assets, the most common forms being subordination, overcollateralization, and excess spread. In the subordination process, each security includes several senior tranches rated AAA (senior), a class of subordinate tranches with a rating below AAA (mezzanine), and an unrated residual equity tranche, as shown in Figure 2. The thickness of the senior and the mezzanine tranches depend on the quality of the underlying assets. The residual equity tranche is usually very small (below 2%). In a typical Alt-A MBS deal, the senior tranches account for 90% of the deal. The subordinate classes serve as credit support for the higher-rated senior bonds and would initially absorb credit losses. The most junior tranche will absorb the first loss, and so on. Under overcollateralization, the face value of the underlying loan pool is larger than the par value of the issued bonds as additional cushion. Excess spread is the additional revenue generated by the difference between the coupon on the underlying collateral (e.g., a

[2]There are also loans sold to another party, like Fannie Mae and Freddie Mac, who in turn securitize those loans.

[3]It can also be provided by third parties.

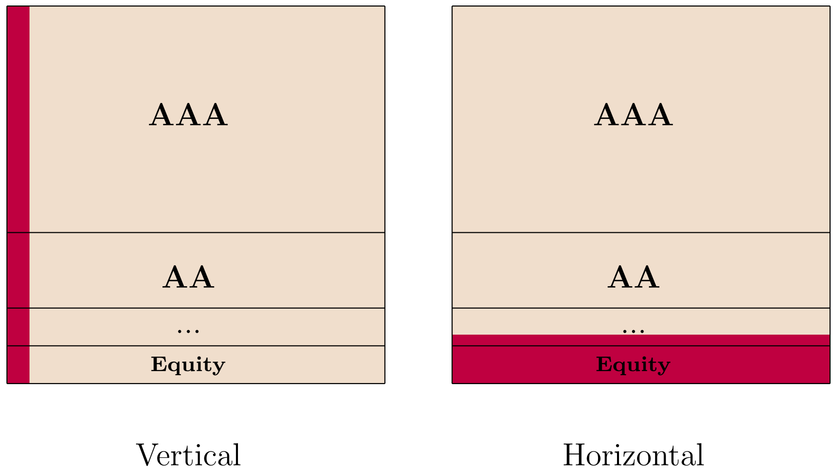

Figure 3: Retention options

mortgage interest rate) and the coupon rate on the securities, which can be used to absorb losses. It is also called credit-enhancing interest-only strips.

Before the mandatory retention rule was introduced, it had long been common practice for securitizers to retain an economic interest, typically first-loss contractual interests mentioned above. In October 2014 after the crisis, to better align the incentives involved in the securitization process, the SEC, FDIC, Federal Reserve, OCC, FHFA, and HUD adopted a final rule (the Final Rule) implementing the requirements of Section 15G of the Exchange Act, which was added pursuant to Section 941 of the Dodd-Frank Act. The Final Rule requires an issuer of ABS to retain at least 5% of the credit risk related to that securitization and restricts the transfer, hedging, or pledge of the risk that the sponsor is required to retain. The issuer must retain either an eligible *vertical interest* (an interest in each class of ABS interests issued as part of the securitization), *horizontal interest* (the issuer holds the most subordinated claim to payments of both principal and interest transactions), as shown in Figure 3, or a *combination of both* so long as the combined retention is not less than 5% of the fair value of the transaction. For the eligible horizontal interest option, the amount of the required risk retention must be calculated under a fair value approach under generally accepted accounting principles (GAAP). The Final Rule came into effect in December 2015 for RMBS and December 2016 for other ABS. A similar retention rule has been introduced by the EU, covered by Article 2(1) of the securitization Regulation. And securitizers have to be in compliance with the rules since January 2019.

Some exemptions exist for particular categories of securitization in the Dodd-Frank Act. For example, sponsors of securitization pools that are solely composed of qualified residential

Table 1: **Measures of risk**

	RMBS securitizers					
	Full sample		Before Dec 2015		After Dec 2015	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
RWA ratio	0.762	0.092	0.749	0.085	0.773	0.097
DEL	0.021	0.027	0.024	0.028	0.019	0.025

Notes: The table describes the RWA ratio and delinquency rate of RMBS securitizers from 2014 to 2017.

mortgages, as defined by the Consumer Finance Protection Bureau (CFPB) under the Truth in Lending Act, are not required to retain any risk; and CLO managers are not subject to risk retention due to a court ruling in 2018. Beyond that, Section 15G permits the agencies to adopt other exemptions from the risk retention requirements for certain types of ABS transactions.

3 Motivating Facts

This section reports on how banks performed before and after the implementation of the risk retention requirement. I use US bank holding company (BHC) data from 2014Q1 to 2017Q4 from Y-9C forms. Risk and capital management are typically carried out at the highest level in bank holding companies. I focus on 35 BHC securitizers that only issue RMBS. Detailed data description is provided in Section 5.

I document bank risk with an observable measure and a not-readily-observable measure. The former is the risk-weighted assets (RWA) ratio, defined as

$$\text{RWA ratio} = \frac{\text{RWA}}{\text{Total assets}}$$

A higher RWA ratio implies that the bank allocates resources toward riskier projects. Hence it is an ex ante measure of bank risk, in particular a measure of risk shifting.[4] The second measure is the securitization delinquency ratio, defined as

$$\text{DEL} = \frac{\text{loans 90 or more days past due} \; + \; \text{(loans charged off)}}{\text{lagged total loans outstanding}}$$

The delinquency rate measures the fraction of non-performing loans, hence it is an ex post index of bank risk.

[4] To my knowledge, there is no other commonly accepted measure of risk shifting.

Table 1 presents the summary statistics on the risk measure of the RMBS BHC securitizers. Notice that, following the implementation of the mandatory risk retention rule, the average RWA ratio rose, but the average delinquency rate fell. On the one hand, roughly, banks shift toward riskier loans. On the other hand, loans are safer ex post, which implies that banks exert more efforts in screening and monitoring those loans. I postpone the formal tests of those arguments to Section 5. The next section develops a model that can generate these observations for a high retention ratio.

4 The Model

This section presents the main model. I consider a bank that securitizes a fraction of its assets and sells to outside investors. Before the securitization transaction, the bank can manipulate the quality of its assets via risk shifting. And after the securitization stage, the bank further affects its risk by choosing whether to monitor the remaining loans and tranches on its book. I illustrate how mandatory retention in the securitization process affects the bank's decisions and market outcomes, and how the above facts represent an overshooting retention ratio. And for exposition and simplicity, the retention form in this section is assumed to be vertical.[5]

4.1 Model Setup

The model has four dates: $t \in \{0, 1, 2, 3\}$. The bank and securitization investors are all risk neutral. At time 0, the bank starts with equity e and raises $d = 1 - e$ from the deposit market. After the financing is done, the bank allocates its assets, wherein it faces a risk shifting problem. Specifically, the bank chooses between two projects: a prudent project,[6] which generates $R_0 > 1$ at time 3, and a gambling project, which generates at time 3

$$\begin{cases} R_b & \text{with prob. } p_b \\ r & \text{with prob. } 1 - p_b \end{cases}$$

where $R_b > R_0$ and $p_b < 1$. I assume that if the bad cash flow r is realized, the bank fails and r can be considered as the liquidation value of the bank. For simplicity, we assume $r = 0$

[5]Discussions of the optimal retention form and the corresponding capital requirement implications are moved to Section 6. My main results of the trade-off between risk shifting and monitoring hold under both forms of retention.

[6]The term "project" should be viewed as a metaphor for the bank's portfolio of projects.

in this section and relax that assumption in Section 6. The values of the projects satisfy

$$R_0 > p_b R_b > 1,$$

that is, the prudent project has higher expected returns, but the gambling project yields higher private returns if the gamble pays off. The risky project still has positive net present value. If the bank is solvent at time 2, it repays d to depositors. The bank has limited liability. In this manner, when the gamble fails, the cost is borne by depositors. Deposits are covered by mispriced insurance, which allows us to omit the analysis of depositor behavior. The presence of deposit insurance is not crucial in this paper. I show in Appendix B that the main results still hold when the bank uses unsecured debt from the credit market. For simplicity, I assume the insurance premium is 0.[7]

At time 2, the bank chooses a monitoring effort $m \in \{0, 1\}$ that also affects its overall credit risk, where 1 indexes for monitoring and 0 indexes for no monitoring. Specifically, for a chosen project (p, R) (with $p_0 = 1$), the ultimate probability of generating R given m is

$$q(p, m) \equiv p - (1 - m)\Delta.$$

If the bank monitors, the probabilities of success stay the same, and the cost associated with monitoring is c. If the bank shirks in monitoring, the probability of success drops by Δ, and even the prudent project becomes risky. The marginal return of monitoring a project is hence ΔR. I make some assumptions on the monitoring technology. First, I assume that $\frac{\Delta_b}{c_b} < \frac{\Delta_0}{c_0}$, which means the efficiency of the monitoring technology is lower for the gambling project in terms of the higher monitoring cost per unit increase in the probability of success. Second, I assume the gross return of monitoring is higher for the gambling project, that is, $\frac{\Delta_b R_b}{c_b} > \frac{\Delta_0 R_0}{c_0}$. Without loss of generality, let $\Delta_0 = \Delta_b = \Delta$; that is, the change in the probability of success is the same for the two projects. The two assumptions can be summarized as follows.

Assumption 1 $1 < \frac{c_b}{c_0} < \frac{R_b}{R_0}$.

At time 1, the bank securitizes a fraction α of its project. The securitized project is repackaged into multiple securities (or "tranches") with different seniorities. The most junior tranche will absorb the first loss, and so on. The security design part, or subordination in practice, is not the focus of this paper; hence it is considered exogenous.[8] The α securitized fraction of the project will be called the *pool*.

[7] I also discuss a fairly priced deposit insurance premium in Appendix B.

[8] Mitchell (2004) provides an overview of tranching and the corresponding literature of security design.

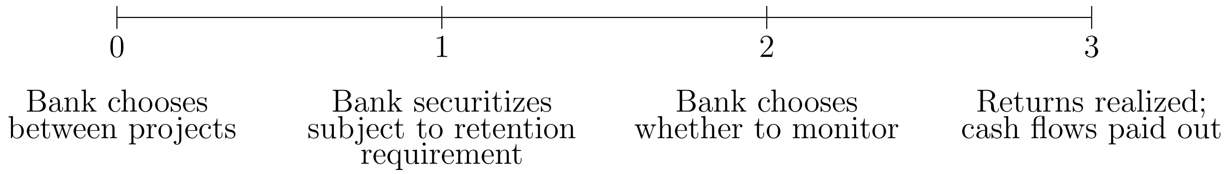

Figure 4: Timeline of the events

The securitization process is regulated by *mandatory retention requirements*: the bank has to retain β fraction of the tranches such that the retained tranche value is *no less than* θ fraction of the market value of the pool. The policymaker sets θ. In the vertical retention we focus on this section, the bank simply holds each class of the tranches, hence the retention constraint is

$$\beta \geq \theta,$$

that is, the bank needs to retain at least θ fraction of the pool.

Competitive investors observe the bank's project choice (p, R) and retention choice β, and pay a price (schedule) P. The project choice being observable to securitization purchasers is in accordance with ABS offering prospectuses that provide detailed information on the underlying collateral assets, as governed by Regulation AB from 2005 and reinforced in Dodd-Frank. Transparency of the retention choice is governed by the retention rules. Competitive investors bid for the securities. The total monetary benefit of securitization for the bank at time 1 is λP, where $\lambda > 1$ is the gain from trade. It can be interpreted, for instance, as lower funding costs in extending new loans, which I do not model precisely,[9] or an increase in reported profit that is related to the compensation of the CEO. Moreover, assuming $\lambda > 1$ is equivalent to assuming the bank discounts future cash flows at a higher rate than investors (e.g. DeMarzo and Duffie, 1999).

The timeline of the model is in Figure 4. I present the definition of the model before proceeding to the analysis.

Definition 4.1 (Equilibrium.) *The equilibrium consists of the bank's project choice* (p, R) *at time 0; the bank's retention choice* β *at time 1; price* P *of the tranches sold to investors; the bank's monitoring choice* m *at time 2 such that* P *reflects seniorities and project and retention choices; investors break even and the bank maximizes its value sequentially subject to retention constraint at time 1.*

[9] One can also assume there is a new project arriving at time 1. The project has an expected return λ. I omit the details of the new project in the analysis.

4.2 Benchmark: No securitization and no limited liability

This section studies a benchmark case where the bank holds the project on the balance sheet to maturity and there is no limited liability. No limited liability is equivalent to the bank being financed completely by equity. I use the bank's project choice and monitoring effort to approximate the socially optimal choices. The bank is maximizing the expected return of the project minus the potential monitoring cost:

$$\max_{(p,R),m} q(p,m)R - c(m)$$

I impose an assumption that the cost of monitoring is relatively low compared to the return of monitoring, $c < \Delta R$. With this assumption, the expected output is maximized when the prudent project is selected and the monitoring effort is exerted.

Proposition 4.1 (Benchmark) *In the benchmark case, the bank chooses the prudent project and monitors.*

4.3 Monitoring decision

The model is solved backwards. In this section, I solve for the bank's monitoring decision at time 2, after the securitization stage. Given the project and retention choices $((p,R),\beta)$, the optimal monitoring decision maximizes the bank's residual profits net of monitoring costs:

$$\Pi_2(\beta,p,R) = \max_{m\in\{0,1\}} q(p,m)\bigg((1-\alpha)R + \beta\alpha R - d\bigg) - c(m).$$

In the expression, when the investment is successful, $(1-\alpha)R$ is the value of the non-securitized part of the project, $\beta\alpha R$ is the value of the tranches retained on the bank's books, and d is the amount of debt the bank repays being solvent. I present the solution.

Proposition 4.2 *The bank chooses $m = 1$ if and only if $\beta \geq \beta^d \equiv \frac{\frac{c}{\Delta}+d-(1-\alpha)R}{\alpha R}$.*

Proof: See Appendix A.

The bank monitors if the loan retained on the balance sheet is large enough, that is, if there is "skin in the game." The threshold β^d is increasing in the inverse of the efficiency of the monitoring technology, measured by $\frac{c}{\Delta}$, and total debts d it owes: when the monitoring technology is highly efficient, the bank will bring forward its monitoring decision to increase profits; similarly, if the bank has to repay large amount of debts (higher leverage), it will hesitate to monitor because the benefits of being solvent may be outweighed by the costs of paying off the debts. In particular, the threshold is decreasing in the project return R, meaning that if the bank chooses the gambling project (with $R_b > R_0$), it has more incentive to monitor. The return on monitoring ΔR is higher for the gambling project: an increase in monitoring effort leads to an increase in the upper side return of the project, and the expected value of the gambling project increases more than that of the prudent project. In this sense, monitoring and gambling are complements.

