Securitization, Ratings, and Credit Supply*

Brendan Daley† Brett Green‡ Victoria Vanasco§

October 6, 2017

Abstract

We develop a framework to explore the interaction between loan origination and securitization. In the model, banks privately screen and originate loans and then issue securities that are backed by loan cash flows. Issued securities are rated and sold to investors. We show that the availability of credit ratings (or other public information) increases the allocative efficiency of cash flows by reducing costly retention, but reduces lending standards and can lead to an oversupply of credit. These findings are in contrast to regulators' view of credit ratings as a disciplining device. Moreover, improved screening does not solve the problem; as banks' screening technology becomes more precise, their lending standards collapse and some (though not all) bad loans are deliberately originated. We use the model to explore several commonly proposed policies and provide conditions under which they increase efficiency. Finally, we consider extensions to allow for ratings shopping and manipulation.

*We thank Barney Hartman-Glaser, Joel Shapiro, and Pablo Ruiz Verdú for their thoughtful discussions, and seminar participants at Stanford GSB, CREI, Universitat Pompeu Fabra, CU Boulder, New York University, London School of Economics, Banco de Portugal, University of North Carolina at Chapel Hill, Bendheim Center for Finance at Princeton University, Washington University at St. Louis and conference participants at MADBAR, EIFE Junior Conference in Finance and, the Economics of Credit Ratings Conference at Carnegie Mellon, the Workshop on Corporate Debt Markets at Cass Business School for helpful feedback and suggestions. Green gratefully acknowledges support from the Fisher Center for Real Estate and Urban Economics, and Vanasco her support from the Clayman Institute for Gender Research.

†Leeds School of Business at The University of Colorado, Boulder.

‡Haas School of Business at UC Berkeley.

§Graduate School of Business at Stanford University.

1 Introduction

Asset-backed securitization is an important driver of credit supply (Loutskina and Strahan, 2009; Shivdasani and Wang, 2011). In the US, since the mid-1990s, there has been substantial growth in the securitization of many asset classes including mortgages, student loans, commercial loans, auto loans, and credit card debt. This practice has financed between 30% and 75% of loan amounts in these consumer lending markets (Gorton and Metrick, 2012), significantly increasing households' access to credit. The development of markets for securitized products has been facilitated in part by credit rating agencies (CRAs), which allowed issuers access to a large pool of investors who would otherwise have perceived these securities as opaque and complex (Coval et al., 2009; Pagano and Volpin, 2010).

In the aftermath of the recent financial crisis, the practice of securitization has been under intense scrutiny. The roles of both originators in screening loans and of rating agencies in evaluating securitized products have come into question.[1] A variety of regulations have been proposed in attempt to discipline loan origination and protect investors. For example, the Dodd-Frank Act imposed a mandatory "skin in the game" rule on securitizers and established disclosure requirements on both securitizers and rating agencies. Clearly, there are important interactions between the accuracy of information available to investors, banks' decisions of which loans to originate, and the market for securities backed by these loan pools. Yet, surprisingly, the academic literature has little to say about these interactions.

In this paper, we propose a stylized model of origination and securitization to study the role of private information (e.g., screening) and of public information (e.g., ratings) and explore the implications for lending standards and the overall supply of credit. Our main finding is that the availability of public information improves the allocation of cash flows by reducing inefficient retention, but reduces lending standards and can lead to an oversupply of credit. Moreover, improvements in screening do not solve the problem; as banks' screening technology becomes more precise, their lending standards collapse and some (though not all) bad loans are deliberately originated. We then explore the effects of common policy proposals, such as those described above from the Dodd-Frank Act.

The model features a continuum of banks and a set of competitive and fully rational investors. Each bank has access to a loan pool, and uses a screening technology to acquire private information about the quality of its loans. Each bank then decides whether to fund its pool—the *origination stage*. After origination, banks have an incentive to reallocate the cash flow rights from their loan pools to investors (e.g., due to capital constraints) and do so by selling securities backed by their loan pool in the secondary market—the *securitization stage*. In this stage,

[1] See Dell'Ariccia et al. (2012), Keys et al. (2010), Jaffee et al. (2009), Mian and Sufi (2009), Agarwal et al. (2012) for how securitization negatively affected lending standards; and Pagano and Volpin (2010) and Benmelech and Dlugosz (2010) for the role, and failures, of CRAs in the securitization process.

the bank's private information hinders the efficient allocation of cash flow rights, which in turn distorts its incentives during the origination stage.

The model admits two channels through which information can be conveyed to investors to mitigate these distortions. First, because it is more costly for a bank to retain bad loans than good ones, retention may serve to signal quality to investors as in Leland and Pyle (1977). Second, information about the pool of loans underlying each security can be conveyed to investors through a noisy public signal, which we refer to as a *rating*—though it can be interpreted more broadly as any form of public information about the value of the security.

In order to understand the role of each channel and the intuition for our main results, it is useful to consider an originate-to-distribute (OTD) environment in which neither channel is present. That is, suppose that banks sell 100% of the loan pools that they originate without obtaining a rating. In this case, the market price for a loan pool in the secondary market is independent of loan quality, which, when combined with no retention, provides no incentive for banks to screen loans during the origination stage. Rather, banks are motivated purely by "volume lending"; a bank originates a loan if the secondary market price is larger than the amount of capital required for origination. In equilibrium, the market price must reflect average quality, and hence the average NPV of loan pools originated in the economy must be zero. Thus, lending standards are too low and too many loans are originated relative to first-best where the marginal loan (instead of the average loan) has zero NPV.

In the OTD environment, because the secondary price is relatively low, banks have an incentive to retain good loans on their balance sheet. Doing so would then reveal fully securitized loans to be of low quality, which would cause the secondary price to fall and the equilibrium to unravel. This observation motivates our exploration of a model with endogenous securitization where banks optimally choose how much of the loan-pool cash flows to retain. Absent ratings or release of other public information, the securitization stage is a standard signaling game where (least-cost) separation is the unique stable outcome. Banks retain a positive fraction if they originated a good pool and sell 100% of originated bad pools. By doing so, investors learn the quality of each loan sold on the secondary market and prices fully reflect all available information. However, because retention is costly, the bank does not realize the full social value of good loans, which leads to inefficiently high lending standards and an undersupply of credit.

We then introduce ratings to the model. After the retention decision, but prior to the sale of a security, a noisy signal about the quality of the underlying collateral is publicly observed. We ask how the presence of ratings affects what loans are originated. One natural intuition is that informative ratings will lead to tighter lending standards and increase the quality of loans made. We confirm this intuition is correct in the OTD environment where the retention channel is absent. That is, if banks securitize and sell all of the loans they originate regardless of

loan quality or rating accuracy, then introducing ratings leads to tighter, more efficient lending standards.

The effect of ratings on lending standards and credit supply is more nuanced when banks optimally choose their retention levels. When ratings are informative, banks with good loans no longer fully separate through retention. Instead there is some degree of pooling at a lower retention level.[2] Since retention is inefficient, ratings improve allocative efficiency in the securitization stage. But, because less is being retained and ratings are imperfect, their introduction actually increases incentives to originate lower quality loans and may induce an oversupply of credit.[3] In essence, when ratings are introduced, the equilibrium of the securitization stage endogenously shifts from a signaling-through-retention equilibrium toward an originate-to-distribute equilibrium. Thus, while introducing noisy public information improves the efficiency of the securitization stage, it does not discipline banks' lending standards during origination.

We then highlight a novel and somewhat perverse interaction between ratings (i.e., public information) and the precision of banks' screening technology (i.e., banks' private information at origination). Without ratings, as the banks' screening technology becomes arbitrarily precise, only good loans are originated. With ratings, however, as the banks' screening technology becomes more precise their lending standard collapses and a non-negligible mass of bad loans are (deliberately) originated.

We use the model to evaluate several different regulations. An intuitive and often proposed regulation is to require banks to retain a fraction of all originated loans. Proponents argue this will provide incentives for banks to make good loans by ensuring that they have some "skin in the game." Critics argue that such regulation may reduce the availability of financing. This trade-off is nicely captured within our framework. In addition, our model suggests a more subtle consideration in the evaluation of skin-in-the-game regulation, which goes as follows. If banks were using retention as a way to signal to investors, then mandated retention will either reduce the information content of the signal or exacerbate the use of retention as a signal of quality. Our model predicts that the latter case obtains and hence skin-in-the-game regulation leads to tighter lending standards and a reduction in credit supply. We identify sufficient conditions under which such a policy increases overall efficiency.

We also investigate policies related to disclosure requirements, both for securitizers and for CRAs. These policies aim to increase the degree of public information, which in our model is equivalent to a more informative rating. Here too we identify sufficient conditions under which such a policy increases overall efficiency, and then discuss situations in which it does

[2] A similar feature is present in Hartman-Glaser (2017), where it is shown that when sellers are able to signal both through retention and reputation (as opposed to with a public signal) the equilibrium is no longer separating.

[3] This result is consistent with empirical evidence that finds that increased third party certification, such as ratings or number of analysts, increases a firm's debt issuances, and sometimes equity issuances (Faulkender and Petersen (2006), Sufi (2007), Derrien and Kecskés (2013)).

not. Finally, motivated by central banks' policy of easing credit constraints in order to promote lending, we study the effect of a decrease in banks' liquidity needs. Surprisingly, we find that significant interventions of this kind may have precisely the opposite effect. That is, reducing banks' liquidity needs makes it cheaper for them to signal through retention, which can lead to increased retention levels and fewer loans being originated.

Our finding, that an oversupply of credit may result from introducing public information, relies on the rating being imperfect at the date of securitization, which seems (to us) a rather natural assumption. As the rating becomes perfectly informative the lending standard and level of credit supply converge to first best. However, our oversupply result does *not* require that ratings are biased nor that investors are somehow misled by their information content.

There is, however, an extensive literature that studies the strategic nature of CRAs and the strength of their incentives to provide unbiased information.[4] Inspired by the CRA models in Skreta and Veldkamp (2009), Sangiorgi and Spatt (2012), Bolton et al. (2012), and Opp et al. (2013), we consider two extensions of the model: ratings shopping and rating manipulation.[5] In both cases, the information content of the rating becomes endogenous. We show that these frictions effectively reduce the information content of ratings and, thus, have an effect similar to a reduction in the informativeness of (exogenously generated) ratings.

Several papers have highlighted the trade-off between productive and allocative efficiency studied in this paper. Parlour and Plantin (2008) study the effect of loan sales on banks' origination and on borrowers' capital structure decisions, while Malherbe (2012) explores the relation between risk-sharing post-origination and market discipline. Chemla and Hennessy (2014) explore a setting in which there is a moral hazard problem followed by a securitization decision. Absent regulation, they show that the incentive to exert effort is too low and an optimal policy to promote effort is forced retention. There is also a rich literature that focuses on optimal contracting with loan sales and moral hazard (Gorton and Pennacchi, 1995; Hartman-Glaser et al., 2012; Vanasco, 2017). None of these papers study the release of public information to investors about the assets being traded.

The approach adopted in this paper builds on our previous work. Daley and Green (2014) consider a signaling model in which receivers observe both the sender's costly signal as well as a stochastic "grade" that is correlated with the sender's type. We enrich this framework by incorporating an ex-ante stage where assets are strategically originated, meaning the distribution

[4]Important considerations include the role of CRA reputation and moral hazard (Mathis et al., 2009; Bar-Isaac and Shapiro, 2013; Fulghieri et al., 2014; Goel and Thakor, 2015; Kashyap and Kovrijnykh, 2016), feedback effects and ratings as coordination devices (Boot et al., 2006; Manso, 2013; Goldstein and Huang, 2017), and the implications of rating-contingent regulation (Opp et al., 2013; Josephson and Shapiro, 2015).

[5]These extensions are also in line with empirical studies on ratings shopping and manipulation: Ashcraft et al. (2011), Griffin and Tang (2011), Griffin et al. (2013), Becker and Milbourn (2011), He et al. (2011), Kraft (2015), Piskorski et al. (2015).

of the quality of assets brought to market is endogenous, similar to Vanasco (2017).

The remainder of the paper is organized as follows. In the next section, we introduce the model and our solution concept. In Section 3, we present several benchmarks. We analyze the equilibrium of the model and its comparative statics in Section 4. In Section 5, we explore the policy implications. Finally, in Section 6, we endogenize the information content of ratings by allowing for ratings shopping and manipulation. Section 7 concludes. All proofs are relegated to the Appendix.

2 The Model

There is a unit mass of loan originators, which we refer to as *banks*, and a competitive market of outside investors. There are two periods. In the first period, each bank makes two decisions: whether to originate a given pool of loans (the *Origination* Stage) and, if originated, what fraction of the loan pool to securitize and sell to outside investors (the *Securitization* Stage)—what is not sold remains on the bank's balance sheet. In the second period, the state of the economy and the cash flows from the originated loans are realized. All agents are risk neutral.

Origination stage. Each bank has access to one potential pool of loans. A loan pool requires one unit of capital to originate and generates a random future cash flow Y that depends on the state of economy, $\omega \in \{\mathbf{S}trong, \mathbf{W}eak\}$, and the pool's type, $t \in \{\mathbf{g}ood, \mathbf{b}ad\}$, which are independent. A *good* loan pool is expected to repay $1+\rho$ in both states of nature. In contrast, a *bad* loan pool is expected to repay $1+\rho$ in a *strong* economy, but only $\lambda\nu+(1-\lambda)(1+\rho) < 1$ if the economy is *weak*. One can interpret $\lambda \in (0,1)$ as the fraction of loans in a bad pool that default in a weak economy and $\nu < 1+\rho$ as the expected recovery given default. Let $\xi \in (0,1)$ denote the proportion of good pools in the economy, $\pi \in (0,1)$ be the probability that the economy is strong, and v_t be the expected repayment of a loan pool of type t.[6] We assume $v_b < 1 < v_g$, meaning only good loan pools create value.

Prior to making origination decisions, banks acquire information about loan pools using their screening technology.[7] The screening technology is a pair of probability density functions, $\{\psi_b, \psi_g\}$, with common support. If a loan pool is of type t, then a bank observes a random variable drawn from ψ_t. Suppose that screening results in a realization s, then the bank's *appraisal* about its loan pool, denoted by p, is given by:

$$p = \Pr(t = good|s) = \frac{\xi\psi_g(s)}{\xi\psi_g(s) + (1-\xi)\psi_b(s)}. \tag{1}$$

[6]The expected repayments are $v_g = 1+\rho$ and $v_b = \pi(1+\rho) + (1-\pi)(\lambda\nu + (1-\lambda)(1+\rho))$.

[7]Evidence of banks having the ability to acquire private information about borrowers can be found in Mikkelson and Partch (1986), Lummer and McConnell (1989), Slovin, Sushka, Polonchek (1993), Plantin (2009), Botsch and Vanasco (2016), among others.

As can be seen from (1), the information content of s is fully captured by its likelihood ratio $L(s) \equiv \psi_b(s)/\psi_g(s)$. We assume that L is a continuous random variable with support $[0, \infty)$.[8] Therefore, across the population of banks, appraisals p are distributed according to a cdf H, with density h that is positive almost everywhere on $[0, 1]$. Since there is a one-to-one match between banks and loan pools, each bank is indexed by its appraisal $p \in [0, 1]$. That is, bank p refers to a bank who observes signal s satisfying (1) when it screens its loan pool.[9]

After observing the realization from the screening technology, each bank decides whether or not to originate the loans in its pool. If the bank chooses not to originate, it has no further actions and earns a payoff of 0. If the bank originates its loans, it has the opportunity to securitize the cash flows from the pool and sell them to the outside investors, which we turn to now.

Securitization stage. Each originating bank has an incentive to raise cash through securitization of the cash flows from its loan pool. This need could arise for a variety of reasons (e.g., credit constraints, binding capital requirements, credit market imperfections combined with profitable investment opportunities). As in DeMarzo and Duffie (1999), we model this incentive in reduced form by assuming that banks discount second-period cash flows by a factor $\delta < 1$, while investors' discount factor is normalized to 1. Because banks are less patient than investors, fixing the origination decisions, the efficient allocation is for all loan cash flow rights to be transferred to investors.

During the securitization process, banks uncover additional information about the quality of their loan pools, which we capture as the bank learning the loan pool type t. For convenience, we focus on a simple securitization structure where banks choose the fraction of the cash flow rights to sell and retain the remaining fraction. Thus, if a bank chooses to sell a fraction $1 - x$ then for any realization of the cash flow y, $(1-x)y$ and xy are the amounts distributed to investors and to the bank respectively in the second period. Choosing a higher x should therefore be interpreted as the bank retaining more, which can serve as a (costly) signal to investors about the quality of the underlying loans (as in Leland and Pyle, 1977).

Remark 1. *In principle, each bank could design and sell a security that is an arbitrary function of its cash flow. In Daley et al. (2016), we study the relevant security design game with ratings and each bank's cash flow being a continuous random variable. Using the results therein, we demonstrate in Appendix B that the main insights of the present paper remain unchanged when we allow banks to design and sell arbitrary securities.*

Ratings. In addition to observing the level of costly retention x, we consider a second channel

[8] This assumption holds if, for example, ψ_t is a Normal density with mean m_t, $m_g \neq m_b$, and variance σ^2.

[9] Rather than specifying a screening technology, one could begin with the distribution of appraisals, H, as the primitive. From Kamenica and Gentzkow (2011), there exists a screening technology that endows this distribution of appraisals provided it satisfies Bayes Plausibility (i.e., $\xi = \int p dH(p)$).

through which information may be conveyed to investors, which we refer to as a *rating*. We start by modeling the rating as an exogenous public signal about the quality of the loan pool backing the security. That is, a rating is a publicly observable random variable R with type-dependent density function f_t on $\mathbb{R}$.[10] In Section 6, we endogenize the distribution of the random variable R by allowing for ratings shopping and manipulation.

The informativeness of a rating realization, r, is captured by the likelihood ratio: $\Gamma(r) \equiv \frac{f_b(r)}{f_g(r)}$.[11] Without loss, order the ratings such that Γ is weakly decreasing. A higher rating therefore corresponds to a "better" signal about the quality of the underlying pool of loans. We assume that ratings are informative, $E[\Gamma(R)|b] > E[\Gamma(R)|g]$, but boundedly so: $\inf_r \Gamma(r) > 0$ and $\sup_r \Gamma(r) < \infty$. To fix ideas and parameterize rating informativeness, we will sometimes refer to a binary-symmetric rating system in which there are two ratings, G and B, with $\gamma = \Pr(G|g) = \Pr(B|b) \in (\frac{1}{2}, 1)$, where higher γ corresponds to more informative ratings.

2.1 Preliminaries

It is useful to cover some preliminary features that must hold in any Perfect Bayesian Equilibrium (PBE) of the model. As is typical, we begin our analysis in the second (i.e., the Securitization) stage and works backward.

At the beginning of the Securitization stage, investors have a (common) prior belief μ_0 about the quality of the loan pool backing each security. Investors then update their belief about a given security based on observing both the bank's retention level x and the rating r to some final belief $\mu_f(x, r)$. This updating can be decomposed into a first update (based on x) and a second update (based on r). The first update results in an *interim belief*, $\mu(x)$. Along the equilibrium path, the interim belief must be consistent with the retention strategy of banks.[12]

The second update is purely statistical; investors update from their interim belief to a final belief based on the rating according to Bayes rule:

$$\mu_f(x, r) = \frac{\mu(x) f_g(r)}{\mu(x) f_g(r) + (1 - \mu(x)) f_b(r)} = \frac{\mu(x)}{\mu(x) + (1 - \mu(x))\Gamma(r)}. \tag{2}$$

Let $P(x, r)$ denote the price of a security as a function of the retention level chosen by the bank and rating. Since investors are risk-neutral and competitive, the price equals the expected

[10] To encompasses a situation with a countable set of ratings $\{y_1, y_2, \dots\}$, with probabilities $q_t(y_n)$, let $f_t(r) = q_t(y_n)$ for $r \in [n, n+1)$ and $f_t(r) = 0$ for all other $r \in \mathbb{R}$.

[11] If $f_g(r) = f_b(r) = 0$, we adopt the convention that $\Gamma(r) = 1$.

[12] A pure strategy for a bank is a type-dependent retention level, and a mixed strategy is a type-dependent probability distribution over retention levels.

value of the cash flows generated by the security given μ_f:

$$P(x,r) = \mathbb{E}[(1-x)Y|x,r] = (1-x)\big(\mu_f(x,r)v_g + (1-\mu_f(x,r))v_b\big). \tag{3}$$

Given a schedule of interim beliefs $\mu(\cdot)$, the expected payoff of a bank that has originated a type-t pool and then chooses retention level x is $u_t(x,\mu(x)) \equiv \mathbb{E}_R[P(x,R)|t] + \delta x v_t$. Equilibrium requires that banks select a retention level that maximizes u_t taking the belief schedule as given. Let u_t^* denote the equilibrium payoff of type t in the continuation game starting from the Securitization stage.

Moving back to the Origination stage, there are two critical links between the two stages. First, given continuation payoffs u_g^*, u_b^* and its appraisal, each bank optimally chooses whether to originate its loan pool, where origination yields an expected profit of $pu_g^* + (1-p)u_b^* - 1$ compared to zero for not originating. Let O^* be the set of loan pools originated. Second, investors' prior belief in the Securitization stage, μ_0, must be consistent with banks' decisions in the Origination stage. Since investors are not privy to the appraisals of individual banks, the belief consistency condition is simply $\mu_0 = E[p|p \in O^*]$.

The Lending Standard. Intuitively, because good pools generate higher returns and better ratings, $u_g^* > u_b^*$ in any PBE. This implies that the origination decision takes a cutoff form, where bank p originates if and only if $p \geq p^*$. We refer to p^* as the equilibrium *lending standard*. To avoid the technicalities associated with corner solutions and guarantee that the lending standard is always interior, we assume the following.

Assumption 1. $\xi v_g + (1-\xi)v_b < 1 < \delta v_g$.

Substantively, the first inequality says that banks have ample access to low quality loans in the aggregate. Hence, if all loan pools were originated, their aggregate NPV would be negative. The second inequality says that banks are patient enough that holding a good loan generates positive NPV for them.

Lemma 1. *In any PBE, the set of originated loan pools is a truncation, $O^* = [p^*, 1]$, where*

$$p^* = \frac{1-u_b^*}{u_g^* - u_b^*} \in (0,1). \tag{4}$$

An immediate corollary is that investors' prior beliefs in the Securitization stage is conditional on the loan pool's appraisal p being above the lending standard p^*. That is, $\mu_0^* = A(p^*) \equiv E[p|p \geq p^*]$. In addition, the total supply of credit is $Q(p^*) \equiv 1 - H(p^*)$.

Collecting these preliminaries, we have the following explicit connection between equilibrium behavior and beliefs across the two stages.

Corollary 1. *Any PBE of the model is characterized by the following.*

1. *In the Securitization stage: Given μ_0, for each originated loan pool, bank retention strategies, investor beliefs, and security prices comprise a PBE of the signaling game.*
2. *In the Origination stage: Given the continuation payoffs implied by the Securitization stage, (u_g^*, u_b^*), the lending standard is p^* as given by (4).*
3. *Belief Consistency: $\mu_0^* = A(p^*)$.*

Finally, as is typical in signaling games, the Securitization stage has multiple PBE due to the flexibility of beliefs off the equilibrium path. To handle this multiplicity, we employ the D1 refinement (Banks and Sobel, 1987; Cho and Kreps, 1987). Roughly, D1 requires investors to attribute an off-path retention choice to the type who is more likely to gain from this deviation. See Appendix A.1 for a formal definition. Hereafter, we use *equilibrium* to refer to a PBE that satisfies D1 in the Securitization stage.

3 Benchmarks

3.1 Full-Information/First-Best (FB)

If the type of each loan pool were publicly observable in the Securitization stage, there would be no incentive for banks to retain any of their cash flow rights, and full *allocative* efficiency would be achieved: $x_b^{FB} = x_g^{FB} = 0$. In addition, prices would perfectly reflect underlying value, so $u_t^* = v_t$. Moving back to the Origination stage, *productive* efficiency is also achieved as loan pools are originated if and only if they generate positive NPV (i.e., if $pv_g + (1-p)v_b - 1 \geq 0$). Hence, the first-best lending standard is

$$p^{FB} = \frac{1 - v_b}{v_g - v_b} \in (0, 1),$$

and the first-best total supply of credit is therefore $Q(p^{FB}) = 1 - H(p^{FB})$.

Remark 2. *Our measure of the first-best lending standard, p^{FB}, implicitly assumes that banks capture all of the surplus from originated loans. This allows us to focus on the distortions arising from information frictions. In our investigation of policy proposals (Section 5), we allow for externalities from origination that are not captured by the bank (e.g., on borrowers or taxpayers).*

3.2 Originate-to-Distribute (OTD)

Suppose that banks are forced to sell 100% of the loan pools they originate. In this case, and perhaps in line with the popular intuition, (i) the lending standard is too lax compared to the

first-best benchmark, leading to an oversupply of credit relative to first-best, and (ii) more informative ratings work to ameliorate (i).

To illustrate these findings, notice that without any retention decision, the price in the Securitization stage is based only on the rating-updated investor belief, $P(r) = \mu_f(r)v_g + (1-\mu_f(r))v_b$, where $\mu_f(r) = \frac{\mu_0}{\mu_0+(1-\mu_0)\Gamma(r)}$. Therefore, for any given investor prior belief $\mu_0 \in (0,1)$, continuation payoffs are

$$u_t^{OTD} = \mathbb{E}_R[\mu_f(R)v_g + (1-\mu_f(R))v_b|t] = \mathbb{E}_R[\mu_f(R)|t](v_g - v_b) + v_b. \tag{5}$$

For any informative (but imperfect) rating system, $0 < \mathbb{E}_R[\mu_f(R)|b] < \mathbb{E}_R[\mu_f(R)|g] < 1$ and therefore $v_b < u_b^{OTD} < u_g^{OTD} < v_g$. On the one hand, originating a good loan is less profitable than v_g, which pushes the lending standard up relative to p^{FB}. On the other hand, originating a bad loan is more profitable than v_b, which pushes the lending standard down relative to p^{FB}. The next result shows that in equilibrium, the second force dominates.

Proposition 1. *For any (imperfect) rating system, the equilibrium lending standard in the OTD setting is too lax, i.e., $p^{OTD} < p^{FB}$.*

Intuitively, since the rating only imperfectly distinguishes good loans from bad ones, without retention, there is not enough discipline on banks during origination. As the informativeness of ratings increase (e.g., as γ increases for binary-symmetric ratings), u_g^{OTD} increases and u_b^{OTD} decreases, leading to an increase in the lending standard and a decrease in credit supply. As ratings become perfectly informative (e.g., as $\gamma \to 1$), $u_g^{OTD} \to v_g$ and $u_b^{OTD} \to v_b$, as they are in the first-best benchmark. Hence, in the limit, $p^{OTD} \to p^{FB}$, but there is always an oversupply of credit if ratings are short of perfectly informative.

At the other extreme, if we take the limit to uninformative ratings (e.g., as $\gamma \to \frac{1}{2}$), then $\mathbb{E}_R[\mu_f(R)|t] \to \mu_0 = A(p^{OTD})$ for either type. Hence, any funded loan pool garners the exact same price, which reflects the average cash flow of all funded loans. In the limit equilibrium, loan pools will be funded up until the average gross return is equal to the cost of funding:

$$A(p^{OTD})v_g + (1 - A(p^{OTD}))v_b - 1 = 0.$$

Thus, in a OTD setting without ratings, the lending standard is set such that the *average* funded loan pool generates zero NPV (whereas efficiency requires the *marginal* funded loan pool to generate zero NPV). Notice that in this case, the secondary price for loan pools is equal to 1 and therefore banks with good loan pools have an incentive to retain them on their balance sheet (since $\delta v_g > 1$). Of course, if banks strategically retain loans then the OTD equilibrium unravels. This observation serves as a motivation for analyzing a model in which we allow banks to make their retention decisions strategically.

3.3 Strategic Model without Ratings (NR)

Consider now the model as described in Section 2, but without informative ratings.[13] In this case, originators of good pools inefficiently retain a portion of their cash flows to signal their quality. This misallocation depresses the value of origination, leading to a lending standard that is too stringent compared to the first-best benchmark, resulting in an undersupply of credit relative to the first-best.

To illustrate, define $\bar{x}$ as the unique solution to

$$\underbrace{u_b(0,0)}_{v_b} \quad = \quad \underbrace{u_b(\bar{x},1).}_{(1-\bar{x})v_g+\delta\bar{x}v_b} \tag{6}$$

That is, the originator of a b-pool is indifferent between efficiently selling all of its cash flow rights at price v_b, and retaining fraction $\bar{x}$ if doing so leads to a price of $(1-\bar{x})v_g$ for the complementary fraction it sells. Therefore, $\bar{x}$ is the minimum amount the g-type must retain to separate from the b-type in the Securitization stage. As seen in similar signaling games, D1 selects this "least-cost separating" equilibrium.

Proposition 2. *Without informative ratings, in any equilibrium, retention levels in the Securitization stage are $x_b = 0$ and $x_g = \bar{x}$. Hence, $u_b^{NR} = v_b$ and $u_g^{NR} = (1-\bar{x})v_g + \delta\bar{x}v_g < v_g$.*

It follows from Lemma 1 that without ratings the equilibrium lending standard, denoted p^{NR}, is higher than in the first-best benchmark. Hence, there are positive expected NPV loans that are not being funded in this economy.

Corollary 2. *Without informative ratings, the equilibrium lending standard is too strict, i.e., $p^{NR} > p^{FB}$.*

4 Equilibrium

We now turn to the equilibrium of the full model in which banks strategically decide on retention/securitization and their issued securities are rated, modeled as the random variable R.[14] Again, we first characterize the equilibrium of the Securitization stage for any investor belief, μ_0 (Section 4.1), and then characterize banks' lending standard in the Origination stage along with the consistent investor belief (Section 4.2). We conclude by exploring the key determinants of the equilibrium lending standard including comparative statics on the precision of the screening technology and the informativeness of ratings (Section 4.3).

[13] That is, $\Gamma(r) = 1$ for all $r \in \mathbb{R}$.

[14] In Section 6, we endogenize the information content of ratings by allowing for ratings shopping and manipulation.

4.1 Securitization stage

Investors can potentially learn about the quality of a bank's pool from both the bank's securitization decision as well as from its rating. Intuitively, an originator of a g-pool would like to use both channels optimally. To this end, consider the following maximization problem:

$$\max_{x,\mu} \quad u_g(x,\mu) \quad s.t. \quad u_b(x,\mu) = v_b. \tag{7}$$

That is, given the rating system, among all retention-level/interim-belief pairs that deliver the b-type its full-information expected payoff, which one delivers the g-type its highest expected payoff? In the Appendix (Lemma A.2) we show that this problem has a unique solution, which we denote $(\tilde{x}, \tilde{\mu})$. The solution can be thought of as a bank with a g-pool making optimal use of the two channels at its disposal, subject to giving the bank with a b-pool its full information payoff. This optimality is a critical part of the equilibrium characterization (and where the D1 refinement plays its role), as Proposition 3 formalizes. We first characterize when the solution to (7) is interior.