4.4 Securitization problem at time 1

In vertical retention, investors purchase from the bank $1-\beta$ fraction of the pool and share the default risk of the underlying project evenly in terms of seniority with the bank. Of course, if the bank goes bankrupt, the deposit insurance company will take over the tranches. Although the transaction takes place before the monitoring stage and the bank cannot commit to monitoring, investors who observe β can later infer the bank monitoring decision by comparing it to β^d, and pay the corresponding adjusted price. Specifically, for a given project, the schedule of prices investors offer based on β and rational expectations is

$$P(\beta|p, R) = \begin{cases} (1-\beta)\alpha pR & \text{if } \beta \geq \beta^d \\ (1-\beta)\alpha(p-\Delta)R & \text{if } \beta < \beta^d \end{cases}$$

The expected value of the pool will be αpR if there is subsequent monitoring and $\alpha(p-\Delta)R$ if there is not.

The bank's Bellman equation at time 1 hence becomes

$$\Pi_1(p, R) = \max_{\beta} \ \lambda P(\beta|p, R) + \Pi_2(\beta|p, R)$$
$$\text{s.t. } \beta \geq \theta$$

The total monetary benefits of securitization is the product of λ and $P(\beta|p, R)$. The constraint is the retention constraint. In the vertical retention, the bank only needs to retain a fraction larger than θ of the pool. We first characterize the shape of the bank's objective function as a function of β.

Proposition 4.3 *The bank's objective function $\lambda P(\beta|p, R) + \Pi_2(\beta|p, R)$ is upper semicontinuous at β^d. Moreover, it is linear and decreasing in β on $[0, \beta^d)$ and $[\beta^d, 1]$.*

Proof: See Appendix A.

To make monitoring indeed a friction, I assume for each project (p, R) the following holds.

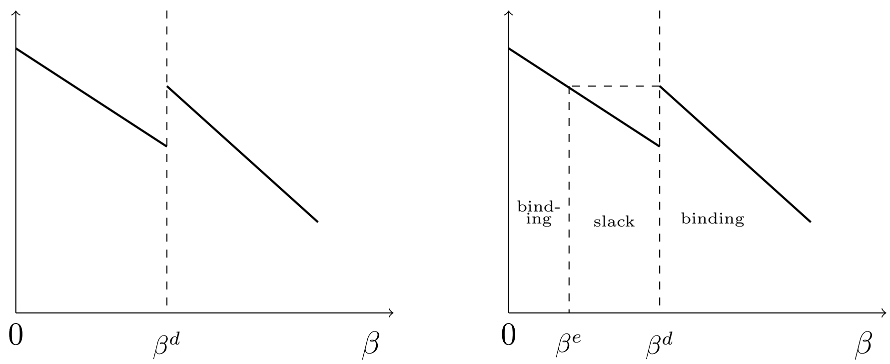

Figure 5: The bank's objective function and retention constraint

Assumption 2 $\lambda\big(\Delta R - pd\big) < c.$

In order to monitor, the bank retains at least β^d of the project. Monitoring improves project quality, but the expense is the lowered securitization revenue due to retention and the monitoring cost. By this assumption, the cost outweighs the benefits. This assumption guarantees that in equilibrium, without mandatory retention rules, the bank chooses zero retention. If the assumption is violated, the bank always chooses $\beta = \beta^d$ and monitors. There is no need to introduce regulation to address the misaligned incentive of monitoring. The left part of Figure 5 portrays the bank's objective function under this assumption. The second piece of the function is steeper than the first piece.

We are now ready to impose the retention constraint. Let β^e be such that

$$\lambda P(\beta^e|p, R) + \Pi_2(\beta^e|p, R) = \lambda P(\beta^d|p, R) + \Pi_2(\beta^d|p, R)$$

According to the shape of the objective function, this β^e exists and is smaller than β^d, as shown in the right part of Figure 5. In fact, the formula of β^e is

$$\beta^e = 1 - \frac{(\lambda - (p - \Delta))(R - d) - (\lambda - 1)\frac{c}{\Delta} - c}{(\lambda - 1)\alpha(p - \Delta)R}. \tag{1}$$

If the required threshold θ is below β^e, the bank chooses the lowest possible retention, implying that the retention constraint is binding. When the threshold θ is between β^e and β^d, the objective function attains the maximum at θ^k: the retention constraint is slack. When θ is above θ^k, the constraint is binding again since the objective function is decreasing. These are summarized by the following result.

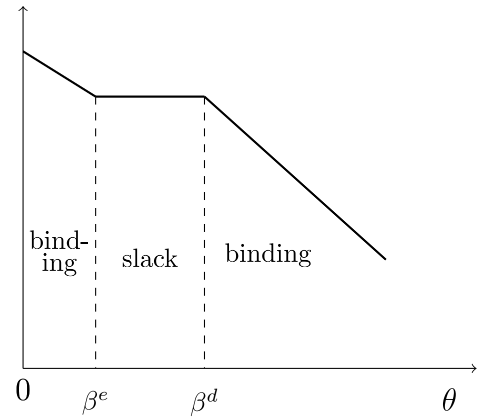

Figure 6: The bank's expected value under the optimal retention choice

Proposition 4.4 (Retention choice.) *The bank's optimal retention choice of β is*

$$\beta(\theta) = \begin{cases} \theta & \text{if } \theta < \beta^e, \\ \beta^d & \text{if } \beta^e \leq \theta < \beta^d, \\ \theta & \text{if } \theta^k \leq \theta. \end{cases}$$

Without the retention requirement, the bank chooses zero monitoring effort. The policymaker can set the lower bound θ of the retention requirement to β^e to induce monitoring efforts from the bank, which as we show later is part of the socially optimal allocation. Note that β^e is smaller than β^d. The policymaker only needs a small lower bound of θ to incentivize efficient monitoring.

Proposition 4.5 (More retention leads to more monitoring.) *Given a project (p, R), the policymaker can set $\theta \geq \beta^e$ to implement monitoring.*

For a given project, under the optimal retention choice, the bank's expected value, as a function of θ, is denoted by

$$\Pi_1(\theta|p, R) \equiv \lambda P(\beta(\theta)|p, R) + \Pi_2(\beta(\theta)|p, R)$$

and graphed in Figure 6. It is constant on $[\beta^e, \beta^d]$ because of the slack retention constraint.

Lastly, I present a side result that will be used later.

Lemma 4.1 $\beta^e_b < \beta^e_0$.

Proof: See Appendix A.

4.5 Risk shifting at time 0 and equilibrium

In this section, I complete the equilibrium characterization by studying the bank's risk shifting motives in the presence of limited liability, which imposes more costs on creditors in bad outcomes. I show that mandatory retention increases the bank's risk shifting propensity, defined by the difference between the values of gambling and being prudent concerning project choice.

At time 0, for a required retention ratio θ (and with a slight abuse of notation), let the bank's expected value of choosing the prudent project be $V_0(\theta) \equiv \lambda P(\beta_0(\theta)|p_0, R_0) + \Pi_2(\beta_0(\theta)|p_0, R_0)$. The retention $\beta_0(\theta)$ comes from Proposition 4.4. Similarly, if the bank shifts risk, its value changes to $V_b(\theta) \equiv \lambda P(\beta_g(\theta)|p_b, R_b) + \Pi_2(\beta_g(\theta)|p_b, R_b)$. The bank's project choice affects both the payments from securitization purchasers and the expected profits from the remaining items on the balance sheet. The gain from risk shifting is defined as

$$G(\theta) \equiv V_b(\theta) - V_0(\theta),$$

and the bank chooses to gamble if $G(\theta) > 0$. The equilibrium depends on the behavior of this gain function under the different requirements of θ. Note that this gain function has four kink points since both $V_0(\theta)$ and $V_b(\theta)$ have two: β_0^e, β_0^d and β_b^e, β_b^d, as shown in Figure 6. We already know that $\beta_b^e < \beta_b^d < \beta_0^d$ and $\beta_b^e < \beta_0^e$. I make the following assumption so that in equilibrium it is possible for the bank to choose the prudent project and monitor it at the same time.

Assumption 3 $\left(\frac{\lambda-1}{\Delta} + 1\right)(c_b - c_0) > \lambda\left(p_b(R_b - d) - (R_0 - d)\right)$.

According to the assumption, when the retention constraint is slack for both projects, the monitoring cost of the gambling project is relatively high compared to the additional securitization revenue generated due to the lower retention at such θ. It guarantees that $V_0(\beta_0^d) > V_b(\beta_b^d)$: that is, when both retention constraints are slack, if the bank starts to monitor the prudent project, it obtains a higher value from that project than from gambling. The main result of the paper is the following.

Proposition 4.6 (Equilibrium) *There exist two thresholds $\underline{\theta} < \bar{\theta}$ such that the bank chooses the prudent project but does not monitor for $\theta < \underline{\theta}$; the bank chooses the prudent project and makes an effort to monitor for $\theta \in [\underline{\theta}, \bar{\theta}]$; the bank gambles and monitors subsequently for $\theta > \bar{\theta}$.*

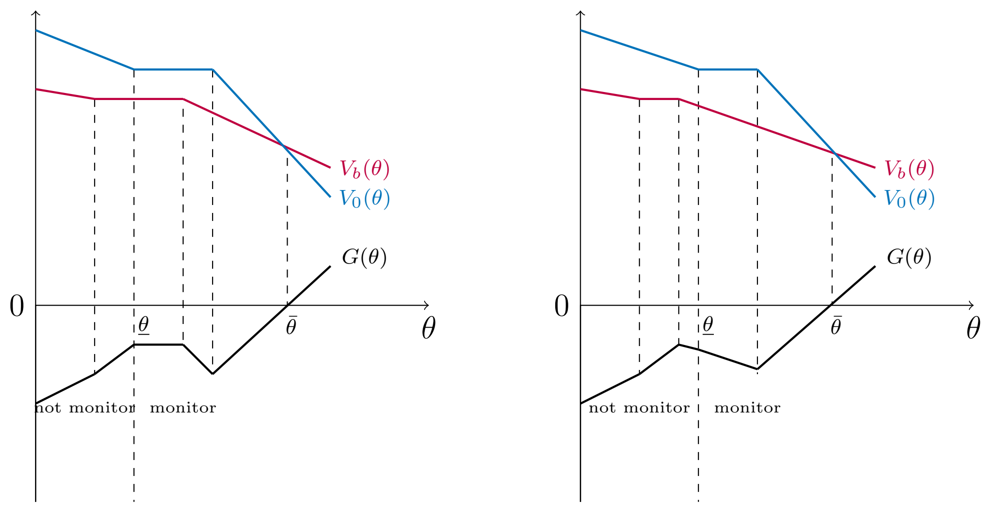

Figure 7: Risk shifting propensity

Proof: See Appendix A.

The above equilibrium characterization implies that at a high retention ratio θ, the bank starts to gamble. Because the gambling project has lower NPV than the prudent project, the total output is lower. This is the unintended consequence of mandatory retention rules.

Corollary 4.1 (Unintended consequence of mandatory retention.) *The bank starts to gamble when* $\theta > \bar{\theta}$.

The two possibilities of $V_0(\theta)$, $V_b(\theta)$, and the corresponding $G(\theta)$ are portrayed in Figure 7. At a low retention ratio, the bank chooses the prudent project because it has favorable market value, hence it is easy to securitize; and this is the driving force of $G(\theta)$ being monotonic increasing on some region. As the retention ratio increases, the value of choosing the prudent project is hurt more because of the higher gain from trade generated. When the retention ratio is high enough, limited liability and higher upper side return R_b make the gambling project more profitable.

It is worth comparing the other two scenarios: output when the prudent project is chosen but not monitored, and output when the gambling project is chosen and monitored. I assume the latter is higher: $p_b R_b - c > (1 - \Delta) R_0$. This assumption implies that

$$p_b > 1 - \Delta.$$

That is, the default rate is lower when there is monitoring even though some risk is shifted. First, this is consistent with the narrative that lax monitoring in securitization leads to bad loans and the crisis. Moreover, it also matches the empirical evidence explored later showing that the mandatory retention rule has caused banks to monitor and gamble simultaneously, but loan delinquencies have dropped.

4.6 Welfare and optimal retention requirement

This section provides the welfare analysis and discusses the optimal retention ratio. The total welfare is defined as the sum of the payoffs of all players[10] as a function of θ. Since investors get zero in expectation, the sum is just the bank's value V^{bank} plus debtholders' expected return, which is $p(\theta)$, the probability of the equilibrium project being successful.[11] In other words,

$$W(\theta) = V^{\text{bank}}(\theta) + p(\theta),$$

where

$$V^{\text{bank}}(\theta) = \begin{cases} V_0(\theta) & \text{if } \theta \leq \bar{\theta}, \\ V_b(\theta) & \text{if } \theta > \bar{\theta}, \end{cases}$$

according to Proposition 4.6 and

$$p(\theta) = \begin{cases} 1 - \Delta & \text{if } \theta \leq \beta_0^e, \\ 1 & \text{if } \theta \in [\beta_0^e, \bar{\theta}], \\ p_b & \text{if } \theta > \bar{\theta} \end{cases}$$

Note that $V^{\text{bank}}(\theta)$ is continuous but $p(\theta)$ is not. After rearrangements, the welfare can be written as

$$W(\theta) = \big(1 + (\lambda - 1)\alpha(1 - \beta(\theta))\big)\bar{R}(\theta) – c(\theta),$$

which is the sum of total output and the net gain from securitization less the potential monitoring cost. The shape is shown in Figure 8.

There are jumps at β_0^e and $\bar{\theta}$ because $p(\theta)$ has jumps at these points. Since securitization brings gains of trade, a higher retention ratio reduces securitization levels and, hence, brings down the total welfare. This is why the welfare is decreasing on each piece, except for $[\beta_0^e, \beta_0^d]$, on which the retention constraint is slack. In fact, if we impose the following assumption, the welfare is maximized at $[\beta_0^e, \beta_0^d]$ and maximized when the prudent project is chosen and monitoring effort is delivered:

[10] That is, equal Pareto weights are attached to each player.

[11] Or the sum of depositors and the deposit insurance company, with zero insurance premium. The depositors get 1 with certainty, and the deposit insurance company expects to pay $(1 - p(\theta))$.