Without ratings, the solution to (7) is $(\tilde{x}, \tilde{\mu}) = (\bar{x}, 1)$. That is, if there are no ratings to convey information to investors, the g-type uses the LCSE retention level to fully establish the superior quality of its cash flows. Add now informative ratings. If the retention-level/interim-belief remains $(\bar{x}, 1)$, then this addition has no effect because investors are completely convinced that $t = g$ even without the rating. Hence, for a g-type to rely on the rating at all, it must have an interim belief below 1. Banks will choose to rely on ratings only when they are sufficiently informative, as precisely captured by the following lemma.

Lemma 2. *In the solution to* (7)*,* $(\tilde{x}, \tilde{\mu}) < (\bar{x}, 1)$ *if and only if*

$$E[\Gamma(R)|b] > \frac{v_g - \delta v_b}{(1-\delta)v_g}. \tag{8}$$

The informativeness of a rating *realization*, r, is captured by its likelihood ratio: $\Gamma(r) = \frac{f_b(r)}{f_g(r)}$. $E[\Gamma(R)|b]$ is a measure of the informativeness of the *rating system*, $\{f_g, f_b\}$.[15] The right-hand side of (8) measures the relative cost advantage of the g-type in retaining cash flows. Thus, the solution to (7) has $(\tilde{x}, \tilde{\mu}) < (\bar{x}, 1)$ if and only if ratings are informative enough relative to the g-type's cost advantage of retention.

Given Lemma 2, it is perhaps not surprising that if (8) does not hold, then ratings are simply too noisy to alter the prediction from the no-ratings benchmark studied in Section 3.3. For

[15] The more informative the rating system, the higher is $E[\Gamma(r)|b]$. This measure is consistent with the notion of informativeness introduced by Blackwell (1951): if one rating system is Blackwell more informative than another, then $E[\Gamma(R)|b]$ is higher under the more informative system. Note that $E[\Gamma(r)|b] \geq E[\Gamma(r)|g] = 1$ for any rating system.

the remainder, we analyze the model in which ratings are informative enough to impact the equilibrium outcome: that is, henceforth we assume (8) holds unless otherwise stated. The equilibrium of the Securitization stage is then characterized as follows.

Proposition 3. *For any $\mu_0 \neq \tilde{\mu}$, there is a unique equilibrium of the Securitization stage. In it*

- *If $\mu_0 < \tilde{\mu}$, there is partial pooling at $\tilde{x} < \bar{x}$. That is, all banks with g-type pools retain $\tilde{x}$, a fraction $\frac{\mu_0(1-\tilde{\mu})}{(1-\mu_0)\tilde{\mu}}$ of banks with b-type pools retain $\tilde{x}$, and a fraction $\frac{\tilde{\mu}-\mu_0}{(1-\mu_0)\tilde{\mu}}$ retain zero. Hence, the interim belief for $x = \tilde{x}$ is $\mu(\tilde{x}) = \tilde{\mu}$.*

- *If $\mu_0 > \tilde{\mu}$, there is full pooling at $x = 0$. That is, all banks retain zero, regardless of type.*

For $\mu_0 = \tilde{\mu}$, there is full pooling in equilibrium, but it can be at any $x \in [0, \tilde{x}]$.

The proposition shows that, with informative ratings, banks with g-pools need not signal as vigorously to convey the quality of their security. Instead, they rely (to some extent) on the rating to convey information to investors. When investors are sufficiently optimistic ($\mu_0 > \tilde{\mu}$), there is full reliance on the rating. That is, banks endogenously choose a policy to sell 100% of the loans they originate. Otherwise, when $\mu_0 < \tilde{\mu}$, banks rely partially on retention and partially on the rating. That is, banks retain enough of g-backed pools to induce an interim belief of $\tilde{\mu}$ and rely on the rating beyond that.

4.2 Origination stage

Having characterized the Securitization stage, we now analyze the Origination stage. This analysis has two components: (i) optimality of the banks' lending standard to originate loan pools given investor beliefs and (ii) consistency of investor beliefs with banks' origination decisions.

Optimal Origination. Recall that given expected payoffs in the Securitization stage of u_g^*, u_b^*, a bank (weakly) prefers to originate if and only if $pu_g^* + (1-p)u_b^* - 1 \geq 0$, or equivalently $p \geq \frac{1-u_b^*}{u_g^*-u_b^*}$. From Proposition 3, u_g^* and u_b^* vary with the investors' belief μ_0 when informative ratings are present—in contrast to the first-best and no-ratings benchmarks. It is therefore useful to define the banks' reaction function as the marginal loan pool a bank is willing to originate (i.e., the lending standard) given investors' beliefs μ_0:

Definition 1. $\Psi(\mu_0) \equiv \left\{ \max\left\{ \frac{1-u_b^*}{u_g^*-u_b^*}, 0 \right\} \;\middle|\; u_g^*, u_b^* \text{ are equilibrium payoffs given } \mu_0 \right\}$.

The max operator in Ψ accounts for the fact that if $\frac{1-u_b^*}{u_g^*-u_b^*} < 0$, then banks will originate all loan pools, which is equivalent to setting the lending standard to 0. Next, from Proposition 3 we have that Ψ is single-valued for all $\mu_0 \neq \tilde{\mu}$. In more detail:

Corollary 3. *For given investor belief μ_0, the equilibrium lending standard with ratings satisfies*

$$p^* \in \Psi(\mu_0) = \begin{cases} \frac{1-v_b}{u_g(\tilde{x},\tilde{\mu})-v_b} & \mu_0 < \tilde{\mu} \\ \left\{ \frac{1-u_b(x,\tilde{\mu})}{u_g(x,\tilde{\mu})-u_b(x,\tilde{\mu})} \mid x \in [0,\tilde{x}] \right\} & \mu_0 = \tilde{\mu} \\ \max\left\{ \frac{1-u_b(0,\mu_0)}{u_g(0,\mu_0)-u_b(0,\mu_0)}, 0 \right\} & \mu_0 > \tilde{\mu}. \end{cases}$$

Figure 1(a) illustrates Ψ, and compares it to the lending standard in the first-best and no-ratings benchmarks, labeled p^{FB} and p^{NR}, respectively. In these two benchmarks, payoffs in the Securitization stage do not depend on investors' prior beliefs, so the lending standards are independent of μ_0. Furthermore, $p^{FB} < p^{NR}$, as documented in Corollary 2.

With ratings, the lending standard adopted by banks depends on the investor belief. When investors are pessimistic about loan pool quality (i.e., when $\mu_0 < \tilde{\mu}$), a b-type earns its full-information payoff ($u_b^* = v_b$), and a g-type optimally relies on both retention and the rating to earn a payoff higher than in the LCSE but below its full-information payoff. Hence, the lending standard with ratings falls in between the two benchmarks ($\Psi(\mu_0) \in (p^{FB}, p^{NR})$, for $\mu_0 < \tilde{\mu}$).

However, when investors are optimistic about loan pool quality (i.e., when $\mu_0 > \tilde{\mu}$), banks eschew inefficient retention, which increases the payoff of both types. Hence, origination is more attractive, and the lending standard drops at $\mu_0 = \tilde{\mu}$. Ψ continues to decrease as μ_0 further increases, as a higher investor belief translates directly into higher security prices for both types. Eventually, u_b^* reaches 1, the cost of origination. We denote this belief level as $\bar{\mu}$ (i.e., $u_b(0,\bar{\mu}) = 1$). Hence, for all investor beliefs $\mu_0 > \bar{\mu}$, banks are willing to originate all loan pools, regardless of their appraisals, since even the pools that turn out to be bad will earn a positive return. Consequently, $\Psi(\mu_0) = 0$ for all $\mu_0 \geq \bar{\mu}$, as seen in Figure 1(a).

Investor Belief Consistency. Finally, in equilibrium, investors' belief that a given loan pool is of high quality must be consistent with the banks' loan appraisal at origination surpassing the lending standard: $\mu_0^* = A(p^*)$. Combining this condition with the banks' optimal origination condition, $p^* \in \Psi(\mu_0)$, we having the following.

Proposition 4. *There exists a unique equilibrium. The equilibrium lending standard is given by the unique p^* satisfying $p^* = A^{-1}(\mu_0) \in \Psi(\mu_0)$.*

Figure 1(b) illustrates how the bank-origination-optimality and investor-belief-consistency conditions pin down the equilibrium lending standard, p^*, and investor beliefs, μ_0^*, as the strictly increasing function A^{-1} intersects Ψ exactly once. Figure 1(b) depicts an example with a lending standard, p^*, that is lower than the first-best benchmark (i.e., an oversupply of credit). However,

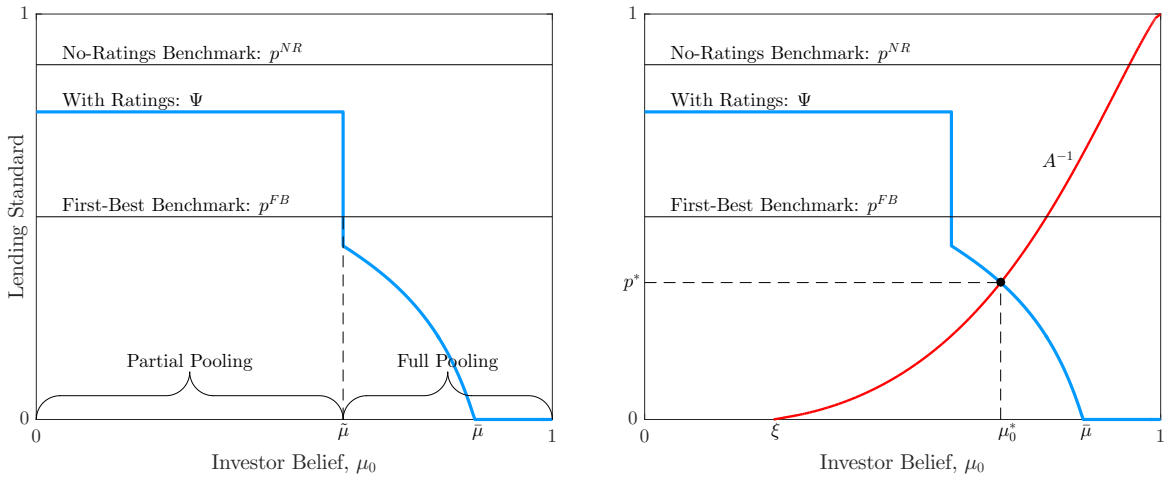


FIGURE 1: Panel (a) illustrates bank's lending as a function of investors' beliefs (Ψ), as well as the lending standard in the First-Best (p^{FB}) and No-Ratings (p^{NR}) benchmarks. Note that for $\mu_0 > \tilde{\mu}$, banks choose to sell 100% of originated loan pools regardless of t; thus the equilibrium payoffs and lending standard for such μ_0 are the same as in the OTD benchmark (Section 3.2). Panel (b) incorporates the belief consistency curve (A^{-1}), and illustrates the equilibrium lending standard, p^*, and investor belief, μ_0^*.

for other parameters the intersection of Ψ and A^{-1} lead to an equilibrium lending standard above the first-best level (see Figure 2, for example).

Corollary 4. *With ratings, the lending standard can be either above or below the first-best benchmark.*

In what follows, we study how changes in the banks' screening technology and/or in the rating informativeness impact banks' origination decisions and the supply of credit.

4.3 Determinants of Credit Supply

Precision of Banks' Screening Technology

A more precise screening technology means that, overall, banks become more certain whether their individual loan opportunities are bad or good before their origination decisions. Analytically, this is captured by mass in the distribution of appraisals shifting toward the extreme values of 0 or 1, which then has implications for the $A(\cdot)$ function that is used to pin down the equilibrium lending standard (as seen in Section 4.2).

Figure 2 illustrates how the precision of banks screening technology affects origination. For this example the screening technology, $\{\psi_g, \psi_b\}$, are Normal density functions with means $m_g > m_b$ and common standard deviation σ. As σ decreases, the screening technology becomes more

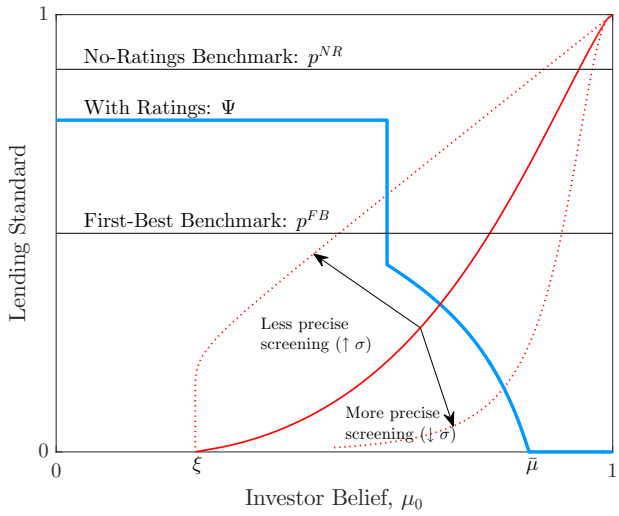


FIGURE 2: This figure illustrates how the precision of the screening technology affects the equilibrium lending standard.

precise and $A^{-1}(\mu_0)$ decreases for all $\mu_0 \in (\xi, 1)$. This is because, for any $p \in (0,1)$, if the loan pool is bad (good) it is becoming more likely that it would have generated an appraisal below (above) p. Consequently, as σ decreases, the equilibrium lending standard falls and the supply of credit increases.

The figure suggests that as σ goes to zero, $A^{-1}(\mu_0)$ tends to zero for all $\mu_0 \in (\xi, 1)$, meaning the equilibrium lending standard p^* would fall to 0. Proposition 5 shows that this result is indeed true and holds for any screening technology that becomes arbitrarily precise as defined below.

Definition 2. *A sequence of screening technologies* $\{\psi_b^n, \psi_g^n\}_{n=1}^{\infty}$ ***limits to perfect screening*** *if* $\lim_{n\to\infty} \Pr(L^n(s) \in (a,b)) = 0$ *for all* $0 < a < b < \infty$.

That is, the screening technology of banks becomes perfect when there is essentially no chance of receiving a signal that does not indicate the pool's underlying quality with arbitrary precision. However, just because banks *can* discern loan quality does not mean they will only originate good loans.

Proposition 5. *With ratings, if* $\{\psi_b^n, \psi_g^n\}_{n=1}^{\infty}$ *limits to perfect screening, then as* $n \to \infty$,

1. *The equilibrium lending standard* p^* *limits to zero.*
2. *The equilibrium supply of credit* $Q(p^*)$ *limits to* $\dfrac{\xi}{\bar{\mu}} > \xi$, *therefore*
3. *The measure of bad loans originated limits to* $\dfrac{\xi(1-\bar{\mu})}{\bar{\mu}} > 0$.

Hence, when banks are very good at appraising which loan opportunities are good or bad, they fund (virtually) all good loan pools as well as a strictly positive amount of loan pools that they are (virtually) certain are bad. This is because there is an incentive to originate until the average quality is driven down to $\bar{\mu}$—the investor belief level at which origination of a bad pool is expected to exactly break even.