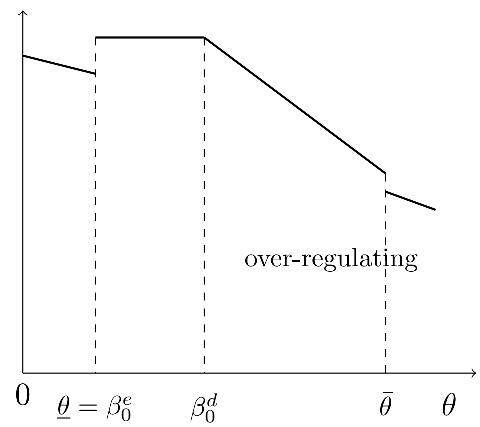

Figure 8: Welfare

Assumption 4 $\big((\lambda - 1)\alpha + 1\big)\Delta R_0 - (\lambda - 1)\alpha(1 + \frac{c_0}{\Delta}) > c_0$.

Similar to the benchmark case, the social monitoring cost and loss of gains from trade due to retention are low compared to the additional social benefits generated from monitoring the prudent project. As λ converges to 1, the above assumption converges to $\Delta R_0 > c_0$. Since $\underline{\theta} = \beta_0^e$, we have the following result:

Proposition 4.7 (Optimal retention.) *The optimal retention ratio is* $[\underline{\theta}, \beta_0^d]$*. On this range of* θ*, the bank invests in the prudent project and monitors and the retention constraint is slack.*

Proof: See Appendix A.

The optimal retention ratio is interior, meaning that a certain degree of retention is the correct policy. Mandatory retention has three effects: it induces monitoring, it encourages risk shifting if the degree of retention is very high, and it restricts securitization activities and hence reduces the gains from trade. What happens when θ is between β_0^d and $\bar{\theta}$? In this region, the bank still chooses the prudent project and delivers monitoring effort according to Proposition 4.6; however, its securitization activities are over-regulated. In summary, welfare is maximized when the prudent project is selected and just monitored, without securitization activities being overly regulated.

Two remarks are worth making. First, if the policymaker puts 100% Pareto weight on debtholders, then the total welfare is simply $p(\theta)$, the probability of success. And identical to the benchmark case, the optimal retention is $[\underline{\theta}, \bar{\theta}]$. Second, the policymaker is imposing an inequality constraint. If this inequality constraint is replaced with an equality constraint, that is, the bank has to retain θ fraction of the pool, the optimal retention is unique: $\theta^* = \beta_0^d$. This is the lowest level of retention to incentivize the bank to monitor. And the policymaker does not want to go beyond that to over-regulate securitization. Since β_0^d is larger than $\underline{\theta}$, based on the equilibrium characterization, β_0^d can be close to $\bar{\theta}$, above which the bank shifts risk. Therefore, choosing an inequality retention constraint leaves leeway for the policymaker to better balance monitoring and risk shifting.

4.7 Model predictions

This section connects the above results with empirical tests in the next part of the paper. The model provides sharp predictions about the link between policy and unintended consequences, which is summarized in Table 2. In particular, if we observe no changes in monitoring and risk shifting, it implies that the current retention ratio is below the optimal level. Similarly, if we observe a simultaneous increase in monitoring and risk shifting, the current retention ratio is too high.

	Monitor	Shift risk
$\theta < \underline{\theta}$	N	N
$\theta \in [\underline{\theta}, \bar{\theta}]$	Y	N
$\theta > \bar{\theta}$	Y	Y

Table 2: Model predictions

Admittedly, since the optimal retention ratio $[\underline{\theta}, \beta_0^d]$ is embedded in $[\underline{\theta}, \bar{\theta}]$, we cannot distinguish between an optimal retention requirement and an over-regulating retention requirement when monitoring is observed and risk shifting is not. Nonetheless, as we will demonstrate in the next section, banks do monitor more and shift more risk after the retention rule's implementation, implying that the current rate of 5% is overshooting.

5 Empirical Evidence

This section provides further evidence for the the model predictions. I use two difference-in-difference (DID) setups to show a significant increase in the RWA ratio for RMBS securitizers and a decrease in the loan delinquency rate after the implementation of the retention rules, implying banks monitor more but shift more risk at the same time. According to the model predictions, the current 5% ratio is overshooting. In Appendix E, using a 3SLS estimation, I demonstrate the presence of a stronger form of risk shifting in the banking system.

As in the motivating facts, I use the risk-weighted assets (RWA) ratio to proxy risk shifting. An increase in this ratio represents risk-increasing change in a bank's asset portfolio. While the downside risk will be absorbed by debtholders, the private returns go to shareholders if the investment strategy pays off. Hence risk is shifted toward debtholders.

While the Dodd-Frank Act was passed in July 2010, the final rule of mandatory retention was, in fact, not agreed upon until October 2014. There was then a two-year delay between agreement and implementation. In December 2015, the new rule took effect for RMBS. One year later, it became effective for all other categories of securitization. Hence, in the first difference-in-difference estimation, I compare US bank holding company (BHC) RMBS securitizers versus other BHC securitizers from 2015Q1 to 2016Q4 from Y-9C forms. However, delinquency takes place more than 90 days after a loan is generated (at least), so a short sample period may not deliver accurate estimates. I therefore consider an alternative difference-in-difference design with a longer horizon (2014–2017) and compare RMBS-only BHC securitizers versus non-securitizers who are mortgage sellers. I use propensity matching to make the two groups more comparable.

A BHC is defined as a securitizer if it reports at least one non-zero outstanding securitization activity in the sample period and has at least two years of existence. From 2001Q2, BHCs were required to detail their securitization activity in their regulatory reports (Schedule HC-S of the Y-9C report). Securitization activity is the outstanding principal balance of assets[12] sold and securitized with servicing retained or with recourse or other seller-provided credit enhancements in millions of US dollars. If another bank acquired a bank, I remove the non-survivor from the sample.

5.1 DID: RMBS securitizers versus other securitizers

In the first DID setup, I use US BHC data from 2015Q1 to 2016Q4 from Y-9C forms, which leads to 52 BHC securitizers and 359 bank-quarter observations. Notice that banks that issue RMBS may also be involved in other categories of securitization businesses. I consider the standard binary treatment as well as the continuous treatment. The focus of this section is the effect of the retention rules on banks' risk shifting behavior according to the reasons

[12] These assets are classified as family residential, home equity lines, credit card receivables, auto loans, other consumer loans, commercial and industrial loans, and all other assets.

mentioned above.

In the binary treatment, banks are split into control and treatment groups. The control group contains 12 BHC securitizers that do not issue RMBS. I list them in Appendix F. The treatment groups are banks that issue RMBS, but with different issuance shares, defined as

$$Share = \frac{\text{RMBS activities}}{\text{Total securitization activities}}$$

The RMBS activities, as well as total securitization activities, do not contain loans sold to other institutions or entities.[13] I consider three definitions of the treatment group: (1) $share > 0.5$ every quarter; (2) $share > 0.8$ every quarter; (3) $share = 1$ every quarter. There are 40 banks in total that are RMBS issuers: 37 of them are in group 1, 36 in group 2, and 33 in group 3. In fact, securitization activities and shares are stable across time for RMBS securitizers. There is one single bank[14] that did not issue RMBS until 2016Q3, and its share is below 0.15. I hence include this bank in the control group. After defining the variable D_{it} to equal 1 if bank i is an RMBS securitizer after December 2015 and 0 otherwise, the benchmark model can be expressed as the following fixed-effect panel regression:

$$RWA_{it} = \beta D_{it} + X_{it}'\delta + \alpha_i + \gamma_t + \epsilon_{it} \tag{2}$$

where subscripts it uniquely identify individual observations for bank i in quarter t. The dependent variable is the RWA ratio. I include bank fixed effects to absorb unobservable differences in bank business models and time fixed effects to absorb macro-economic shocks such as quantitative easing. The independent variable γ_t is time fixed effects, X_{it} is the set of bank-specific control variables, α_i is the individual fixed effects, and $\epsilon_i t$ is the error term. Standard errors are clustered at the bank level. The coefficient of interest is β. I estimate (2) with and without bank-specific controls.

To shed light on the dynamics of the average treatment effect before and after the passage of the mandatory retention rule, I replace the dummy D_{it} in specification (2) with leads and lags, which I name as $Before$ and $After$. The new regression model is

$$RWA_{it} = \alpha_i + \sum_{N=2}^{4} \beta_{5-N} B_{N,it} + \sum_{M=0}^{3} \alpha_{3-M} A_{M,it} + X_{it}'\delta + C_t + \epsilon_{it} \tag{3}$$

[13]Examples of other institutions and entities are the Federal National Mortgage Association (Fannie Mae) or the Federal Home Loan Mortgage Corporation (Freddie Mac). These government-sponsored agencies, in turn, securitize these loans and are not subject to the mandatory retention. These items are reported separately in Item 11 and 12 in Schedule HC-S.

[14]Santander Holdings USA, Inc, RSSD ID 3981856.

where B_N denotes N periods before the treatment and A_M denotes M periods after treatment. The quarter before treatment is used as the benchmark. The coefficients β to α estimate the average changes in banks' risk in the quarters preceding and following the Final Rule's implementation. The aim here is to test whether the effect is isolated to periods occurring only after the onset of the implementation.

The set of bank-specific control variables includes on-balance-sheet and off-balance-sheet items. On-balance-sheet items include some standard explanatory variables in banking literature, like size (measured by the log of total assets), profitability (measured by return on assets), capital buffer (measured by the difference between banks' regulatory risk-based capital and the minimum required capital ratio), and liquidity ratio (over total assets). The trading ratio that measures the volatility of a bank's activity is also considered. Banks of different sizes are regulated with different stringencies. For example, banks with assets of more than $50 billion are subject to stress tests, must submit resolution plans, and have tighter liquidity requirements. The capital buffer and liquidity buffer are a sign of bank solvency and stability. The off-balance-sheet item I use is securitization activities scaled by total assets.

In the US Basel framework Final Rule, there are two implementation approaches to calculate RWA. The standard approach, which applies to all banks, attaches fixed risk weights to multiple exposure types. In the advanced approach, which applies to banks with consolidated assets greater than $250 billion, exposures are broadly classified into four categories: retail, wholesale, securitization, and equity. In addition, the internal ratings-based (IRB) formula applies to retail and wholesale exposures. The standard and advanced approaches took effect in January 2015 and January 2014, respectively. Banks with total assets greater than $250 billion have to calculate RWA using both standardized and advanced approaches and may act differently from the rest of the banks. I first estimate the coefficient by removing those banks and add them back as a robustness check.

The summary statistics for the treatment and control group are reported in Table 3. I report treatment groups with RMBS activity share greater than 0.8 as well as equal to 1. The average total assets of banks in the control group is $60.8 billion, which is greater than banks in the treatment group, whose average total assets is under $40 billion. Moreover, as I increase the share of RMBS activities, the average size drops. The average size becomes $26.6 billion when banks only issue RMBS. Banks in both groups participate in securitization activities each quarter at levels about 13% to 14% of their size.

Regarding on-balance-sheet items, banks in the treatment group have a slightly higher

Table 3: **Summary statistics**

Variables	Control group		RMBS share≥0.8		RMBS share=1	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Total assets ($ billions)	66.897	63.088	39.754	58.132	26.591	37.946
RWA ratio	0.780	0.119	0.749	0.092	0.748	0.089
RWA ratio (exclude securitization exposure)	0.776	0.119	0.748	0.091	0.748	0.089
Loan ratio (of total assets)	0.646	0.165	0.691	0.100	0.698	0.089
Home mortgage (of total loans)	0.226	0.151	0.322	0.171	0.325	0.180
Consumer loans (of total loans)	0.168	0.162	0.081	0.099	0.080	0.104
C&I loans (of total loans)	0.232	0.124	0.180	0.102	0.177	0.105
Capital buffer	0.101	0.134	0.138	0.915	0.149	0.982
Deposit ratio	0.419	0.186	0.587	0.112	0.597	0.101
Liquidity ratio	0.207	0.104	0.216	0.075	0.213	0.077
ROA	0.009	0.013	0.006	0.005	0.006	0.005
Securitization-assets ratio	0.142	0.340	0.129	0.161	0.122	0.161
# of banks	12		40		33	

Notes: The table presents descriptive statistics of the control group and treatment groups from 2015Q1 to 2016Q4. I report treatment groups with RMBS activity share greater than 0.8 as well as equal to 1. The table contains means and standard deviations of bank characteristics. The RWA ratio is the risk-weighted assets to total assets ratio. ROA is the ratio of net income to total assets. The securitization-assets ratio is the ratio of securitization activities to total assets. Credit enhancements is the ratio of total credit enhancements provided by the securitizer to total assets.

loan ratio (69% versus 65%) and a higher home mortgage ratio over total loans (32% versus 23%), which is no surprise. Banks in the control group focus more on consumer loans and C&I loans. Banks in both groups hold 21% of their total assets in the form of liquid assets.

The average RWA ratio in the full sample is 0.76. Banks in the control group have a higher RWA ratio. The difference between the two groups is around 0.03 (0.78 versus 0.75). I also report the RWA ratio excluding securitization exposure. The numbers do not change much. Though operating at a higher asset level, banks in the control group have more profitable projects than banks in the treatment group (0.009 versus 0.006). As a sign of bank solvency and stability, the average capital buffer in the sample is 0.11. The treatment group banks are more capitalized than control group banks (0.14 versus 0.10).

Table 4 reports the estimation of equation (2) regarding the implementation of the mandatory risk retention rule on RMBS on banks' risk shifting. The first two columns report the results when no other control variables are included. The treatment groups are defined as RMBS activity share ≥ 0.5 and 0.8. The RWA ratio of the treatment banks increases 2.3 percentage points on average and is significantly positive at the 10% level. Column 4 of the table reports the result when I control for bank-specific variables. I include size, ROA, capital buffer, liquidity ratio, and the two off-balance-sheet variables. I increase the activity share threshold to 1 in column 4. While securitization activities and credit enhancements

Table 4: **The effect of mandatory retention rule on banks' risk shifting**

	Dependent variable: RWA ratio				
	(1)	(2)	(3)	(4)	(5)
	RMBS≥0.5	RMBS≥0.8	RMBS≥0.8	RMBS=1	RMBS=1
					>= $250 billion included
Treatment	0.023*	0.023*	0.021*	0.024*	0.020*
	(0.013)	(0.013)	(0.011)	(0.012)	(0.011)
Capital buffer			0.001*	0.001*	-0.000
			(0.000)	(0.000)	(0.000)
Size			-0.097*	-0.085	0.020
			(0.049)	(0.053)	(0.040)
ROA			0.720**	0.625*	0.728**
			(0.332)	(0.342)	(0.294)
Liquidity ratio			-0.580***	-0.607***	-0.688***
			(0.136)	(0.119)	(0.118)
Securitization-assets ratio			0.077***	0.069**	0.071***
			(0.026)	(0.028)	(0.024)
Observations	333	325	325	293	325
R-squared	0.083	0.084	0.221	0.208	0.478
Number of bank	49	48	48	45	50

Notes: This table reports the effects of mandatory retention rules on banks' risk-weighted assets (RWA) ratio, specified by

$$RWA_{it} = \beta D_{it} + X'_{it}\delta + \alpha_i + \gamma_t + \epsilon_{it}$$

Treatment equals 1 if bank i is an RMBS securitizer after December 2015 and 0 otherwise. In the first 4 columns, I increase the threshold of RMBS activity share from 0.5 to 1. In the first two columns, only bank and time fixed effects are considered. In columns 3 and 4, bank-specific control variables are added. Banks with total assets more than $250 billion are removed in the first 4 columns and added in column 5. Standard errors are in parentheses. The superscripts ***, **, and * represent 1%, 5%, and 10% significance levels, respectively.