It is worth noting that (informative) ratings are critical for this result. In the no-ratings benchmark, the lending standard is p^{NR} regardless of the screening technology. Further, since bad pools are sold for $v_b < 1$, it is not profitable to originate pools with low appraisals. Hence, without ratings, if $\{\psi_b^n, \psi_g^n\}_{n=1}^{\infty}$ limits to perfect screening then only good loan pools will be originated in the limit (i.e., the supply of credit tends to ξ—the mass of good loan opportunities).[16]

Informativeness of Ratings

We next analyze how changes in rating informativeness affect origination and securitization decisions. To sharpen our predictions, we focus on the binary-symmetric rating system (introduced in Section 2): $P(R = G|g) = P(R = B|b) = \gamma \in (\frac{1}{2}, 1)$, where higher γ implies more informative ratings. To begin, we examine how an increase in rating informativeness affects the Securitization stage, and consequently, the banks' reaction function for origination, Ψ.

Lemma 3. *As the informativeness of ratings (γ) increases,*

1. *Both $\tilde{\mu}$ and $\tilde{x}$ decrease, implying lower retention levels for all μ_0.*

2. *Letting $\hat{\mu} \equiv \max\{\tilde{\mu}, p^{FB}\}$, Ψ decreases for $\mu_0 < \hat{\mu}$ and Ψ increases for $\mu_0 > \hat{\mu}$.*

From statement (2) of the lemma, it is not surprising that more informative ratings can increase or decrease the lending standard/credit supply (as illustrated in Figure 3). There is however structure to these possibilities.

Proposition 6. *If the equilibrium lending standard is at least as high as the first-best benchmark ($p_\gamma^* \geq p^{FB}$), then the lending standard is strictly decreasing in rating informativeness (γ).*

Hence, starting from no-ratings/completely uninformative ratings (where $p^* = p^{NR} > p^{FB}$), increasing informativeness decreases the lending standard and increases the supply of credit. Will further increases in rating informativeness eventually lead to $p^* < p^{FB}$? In general the answer may depend on banks' screening technology. However, an unambiguous result can be obtained if overall loan opportunities are not too valuable.

[16]The screening technology affects the equilibrium lending standard/credit supply only by affecting A^{-1}. We can also note that the only other ingredient that determines A^{-1} is the proportion of good loans, ξ. Increasing ξ shifts A^{-1} to the right, leading to a decrease in the lending standard.

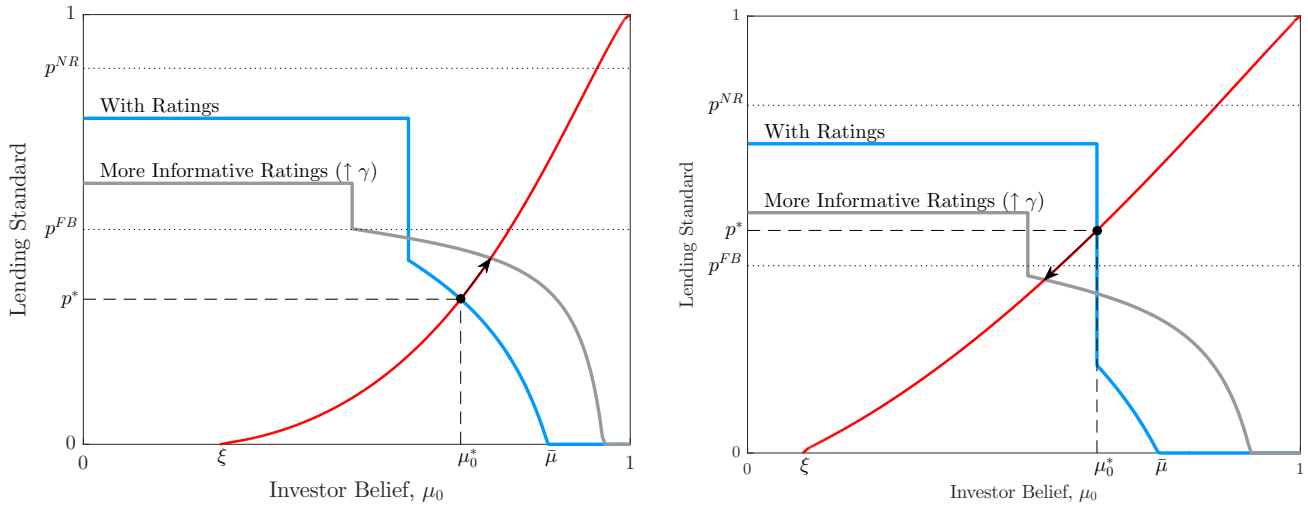


FIGURE 3: This figure illustrates how the informativeness of the rating technology (γ) affects the equilibrium lending standard and investor belief. In panel (a), an increase in rating informativeness leads to a higher lending standard, whereas in panel (b) the lending standard decreases.

Proposition 7. *If $v_g \cdot v_b \leq 1$, then for any screening technology, there exists $\hat{\gamma} \in (\frac{1}{2}, 1)$ such that the equilibrium lending standard is below the first-best benchmark ($p^*_\gamma < p^{FB}$) if and only if $\gamma \in (\hat{\gamma}, 1)$.*

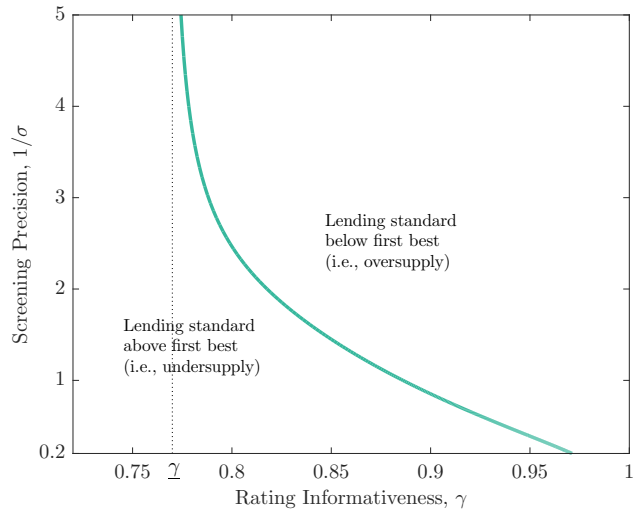
Notice that the condition $v_g \cdot v_b \leq 1$ is *not* independent of Assumption 1, as both restrict how valuable loan opportunities are in the aggregate. For example, if $\xi \leq \frac{1}{2}$, then Assumption 1 implies $v_g \cdot v_b \leq 1$, and sufficiently informative ratings always lead to an oversupply of credit. Graphically, as the rating becomes more informative, Ψ converges pointwise to p^{FB}, but it does so from above to the left of $\tilde{\mu}$ and from below to the right. The condition $v_g \cdot v_b \leq 1$ implies that $\lim_{\gamma \to 1} \tilde{\mu} < p^{FB}$, which ensures that any intersection with A^{-1} must occur at a lending standard below p^{FB}.[17]

Finally, and perhaps unsurprisingly, as the rating becomes perfectly informative, any mismatch between equilibrium and first-best origination (be it under- or oversupply) disappears.

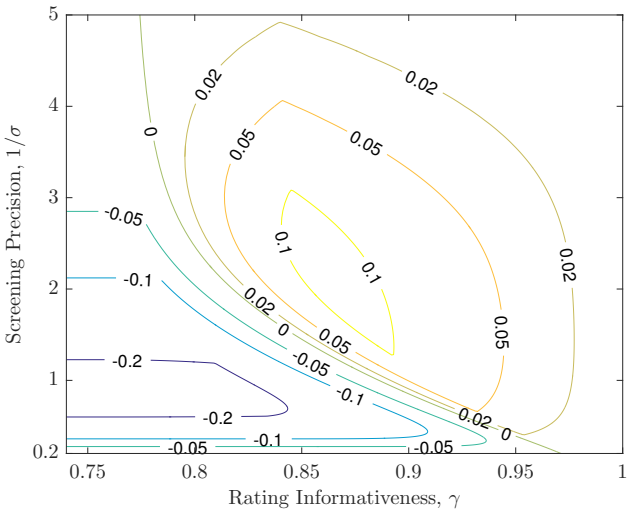
Proposition 8. *As ratings become perfectly informative ($\gamma \to 1$), the equilibrium lending standard tends to the first-best benchmark ($p^*_\gamma \to p^{FB}$).*

Having analyzed the effects of the precision of banks' screening technology and the informativeness of ratings, Figure 4 depicts the two in conjunction. Panel (a) illustrates when the equilibrium lending standard is above, equal to, or below the first-best benchmark. Recall from Lemma 2

[17]Recall that for any screening technology, A^{-1} lies weakly below the 45-degree line (i.e., the average loan above a threshold is always greater than the threshold).



(a) Lending standard



(b) Contour plot of origination quantity minus the first best level of origination

FIGURE 4: This figure illustrates the interaction between the precision of screening technology and rating informativeness on the equilibrium lending standard and quantity of origination. To illustrate Proposition 7, the parameters are such that $v_g \cdot v_b < 1$.

that there is a minimum level of ratings informativeness, labeled $\underline{\gamma}$ in the figure, required to alter the equilibrium predictions from the no-ratings benchmark in which the lending standard is $p^{NR} > p^{FB}$. Hence, if $\gamma < \underline{\gamma}$ there is an undersupply of credit, regardless of the screening precision. In contrast, for ratings informativeness above $\underline{\gamma}$, there is a strictly decreasing threshold for screening precision above which the lending standard is below first-best (in accordance with Proposition 5).[18] As ratings become more informative, less screening precision is required for the equilibrium to exhibit oversupply. In this example $v_g \cdot v_b < 1$, and thus oversupply always obtains for any screening precision when γ is large enough (Proposition 7).

While Figure 4(a) shows the under/oversupply regions, Figure 4(b) shows the quantity under/oversupplied in equilibrium to give a sense of when the mismatch is most pronounced (i.e., the "0.1"-contour implies there is an oversupply of credit equal in size to 10% of all loan opportunities).[19] Notice that the "0"-contour is identical to the single under/oversupply threshold in Figure 4(a). In accordance with Proposition 8, the supply of credit tends to the first-best level as rating become perfectly informative ($\gamma \to 1$). The configurations with the largest supply of credit occur for intermediate levels of both γ and σ (roughly around $(\gamma, \frac{1}{\sigma}) \approx (0.875, 2)$), which highlights a non-monotonicity of credit supply in both parameters.

[18]This threshold asymptotes to ∞ as $\gamma \to \underline{\gamma}$.

[19]The quantity of origination above first best is given by $Q(p^*) - Q(p^{FB}) = H(p^{FB}) - H(p^*)$.

5 Efficiency and Policy Analysis

After the recent financial crisis, both the US and Europe introduced a number of reforms to the securitization and rating industries. These regulatory responses conceptually fell into four categories: requiring risk-retention, increasing information disclosure, reforming rating agencies, and imposing capital requirements.[20] In addition, central banks intervened in a variety of ways to provide liquidity to banks, both before and after the crisis. Motivated by these regulatory responses, in this section we analyze the effect of forced risk-retention via "skin-in-the-game" rules (Section 5.1), disclosure requirements and regulation of CRAs (Section 5.2), and liquidity provision policies (Section 5.3).

In order to do so, we will consider not only the effect on the value of the banking sector, but also allow for additional surplus (positive or negative) from originating a t-type loan pool that is not capture by banks, which we denote by s_t. Such "externalities" could capture, for example, surplus to borrowers from access to credit, default costs, or systemic risks associated with the origination of bad loans. To fix ideas assume that $s_g \geq 0$, and to avoid trivial cases we maintain the assumption that it is inefficient to originate a bad loan (i.e., $v_b + s_b - 1 < 0$).

The total surplus generated in equilibrium is

$$\int_{p^*}^{1} \mathbb{E}[u_t^* + s_t - 1|p] dH(p) = \left[\mu_0^*(u_g^* + s_g) + (1 - \mu_0^*)(u_b^* + s_b) - 1\right] Q(p^*), \tag{9}$$

where $\mu_0^* = A(p^*)$ is the fraction of originated loan pools that are good and $Q(p^*) = 1 - H(p^*)$ is the total quantity of loan pools originated. By way of terminology, we say that there are positive (negative) externalities *on the margin* if $p^* s_g + (1-p^*)s_b > (<)\ 0$, and that the economy is *absent externalities* if $s_g = s_b = 0$. Absent both externalities and costly retention, efficiency is maximized at the first-best lending standard, p^{FB}. The socially optimal lending standard is higher (lower) than the first-best lending standard if the negative externalities associated with making bad loans are sufficiently large (small).[21]

The effect of the policies under consideration will depend on the equilibrium of the economy in which it is introduced. Recall from Section 4 that in the unique equilibrium, the nature of securitization/retention depends on the precision of their screening technology and the informativeness of ratings. Roughly, the equilibrium involves some degree of signaling through retention (which we refer to as a *signaling equilibrium*) when screening precision and ratings informativeness are low. Improvements in screening precision and/or ratings informativeness push equilibrium retention levels to zero (which we refer to as an *OTD equilibrium*).

[20] See Schwarcz (2015) for an analysis of the regulatory changes in securitization in response to the financial crisis, both in the US and in Europe.

[21] That is, if $s_b < (>)\ \frac{s_g(v_b-1)}{v_g-1}$.

Efficiency in OTD Equilibria. If the economy is in an equilibrium with zero retention, total surplus is given by

$$\int_{p^*}^{1} \mathbb{E}[v_t + s_t - 1|p]dH(p) = [\mu_0^*(v_g + s_g) + (1 - \mu_0^*)(v_b + s_b) - 1]\, Q(p^*). \tag{10}$$

We have already seen that the lending standard in an *OTD equilibrium* is lower than in the first-best (i.e., $p^* < p^{FB}$, see Section 3.2). Thus, absent externalities, OTD equilibria feature too much credit from a social perspective. The reason is that banks do not internalize the effect that their origination decision has on investors' equilibrium belief, μ_0^*, and therefore equilibrium payoffs. Naturally, the degree of inefficiency is larger (smaller) if there are negative (positive) externalities on the margin.

Efficiency in Signaling Equilibria. If the economy is in an equilibrium where banks with good pools are using retention to partially separate from those with bad pools, total surplus is given by

$$[\mu_0^*(u_g(\tilde{x}, \tilde{\mu}) + s_g) + (1 - \mu_0^*)(v_b + s_b) - 1]\, Q(p^*). \tag{11}$$

In contrast to *OTD equilibria*, absent externalities and taking the equilibrium retention level as given, the equilibrium lending standard is at its social optimum, with $p^* > p^{FB}$. This is because bank's payoff, u_t^*, are independent of μ_0^* in *signaling equilibria*. Given that banks are already optimizing on the choice of lending standards, overall efficiency can only be increased by reducing retention levels, and lending standards are inefficiently low if and only if there are negative externalities.

In what follows, we use these results to analyze the effect of different policies on lending standards, credit supply, and overall efficiency in detail.