Table 5: **The dynamics of treatment effects**

	Dependent variable: RWA ratio		
	(1)	(2)	(3)
	RMBS $\geq$ 0.5	RMBS $\geq$ 0.8	RMBS = 1
Before4	0.003	0.003	0.001
	(0.016)	(0.016)	(0.016)
Before3	0.001	-0.001	-0.004
	(0.008)	(0.008)	(0.008)
Before2	0.010	0.007	0.004
	(0.010)	(0.010)	(0.010)
After0	0.021*	0.021*	0.021*
	(0.012)	(0.012)	(0.012)
After1	0.028**	0.027**	0.025**
	(0.012)	(0.012)	(0.012)
After2	0.022	0.022	0.022
	(0.014)	(0.014)	(0.015)
After3	0.034**	0.034**	0.033**
	(0.014)	(0.014)	(0.014)
Observations	333	325	293
R-squared	0.091	0.092	0.100
Number of bank	49	48	45

Notes: This table reports the dynamics of average treatment effects estimated by

$$RWA_{it} = \alpha_i + \sum_{N=2}^{4} \beta_{5-N} B_{N,it} + \sum_{M=0}^{3} \alpha_{3-M} A_{M,it} + X'_{it}\delta + C_t + \epsilon_{it}$$

where B_N denotes N periods before the treatment and A_M denotes M periods after treatment. The quarter before treatment is used as the benchmark. Banks with total assets more than \$250 billion are removed. Standard errors are in parentheses. The superscripts ***, **, and * represent 1%, 5%, and 10% significance levels, respectively.

that banks provide significantly contribute to the overall risk of banks, there is almost no discrepancy between the treatment effect estimates. As a robustness check, I add back banks with total assets greater than \$250 billion and obtain a similar result, which is reported in column 5.

I report in Table 9 in Appendix D that coefficients are similar when the dependent variable is replaced with the RWA ratio excluding securitization exposure. Hence, bank portfolio risk increased outside retained tranches in securitization.

In Table 5, I report the estimation of equation (3) regarding the dynamics of banks' risk shifting with bank-specific controls. The coefficients on $B_{N,it}$ are all insignificant across all treatment groups, some of them being negative, indicating that prior to the Final Rule's implementation there is no difference in risk shifting measured by the RWA ratio, hence there are no pre-existing trends. I find positive and significant coefficients on A_0: the results

Figure 9: Time passage relative to the mandatory retention rule's implementation

in the previous table occur after the Final Rule's implementation.

Figure 9 plots the treatment effect coefficients and graphically illustrates the observed time pattern of the RWA ratio around the implementation of the mandatory retention rule. Each point on the graph reflects the average difference in the RWA ratio for treatment BHCs and control BHCs, netting out bank and time fixed effects, where the treatment group consists of all RMBS-only BHC securitizers. I use the quarter before the treatment (before1) as a base, represented by the dashed line. Bands represent 1.96 standard error of each point estimate.

As another robustness check, I consider the impacts of the RMBS retention rule in a continuous way: banks have different issuance shares of RMBS securitization activities, hence they have different treatment intensity. In this way, instead of defining thresholds in the treatment group and dropping banks not meeting the threshold, I define the continuous treatment as the product of RMBS share and a time dummy that equals 1 if time is after December 2015. The econometric specification is

$$RWA_{it} = \beta Share_{it} * D_t + X'_{it}\delta + \alpha_i + \gamma_t + \epsilon_{it} \tag{4}$$

The coefficient of interest is β. Banks' specific controls are the same as described in the previous part. The dynamics of the continuous treatment effects is also considered and reported in Table 10 in Appendix D.

The regression results of equation (4) are reported in Table 6. Across specifications, the coefficient is significant and on the order of 0.02, meaning that banks with a 10 percentage

Table 6: **The effect of the mandatory retention rule on banks' risk shifting**

	Dependent variable: RWA ratio		
	(1)	(2)	(3)
Treatment	0.022*	0.020**	0.018*
	(0.011)	(0.010)	(0.010)
On-balance-sheet items	No	Yes	Yes
Off-balance-sheet items	No	No	Yes
Observations	359	359	359
R-squared	0.085	0.449	0.473
Number of bank	52	52	52

Notes: This table reports the effects of mandatory retention rules on banks' risk-weighted assets (RWA) ratio using continuous treatment specification

$$RWA_{it} = \beta Share_{it} * D_t + X_{it}'\delta + \alpha_i + \gamma_t + \epsilon_{it}$$

Treatment is the product of RMBS share and a dummy that equals 1 if time is after December 2015 and 0 otherwise. In the first column, only bank and time fixed effects are considered. In columns 2 and 3, bank-specific control variables are added. Banks with total assets more than $250 billion are removed. Standard errors are in parentheses. The superscripts ***, **, and * represent 1%, 5%, and 10% significance levels, respectively.

point higher share in RMBS activities will increase their RWA ratio by 0.2 percentage points after the implementation. The coefficient does not change too much across the binary and continuous treatment, since the distribution of RMBS activity shares has two mass points at 0 and 1, which is shown in Figure 17 in Appendix D.

Lastly, I also use the above DID design to study the impact of the retention rules on bank delinquency rates. The delinquency rate is defined in the same way as in the motivating facts section. While the details are provided in Tables 11 and 12 in Appendix D, I show that the delinquency rate declined about 0.3 percentage points across all specifications. However, delinquency takes place more than 90 days after a loan is generated (at least). Hence, a short sample period may not deliver accurate estimates, which motivates me to consider an alternative DID setup, as illustrated in the next section.

5.2 DID: RMBS-only securitizers versus mortgage sellers

This section employs an alternative DID method to study the impact of the retention rules on banks' delinquency rates as well as RWA ratios. The new sample range is from 2014Q1 to 2017Q4. The treatment group consists of RMBS-only BHC securitizers (corresponding to RMBS securitizers with activity share equal to 1 in the previous DID setup), and the control

Table 7: **Summary statistics**

Variables	Control group		Treatment group	
	Mean	Std Dev	Mean	Std Dev
Total assets ($ billions)	6.579	16.649	24.738	36.819
DEL rate	0.021	0.027	0.010	0.015
Loan ratio (of total assets)	0.701	0.089	0.690	0.127
Home mortgage (of total loans)	0.284	0.165	0.280	0.165
Consumer loans (of total loans)	0.080	0.100	0.048	0.072
C&I loans (of total loans)	0.192	0.109	0.156	0.099
Capital buffer	0.111	0.667	0.073	0.040
Deposits ratio	0.619	0.111	0.658	0.114
Core deposits ratio	0.565	0.114	0.523	0.078
Liquidity ratio	0.218	0.080	0.250	0.122
ROA	0.006	0.005	0.006	0.005
# of banks	194		35	

Notes: The table describes the control group and treatment group from 2014Q1 to 2017Q4. The treatment group consists of RMBS-only BHC securitizers (corresponding to RMBS securitizers with activity share equal to 1 in the previous DID setup), and the control group contains non-securitizers who are mortgage sellers.

group contains non-securitizers who are mortgage sellers. I focus on RMBS-only securitizers because otherwise some banks are affected by the mandatory retention rule twice. Similar to securitizers, I define a BHC as a mortgage seller if it has at least one quarter of non-zero mortgage sales. This item is reported separately in the FRY-9C form. The process generates 35 banks in the treatment group and 194 banks in the control group. Note that mega banks are also removed in this setup originally and are added back for a robustness check in the delinquency rate part.

Summary statistics of the new control and treatment groups are presented in Table 7. Banks in the treatment group are bigger than banks in the control group ($25 billion versus $7 billion). Banks in the two groups have similar asset compositions: 70% of the total assets are in the form of loans, and 20% are in the form of liquid assets. The loan types are also similar across the two groups. Moreover, the two groups have the same level of profitability (0.006). On the liability side, the treatment group is more capitalized (11% versus 7% in terms of capital buffer) and borrows less from the deposit market (62% versus 66%).

In the benchmark, I consider empirical specifications similar to the study of risk shifitng

$$Y_{it} = \alpha_i + \beta D_{it} + Y_{it}'\delta + C_t + \epsilon_{it} \tag{5}$$

where Y_{it} is the delinquency rate or RWA ratio. Variable D_{it} equals 1 if bank i is an RMBS-only securitizer after December 2015 and 0 otherwise. Since the sample is large, I also use the propensity score matching to reduce baseline bias. To estimate propensity score, I choose a set of observable variables that will be used in the logit group predicting function. Each

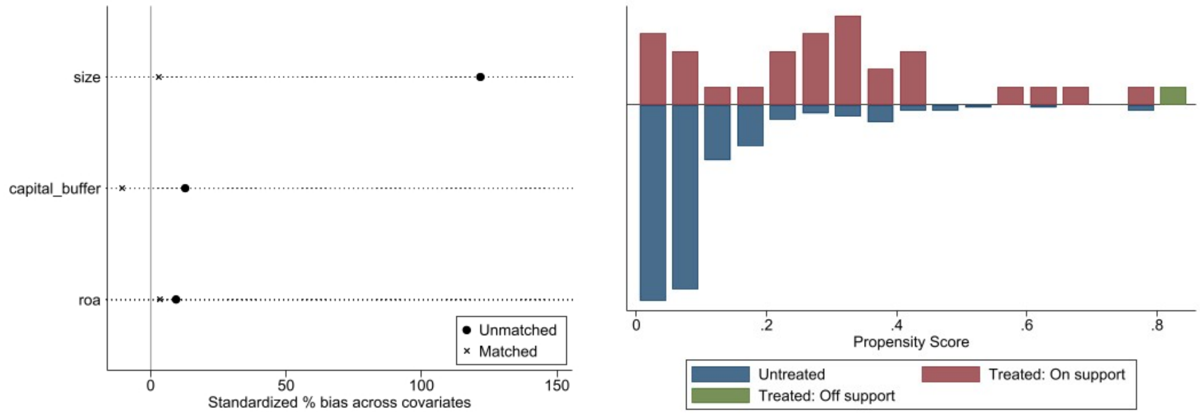

Figure 10: Matching efficiency

bank will be distributed a score, which is the probability in the treatment group. Banks in the treatment group will receive a weight of 1, and banks in the control group receive a weight that is proportional to the probability of them being in the treatment group relative to the probability of them being in the control group, which they were actually in. I then estimate

$$\Delta Y_i = \beta T_i + \epsilon_i \tag{6}$$

using the matched sample. Here Y_i is the delinquency rate, and ΔY_i is defined as

$$\Delta Y_i = \bar{Y}_{i,after} - \bar{Y}_{i,before}$$

and T_i equals 1 if bank i is in the control group.

The set of covariates in the matching includes the common ones already mentioned: size, capital buffer, and ROA (profitability). I also take into account the core deposits ratio, which measures the fraction of stable sources of funding, and the home mortgage sale share, which is defined as the fraction of mortgage sales to total loan sales. Large banks tend to engage in more securitization activities. Securitization choice is also closely related to bank profitability, how tight of a capital constraint the bank faces, and how much stable funding the bank can obtain. The distributions of the covariates in the two groups are reported in Appendix D. Figure 10 illustrates the matching efficiency: selection bias (in terms of measured and tested covariates) in size and ROA is reduced by matching. It also portrays the Kernel distribution of propensity scores when banks are matched according to the common covariates above.

Table 8: **The effects of the mandatory retention rule on bank risk**

	Dependent variable: RWA ratio			
	(1)	(2)	(3)	(4)
	OLS	OLS	PSM	PSM
Treatment	-0.011	0.001	0.004	0.021*
	(0.015)	(0.006)	(0.014)	(0.011)
Bank controls	No	Yes		
Common covariates			Yes	No
All covariates			No	Yes

	Dependent variable: Delinquency rate			
	(1)	(2)	(3)	(4)
	OLS	OLS	PSM	PSM
Treatment	-0.003	-0.003*	-0.002**	-0.002
	(0.002)	(0.002)	(0.001)	(0.002)
Bank controls	No	Yes		
Common covariates			Yes	Yes
All covariates			No	Yes
Number of bank	229	228	220	140

Notes: This table reports the effects of mandatory retention rules on banks' RWA ratios and delinquency rates under propensity score matching. Columns 1 and 2 are the benchmark cases without matching. Columns 3 and 4 report the PSMDID estimates using different sets of covariates. *Treatment* equals 1 if bank i is an RMBS-only securitizer after December 2015 and 0 otherwise. Standard errors are in parentheses. The superscripts ***, **, and * represent 1%, 5%, and 10% significance levels, respectively.

Table 8 reports the results. First, regarding the RWA ratio, in OLS the treatment coefficient is not significant and is inconsistent depending on whether bank-specific controls are added or not. Under matching, the coefficient becomes positive when covariates contain size, capital buffer and ROA. The major factor affecting this result is size. The large difference between the sizes of the two groups makes them incomparable. And the coefficient becomes significant at the 10% level when the core deposits ratio or mortgage sale share is included. The increase in the RWA ratio is 2.1 percentage points with the core deposits ratio included, similar to the result in the previous setup. One detail is that there are missing values in the core deposits ratios, and they are dropped rather that imputed in the matching. Additional PSMDID results using a subset of the above covariates are presented in Table 13 in Appendix D.

Second, there is a decline in the delinquency rate with and without matching techniques. In OLS, the drop is significant when bank controls are added. Under matching, the drop is 0.2 percentage points and significant at the 5% level using common covariates. I also portray the dynamics of the treatment effects in Figure 20 in Appendix D. The results take effect after the effective date of the Dodd-Frank retention rules in RMBS.

In sum, the above two sections show that loans are riskier ex ante and safer ex post. Consequently, banks shifted toward riskier investing strategies and must have engaged in more efforts to gather information and monitor their borrowers. According to the model prediction, the current 5% ratio is too high relative to the optimal.

6 Extension: Optimal Retention Form

The model assumes that the bank retains the vertical interest. This section studies three questions: What are the bank's behaviors under the horizontal retention? Which retention form is optimal from the bank's perspective? And which retention form is optimal from the welfare perspective? I show that the trade-off between monitoring and gambling also exists when the bank retains horizontal interest. When the capital requirement is not binding for a horizontal retention, the bank will choose such form. From the efficiency perspective, there is also a trade-off between the two retention options.

6.1 Horizontal retention

Recall that r is the realization in the bad state. In this section, assume $r \geq 0$. As a starting point, the market value of the pool given m can be rewritten as

$$\alpha R - \big(1 - p(m)\big)\frac{R-r}{R}\alpha R,$$

which is the promised payment net of expected loss. The realization of r means a loss, and $\frac{R-r}{R}$ measures the loss severity as the percentage lost in the event of default. In a horizontal retention, the analysis of monitoring choice at time 2 remains the same. The bank retains β fraction of the most junior tranches in the pool. The default risk of the underlying project is then disproportionally distributed between investors and the bank. By seniority, investors who start to bear loss only if the loss severity $\frac{R-r}{R}$ is above β, or

$$\beta \leq \frac{R-r}{R}. \tag{7}$$

We make the following assumption that r is not large.

Assumption 5 $r < R - d - \frac{c}{\Delta}$.

Under this assumption, the loss severity is always larger than β^d. The price schedule investors provide is the following

$$P_h(\beta|p,R) = \begin{cases} (1-\beta)\alpha R & \text{if } \beta \geq \frac{R-r}{R} \\ (1-\beta)\alpha R - (1-p)\alpha\big(\frac{R-r}{R}-\beta\big)R & \text{if } \beta^d \leq \beta \leq \frac{R-r}{R} \\ (1-\beta)\alpha R - (1-p+\Delta)\alpha\big(\frac{R-r}{R}-\beta\big)R & \text{if } \beta < \beta^d \end{cases}$$

The bank's monitoring decision affects the probability of investors receiving losses, which is $1-p(m)$. But they assume losses only amount to $\frac{R-r}{R}-\beta$. Let $\bar{R} \equiv pR + (1-p)r$ and $\underline{R} \equiv (p-\Delta)R + (1-p+\Delta)r$ denote the expected returns of the project under monitoring and no monitoring. The value of the retained tranches, denoted by $V_h(\beta)$, is the total value of the pool minus the value of the tranches sold, which equals the price schedule above.

$$V_h(\beta|p,R) = \begin{cases} \alpha(p+\beta-1)R + \alpha(1-p)r & \text{if } \beta \geq \frac{R-r}{R} \\ p\beta\alpha R & \text{if } \beta^d \leq \beta \leq \frac{R-r}{R} \\ (p-\Delta)\beta\alpha R & \text{if } \beta < \beta^d \end{cases}$$

When $\beta \geq \beta^d$, the pool value is $V^{\text{pool}} = \bar{R}$; and when β is below β^d, the pool value drops to $V^{\text{pool}} = \underline{R}$. Hence, the bank's retention constraint, $\frac{V_h(\beta|p,R)}{V^{\text{pool}}} \geq \theta$ can be written as

$$\beta \geq \begin{cases} \frac{\theta\bar{R}+(1-p)(R-r)}{R} & \text{if } \beta \geq \frac{R-r}{R} \\ \frac{\theta\bar{R}}{pR} & \text{if } \beta^d \leq \beta \leq \frac{R-r}{R} \\ \frac{\theta\underline{R}}{(p-\Delta)R} & \text{if } \beta < \beta^d \end{cases}$$

For any given θ, the retention constraint imposes a lower bound for β that is strictly larger than θ, due to the disproportional risk sharing under the horizontal retention. When risk is mostly absorbed by the junior tranches, the value of the tranches is discounted, or smaller than θ fraction of the pool value. The bank thus has to retain more to meet the value requirement.

One result follows from the above expressions.

Proposition 6.1 *If $r = 0$, the vertical and horizontal retentions are identical.*

Proof: See Appendix A.

When $r > 0$, the bank's objective function at time 1, $\lambda P_h(\beta|p,R) + \Pi_2(\beta|p,R)$, has the general shape in Figure 11. The objective function decreases on $[\beta^d, 1]$ but decreases faster when $\beta \geq \frac{R-r}{R}$ because the junior tranches are depleted on this region.

Let $\beta^{e'}$ be such that $\lambda P_h(\beta^{e'}|p,R) + \Pi_2(\beta^{e'}|p,R) = \lambda P_h(\beta^d|p,R) + \Pi_2(\beta^d|p,R)$, similar to β^e in the vertical retention. The solution to the bank's retention problem is the following:

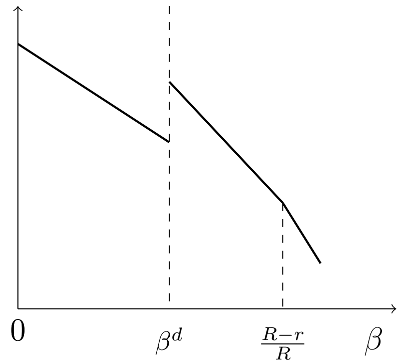

Figure 11: The bank's objective function

Proposition 6.2 (Retention choice.) *Under horizontal retention, the bank's optimal retention choice of β is*

$$\beta_h(\theta) = \begin{cases} \frac{\theta \underline{R}}{(p-\Delta)R} & \text{if } \theta < \frac{\beta^{e'}(p-\Delta)R}{\underline{R}} \\ \beta^d & \text{if } \frac{\beta^{e'}(p-\Delta)R}{\underline{R}} \leq \theta < \frac{\beta^d pR}{\bar{R}} \\ \frac{\theta \bar{R}}{pR} & \text{if } \theta^k \leq \beta \leq \frac{R-r}{p\bar{R}} \\ \frac{\theta \bar{R}+(1-p)(R-r)}{R} & \text{if } \beta \geq \frac{R-r}{p\bar{R}} \end{cases}$$

Moreover, $\beta_h(\theta) \geq \theta$. The equality holds only at 0 and 1.

Proof: See Appendix A.

Again, due to seniority, the junior tranches bear most of the risk, which discounts their value. The bank has to retain a higher fraction of the pool to meet the requirement. The bank's expected value, denoted by

$$\Pi_1(\theta|p, R) \equiv \lambda P_h(\beta_h(\theta)|p, R) + \Pi_2(\beta_h(\theta)|p, R),$$

is portrayed in Figure 12. Besides the positions of the kink points, the only difference between this value under horizontal retention and vertical retention occurs for very large θ. Similar analysis can be conducted for the project choice stage, although the more retention leading to more risk shifting argument is omitted here.

6.2 The banks' optimal choice of retention form

Based on the results above, I argue here that given a project (p, R), if the capital requirement is slack, then the expected value under horizontal retention is higher than under vertical retention. In practice, with horizontal retention, higher risk weights will be attached to the

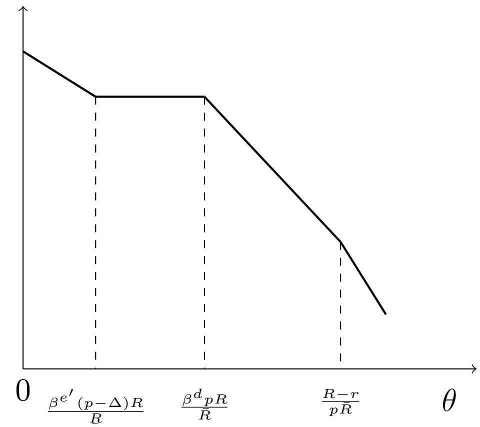

Figure 12: The bank's expected value under an optimal horizontal retention choice

junior tranches retained by the bank, which tightens the capital requirement constraint. We define the capital requirement for a project as being "slack" at a fixed retention ratio θ if the capital requirement is met under horizontal retention without additional equity issuance.

Proposition 6.3 *If the capital requirement is slack at retention requirement θ, the bank chooses horizontal retention.*

Proof: See appendix A.

Horizontal retention requires the bank to retain more on the balance sheet due to more involved risk, leading to more gains in the good states and more losses in default. Since the bank already loses everything in bankruptcy by limited liability, the net expected gains are higher. This result is better illustrated in Figure 13. If the bank is unable to meet the capital requirement by holding a horizontal piece, and the cost of issuing additional equity is too high, the bank has to choose the vertical form.

In practice, banks are heterogeneous and retain differently. For example, Flynn, Ghent, and Tchistyi (2020) examine CMBS deals from 2017 to 2019 and show that 45% of the deals involve horizontal retention and 40% involve vertical retention (though the issuers may not be bank holding companies). Other reasons to hold a vertical piece include consolidation concerns[15] and opportunity cost, because senior tranches can be collateral to borrow Repo

[15]There can be the potential impact of accounting standard FAS 166 and FAS 167. Effective as of 2010, the two accounting standards tightened the accounting for securitizations and the consolidation of variable interest entities (VIEs). If a securitizer is identified as the primary beneficiary for the special-purpose entity used to issue ABS, it must consolidate the securitized assets. The accounting guidance does not provide a quantitative bright-line threshold for determining what "could potentially be significant" for purposes of

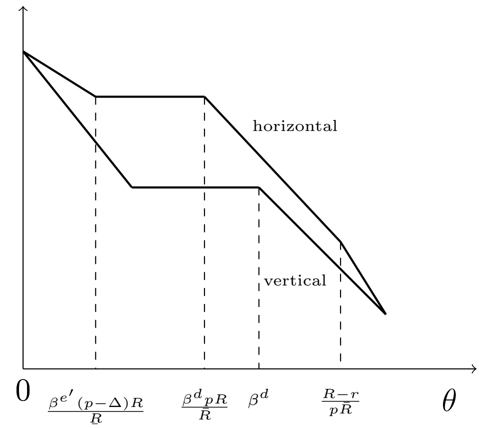

Figure 13: The bank's expected value under horizontal and vertical retention choices

and other short-term funding. The discussion above hence predicts that banks that retain vertically should perform better and accumulate less risk.

6.3 Optimal retention form

The remaining question is, what is the optimal retention form from the welfare perspective? In fact, policymakers expect that incentives are better aligned under horizontal risk retention. For example, the Department of the Treasury and Federal Reserve System and Federal Deposit Insurance Corporation and Federal Housing Finance Agency and Securities and Exchange Commission and Department of Housing and Urban Development (2014) (p. 486) states

> "*a sponsor choosing to retain risk in a more expensive horizontal form over a vertical form would have greater exposure to credit risk, and that sponsor's incentives should be better aligned with investors.*"

This is true with regard to monitoring, according to the observations in the previous section. Recall that the welfare-maximizing retention ratio is $[\beta_0^e, \beta_0^d]$ under vertical retention and $[\frac{\beta_0^{e'}(p-\Delta)R}{\underline{R}}, \frac{\beta_0^d pR}{\bar{R}}]$ under horizontal retention. Since $\frac{\beta_0^{e'}(p-\Delta)R}{\underline{R}} < \beta_0^e$ and $\frac{\beta_0^d pR}{\bar{R}} < \beta_0^d$, as shown in Figure 13, it is easier to induce monitoring under horizontal retention but also easier to over-regulate securitization. Moreover, concerning risk shifting, the threshold above which banks

determining whether an entity is required to consolidate. The 5% retention may meet the threshold of an interest that could potentially be significant. Furthermore, the horizontal strip of risk is more likely to be significant from the results in the last section. If the consolidation cost is high enough, retaining a horizontal component is no longer optimal if holding a vertical strip can avoid such consolidation.

start to gamble is also lower under horizontal retention, implying that banks are more likely to shift risk under horizontal retention. The reason is that the horizontal piece requires the bank to retain a higher fraction of the pool, which resembles increasing the retention ratio to some extent, under which gambling is more attractive. Therefore, there is also a trade-off between the two retention forms. This is in the spirit of Fender and Mitchell (2009). But due to the costly equity issuance and accounting issues about consolidation, policymakers should give discretion to banks.

7 Concluding Remarks

Policymakers expect the new risk retention rules to coordinate incentives and reduce banks' risk. However, I show that the new rules involve trade-offs among various important dimensions of bank risk taking. Under mandatory retention, banks become more prudent in screening and monitoring; but they may also choose a riskier investment strategy. The product and by-product, as I demonstrate, are both observed in the US BHC data. Banks' portfolios had become riskier ex ante and had fewer delinquencies ex post after the implementation of the new rules. More critically, according to my theory, the simultaneous increase in monitoring and risk shifting occurs only when the retention ratio exceeds the optimal level. This indicates that the current 5% ratio is too high. I show that the optimal level is an interior number, but the structural model needed to precisely quantify it is left for future research.

References

Agarwal, Sumit, Brent W Ambrose, Yildiray Yildirim, and Jian Zhang. 2019. "Risk retention and qualified commercial mortgages." *Working Paper, National University of Singapore* .

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff. 2011. "The role of securitization in mortgage renegotiation." *Journal of Financial Economics* 102:559–578.

Chemla, Gilles and Christopher A. Hennessy. 2014. "Skin in the game and moral hazard." *The Journal of Finance* 69 (4):1597–1641.

DeMarzo, Peter and Darrell Duffie. 1999. "A Liquidity-Based Model of Security Design." *Econometrica* 67 (1):65–100.

Demsetz, Rebecca S., Marc R. Saidenberg, and Philip E. Strahan. 1996. “Banks with something to lose: The disciplinary role of franchise value.” *FRBNY Economic Policy Review* 2 (2).

Department of the Treasury and Federal Reserve System and Federal Deposit Insurance Corporation and Federal Housing Finance Agency and Securities and Exchange Commission and Department of Housing and Urban Development. 2014. “Credit Risk Retention.” .

Duran, Miguel and Ana Lozano-Vivas. 2014. “Risk shifting in the US banking system: An empirical analysis.” *Journal of Banking and Finance* 13 (C):64–74.

Fender, Ingo and Janet Mitchell. 2009. “Incentives and tranche retention in securitisation: a screening model.” *Working paper, Centre for Economic Policy Research* .