5.1 Skin-in-the-Game Requirements

In October, 2014, as part of the Dodd-Frank Wall Street Reform and Consumer Protection Act, the US passed a skin-in-the-game rule requiring sponsors of securitization transactions to retain risk in those transactions. The regulation requires sponsors of asset-backed securities to retain at least 5 percent of the credit risk of the assets collateralizing the issuance. The rule also sets forth prohibitions on transferring or hedging the credit risk that the sponsor is required to retain. This rule aims to align incentives between the originators of assets and the investors who end up holding these assets. A similar rule has been imposed in Europe in the Capital Requirements Regulation (CRR).[22]

[22]In contrast to the skin-in-the-game rule in the US, the CRR requires regulated banks to only purchase asset-backed securities for which the originator explicitly discloses that it will retain, on an ongoing basis, a material

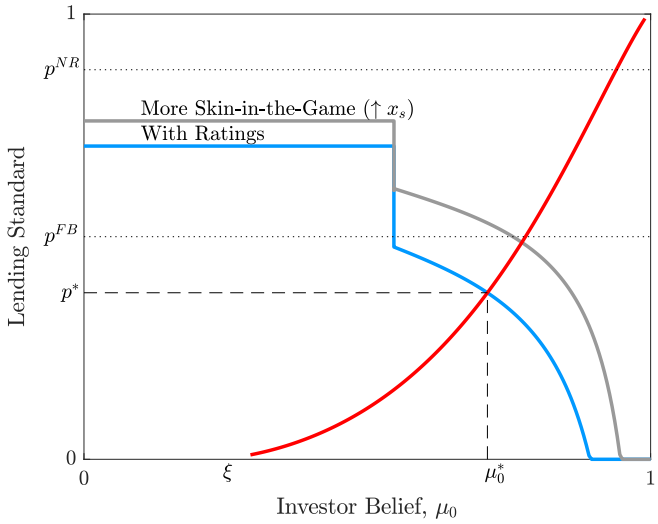


FIGURE 5: The Effect of a Skin-in-the-Game Rule.

We study the impact of retention rules by considering a policy in which banks are forced to retain an exposure to their loan pool of at least x_s. As in practice, the retention requirement is not contingent on the choice of security, the rating, nor on other measures of quality of the underlying cash flows.[23] Risk-retention rules hinder banks' ability to signal the quality of their underlying loans to investors through retention. But they also tighten the lending standard and reduce credit supply, which can increase efficiency. The following proposition summarizes.

Proposition 9. *Imposing a retention requirement of $x_s > 0$ increases $\tilde{x}$ (and $\tilde{x} > x_s$), but does not affect $\tilde{\mu}$. As a result, the lending standard increases and aggregate credit supply falls. Furthermore,*

1. *In an OTD equilibrium, there exists a retention requirement $x_s > 0$ that increases overall efficiency if*

$$\underbrace{\left(p^*(v_g + s_g) + (1-p^*)(v_b + s_b) - 1\right)}_{\text{Social value of marginal loan}} \underbrace{\left(-h(p^*)\frac{dp^*}{dx_s}\bigg|_{x_s=0}\right)}_{\substack{\text{Reduction in credit} \\ \text{supply for margin} \uparrow x_s}} > \underbrace{(1-\delta)\left(\mu_0^* v_g + (1-\mu_0^*)\, v_b\right) Q(p^*)}_{\text{Social cost of marginal} \uparrow x_s} \tag{12}$$

2. *In a signaling equilibrium, retention requirements increase overall efficiency if and only if negative externalities on the margin are sufficiently large.*

net economic interest in the securitization of at least five percent for the life of the transaction.

[23]The present regulation does make exceptions for particular asset classes. However, for a given asset class, retention rules are equal for all asset qualities.

Forced cash flow retention lowers the full-information payoff of banks originating bad loan pools because some gains from trade are now necessarily forgone. As a result, it is more costly for banks with good pools to signal their quality to investors: retention levels have to increase in order to signal the same information, and thus $\tilde{x}$ increases to levels above x_s. In turn, the reduction in expected payoffs from securitization due to forced retention reduces the profitability of origination, increasing the lending standard and decreasing aggregate credit supply. This effect is illustrated in Figure 5.

In *OTD equilibria* retention policies can increase efficiency if the marginal gain from increasing the lending standard more than compensates for the marginal increase in the cost of retention, as stated in (12). Note that even absent externalities, there can be efficiency gains from imposing such a policy since the social value of the marginal loan is negative in *OTD equilibria*.

In *signaling equilibria* the lending standard is optimal given equilibrium retention levels (absent externalities), since changes in the lending standard do not impact u_t^*. As a result, an increase in retention levels can improve efficiency only if there are sufficient negative externalities on the margin.

5.2 Disclosure Requirements and CRA Regulation

Rules have been adopted in both the US and Europe to improve the disclosure, reporting, and offering process of securitized products. Regulations now require that securitizers disclose standardized, detailed, loan-level information as well as the risk models used to analyze it. They also mandate a minimum amount of time that must be given to investors to process and analyze these disclosures. In addition, the Dodd-Frank Act mandated the creation of the Office of Credit Ratings (OCR) to conduct oversight of the "nationally recognized statistical rating organizations" (NRSROs). The role of the OCR is to monitor and report on the NRSROs internal control structures, rating methodologies and models, conflicts of interest, quality of information disclosure, etc.[24] Similarly, the European Securities and Markets Authority (ESMA) was created to supervise CRAs in the European Union.

The overarching goal of these policies, be it through mandatory disclosures, additional time for investors, or oversight of CRAs, seems to be aimed at increasing the quality of public information available to investors.[25] In our model, the "rating" stands in for any release of public information. Hence, an increase in the quality of public information corresponds to an increase in

[24]For a more detailed description of the OCR mandate, see the *2016 Summary Report of Commission Staff's Examinations of Each Nationally Recognized Statistical Rating Organization* prepared by the SEC.

[25]In Section 6, we allow for behavior such as ratings shopping or rating manipulation and show that it effectively reduces the informativeness of ratings in equilibrium. As a result, oversight of CRAs which limits the scope for such behavior would also serve to increase the quality of public information.

the informativeness of the rating. We analyzed the effect of increasing rating informativeness on origination and securitization decisions in Section 4.3. Below, we restate some of the key findings together with their implications for efficiency using the case of binary-symmetric ratings.

Proposition 10. *Absent externalities, a marginal increase in the informativeness of public information (i.e., a marginal increase in γ)*

1. *Increases both the lending standard and overall efficiency if the economy is in an OTD equilibrium,*

2. *Reduces average retention levels and the lending standard, but increases overall efficiency, if the economy is in a signaling equilibrium.*

In an *OTD equilibrium*, improving the accuracy of public information brings the lending standard closer to its first-best level without increasing equilibrium retention, which clearly increases overall efficiency. The efficiency gain is even larger if there are negative externalities on the margin.

In a *signaling equilibrium*, improving the accuracy of public information also increases overall efficiency, but via a different channel. In this case, more accurate information reduces average retention levels, improving the efficiency of the allocation of cash flows. Recall that in a *signaling equilibrium* the lending standard is socially optimal *given* the equilibrium level of retention. Therefore any marginal reduction in retention levels leads to an increase in efficiency, as the lending standard simply adjusts to the optimal level give this new lower (more efficient) level of equilibrium retention.

Interestingly, if the equilibrium involves full pooling at some positive retention level, an increase in informativeness can decrease overall efficiency. When banks pool at a positive retention level both margins of inefficiency are at play: retention is too high and the lending standard is too low given the level of retention. An increase in the informativeness of public information lowers both retention and the lending standard, and can therefore increase or decrease efficiency depending on whether the benefits of decreased retention outweigh the costs of a lower lending standard.[26] Consistent with this finding, a large improvement in the quality of public information can reduce efficiency if it moves the economy to an *OTD equilibrium* (as in Figure 3(b)) since the benefits of reducing retention are accompanied by the costs of an inefficiently low lending standard.

[26] A formal proof of these statements is available upon request.

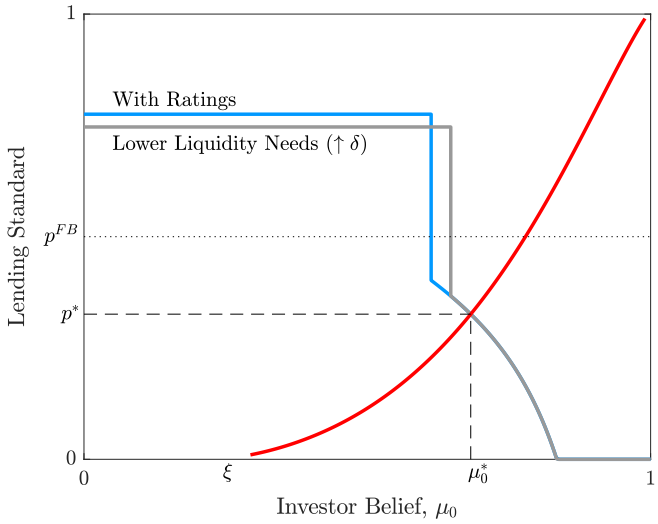


FIGURE 6: The Effect of a Decrease in Liquidity Needs.

5.3 Liquidity Needs

Central banks often undertake policies aimed at easing credit constraints of distressed financial institutions, in attempt to stimulate the economy by inducing banks to lend more.[27] Within our model, such policies can be interpreted as increasing δ.

Proposition 11. *As liquidity needs are reduced (i.e., as δ increases) $\tilde{\mu}$ and $\tilde{x}$ increase. A marginal increase in δ*

1. *Has no effect if the economy is in an OTD equilibrium.*
2. *Leads to a decrease in the lending standard and an increase in overall efficiency if the economy is in a signaling equilibrium (provided negative externalities on the margin are not too large).*

As liquidity needs decrease, retention (weakly) increases for all banks, but the cost of retention is also lower. The second effect dominates when equilibrium retention levels are already relatively high (e.g., in *signaling equilibria*). In this case, the reduction in retention costs more than compensates for the increase in equilibrium retention levels, increasing efficiency and the value of origination, which in turn reduces the lending standard toward the first-best level.

In an *OTD equilibria*, a small change in δ has no effect since banks are not retaining anything in the first place. This is the case illustrated in Figure 6. However, large enough increases in δ

[27]For example, in March 2008, the Federal Reserve announced the Term Securities Lending Facility that enabled banks to use MBS as collateral for short-term loans, which naturally reduced their need to sell such securities. Later, during quantitative easing, the Federal Reserve purchased outright billions of dollars in MBS. The European Central Banks adopted similar policy measures during the European Financial Crisis.

can cause the economy to shift to an equilibrium with positive retention levels, which reduces the value of origination and causes banks to lend less. Such a policy may increase efficiency if the gain from increasing the lending standard more than compensates for the increase in costly retention (similar to (12)).

6 Ratings Manipulation and Shopping

In this section, we consider extensions of the Securitization stage analyzed in Section 4 (hereafter, the *baseline* model) in order to investigate several realistic aspects of the CRA industry. First, we allow for rating manipulation by supposing that banks can inflate their rating by incurring a cost. The cost can be interpreted either as a side payment made to the CRA or as the cost of effort required to obscure the value of the underlying loan pool.[28] Second, we allow for ratings shopping by supposing that, after privately learning its rating r, a bank must pay the CRA to publish r publicly.[29] In both cases, the information content of the rating is determined endogenously, which we characterize below. Our main insight is that, provided investors are rational, the ability to shop or manipulate has an effect similar to a reduction in rating informativeness.

6.1 Manipulable Ratings

Suppose that after making their retention decision in the securitization stage, but before the rating r is realized, banks with a b-pool can manipulate ratings upward by incurring a cost m. To keep this extension simple, assume that when a bank with a b-type loan pool incurs the cost of manipulation, it obtains a rating from the distribution corresponding to a g-type loan pool. We wish to consider two cases. In the first case, investors are fully rational and understand that banks have the ability to manipulate their rating. In the second case, investors are "naïve" regarding the fact that banks can manipulate ratings.[30]

In order to do so, let x^* be the retention level on which banks pool (either partially or fully) in the model without manipulable ratings and let $\mu^* = \mu(x^*)$ denote investor's corresponding interim belief (see Propositions 3 and 4). Let $\alpha_t(\mu) = \mathbb{E}_R\left[\frac{\mu}{\mu+(1-\mu)\Gamma(r)}|t\right]$ denote the expected

[28]Piskorski et al. (2015) document that financial intermediaries disclosed false information about loans during the sale of residential mortgage and that misrepresented loans were significantly more likely to default. Findings by Ashcraft et al. (2010) and Ashcraft et al. (2011) suggest that CRAs ignored some types of information when assigning ratings to private-label MBS.

[29]Becker and Milbourn (2011) show that the greater competitive threat posed by Fitch in the corporate bond market coincides with an increase in rating levels from the incumbent agencies, consistent with the view that a larger scope for ratings shopping induces rating inflation.

[30]One can think of naïve investors as utility maximizing agents with an incorrect prior belief that $m = +\infty$.

posterior of investors given an interim belief μ and conditional on type t and define

$$\Delta \equiv (1 - x^*)(\alpha_g(\mu^*) - \alpha_b(\mu^*))(v_g - v_b). \tag{13}$$

Notice that Δ captures the benefit of manipulation if investors, correctly or not, believe that ratings are not manipulated.

Fully Rational Investors. Suppose first that investors are fully rational and thus aware of banks' ability to manipulate their ratings (though they do not actually observe whether manipulation occurred). If no manipulation takes place in equilibrium, then the b-type's net-benefit of deviating and manipulating its rating is $\Delta - m$. If $m \geq \Delta$ then the benefit of manipulation is not worth the cost and the ability to manipulate does not affect the equilibrium predictions. If instead $m < \Delta$, then manipulation must take place in equilibrium.

Proposition 12. *If ratings are manipulable and $m < \Delta$, then a fraction $\beta \in (0, 1)$ of banks who originate bad loan pools and choose the pooling retention level in the securitization stage will manipulate the rating.*

To understand the intuition for this result, consider the information content of a rating r given a fraction β of b-types are manipulating

$$\Gamma(r, \beta) \equiv \frac{\beta f_g(r) + (1 - \beta) f_b(r)}{f_g(r)} = \Gamma(r) + \beta(1 - \Gamma(r)). \tag{14}$$

Suppose investors believe that $\beta = 1$. Then ratings are completely uninformative (i.e., $\Gamma(r, 1) = 1$ for all r). Therefore, investors will disregard the rating and banks with g-type pools will choose the separating retention level as in the no-ratings benchmark (Section 3.3). In this case, banks with b-type pools will have no incentive to manipulate, which contradicts $\beta = 1$. On the other hand, if investors believe that $\beta = 0$, then the benefit of manipulating is strictly positive and all b-type banks will manipulate, which contradicts $\beta = 0$. Therefore, it must be that β is interior and such that the net-benefit of manipulating is zero.

In essence, Proposition 12 says that when ratings are manipulable and the cost of manipulation is not too large, then some manipulation will take place in equilibrium. From equation (14), manipulation causes each rating to become less informative (i.e., the likelihood ratio of each rating moves closer to 1). Thus, the ability to manipulate ratings leads to a rating system that is less informative and the comparative statics on the lending standard, credit supply and overall efficiency are similar to those already presented in Sections 4.3 and 5.2.

Naïve Investors. If investors are not aware of banks ability to manipulate ratings and $\Delta >$

m, then all banks with b-type pools will manipulate and the expected payoff associated with originating a b-type loan pool will increase by $\Delta - m$, while that of originating a g-type loan remains unchanged. Thus, when investors are not aware of manipulation, they transfer value to banks selling b-backed securities. As a result, banks increase their willingness to originate loans. In this case, the ability to manipulate ratings leads to a lower lending standard, higher credit supply, a reduction in overall efficiency, and losses for investors. Therefore, disclosure requirement policies designed to improve the transparency of the ratings and/or eliminate the scope for manipulation help protect investors from incurring losses (in addition to increasing the lending standard and overall efficiency).

6.2 Ratings Shopping

We now consider an alternative extension of the baseline model to allow for ratings shopping. After choosing the fraction of the loan pool to securitize, a bank approaches the CRA to get a rating for its issuance. The CRA generates a preliminary rating R with type-dependent density f_t, as in the baseline model. The CRA reports the preliminary rating to the bank, which then chooses whether to pay the CRA a fee in order to have the rating made public to investors.