Flynn, Jr., Sean J, Andra C Ghent, and Alexei Tchistyi. 2020. “Informational efficiency in securitization after Dodd-Frank.” *The Review of Financial Studies* 33 (11):5131–5172.

Furfine, Craig. 2020. “The Impact of Risk Retention Regulation on the Underwriting of Securitized Mortgages.” *Journal of Financial Services Research* 58 (2):91–114.

Gorton, Gary and George Pennacchi. 1995. “Banks and loan sales Marketing nonmarketable assets.” *Journal of Monetary Economics* 35 (3):389–4110.

Gorton, Gary and Andrew Winton. 2003. “Financial intermediation.” In *Handbook of the Economics of Finance*, vol. 1, Part 1, edited by G.M. Constantinides, M. Harris, and R. M. Stulz, chap. 08. Elsevier, 1 ed., 431–552.

Guo, Guixia and Ho-Mou Wu. 2014. “A study on risk retention regulation in asset securitization process.” *Journal of Banking and Finance* 45 (C):61–71.

Jensen, Michael and William H Mecking. 1976. “Theory of the firm: Managerial behavior, agency costs and ownership structure.” *Journal of Financial Economics* 3 (4):305–360.

Keeley, Michael C. 1990. “Deposit Insurance, Risk, and Market Power in Banking.” *American Economic Review* 80 (5):1183–1200.

Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig. 2010. “Did securitization lead to lax screening? Evidence from subprime loans.” *The Quarterly Journal of Economics* 125 (1):307–362.

Kiff, John and Michael Kisser. 2014. "A shot at regulating securitization." *Journal of Financial Stability* 10 (C):32–49.

Leland, Hayne and David H Pyle. 1977. "Informational Asymmetries, Financial Structure, and Financial Intermediation." *Journal of Finance* 32 (2):371–87.

Malekan, Sara and Georges Dionne. 2014. "Securitization and optimal retention under moral hazard." *Journal of Mathematical Economics* 55 (C):74–85.

Mian, Atif and Amir Sufi. 2009. "The consequences of mortgage credit expansion: evidence from the U.S. mortgage default crisis." *The Quarterly Journal of Economics* 124 (4):1449–1496.

Mitchell, Janet. 2004. "Financial intermediation theory and the sources of value in structured finance markets." *Working Paper No. 71, National Bank of Belgium* .

Pagès, Henri. 2013. "Bank monitoring incentives and optimal ABS." *Journal of Financial Intermediation* 22 (1):30–54.

Parlour, Christine A. and Guillaume Plantin. 2008. "Loan sales and relationship banking." *The Journal of Finance* 63 (3):1291–1314.

Pennacchi, George. 1988. "Loan sales and the cost of bank capital." *The Journal of Finance* 43 (2):375–396.

Petersen, Mitchell A. and Raghuram G. Rajan. 2002. "Does distance still matter? The information revolution in small business lending." *The Journal of Finance* 57 (6):2533–2570.

Piskorski, Tomasz, Amit Seru, and Vikrant Vig. 2010. "Securitization and distressed loan renegotiation: Evidence from the subprime mortgage crisis." *Journal of Financial Economics* 97:369–397.

Purnanandam, Amiyatosh. 2011. "Originate-to-distribute Model and the Subprime Mortgage Crisis." *The Review of Financial Studies* 24 (6):1881–1915. URL https://doi.org/10.1093/rfs/hhq106.

Repullo, Rafael. 2004. "Capital requirements, market power, and risk-taking in banking." *Journal of Financial Intermediation* 13 (2):156–182.

Sarkisyan, Anna and Barbara Casu. 2013. "Retained interests in securitizations and implications for bank solvency." *Working Paper Series 1538, European Central Bank* .

Shrieves, Ronald and Drew Dahl. 1992. "The relationship between risk and capital in commercial banks." *Journal of Financial Stability* 16 (2):439–457.

Appendix A Proofs

A.1 Proof of Proposition 4.2

Given (β, p, R), the bank monitors if

$$p\big((1-\alpha)R + \alpha\beta R - d\big) - c \geq (p-\Delta)\big((1-\alpha)R + \alpha\beta R - d\big).$$

The result follows. □

A.2 Proof of Proposition 4.3

For $\beta < \beta^d$, the bank does not monitor. The bank's objective function is

$$\lambda(1-\beta)(p-\Delta)R + (p-\Delta)\Bigg((1-\alpha)R + \beta\alpha R - d\Bigg)$$

and is linear and decreasing since $\lambda > 1$. For $\beta \geq \beta^d$, the bank monitors, and its objective function is

$$\lambda(1-\beta)\alpha pR + p\Bigg((1-\alpha)R + \beta\alpha R - d\Bigg) - c$$

and is linear and decreasing by the same token. To show the upper semicontinuity, note that $\Pi_2(\beta|(p,R))$ is continuous at β^d, and

$$(1-\beta^d)\alpha pR > (1-\beta^d)\alpha(p-\Delta)R$$

hence the value of the objective function at β^d is strictly larger than its left limit. □

A.3 Proof of Lemma 4.1

The result follows from (1), Assumption 1, and the facts that $R_b > R_0$ and $(p_b - \Delta)R_b < (1-\Delta)R_0$. □

A.4 Proof of Proposition 4.6

The proof is divided into three steps.

Step 1. In this step, we show that $G(\theta)$ is increasing in θ when the retention constraints are binding for both projects; that is, when θ is below β_b^e and when θ is above β_0^d. In the first region, there is no monitoring,

$$\begin{aligned} G(\theta) =& \lambda(1-\theta)\alpha((p_b-\Delta)R_b - (1-\Delta)R_0) \\ &+ (p_b-\Delta)\big((1-\alpha)R_b + \theta\alpha R_b - d\big) - (1-\Delta)\big((1-\alpha)R_0 + \theta\alpha R_0 - d\big) \end{aligned}$$

It follows that $G'(\theta) = (\lambda - 1)\alpha((1-\Delta)R_0 - (p_b - \Delta)R_b) > 0$. In the second region, the bank monitors for both projects, and

$$G(\theta) = \lambda(1-\theta)\alpha(p_b R_b - R_0) + p_b\big((1-\alpha)R_b + \theta\alpha R_b - d\big) - \big((1-\alpha)R_0 + \theta\alpha R_0 - d\big).$$

By the same token, $G'(\theta) = (\lambda - 1)\alpha(R_0 - p_b R_b) > 0$.

Step 2. In this step, we pin down $\bar{\theta}$ and show that $G(\theta)$ is positive if and only if θ is above $\bar{\theta}$. In fact, $\bar{\theta}$ is such that

$$G(\bar{\theta}) = 0$$

when θ is above β_0^d. Specifically,

$$\bar{\theta} = \frac{(\lambda\alpha + 1 - \alpha)(R_0 - p_b R_b) - (1-p_b)d}{(\lambda - 1)\alpha(R_0 - p_b R_b)}. \tag{8}$$

According to step 1, when θ is above $\bar{\theta}$, $G(\theta)$ is positive. We show that $G(\theta) < 0$ when θ is below $\bar{\theta}$. There are two scenarios to consider. First, consider $\beta_0^e < \beta_b^d$. In this case, $G(\theta)$ is increasing on $[\beta_k^e, \beta_0^e]$ because $V_b(\theta)$ is constant, constant on $[\beta_0^e, \beta_b^d]$ because both V_0 and V_b are constant, and decreasing on $[\beta_b^d, \beta_0^d]$ because V_0 is constant. The maximum of $G(\theta)$ on this region is achieved on $[\beta_0^e, \beta_b^d]$. According to Assumption 3, this value is negative. Second, consider $\beta_0^e > \beta_b^d$. In this case, $G(\theta)$ is increasing on $[\beta_k^e, \beta_b^d]$ because $V_b(\theta)$ is constant, and decreasing on $[\beta_0^e, \beta_0^d]$ because V_0 is constant. On $[\beta_k^e, \beta_0^e]$ $G(\theta)$ can be increasing or decreasing, depending on the comparison of $p_b R_b$ and $(1-\Delta)R_0$, because the bank monitors the gambling project but not the prudent one. We will assume later that $p_b > 1 - \Delta$, hence $p_b R_b > (1-\Delta)R_0$, and the maximum is achieved at β_b^d. At this point,

$$V_b(\beta_b^d) - V_0(\beta_b^d) < V_b(\beta_b^d) - V_0(\beta_0^d) < 0.$$

This step is completed.

Step 3. Let $\underline{\theta} \equiv \beta_0^e$. Assumption 3 also guarantees that $\underline{\theta} < \bar{\theta}$. When $\theta \in [\underline{\theta}, \bar{\theta}]$, the bank chooses the prudent project and monitors. The proof is completed. □

A.5 Proof of Proposition 4.7

One needs to compare the welfare at $\theta = 0$ and β_0^e. The latter is higher under the above assumption. □

A.6 Proof of Proposition 6.1

If $r = 0$, $\bar{R} = pR$ and $\underline{R} = (1-\Delta)R$. The rest follows. □

A.7 Proof of Proposition 6.2

By the definition of $\beta^{e'}$, the retention constraint is not binding only when β is between $\beta^{e'}$ and β^d. The results follow. □

A.8 Proof of Proposition 6.3

When the retention constraint is binding under both retention forms, $P_h(\beta_h(\theta)|p, R)$ is equal to $P_v(\beta_v(\theta)|p, R)$, which are just $1-\theta$ fraction of the total pool value, but $\beta_h(\theta) > \theta = \beta_v(\theta)$ at non-endpoints. Hence the retained tranches have higher expected value under horizontal retention. These together imply that a horizontal component generates higher expected value. This argument holds at β^d. With the fact that the value function is decreasing faster under vertical retention for θ below $\frac{R-r}{p\bar{R}}$, and the two value functions coincide at 0 and 1, the results follow. □

Appendix B Other sources of funding

In the baseline model there is deposit insurance. In this section I consider a different form of funding for the bank: the bank borrows $1-e$ from competitive lenders using unsecured debt, after which the bank initiates the project. The debt contract pays interest rate d if the bank succeeds, plus the transfer of control rights: if the bank fails, the lenders will claim the residual. Following the main model, the payoff in the bad state is $r = 0$; hence there is no residual left, and the date 2 incomes are verifiable:

$$(1-\Delta)\underline{d}_0 \geq 1-e \tag{9}$$

and (9) binding generates

$$\underline{d}_0 = \frac{d}{1-\Delta}$$

The equilibrium exists when

$$\theta < \underline{\theta}' \equiv 1 - \frac{(\lambda-(1-\Delta))(R_0-\underline{d}_0)-(\lambda-1)\frac{c}{\Delta}-c}{(\lambda-1)\alpha(1-\Delta)R_0}$$

where the right-hand side is when we replace d with $\underline{d}_0$ in $\underline{\theta}$. This threshold is larger than $\underline{\theta}$. Similarly, when the bank chooses the prudent project and monitors, the interest rate $\bar{d}_0$

will be

$$\bar{d}_0 = 1 - e$$

This equilibrium exists if $\theta \in [\underline{\theta}, \bar{\theta}]$, as in the main model. When the bank gambles but monitors, the interest rate will be

$$\bar{d}_b = \frac{1-e}{p_b}$$

The equilibrium exists if

$$\theta > \bar{\theta}' \equiv \frac{(\lambda\alpha + 1 - \alpha)(R_0 - p_b R_b) - (1 - p_b)\frac{1-e}{p_b}}{(\lambda - 1)\alpha(R_0 - p_b R_b)}$$

The $\bar{\theta}'$ on the right-hand side is $\bar{\theta}$ with d replaced with $\bar{d}_b$ and is smaller than $\bar{\theta}$.

Compared to the main model, the difference is that there exist two optimal regions of θ in which multiple equilibria arise: $[\underline{\theta}', \underline{\theta})$ and $(\bar{\theta}', \bar{\theta}]$.

As a last remark, I show that this debt contract is equivalent to an actuarially fair deposit insurance where a premium is paid ex post. For example, consider the equilibrium in which there is no screening yet no risk shifting. Let the premium be $A_0(0)$. The bank pays this premium only when being solvent. If the bank fails, the insurance authority takes over the bank as the lenders under a debt contract and repays d to depositors. For $A_0(0)$ to be actuarially fair, the authority will break even, which means

$$(1 - \Delta)A_0(0) + \Delta(-d) = 0$$

It follows that the only difference in this scheme is that the bank will be taken Δd in expectation instead of d.

Appendix C Ex ante screening

In this section, I present a different version of the model that involves ex ante screening.

C.1 Setup

Before the securitization transaction, the bank can manipulate the quality of its assets via its screening effort $s \in \{0, 1\}$. The screening technology is the same as monitoring in the main model. The difference is that screening effort is unobservable to investors. Investors are, however, sophisticated enough to take into account the bank's incentive problem. With this assumption, the model distinguishes itself from signaling models in which securitizers signal the quality of the underlying assets by choosing different retention fractions or forms; for instance, see DeMarzo and Duffie (1999) and Flynn, Ghent, and Tchistyi (2020). Denote the investors' belief on the level of the bank's screening effort by s'.

I keep the assumption that $r = 0$ so that the two retention forms are identical and the retention constraint is just $\beta \geq \theta$. I reorder the dates to $t \in \{-1, 0, 1, 2\}$, and the timeline of events is

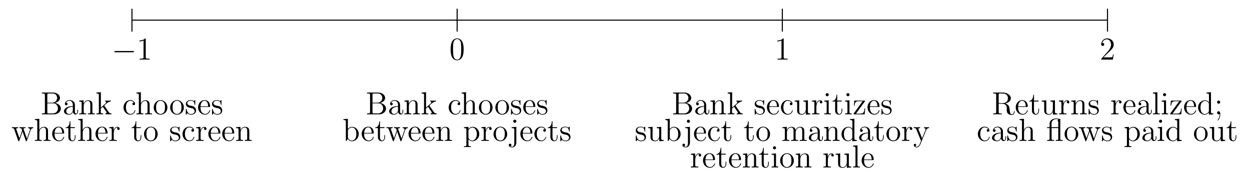

Figure 14: Timeline of the events

The equilibrium definition in the sense of subgame perfect equilibrium is the following.