We assume that the fee charged by the CRA is proportional to the cash flows being issued, $1 - x$, and we normalize it by the difference in expected payoffs from a good versus a bad pool, which is without loss of generality. That is, given a CRA fee $\phi \geq 0$, a bank that issues fraction $1 - x$ has to pay $\frac{\phi}{v_g - v_b}(1 - x)$ in order to have its rating published. We denote by $\mu_f(\mu, r; \phi)$ the posterior belief of investors after the rating r is published, given an interim belief $\mu \in (0, 1)$ and a rating fee, and by $\mu_n(\mu; \phi)$ the posterior belief associated with an unrated issuance.

Consider the decision of a bank with a t-type loan pool that has chosen to sell fraction $1 - x(t)$ and that has been proposed rating r by the CRA. The payoff from paying the CRA to have rating r published is

$$u_t\left(\mu, r; \phi\right) = \left(1 - x\left(t\right)\right)\left[\left(\mu_f\left(\mu, r; \phi\right) - \phi\right)\left(v_g - v_b\right) + v_b\right] + \delta x(t) v_t \tag{15}$$

while the payoff from staying unrated is:

$$u_t(\mu, n; \phi) = \left(1 - x(t)\right)\left[\mu_n(\mu; \phi)(v_g - v_b) + v_b\right] + \delta x(t) v_t \tag{16}$$

The bank chooses to hire the CRA if:

$$u_t\left(\mu, r; \phi\right) \geq u_t(\mu, n; \phi) \iff \mu_f(\mu, r; \phi) - \phi \geq \mu_n\left(\mu; \phi\right) \tag{17}$$

Notice from (17) that the decision to hire the CRA is independent of t. Therefore, the fact that a rating r is published does not convey information to investors about t beyond the information contained in the rating. Let $\mathcal{R}(\mu;\phi)$ denote the set of ratings that are published when the interim belief is μ. Then for any $r \in \mathcal{R}(\mu;\phi)$, Bayes rule requires that

$$\mu_f(\mu, r;\phi) = \frac{\mu}{\mu + (1-\mu)\Gamma(r)}. \tag{18}$$

Of course, for $r \notin \mathcal{R}(\mu;\phi)$, Bayes rule does not apply, and thus μ_f is not pinned down if such a rating is observed. To sharpen predictions, we refine off-path beliefs by specifying (18) also holds for all $r \notin \mathcal{R}(\mu;\phi)$. Clearly then, μ_f is decreasing in Γ. Recall that (without loss) Γ is decreasing in r (i.e., a higher rating is a better rating). Quite naturally, a cutoff strategy emerges: a bank pays the CRA to publish its rating if and only if the rating is above some threshold $\bar{r}(\mu;\phi)$ (i.e., $\mathcal{R}(\mu;\phi) = \{r : r \geq \bar{r}(\mu;\phi)\}$).

Therefore, the belief assigned to an unrated issuance is given by

$$\mu_n(\mu;\phi) = \frac{\mu F_g(\bar{r}(\mu;\phi))}{\mu F_g(\bar{r}(\mu;\phi)) + (1-\mu)F_b(\bar{r}(\mu;\phi))}. \tag{19}$$

Proposition 13. *A bank pays the CRA to publish its rating r if and only if $r \geq \bar{r}(\mu;\phi)$, where the threshold $\bar{r}(\mu;\phi)$ is characterized by*

$$\mu_f(\mu, \bar{r};\phi) - \mu_n(\mu;\phi) = \phi. \tag{20}$$

Furthermore, the set of ratings for which a bank chooses not to publish, $\{r : r < \bar{r}(\mu;\phi)\}$, has positive measure if and only if $\phi > 0$.

When ratings are costly ($\phi > 0$) and banks can decide whether or not to publish their assigned rating, the ratings observed by investors are truncated from below. As a result, the average rating observed by investors is "inflated" relative to the baseline model. Moreover, less information is conveyed to investors through the rating since a positive measure set of ratings is not disclosed. Therefore, similar to when ratings are manipulable, introducing a strategic decision of whether to publish a costly rating has an effect similar to reducing the informativeness of the ratings.

7 Conclusion

We have studied the effect of both ratings (i.e., public information) and screening precision (i.e., private information) on loan origination and securitization decisions. Informative ratings increase market liquidity and improve allocative efficiency; but reduce lending standards and may lead

to an oversupply of credit. We illustrate a novel and somewhat perverse interaction between the presence public information and the degree of banks' private information during origination; as the banks' screening technology becomes arbitrarily precise, the lending standard collapses to zero and some bad loans are (deliberately) originated by banks.

We use our model to explore the implications of policies, such as mandatory "skin-in-the-game" or disclosure requirements for CRAs and identify conditions under which such policies are welfare improving. We also consider several extensions to allow for ratings manipulation and shopping. Provided investors are fully rational, the possibility of such behavior has an effect similar to reducing the informativeness of ratings in equilibrium.

References

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

Ashcraft, A., P. Goldsmith-Pinkham, P. Hull, and J. Vickery (2011): "Credit ratings and security prices in the subprime MBS market," *The American Economic Review*, 101, 115–119.

Ashcraft, A. B., P. Goldsmith-Pinkham, and J. I. Vickery (2010): "MBS ratings and the mortgage credit boom," *New York Federal Reserve Staff Report no. 449*.

Banks, J. and J. Sobel (1987): "Equilibrium Selection in Signaling Games," *Econometrica*, 55, 647–662.

Bar-Isaac, H. and J. Shapiro (2013): "Ratings quality over the business cycle," *Journal of Financial Economics*, 108, 62–78.

Becker, B. and T. Milbourn (2011): "How did increased competition affect credit ratings?" *Journal of Financial Economics*, 101, 493–514.

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

Blackwell, D. (1951): "Comparison of Experiments," *Proceedings of the Second Berkeley Symposium on Mathematics, Statistics and Probability, (1950)*, 93–102.

Bolton, P., X. Freixas, and J. Shapiro (2012): "The credit ratings game," *The Journal of Finance*, 67, 85–111.

Boot, A. W., T. T. Milbourn, and A. Schmeits (2006): "Credit ratings as coordination mechanisms," *Review of Financial Studies*, 19, 81–118.

Chemla, G. and C. Hennessy (2014): "Skin in the game and moral hazard," *The Journal of Finance*.

Cho, I.-K. and D. M. Kreps (1987): "Signaling games and stable equilibria," *The Quarterly Journal of Economics*, 179–221.

Coval, J., J. Jurek, and E. Stafford (2009): "The Economics of Structured Finance," *The Journal of Economic Perspectives*, 23, 3–26.

Daley, B., B. Greeen, and V. Vanasco (2016): "Security Design with Ratings," *Working Paper*.

Daley, B. and B. Green (2014): "Market signaling with grades," *Journal of Economic Theory*, 151, 114–145.

Dell'Ariccia, G., D. Igan, and L. a. Laeven (2012): "Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market," *Journal of Money, Credit and Banking*, 44, 367–384.

DeMarzo, P. M. (2005): "The pooling and tranching of securities: A model of informed intermedia-

tion," *Review of Financial Studies*, 18, 1–35.

DeMarzo, P. M. and D. Duffie (1999): "A Liquidity-Based Model of Security Design," *Econometrica*, 67, 65–99.

Derrien, F. and A. Kecskés (2013): "The real effects of financial shocks: Evidence from exogenous changes in analyst coverage," *The Journal of Finance*, 68, 1407–1440.

Faulkender, M. and M. A. Petersen (2006): "Does the source of capital affect capital structure?" *Review of financial studies*, 19, 45–79.

Fulghieri, P., G. Strobl, and H. Xia (2014): "The Economics of Solicited and Unsolicited Credit Ratings," *Review of Financial Studies*, 27, 484–518.

Goel, A. M. and A. V. Thakor (2015): "Information reliability and welfare: A theory of coarse credit ratings," *Journal of Financial Economics*, 115, 541–557.

Goldstein, I. and C. Huang (2017): "Credit Rating Inflation and Firms' Investments," *Working Paper*.

Gorton, G. and A. Metrick (2012): "Securitization," Tech. rep., National Bureau of Economic Research.

Gorton, G. B. and G. G. Pennacchi (1995): "Banks and Loan Sales Marketing Nonmarketable Assets," *Journal of Monetary Economics*, 35, 389–411.

Griffin, J. M., J. Nickerson, and D. Y. Tang (2013): "Rating shopping or catering? An examination of the response to competitive pressure for CDO credit ratings," *Review of Financial Studies*, 26, 2270–2310.

Griffin, J. M. and D. Y. Tang (2011): "Did credit rating agencies make unbiased assumptions on CDOs?" *The American Economic Review*, 101, 125–130.

Hartman-Glaser, B. (2017): "Reputation and signaling in asset sales," *Journal of Financial Economics*, 125, 245–265.

Hartman-Glaser, B., T. Piskorski, and A. Tchistyi (2012): "Optimal securitization with moral hazard," *Journal of Financial Economics*, 104, 186–202.

He, J., J. Qian, and P. E. Strahan (2011): "Credit ratings and the evolution of the mortgage-backed securities market," *The American Economic Review*, 101, 131–135.

Jaffee, D., A. Lynch, M. Richardson, and S. Van Nieuwerburgh (2009): "Mortgage Origination and Securitization in the Financial Crisis," in *Restoring Financial Stability*, ed. by V. Acharya and M. Richardson, chap. 1.

Josephson, J. and J. Shapiro (2015): "Credit Ratings and Structured Finance," *Working Paper*.

Kamenica, E. and M. Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

Kashyap, A. K. and N. Kovrijnykh (2016): "Who Should Pay for Credit Ratings and How?" *Review of Financial Studies*, 29, 420–456.

Keys, B. J., T. Mukherjee, A. Seru, and V. Vikrant (2010): "Did securitization lead to lax

screening? Evidence from Subprime Loans," *The Quarterly Journal of Economics*.

Kraft, P. (2015): "Do rating agencies cater? Evidence from rating-based contracts," *Journal of Accounting and Economics*, 59, 264–283.

Leland, H. E. and D. H. Pyle (1977): "Informational Asymmetries, Financial Structure, and Financial Intermediation," *The Journal of Finance*, 32, 371–387.

Loutskina, E. and P. E. Strahan (2009): "Securitization and the Declining Impact of Bank Finance on Loan Supply: Evidence from Mortgage Originations," *The Journal of Finance*, 64, 861–889.

Malherbe, F. (2012): "Market Discipline and Securitization," .

Manso, G. (2013): "The Feedback Effects of Credit Ratings," *Journal of Financial Economics*, 109, 535–548.

Mathis, J., J. McAndrews, and J.-C. Rochet (2009): "Rating the raters: Are reputation concerns powerful enough to discipline rating agencies?" *Journal of Monetary Economics*, 56, 657–674.

Mian, A. and A. Sufi (2009): "The Consequences of Morgage Credit Expantion : Evidence from the U.S. Mortage Default Crisis," .

Opp, C. C., M. M. Opp, and M. Harris (2013): "Rating agencies in the face of regulation," *Journal of Financial Economics*, 108, 46–61.

Pagano, M. and P. Volpin (2010): "Credit ratings failures and policy options," *Economic Policy*, 25, 401–431.

Parlour, C. A. and G. Plantin (2008): "Loan Sales and Relationship Banking," *The Journal of Finance*, LXIII, 1291–1314.

Piskorski, T., A. Seru, and J. Witkin (2015): "Asset quality misrepresentation by financial intermediaries: evidence from the RMBS market," *The Journal of Finance*, 70, 2635–2678.

Sangiorgi, F. and C. Spatt (2012): "Opacity, credit rating shopping and bias," *Unpublished working paper. Stockholm School of Economics and Carnegie Mellon*.

Schwarcz, S. L. (2015): "Securitization and post-crisis financial regulation," *Cornell L. Rev. Online*, 101, 115.

Shivdasani, A. and Y. Wang (2011): "Did Structured Credit Fuel the LBO Boom?" *The Journal of Finance*, 66, 1291–1328.

Skreta, V. and L. Veldkamp (2009): "Ratings shopping and asset complexity: A theory of ratings inflation," *Journal of Monetary Economics*, 56, 678–695.

Sufi, A. (2007): "Information Asymmetry and Financing Arrangements: Evidence fromSyndicated Loans," *The Journal of Finance*, LXII, 629–668.

Vanasco, V. (2017): "The Downside of Asset Screening for Market Liquidity," *forthcoming, Journal of Finance*.

A Appendix

A.1 Preliminaries and Definitions

Let $\alpha_t(\mu) \equiv E_R[\mu_f(\mu, r)|t]$ be the expected posterior belief for a t-type bank given investor's interim belief μ. The following claims are rudimentary or, in the case of Fact A.1(3), have been established previously.

Fact A.1. *For any $t \in \{b, g\}$,*

1. *$\alpha_t(\mu)$ is strictly increasing in μ for any $x \in (0, 1]$.*
2. *$\alpha_g(\mu) - \alpha_b(\mu)$ is concave and achieves a unique maximum at $\mu_{\max} \in (0, 1)$.*
3. *$\frac{\partial}{\partial \mu}\left(\frac{\alpha_g'(\mu)}{\alpha_b'(\mu)}\right) < 0$ for all $\mu \in (0, 1)$ (shown in Lemma A.1. of Daley and Green (2014)).*
4. *$E[(1-x)Y|g] > E[(1-x)Y|b]$ for any $x \in (0, 1]$*
5. *$u_t(x, \mu)$ is strictly increasing in μ for any $x \in (0, 1]$.*
6. *$u_b(x, \mu)$ is strictly decreasing in x for any $\mu \in [0, 1]$.*

Fact A.2. *In any PBE, $u_t \in [v_b, v_g)$ for any $t \in \{b, g\}$.*

The D1 Refinement

Definition A.1. *We define $b_t(x, v)$ as the belief necessary to provide the t-type utility v if retention is x; that is, $u_t(b_t(x, v), x) = v$, and by $B_t(x, v) = (b_t(x, v), 1]$ the set of beliefs for which the t-type obtains strictly higher utility than v when retention is x.*

Fix $k \in [v_g, v_g)$ and $x \in [0, 1]$, and consider the belief $b_t(x, k)$ as defined in Definition A.1. By Fact A.1(5), there exists at most one $b_t(x, k)$ such that $u_t(b_t(x, k), x) = k$. Furthermore, the connection between b_t and B_t is immediate: if $b_t(x, k)$ exists, then $B_t(x, k) = (b_t(x, k), 1]$. If $b_t(x, k)$ fails to exist, then either $B_t(x, k) = [0, 1]$ or $B_t(x, k) = \emptyset$.

In our model, the D1 refinement can be stated as follows. Fix an equilibrium endowing expected payoffs $\{u_b, u_g\}$. Consider a retention choice x that is not in the support of either type's strategy. If $B_L(x, u_b) \subset B_H(x, u_g)$, then D1 requires that $\mu(x) = 1$ (where $\subset$ denotes strict inclusion). If $B_H(x, u_g) \subset B_L(x, u_b)$, then D1 requires that $\mu(x) = 0$.

A.2 Proofs for Section 2

Proof of Lemma 1. An originating bank always has the option to retain its loan pool. So, in any PBE, $u_t^* \geq \delta v_t$. Let $\bar{p}$ satisfy, $\bar{p}\delta v_g + (1-\bar{p})\delta v_b = 1$. Hence, any bank with loan opportunity $p \in (\bar{p}, 1]$ originates in any PBE (and $\bar{p} < 1$ follows from Assumption 1). So the set of originated pools, O^*, has positive measure under H, and $\mu_0 = E[p|p \in O^*] \in (0,1)$, by the belief consistency condition of PBE.