Definition C.1 (Equilibrium.) *The equilibrium consists of the bank's screening choice s at time -1; the bank's project choice (p, R) at time 0; the bank's retention choice β at time 1; prices P of the tranches sold to investors and investors' belief on screening effort s' such that P reflects project choice and retention choice, and investors' belief; investors break even; the bank maximizes its value sequentially subject to the retention constraint at time 1; investors' belief is correct: $s' = s$.*

C.2 Equilibrium

At time 1, the price schedule offered by investors is

$$P(\beta, p, R, s') = (1 - \beta)\alpha p(s')R$$

The bank hence solves at time 1

$$\Pi_1(p, R, s; \theta) = \max_{\beta} \ \lambda P(\beta, p, R, s') + p(s)\Big((1-\alpha)R + \alpha\beta R - d\Big) - c(s)$$
$$\text{s.t. } \beta \geq \theta$$

Proposition C.1 (Retention choice under vertical component.) *If $s' \geq s$, the bank chooses $\beta = \theta$. If $s' = 0$ and $s = 1$, the bank chooses*

$$\beta = \begin{cases} \theta & \text{if } 1 < \frac{(\lambda-1)p}{\lambda\Delta} \\ 1 & \text{if } \frac{(\lambda-1)p}{\lambda\Delta} < 1 \end{cases}$$

Proof: The partial derivative with respect to β is

$$-\alpha\big(\lambda p(s)R - p(s)R + \lambda(s' - s)\Delta R\big)$$

The result follows. □

In equilibrium, the investors have correct beliefs that the retention constraint will be binding: $\beta = \theta$. Securitization provides the bank with the full market value plus the securitization benefit, which overrides the expected project gain from retention.

At time 0, similar to the main model, I define the risk shifting propensity as

$$G(\theta, s) = \Pi_1(p_b, R_b, s; \theta) - \Pi_1(p_0, R_0, s; \theta)$$

The bank chooses to gamble if this gain is positive.

Proposition C.2 (More retention leads to more bank risk shifting.) *In equilibrium, $G(\theta, s)$ is increasing in θ for any s.*

Proof: In equilibrium, $s' = s$. The bank retains $\beta = \theta$. We have

$$\frac{\partial G}{\partial \theta} = (\lambda - 1)(p_0(s)R_0 - p_b(s)R_b) > 0$$

because $p_0(s)R_0 > p_b(s)R_b$ for any s. □

The intuition is similar to the main model. At a low retention ratio, the bank chooses the prudent project because it has favorable market value; hence it is easy to securitize. At a high retention ratio, limited liability and higher upper side return R_b make the gambling project more profitable. And let $\bar{\theta}$ be such that $G(\bar{\theta}, 1) = 0$. In equilibrium, this is the threshold above which the bank shifts risk when it does not shirk in screening.

The next result presents a crowding-out theory of screening and risk shifting. I show that screening and choosing the prudent project can be substitutes.

Proposition C.3 (More screening leads to more bank risk shifting.) *In equilibrium, $G(\theta, 1) > G(\theta, 0)$.*

Proof: In equilibrium, we have

$$G(\theta, 1) - G(\theta, 0) = \bigg(\lambda\alpha(1 - \theta) + \alpha\theta + 1 - \alpha\bigg)\Delta(R_b - R_0) > 0$$

which completes the proof. □

The return to screening, ΔR, is higher for the gambling project: an increase in screening effort leads to an increase in the upper side return of the project,[16] and the expected value of the gambling project increases more than that of the prudent project. Screening and gambling are complements. In this sense, screening crowds out the prudent project.

I now examine the bank's screening motive and complete the characterization of the equilibrium. I already show that the risk shifting propensity is increasing in screening effort. Specifically, $G(\theta, 1)$ is above $G(\theta, 0)$. I have also defined $\bar{\theta}$ as the threshold satisfying $G(\bar{\theta}, 1)$. Above this threshold, the bank always chooses to gamble if it has screened the pool of projects.

At the boundary $\theta = 1$, there are two possibilities for $G(\theta = 1, 0)$. I first discuss the scenario when $G\big(1, 0\big) \leq 0$, which means that gambling will not happen without screening. As I mentioned above, a lack of due diligence in screening causes the value of the gambling project to drop more. This scenario is an extension of this result. The analysis is then divided into two parts: for $\theta \leq \bar{\theta}$ and for $\theta \in (\bar{\theta}, 1]$. In the first part, the bank always chooses the prudent project, due to the fact that

$$G(\theta, 0) < G(\theta, 1) < G(\bar{\theta}, 1) = 0.$$

That is, the gain from risk shifting propensity is always negative, regardless of the bank's screening effort. Given the choice of the prudent project, the bank screens if

$$(1-\alpha)R_0 + \alpha\beta(\theta, 1, s')R_0 - d - c \geq (1-\Delta)\big((1-\alpha)R_0 + \alpha\beta(\theta, 0, s')R_0 - d\big) \qquad (10)$$

The bank is comparing profits conditional on success. The payment from securitization purchasers is independent of the bank's screening effort. Therefore, it does not show up in the comparison. The bank's expected profits, or its skin in the game, are increasing in the retention ratio. In equilibrium, $\beta = \theta$, the above condition (10) simplifies as

$$\theta \geq \underline{\theta} \equiv \frac{\frac{c}{\Delta} + d - (1-\alpha)R_0}{\alpha R_0}$$

This is the intended consequence of mandatory retention: the bank increases its effort in screening as the retention ratio rises. I summarize the equilibrium for $\theta \leq \bar{\theta}$.

Proposition C.4 *When $\theta < \underline{\theta}$, the bank chooses the prudent project but does not screen ex ante. When $\underline{\theta} \leq \theta \leq \bar{\theta}$, the bank chooses the prudent project and screens.*

[16] If the screening technology has different implications for different projects, say Δ_0 and Δ_b, then for the risk shifting propensity to be increasing in s, I need $\Delta_b R_b > \Delta_0 R_0$. A sufficient condition will be $\Delta_b > \Delta_0$.

Figure 15 portrays the risk shifting propensities under different screening efforts on the θ space. Thanks to the results in the previous section, both propensities are increasing in θ, and $G(\theta, 1)$ is above $G(\theta, 0)$. Note that $G(\theta, 0)$ is steeper than $G(\theta, 1)$.

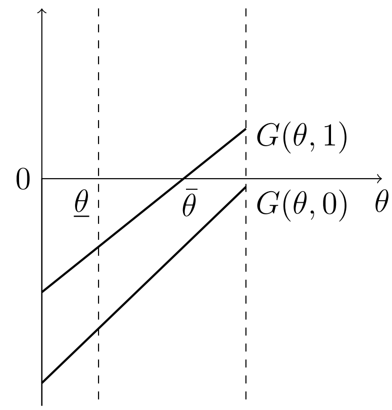

Figure 15: Risk shifting propensity under screening choice

For θ greater than $\bar{\theta}$ but smaller than θ_k, the bank chooses the gambling project if there is screening and chooses the prudent project if there is no screening due to the fact that

$$G(\theta, 0) < G\big(1, 0\big) < 0 = G(\bar{\theta}, 1) < G(\theta, 1)$$

It turns out that the bank will screen and gamble thereafter.

Proposition C.5 *When $\theta \in (\bar{\theta}, 1]$, the bank screens and gambles.*

Proof: Since $\theta > \underline{\theta}$, screening generates a higher value for the bank undertaking the prudent project. And under screening, the gambling project dominates the prudent project because $\theta > \bar{\theta}$. $\square$

In sum, the equilibrium is similar to the one in the main model. The welfare analysis hence follows. Lastly, if instead $G(1, 0) > 0$, we have Figure 16.

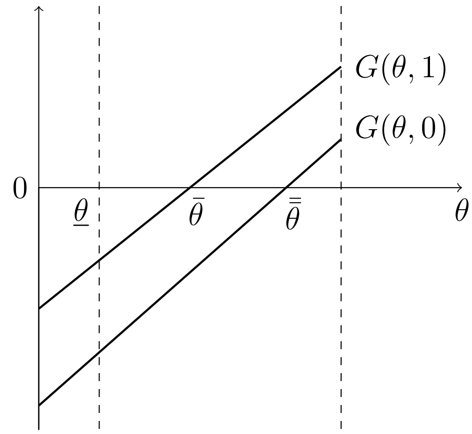

Figure 16: Risk shifting propensity under screening choice

Let $\bar{\bar{\theta}}$ be such that $G(\bar{\bar{\theta}}, 0) = 0$. Then the analysis for $\theta \leq \bar{\bar{\theta}}$ is the same as the main model. The bank screens and gambles simultaneously on $(\bar{\theta}, \bar{\bar{\theta}}]$. On $(\bar{\bar{\theta}}, \theta_k]$, the bank always gambles since both G functions are positive. It screens if

$$\theta \geq \frac{\frac{c}{\Delta} + d - (1-\alpha)R_b}{\alpha R_b}$$

Since $R_b > R_0$, this threshold is smaller than $\underline{\theta}$, hence smaller than $\bar{\bar{\theta}}$, implying that the bank always screens on this interval. As a result, the threshold $\bar{\bar{\theta}}$ does not matter, and the bank screens and gambles on $(\bar{\theta}, 1]$. The equilibrium is the same.

Appendix D

Table 9: **The effect of the mandatory retention rule on banks' risk shifting**

	Dependent variable: RWA ratio excluding securitization exposure			
	(1)	(2)	(3)	(4)
	RMBS≥0.5	RMBS≥0.8	RMBS=1	RMBS=1
Treatment	0.023*	0.023*	0.024*	0.023*
	(0.013)	(0.013)	(0.013)	(0.012)
Bank controls	No	No	No	Yes
Observations	331	323	291	291
R-squared	0.080	0.082	0.091	0.464
Number of bank	49	48	45	45

Notes: This table reports the effects of mandatory retention rules on banks' risk-weighted assets (RWA) ratio excluding securitization exposure, specified by

$$RWA_{it} = \beta D_{it} + X_{it}'\delta + \alpha_i + \gamma_t + \epsilon_{it}$$

Treatment equals 1 if bank i is an RMBS securitizer after December 2015 and 0 otherwise. In the first 4 columns, I increase the threshold of RMBS activity share from 0.5 to 1. In the first two columns, only bank and time fixed effects are considered. In columns 3 and 4, bank-specific control variables are added. Banks with total assets more than $250 billion are removed in the first 4 columns and added in column 5. Standard errors are in parentheses. The superscripts ***, **, and * represent 1%, 5%, and 10% significance levels, respectively.

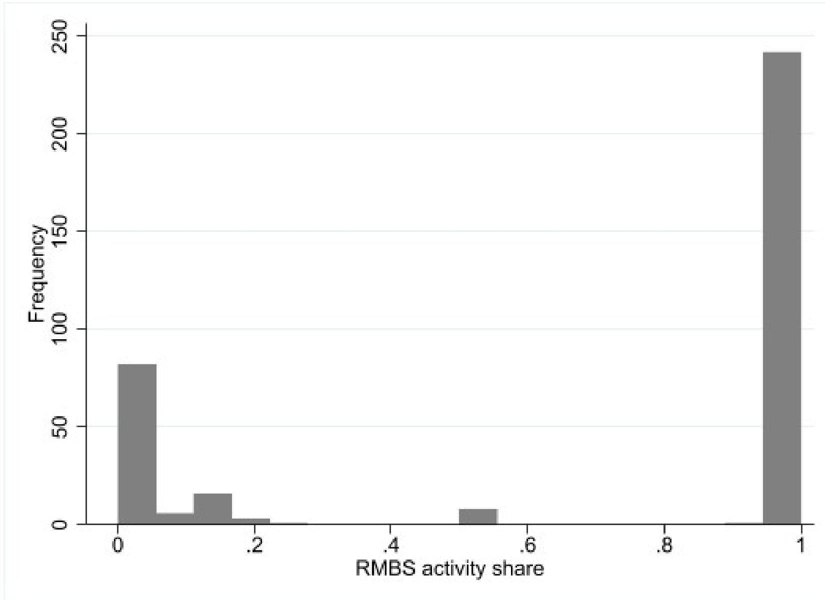

Figure 17: Distribution of RMBS activity share across BHC securitizers

Table 10: **The dynamics of continuous treatment effects**

	(1)	(2)
Before4	0.003	0.003
	(0.015)	(0.016)
Before3	-0.004	-0.005
	(0.008)	(0.009)
Before2	0.000	-0.001
	(0.009)	(0.009)
After0	0.015	0.011
	(0.010)	(0.008)
After1	0.017	0.015
	(0.011)	(0.009)
After2	0.020*	0.014
	(0.012)	(0.011)
After3	0.036***	0.031**
	(0.013)	(0.013)
Observations	359	359
R-squared	0.101	0.170
Number of bank	52	52

Notes: This table reports the dynamics of average continuous treatment effects according to

$$RWA_{it} = \alpha_i + \sum_{N=2}^{4} \beta_{5-N} Share_{it} * B_{N,it} + \sum_{M=0}^{3} \alpha_{3-M} Share_{it} * A_{M,it} + X_{it}'\delta + C_t + \epsilon_{it} \tag{11}$$

where B_N denotes N periods before the treatment and A_M denotes M periods after treatment. The quarter before treatment is used as the benchmark. Banks with total assets more than \$250 billion are removed. Standard errors are in parentheses. The superscripts ***, **, and * represent 1%, 5%, and 10% significance levels, respectively.

Table 11: **The effect of the mandatory retention rule on banks' loan delinquency rate**

	(1) RMBS≥0.5	(2) RMBS=1	(3) RMBS=1
Treatment	-0.003*	-0.003*	-0.004*
	(0.002)	(0.002)	(0.002)
Capital buffer			-0.000***
			(0.000)
Size			0.012
			(0.016)
ROA			0.003
			(0.095)
Liquidity ratio			-0.020
			(0.024)
Securitization-asset ratio			0.001
			(0.006)
Credit enhancements			-0.140
			(0.203)
Observations	333	293	293
R-squared	0.089	0.107	0.161
Number of bank	49	45	45

Notes: This table reports the effects of mandatory retention rules on banks' loan delinquency ratio. The econometric specification is binary treatment

$$DEL_{it} = \beta D_{it} + X'_{it}\delta + \alpha_i + \gamma_t + \epsilon_{it}$$

The sample period is from 2015Q1 to 2016Q4. *Treatment* equals 1 if bank i is an RMBS securitizer after December 2015 and 0 otherwise. In the first 2 columns, I increase the threshold of RMBS activity share from 0.5 to 1. In the first two columns, only bank and time fixed effects are considered. In column 3, bank-specific control variables are added. Banks with total assets more than $250 billion are removed. Standard errors are in parentheses. The superscripts ***, **, and * represent 1%, 5%, and 10% significance levels, respectively.

Table 12: **The effect of the mandatory retention rule on banks' loan delinquency rate (continuous treatment)**

	(1)	(2)	(3)
Treatment	-0.004*	-0.003*	-0.003*
	(0.002)	(0.002)	(0.002)
Capital buffer			-0.000***
			(0.000)
Size			0.011
			(0.015)
ROA			-0.023
			(0.101)
Liquidity ratio			-0.004
			(0.022)
Securitization-asset ratio			-0.005
			(0.008)
Credit enhancements			-0.019
			(0.111)
Observations	333	293	293
R-squared	0.089	0.107	0.161
Number of bank	49	45	45

Notes: This table reports the effects of mandatory retention rules on banks' loan delinquency ratio. The econometric specification is continuous treatment

$$DEL_{it} = \beta Share_{it} * D_t + X_{it}'\delta + \alpha_i + \gamma_t + \epsilon_{it}$$

The sample period is from 2015Q1 to 2016Q4. D_t equals 1 if time t is after December 2015 and 0 otherwise. In the first 2 columns, I increase the threshold of RMBS activity share from 0.5 to 1. In the first two columns, only bank and time fixed effects are considered. In column 3, bank-specific control variables are added. Banks with total assets more than $250 billion are removed. Standard errors are in parentheses. The superscripts ***, **, and * represent 1%, 5%, and 10% significance levels, respectively.

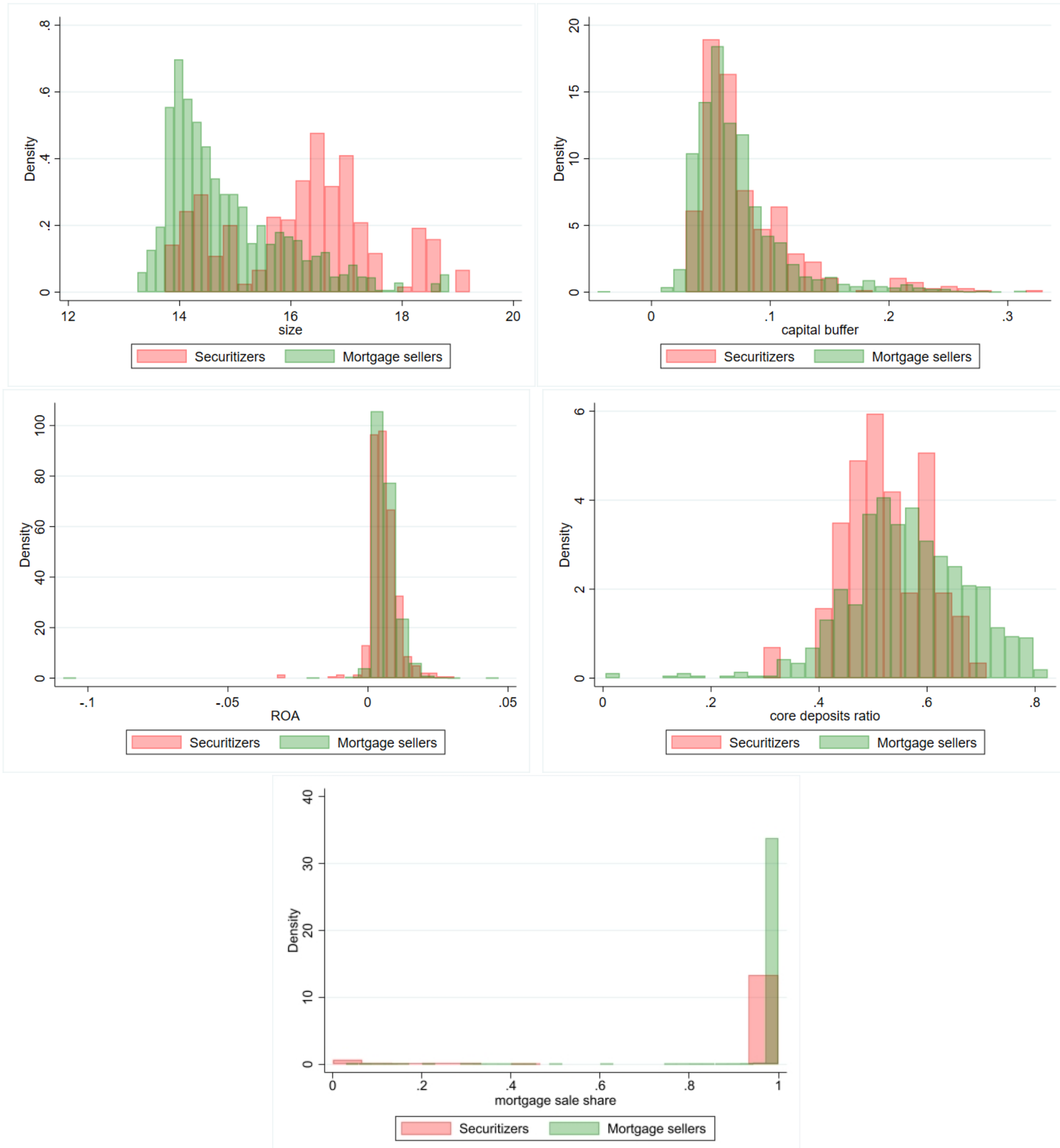

Figure 18: Pre-matching distributions of the covariates

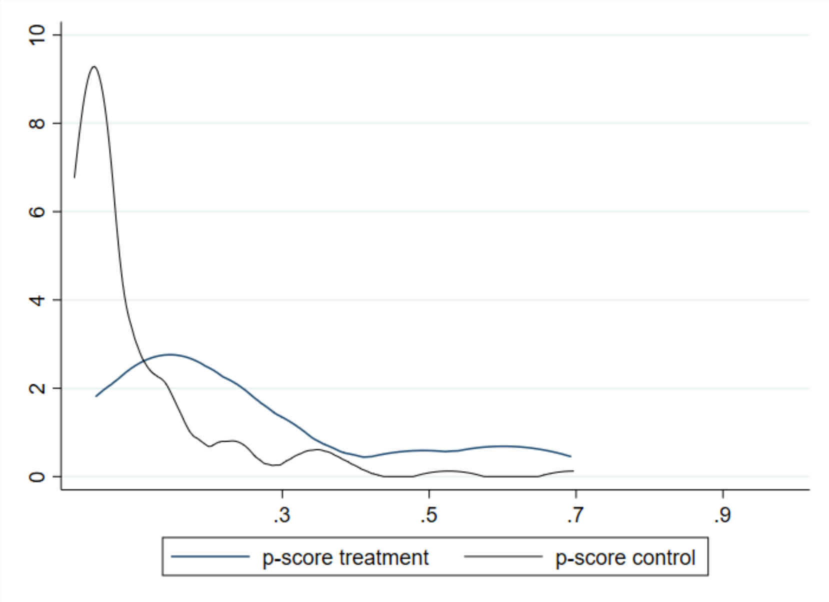

Figure 19: Common support

Table 13: **PSMDID**

	Dependent variable: RWA ratio			
	(1)	(2)	(3)	(4)
Treatment	0.04	0.021*	0.047**	0.040**
	(0.014)	(0.011)	(0.015)	(0.017)
Core deposits ratio	No	Yes	No	Yes
Mortgage sale share	No	No	Yes	Yes
Observations	220	141	207	133

Notes: This table reports the effects of mandatory retention rules on banks' RWA ratio under different propensity score matching covariates. In all specifications, size, capital buffer, and ROA are included in matching. *Treatment* equals 1 if bank *i* is an RMBS-only securitizer and 0 otherwise. Standard errors are in parentheses. The superscripts ***, **, and * represent 1%, 5%, and 10% significance levels, respectively.

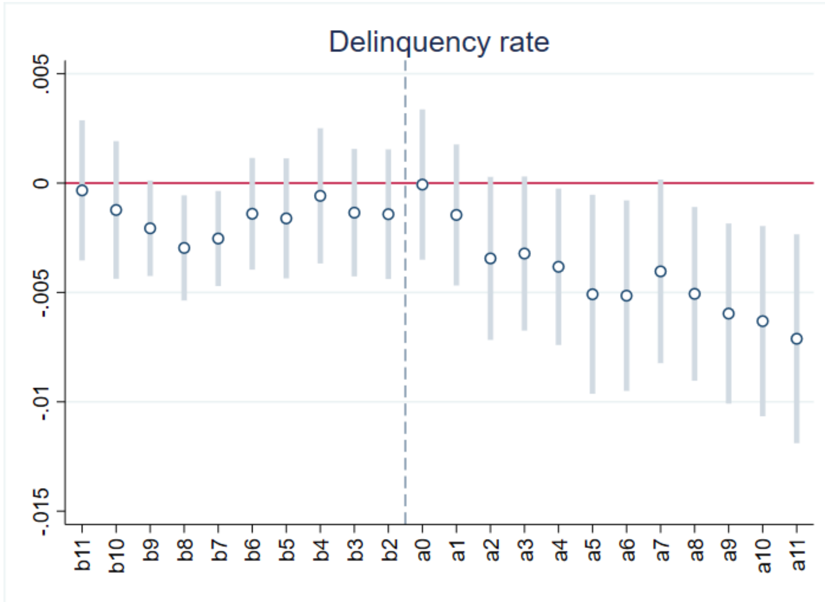

Figure 20: Time passage relative to the mandatory retention rule's implementation (from 2013Q1 to 2018Q4)

Appendix E The presence of a stronger form of risk shifting

In the previous discussion, investing resources in a riskier portfolio implies risk shifting. However, shareholders will have an even stronger incentive to shift risk when their share in the bank decreases. In this section, I add an exercise to test whether this stronger form of risk shifting is present among RMBS securitizers after the implementation of the retention rule on December 31, 2015.

Following Shrieves and Dahl (1992) and Duran and Lozano-Vivas (2014), among others, I assume that changes in equity share and risk level of bank i are simultaneous. Admittedly, on the one hand, banks will manage their equity ratio by accounting primarily for the risk of default. On the other hand, risk shifting will depend on how close the capital buffer is to the minimum requirement.[17] The changes consist of a decision component and a random part:

$$\Delta Y_{it} = \Delta Y_{it}^{*} + \epsilon_{it}$$

The decision part is assumed to be a partial adjustment

$$\Delta Y_{it}^{*} = c(Y_{it}^{*} - Y_{i,t-1})$$

where Y_{it}^{*} is the target level, $Y_{i,t-1}$ is the observable beginning-of-the-period level, and c is some adjustment speed. Specifically, the adjustment framework is the following two equations:

$$\Delta RWA_{i,t} = \alpha(RWA_{i,t}^{*} - RWA_{i,t-1}) + u_{i,t} \quad (12)$$

$$\Delta e_{i,t} = \beta(e_{i,t}^{*} - e_{i,t-1}) + w_{i,t} \quad (13)$$

where $\Delta RWA_{i,t} = RWA_{i,t} - RWA_{i,t-1}$ and $\Delta e_{i,t} = e_{i,t} - e_{i,t-1}$ are the change in bank i's RWA level and equity ratio; $RWA_{i,t}^{*}$ and $e_{i,t}^{*}$ are targeted values of the RWA and equity ratios, while $RWA_{i,}$ and $e_{i,t}$ are observed values; α and β are adjustment speed. Target levels are not observable and can vary across banks.

The target levels are approximated by a linear function that depends upon a set of explanatory variables. I include variables that I anticipate to affect both banks' capital structure and risk level. Specifically, I include size, profitability, and capital buffer, which I also use in the previous DID designs. The system of equations (12) and (13) after target

[17] The current version of my model does not feature this framework, but can be relaxed to incorporate it.

levels are replaced becomes

$$\Delta RWA_{i,t} = \alpha_0 + \alpha_e \Delta e_{i,t} + \alpha_1 Size_{i,t} + \alpha_2 ROA_{i,t} + \alpha_3 Buffer_{i,t} - \alpha_4 RWA_{i,t-1} + u_{i,t} \quad (14)$$

$$\Delta e_{i,t} = \beta_0 + \beta_r \Delta RWA_{i,t} + \beta_1 Size_{i,t} + \beta_2 ROA_{i,t} + \beta_3 Buffer_{i,t} - \beta_4 e_{i,t-1} + w_{i,t} \quad (15)$$

Since I have already identified a rise in risk shifting after the implementation of the mandatory rule in Section 5.1, I focus on risk-increasing banks from 2016 to 2018. The stronger form of risk shifting is present if the coefficients α_e and β_r are negative and significant. That is, banks that lower the capital ratio increase their portfolio risk at the same time.

The model is estimated by a three-stage least square estimator. As in Table 14, the coefficients α_e and β_r are indeed negative and significant at the 1% level, demonstrating the presence of a stronger risk shifting pattern. Banks that reduce the stake of equity also take more risk. The result is robust to a longer time horizon after 2015.

Table 14: **3SLS result**

Dependent variable	ΔRWA	Δe
Δe	-1.528***	
	(0.402)	
ΔRWA		-0.279***
		(0.059)
Size	-0.003***	-0.002***
	(0.001)	(0.000)
ROA	-0.565**	-0.212*
	(0.234)	(0.109)
Capital buffer	-0.001	0.001***
	(0.000)	(0.000)
Lagged RWA	-0.072***	
	(0.013)	
Lagged capital		-0.234***
		(0.038)
Observations	391	391

This table presents the 3SLS estimation of the presence of risk shifting for RMBS securitizers in 2016–2018. The second and third columns refer to equations (14) and (15), respectively. Standard errors are in parentheses. The superscripts ***, **, and * represent 1%, 5%, and 10% significance levels, respectively.

Among those control variables, capital buffer is shown to lower risk-taking incentive, because it affects Δe in a positive and significant way and affects ΔRWA in a negative and significant way. This is not surprising since the capital buffer is considered a sign of bank solvency. Other control variables have ambiguous effects on risk shifting, however.

Appendix F BHC securitizers that do not issue RMBS in the sample

RSSD ID	Name	Location	Securitization Type
1026801	Fremont Bank Corp.	Fremont, CA	Others
1066209	Lauritzen Corp.	Omaha, NE	Others
1068025	Key Corp.	Cleveland, OH	Other consumer; Others
1068191	Huntington Bancshares Inc.	Columbus, OH	Auto; Others
1070345	Fifth Third Bankcorp	Cincinnati, OH	Auto
1562859	Ally Financial Inc.	Detroit, MI	Auto
2367921	Tompkins Financial Corp.	Ithaca, NY	Credit card
2389941	TCF Financial Corp.	Wayzata, MN	Auto
3226762	RBC USA Holdco Corp.	New York, NY	Others
3981856	Santander Holdings USA Inc.	Boston, MA	Home equity lines; Auto
4346751	California Republic Bankcorp	Irvine, CA	Auto
4759669	EB Acquisition Company, LLC	Dallas, TX	Auto