Next, given any $\mu_0 \in (0,1)$, in any PBE of the Securitization stage, $u_g^* > u_b^*$. To see, consider first any separating equilibrium: then $u_b^* = v_b < \delta v_g \leq u_g^*$, as no portion of a bad loan is retained, whereas the originator of a good loan must earn at least its full-retention payoff. If instead, the equilibrium has any degree of pooling on some retention level x, then

$$u_t^* = u_t(x, \mu(x)) = \mathbb{E}_R[P(x,R)|t] + \delta x v_t,$$

where both $\mathbb{E}_R[P(x,R)|g] > \mathbb{E}_R[P(x,R)|b]$ and $\delta x v_g > \delta x v_b$. The second inequality is immediate, and the first inequality follows from $\mu(x)$ being non-degenerate since there is (some) pooling on x, and the rating being informative.

Given that $u_g^* > u_b^*$, a bank's expected payoff from origination, $pu_g^* + (1-p)u_b^*$, is strictly increasing in p, implying that there is a cutoff lending standard, $p^* \leq \bar{p} < 1$. Finally, to see that $p^* > 0$, suppose not. Then $\mu_0^* = A(0) = \xi$. We already have that in any PBE, $u_b^* < u_g^*$. Hence, $u_b^* < \xi u_g^* + (1-\xi)u_b^* \leq \xi v_g + (1-\xi)v_b < 1$ since $\xi v_g + (1-\xi)v_b$ is the highest possible average bank payoff in the Securitization stage given $\mu_0^* = \xi$, and the last inequality is Assumption 1. But if $u_b^* < 1$, then a bank with p sufficiently close to 0 earns negative expected profit by originating, contradicting $p^* = 0$. ■

A.3 Proofs for Section 3

Definition A.2. *Let* $\tilde{\Psi}(\mu_0) \equiv \max\left\{\frac{1-u_b(0,\mu_0)}{u_g(0,\mu_0)-u_b(0,\mu_0)}, 0\right\}$.

Lemma A.1. *Under any rating system,*

1. *If* $\tilde{\Psi}(\mu_0) \in (0,1)$, *then* $\tilde{\Psi}'(\mu_0) < 0$.
2. $\tilde{\Psi}(p^{FB}) = p^{FB}$.
3. *With binary-symmetric ratings, if* $\tilde{\Psi}(\mu_0) > 0$, *then* $\text{sign}\left(\frac{d\tilde{\Psi}(\mu_0)}{d\gamma}\right) = \text{sign}(\mu_0 - p^{FB})$.

Proof. For (1), first $0 < u_b(0,\mu_0) < u_g(0,\mu_0)$ for any $\mu_0 \in (0,1)$. Hence, if $\tilde{\Psi}(\mu_0) \in (0,1)$, then $0 < u_b(0,\mu_0) < 1 < u_g(0,\mu_0)$. Given this, it is straightforward to show that $\tilde{\Psi}$ is decreasing in u_t for $t = b, g$. The result then follows from $u_t(0,\mu_0)$ being increasing in μ_0 for $t = b, g$.

For (2), because investor earn zero expected profit, expected bank gross return in the Securitization stage must equal the expected value of their offering. Therefore, for any μ_0,

$$\begin{aligned}\mu_0 u_g(0,\mu_0) + (1-\mu_0)u_b(0,\mu_0) &= \mu_0 v_g + (1-\mu_0)v_b \\ u_g(0,\mu_0) &= v_g + \frac{1-\mu_0}{\mu_0}(v_b - u_b(0,\mu_0))\end{aligned} \tag{21}$$

Substituting (21) into the definition for $\tilde{\Psi}$,

$$\tilde{\Psi}(\mu_0) = \max\left\{\frac{1-u_b(0,\mu_0)}{v_g + \frac{1-\mu_0}{\mu_0}(v_b - u_b(0,\mu_0)) - u_b(0,\mu_0)}, 0\right\}$$

Finally, using that $p^{FB} = \frac{1-v_b}{v_g - v_b}$,

$$\begin{aligned}\tilde{\Psi}(p^{FB}) &= \max\left\{\frac{1-u_b(0,p^{FB})}{v_g + \frac{1-p^{FB}}{p^{FB}}(v_b - u_b(0,p^{FB})) - u_b(0,p^{FB})}, 0\right\} \\ &= \max\left\{\frac{1-u_b(0,p^{FB})}{v_g + \frac{v_g-1}{1-v_b}(v_b - u_b(0,p^{FB})) - u_b(0,p^{FB})}, 0\right\} = p^{FB} > 0.\end{aligned}$$

For (3), for $\tilde{\Psi} > 0$, compute the derivative directly as

$$\frac{\partial\tilde{\Psi}(\mu_0)}{\partial\gamma} = \frac{\mu_0 v_g + (1-\mu_0)v_b - 1}{(2\gamma-1)^3(1-\mu_0)\mu_0(v_g - v_b)}. \tag{22}$$

The sign of this derivative is given by the sign of its numerator. Since the numerator is strictly increasing in μ_0 and it takes value zero when $\mu_0 = p^{FB}$, the result follows. ■

Proof of Proposition 1. Lemma A.1(1) and (2) yield the following equivalency:

$$\mu_0 > \tilde{\Psi}(\mu_0) \iff \mu_0 > p^{FB} \iff \tilde{\Psi}(\mu_0) < p^{FB} \tag{23}$$

Equilibrium in the OTD model requires

$$\mu_0^{OTD} = A(p^{OTD}) > p^{OTD} = \tilde{\Psi}(\mu_0^{OTD}), \tag{24}$$

where the first equality is investor-belief consistency, the inequality is by the definition of A, and (using Definition A.2) the second equality is optimal bank origination. Combining (23) and (24) establishes that $p^{OTD} < p^{FB}$. ■

Proof of Proposition 2. To check that this is a PBE, we need to check that neither type wishes

to deviate at any stage.

Securitization stage: First, an originator of a bad loan pool (b-type) does not profit from deviating since the retention of those holding good loan pools (g-types), x_g, is chosen so that the incentive compatibility (IC) for b-type binds. Second, a binding IC for the b-type implies a slack IC for the g-type since $v_g > v_b$ and thus $v_b < x_g v_g + \delta(1 - x_g)v_g$. The following off-equilibrium beliefs: $\mu(x) = 0$ for all $x < x_g$ and $\mu(x) = 1$ for all $x \geq x_g$ satisfy D1 and support this equilibrium. Single-crossing ensures that the LCSE is the unique equilibrium that satisfies D1 (see DeMarzo (2005)).

Origination stage: From the previous results, the payoffs associated with originating a g- and a b-type loan pool, respectively, are:

$$u_g^{LC} = (1 - x_g)v_g + \delta x_g v_g \tag{25}$$

$$u_b^{LC} = v_b \tag{26}$$

where $u_b^{LC} < 1 < u_g^{LC}$. Most importantly, these payoffs are independent of the actual lending standard chosen by banks at the Origination stage, p^*. Since there is a continuum of banks and lending standards are not observable by investors, deviations in individual lending standards do not impact Securitization stage payoffs. As a result, banks choose lending standard p^* to maximize their $t = 0$ value, and there are no profitable deviations at the Origination stage either. ■

A.4 Proofs for Section 4

Lemma A.2. *The solution to the following $\mathcal{M}(k)$ problem:*

$$\max_{\mu,x} \; u_g(x, \mu) \quad s.t. \quad u_b(x, \mu) = k \tag{27}$$

denoted by $\{\mu(k), x(k)\}$ is unique and characterized by the problem's first-order conditions. In addition, $\mu(k)$ is independent of k and $x(k)$ is decreasing in k.

Proof. We write expected Securitization stage payoffs as a function of retention level and investors beliefs as follows

$$u_t(x, \mu) = (1 - x)(\alpha_t(\mu)(v_g - v_b) + v_b) + \delta x v_t \tag{28}$$

where $\alpha_t(\mu) \equiv E_R[\mu_f(\mu, r)|t]$. Let $\alpha(\mu) \equiv \alpha_g(\mu) - \alpha_b(\mu)$ be the difference between expected posteriors for prior beliefs μ. It will be useful to re-state the $\mathcal{M}(k)$ problem as follows:

$$\max_{\mu,x} \; u_g(\mu, x) - k \quad s.t. \quad u_b(\mu, x) = k \tag{29}$$

By plugging in the corresponding expressions and the binding constraint, we obtain

$$\max_{\mu,x} \ (1-x)(\alpha(\mu)-\delta)(v_g-v_b)+\delta(v_g-v_b) \tag{30}$$

$$s.t. \quad (1-x)(\alpha_b(\mu)(v_g-v_b)+v_b)+\delta x v_b = k \tag{31}$$

Let $\{\mu(k), x(k)\}$ satisfy the problem's first-order conditions

$$\alpha(\mu)-\delta-\frac{\alpha'(\mu)}{\alpha_b'(\mu)}\alpha_b(\mu) = \frac{\alpha'(\mu)}{\alpha_b'(\mu)}\frac{(1-\delta)v_b}{v_g-v_b} \tag{32}$$

$$(1-x)(\alpha_b(\mu)(v_g-v_b)+v_b)+\delta x v_b = k \tag{33}$$

First, $\mu(k)$ is given by (32) and is thus independent of k. To see that $\mu(k)$ is unique, we analyze the left-hand side (LHS) and right-hand side (RHS) of (32) separately. From Facts A.1(3) and A.1(2), we have that the RHS is strictly decreasing in μ, positive for $\mu < \mu_{\max}$, zero for $\mu = \mu_{\max}$, and negative otherwise. Also from Fact A.1(3) we have that the LHS is strictly increasing in μ, and negative ($-\delta$) for $\mu = 0$. Therefore, if a solution to (32) exists, it is unique. Otherwise, the solution is given by the corner $\mu = 1$ (which by (33) implies a retention of x^{LC}). Finally, note that $x(k)$ is strictly decreasing in k and given by (33).

It remains to verify the second order conditions. We verify that the determinant of the Bordered Hessian is negative at our interior candidate $\{\mu(k), x(k)\}$:

$$BH = \begin{bmatrix} 0 & \frac{\partial u_b(x,\mu)}{\partial x} & \frac{\partial u_b(x,\mu)}{\partial \mu} \\ \frac{\partial u_b(x,\mu)}{\partial x} & L_{xx} & L_{x\mu} \\ \frac{\partial u_b(x,\mu)}{\partial \mu} & L_{\mu x} & L_{\mu\mu} \end{bmatrix}$$

where $L(x,\mu) = u_g(x,\mu) - \lambda\left(u_b(x,\mu)-k\right)$ where λ is the Lagrange multiplier.

$$\begin{aligned} L_{xx} &= 0 \\ L_{\mu\mu} &= \left(\alpha_g''(\mu) - \lambda\alpha_b''(\mu)\right)(1-x)(v_g-v_b) \\ L_{x\mu} &= L_{\mu x} = -(v_g-v_b)(\alpha_g'(\mu) - \lambda\alpha_b'(\mu)) = 0 \end{aligned}$$

A sufficient condition for our solution to be a local maximum is that the bordered Hessian is negative definite when evaluated at $\{x(k), \mu(k), \lambda(k)\}$, where $\lambda(k) = -\frac{\alpha'(\mu(k))}{\alpha_b'(\mu(k))}$. That is, we need $|BH_1| < 0$ and $|BH_2| > 0$. It is easy to see that $|BH_1| = -(\frac{\partial u_b(x,\mu)}{\partial x})^2 < 0$ and that $|BH_2| = -\left(\frac{\partial u_b(x,\mu)}{\partial x}\right)^2 L_{\mu\mu} > 0$ since $L_{\mu\mu}|_{\{x(k),\mu(k),\lambda(k)\}} < 0$ from $\frac{\partial}{\partial \mu}\left(\frac{\alpha_g'(\mu)}{\alpha_b'(\mu)}\right) < 0$. Thus, SOC are satisfied ■

Proof of Lemma 2. First see Lemma A.2 and its proof, and let $\tilde{\mu} = \mu(v_b)$ and $\tilde{x} = x(v_b)$. The

solution to (30) is interior with $\tilde{\mu} < 1$ (and thus $\tilde{x} < \bar{x}$) if and only if condition (32) holds for an interior μ, which requires that the RHS and the LHS of (32) intersect at $\mu < 1$. In the proof of Lemma A.2 we show that the LHS is negative and the RHS positive at $\mu = 0$, and that the LHS is strictly increasing while the RHS is strictly decreasing in μ for all $\mu \in (0, 1)$. Thus, a necessary and sufficient condition for $\tilde{\mu} < 1$ is that LHS>RHS at $\mu = 1$:

$$\alpha(1) - \delta - \frac{\alpha'(1)}{\alpha'_b(1)}\alpha_b(1) > \frac{\alpha'(1)}{\alpha'_b(1)}\frac{(1-\delta)v_b}{v_g - v_b} \tag{34}$$

$$-\delta - \frac{\alpha'_g(1) - \alpha'_b(\mu)}{\alpha'_b(1)} > \frac{\alpha'_g(1) - \alpha'_b(\mu)}{\alpha'_b(1)}\frac{(1-\delta)v_b}{v_g - v_b} \tag{35}$$

$$\iff \frac{\alpha'_b(1)}{\alpha'_g(1)} > \frac{v_g - \delta v_b}{(1-\delta)v_g} \tag{36}$$

Since $\alpha'_b(1) = E[\Gamma(R)|b]$ and $\alpha'_g(1) = E[\Gamma(R)|g] = 1$ the result follows. This condition is a statement about the slope of the indifference curves at the LCSE outcome $\{\mu = 1, \bar{x}\}$, and it states that an interior solution exists iff:

$$\left(\frac{\partial u_g(\mu, x)}{\partial x} \Big/ \frac{\partial u_g(\mu, x)}{\partial \mu}\right)|_{\mu=1,\bar{x}} > \left(\frac{\partial u_b(\mu, x)}{\partial x} \Big/ \frac{\partial u_b(\mu, x)}{\partial \mu}\right)|_{\mu=1,\bar{x}}$$

That is, if the slope of the indifference curve at the LCSE outcome is steeper for the g-type than it is for the b-type, breaking the single-crossing condition necessary for separation. ■

Proof of Proposition 3. Let $\{\mu(k), x(k)\}$ be the solution to the constrained maximization problem $\mathcal{M}(k)$ in Lemma A.2. From that same Lemma, this solution is unique, $\mu(k)$ is constant and $x(k)$ is continuous and decreasing in $k \in [v_g, v_g]$. Let $\tilde{\mu} = \mu(v_b)$ and $\tilde{x} = x(v_b)$. By Lemma 2, the equilibrium is not separating if ratings are sufficiently informative in the following sense: $E[\Gamma(R)|b] > \frac{v_g - \delta v_b}{(1-\delta)v_g}$.

The next step is to show that the equilibrium proposed in Proposition 3 is a PBE that satisfies D1. Before doing so, we introduce the following definitions. Let S_t be the support of the t-type's strategy. In the proposed unique equilibrium, the good type plays a pure strategy, denoted it x_g, so $S_g = \{x_g\}$, while the bad type could mix, and thus $S_b \subseteq \{0, x_g\}$. For completeness, we must specify the off-path beliefs: $\mu(x) = 0$ for all $x \neq x_g$.

Verifying that the proposed profile is a PBE is straightforward. To see that it satisfies D1, fix a μ_0 and consider the Proposition's unique equilibrium candidate. Denote the good type's equilibrium payoff u_g^e and the bad type's equilibrium payoff k, so $x_g = \tilde{x}(k)$. Let x be an arbitrary retention level in $[0, 1]$ such that $x \neq \tilde{x}(k)$. First, if $B_b(x, k) = [0, 1]$, then the low type could deviate to x and obtain a payoff strictly greater than k, regardless of $\mu(x)$, breaking the PBE. Hence, either $b_b(x, k) \in (0, 1]$ exits or $u_b(x, 1) < k$. If $b_b(x, k)$ exits, then since $\{\tilde{x}(k), \tilde{\mu}(k)\}$

is the unique solution to (27), $u_g(x, b_b(x,k)) < u_g(\tilde{x}(k), \tilde{\mu}(k)) = u_g^e$. By Fact A.1(5) then, $b_g(x, u_g^e) > b_b(x,k)$ implying $B_g(x, u_g^e) \subseteq B_b(x,k)$. So, $\mu(x) = 0$ is consistent with D1. If instead $u_b(x,1) < k$ (so $B_b(x,k) = \emptyset$), then there exists a unique $\epsilon > 0$ such that $u_b(x-\epsilon, 1) = k$. Since $\{\tilde{x}(k), \tilde{\mu}(k)\}$ solves (27), $u_g(\tilde{x}(k), \tilde{\mu}(k)) \geq u_g(x-\epsilon, 1) > u_g(x,1)$. Hence, $B_g(x, u_h^e) = \emptyset$ as well, and D1 places no restriction on $\mu(x)$.

We now establish uniqueness. Fix an equilibrium with $u_g = u_g^e$ and $u_b = k$. Since banks with bad loan pools have the option to choose the same retention as banks with good loan pools, $u_b(x, \mu(x)) \leq k$ for all $x \in S_g$. Fix now $x \in S_g$ and suppose that $u_b(x, \mu(x)) < k$. Then $x \notin S_b$, so $\mu(x) = 1 = b_g(x, u_g^e)$ and $B_b(x,k) = \emptyset$. Further, it must be that $x \neq 0$ since $u_b(0,1) = v_g > k$. Then for $\epsilon > 0$ small enough $b_g(x-\epsilon, u_g^e) \in (0,1)$ and $B_b(x-\epsilon, k) = \emptyset$. Therefore, $x - \epsilon \notin S_b$ and $\mu(x-\epsilon) = 1$ (by belief consistency if $x - \epsilon \in S_g$, by D1 if not). Since $u_g(x-\epsilon, 1) > u_g(x,1) = u_g^e$, the high type would gain by deviating to $x - \epsilon$, breaking the equilibrium. Therefore, $u_b(x, \mu(x)) = k$, or equivalently $\mu(x) = b_b(x,k)$, for all $x \in S_g$.

Suppose now there exists $x \in S_g$ such that $x \neq \tilde{x}(k)$. Then

$$u_g(x, \mu(x)) = u_g(x, b_b(x,k)) < u_g(\tilde{x}(k), \tilde{\mu}(k)) = u_g(\tilde{x}(k), b_b(\tilde{x}(k), k)),$$

and thus $b_g(\tilde{x}(k), u_h^e) < \tilde{\mu}(k) = b_b(\tilde{x}(k), k)$. D1 then implies that $\mu(\tilde{x}(k)) = 1$, meaning that deviating to $\tilde{x}(k)$ is profitable for the high type and breaking the equilibrium. Hence, if the b-type's equilibrium payoff is k, then $S_g = \{\tilde{x}(k)\}$ and $\mu(\tilde{x}(k)) = \tilde{\mu}(k)$.

Further, if the b-type selects $x \notin S_g \cup \{0\}$, then $\mu(x) = 0$, and $u_b(x,0) < u_b(0,0) \leq u_b(0, \mu(0))$ for any value of $\mu(0)$. It could therefore profitably deviate to $x = 0$. Hence, $S_b \subseteq S_g \cup \{0\}$.

The final step is to characterize which values of $u_b = k$ are consistent with equilibrium, which depends on the prior, μ_0. Recall that $\tilde{\mu}$ does not vary with k. First, let $\mu_0 < \tilde{\mu}$, and let $u_b = k$. Therefore, $S_g = \{\tilde{x}(k)\}$ and $\mu(\tilde{x}(k)) = \tilde{\mu} > \mu_0$. For this belief to be consistent with seller strategies, $S_b \neq \{\tilde{x}(k)\}$. Hence, $S_b = \{\tilde{x}(k), 0\}$ and $k = v_b$. The precise mixing probabilities given in the proposition are required for the Bayesian consistency: $\mu(\tilde{x}(v_b)) = \tilde{\mu}(v_b)$. Second, let $\mu_0 \geq \tilde{\mu}$. Then $S_g = S_b = \{\tilde{x}(k)\}$ is consistent with $\tilde{\mu}(\tilde{x}(k)) = \tilde{\mu}(k) = \mu_0$, as stated in the Proposition. ■

Proof of Proposition 4. From Proposition 3, Corollary 3, and the equilibrium belief-consistency requirement (Corollary 1(3)), an equilibrium is pinned down by μ_0^* such that $A^{-1}(\mu_0^*) \in \Psi(\mu_0^*)$. First, note that A^{-1} is continuous and strictly increasing, with $A^{-1}(\xi) = 0$ and $A^{-1}(1) = 1$. Second, $\Psi(\mu_0)$ is constant for $\mu_0 < \tilde{\mu}$ at $\bar{p} \equiv \frac{1-v_b}{u_g(\tilde{x},\tilde{\mu}) - v_b} > p^{FB}$. At $\tilde{\mu}$, $\Psi(\tilde{\mu}) = [\frac{1-u_b(0,\tilde{\mu})}{u_g(0,\tilde{\mu}) - u_b(0,\tilde{\mu})}, \bar{p}]$. And for $\mu_0 > \tilde{\mu}$, $\Psi(\mu_0) = \tilde{\Psi}(\mu_0) = \max\left\{\frac{1-u_b(0,\mu_0)}{u_g(0,\mu_0) - u_b(0,\mu_0)}, 0\right\}$ (Definition A.2). $\tilde{\Psi}$ continuous and strictly decreasing at all $\mu_0 \in (0,1)$ (Lemma A.1) then guarantees existence and uniqueness of μ_0^* (and therefore equilibrium). ■

Proof of Proposition 5. Let $\{\psi_b^n, \psi_g^n\}_{n=1}^{\infty}$ limit to perfect screening. From (1), p is a monotone, bijective function of the likelihood ratio $L(s)$. By Definition 2 then, as $n \to \infty$,

$$H^n(d) - H^n(c) = \Pr(p \in (c, d]) \to 0$$

for all $0 < c < d < 1$. Therefore, $H^n(p)$ limits to some constant K for all $p \in (0, 1)$. For any screening technology, $E[p] = \xi$ by the Law of Iterated Expectation, implying K must be $1 - \xi$.

Next, by definition, for any screening technology $A(1) = 1$. Fix $p \in (0, 1)$, and $a \in (p, 1)$.

$$\begin{aligned} A^n(p) = E[p'|p' \geq p] &\geq \Pr(p' \in [p, a)|p' \geq p) \cdot p + \Pr(p' \geq a|p' \geq p) \cdot a \\ &= \frac{H^n(a) - H^n(p)}{1 - H^n(p)} \cdot p + \frac{1 - H^n(a)}{1 - H^n(p)} \cdot a. \end{aligned} \tag{37}$$

Since $H^n(p) \to 1 - \xi$ for all $p \in (0, 1)$, the expression in (37) limits to a as $n \to \infty$. Therefore, as $n \to \infty$, $A^n(p)$ grows at least as large as any $a \in (p, 1)$. Since, $A^n(p) \leq 1$ for all n, $A^n(p) \to 1$ for all $p > 0$. Equivalently, $(A^n)^{-1}(\mu_0) \to 0$ for all $\mu_0 \in (\xi, 1)$. Hence, the single intersection of $(A^n)^{-1}$ and Ψ, $(\mu_0^{*,n}, p^{*,n}) \to (\bar{\mu}, 0)$. In the limit,then, $\Pr(t = g|origination) = \frac{\xi}{Q(p^*)}$ which must equal $\bar{\mu}$. Hence, $Q(p^*)$ limits to $\frac{\xi}{\bar{\mu}}$, which must also equal $\xi + q_b$, where q_b is the measure of bad loans originated in the limit. Hence, $q_b = \frac{\xi(1-\bar{\mu})}{\bar{\mu}}$. ■

Proof of Lemma 3. For statement (1) in the lemma, we analyze how $\{\tilde{x}, \tilde{\mu}\}$ from Proposition 3 change with rating informativeness γ. After some algebra, we have that $\{\tilde{x}, \tilde{\mu}\}$ solve

$$\frac{(1-\delta)\, v_b}{v_g - v_b} = (\alpha(\tilde{\mu}) - \delta) \frac{\alpha_b'(\tilde{\mu})}{\alpha'(\tilde{\mu})} - \alpha_b(\tilde{\mu}) \tag{38}$$

$$\tilde{x} = \frac{\alpha_b(\tilde{\mu})(v_g - v_b)}{\alpha_b(\tilde{\mu})(v_g - v_b) + (1-\delta)v_b} \tag{39}$$

We proceed to characterize how this solution changes with γ. We have that:

$$\alpha_g(\mu) = \frac{\mu\gamma^2}{\mu\gamma + (1-\mu)(1-\gamma)} + \frac{\mu(1-\gamma)^2}{\mu(1-\gamma) + (1-\mu)\gamma} \tag{40}$$

$$\alpha_b(\mu) = \frac{\mu\gamma(1-\gamma)}{\mu\gamma + (1-\mu)(1-\gamma)} + \frac{\mu\gamma(1-\gamma)}{\mu(1-\gamma) + (1-\mu)\gamma} \tag{41}$$

$$\alpha(\mu) = \frac{(2\gamma-1)^2\mu(1-\mu)}{(\mu\gamma + (1-\mu)(1-\gamma))(\mu(1-\gamma) + (1-\mu)\gamma)} \tag{42}$$

After some algebra, from the RHS of condition (38) we have that:

$$\frac{\partial RHS}{\partial \gamma}|_{\mu=\tilde{\mu}} = \frac{\delta}{(2\gamma-1)^3(1-2\tilde{\mu})}$$
$$\frac{\partial RHS}{\partial \mu}|_{\mu=\tilde{\mu}} = -\frac{\alpha(\tilde{\mu})-\delta}{\alpha(\tilde{\mu})^2}\frac{\partial}{\partial \mu}\left(\frac{\alpha_g'(\tilde{\mu})}{\alpha_b'(\tilde{\mu})}\right)$$

If $\alpha(\tilde{\mu})-\delta<0$, then from FOC the solution requires $\alpha'(\tilde{\mu})<0 \iff \tilde{\mu} > \arg\max_\mu \alpha(\mu) = \frac{1}{2}$ (see Fact 2). As a result, $\frac{\partial RHS}{\partial \gamma} < 0$ and $\frac{\partial RHS}{\partial \tilde{\mu}} < 0$. Otherwise, $\alpha(\tilde{\mu})-\delta > 0$, which requires $\alpha'(\tilde{\mu}) > 0$, that is, $\tilde{\mu} < \frac{1}{2}$. Thus, $\frac{\partial RHS}{\partial \gamma} > 0$ with $\frac{\partial RHS}{\partial \tilde{\mu}} > 0$. If follows that as γ increases, $\tilde{\mu}$ has to decrease.

We have established that $\tilde{\mu}$ decreases in γ. It remains to characterize how $\tilde{x}$ changes in γ. Let RHS_c denote the left-hand side of the constraint (39). Then, we have that:

$$\frac{d\tilde{x}}{d\gamma} = \frac{\partial RHS_c}{\partial \tilde{\mu}}\frac{d\tilde{\mu}}{d\gamma} + \frac{\partial RHS_c}{\partial \gamma} < 0 \tag{43}$$

Where the results follows from (i) $\frac{\partial \alpha_b(\mu)}{\partial \gamma} < 0$ for all $\mu \in (0,1)$; (ii) $\alpha_b'(\cdot) > 0$ for all $\gamma \in (\frac{1}{2}, 1)$; and (iii) RHS_c being increasing in $\alpha_b(\mu)$ for all $\mu \in (0,1)$ since $v_b > 0$, which are all easy to check.

For statement (2) in the lemma, let $\tilde{\mu}'$ denote the $\tilde{\mu}$ after the increase in γ, which from (i) implies $\tilde{\mu}' < \tilde{\mu}$. First consider the case where $p^{FB} < \tilde{\mu}$. For all $\mu_0 > \tilde{\mu}, \Psi(\mu_0) = \tilde{\Psi}(\mu_0)$, which by Lemma A.1 is increasing in γ since $\mu_0 > p^{FB}$. For $\mu_0 \in [\tilde{\mu}', \tilde{\mu}]$, $\Delta\Psi(\mu_0) = \tilde{\Psi}(\mu_0) - \frac{1-v_b}{u_g(\tilde{x},\tilde{\mu})-v_b} < 0$. For $\mu_0 < \tilde{\mu}'$, $\Psi(\mu_0) = \frac{1-v_b}{u_g(\tilde{x},\tilde{\mu})-v_b}$ which decreases in γ since $u_g(\tilde{x},\tilde{\mu})$ increases in γ. Now consider the case with $p^{FB} > \tilde{\mu}$. As before, for all $\mu_0 > p^{FB}$, we know that $\Psi(\mu_0) = \tilde{\Psi}(\mu_0)$ which increases in μ_0. For $\mu_0 < p^{FB}$, if $\mu_0 \in [\tilde{\mu}, p^{FB})$, $\Psi(\mu_0) = \tilde{\Psi}(\mu_0)$, which by Lemma A.1 is now decreasing in γ since $\mu_0 - p^{FB} < 0$. For $\mu_0 < \tilde{\mu}$, we have already shown that Ψ decreases in γ. ■

Proof of Proposition 6. By definition, $A^{-1}(\mu_0) < \mu_0$ for all $\mu_0 \in [\xi, 1)$. Suppose that $p^* \geq p^{FB}$. By Lemma A.1(1)-(2) then, $p^* \neq \tilde{\Psi}(\mu_0^*)$. From the structure of Ψ in Corollary 3 it then follows that $\tilde{\mu} > p^{FB}$ and $\mu_0^* \in (p^{FB}, \tilde{\mu}]$. An increase in γ implies a decrease in both $\tilde{\mu}$ and $\Psi(\mu_0)$ for $\mu_0 < \tilde{\mu}$ (Lemma 3), and therefore a decrease in the point of intersection of Ψ and the strictly increasing A^{-1} function, meaning a lower p^*. ■

Proof of Proposition 7. Let μ_1 be the unique solution to $A^{-1}(\mu_1) = p^{FB}$. By definition, $A^{-1}(\mu_0) < \mu_0$ for all $\mu_0 \in [\xi, 1)$, meaning $\mu_1 > p^{FB}$. Suppose now that $\tilde{\mu} < \mu_1$. We claim that $p^* < p^{FB}$ is implied. To see this, recall that (μ_0^*, p^*) is the unique intersection of A^{-1} and Ψ. From Corollary 3 and A^{-1} strictly increasing, $\Psi(\mu_0) > p^{FB} > A^{-1}(\mu_0)$ for all $\mu_0 < \tilde{\mu} < \mu_1$. So $\mu_0^* \geq \tilde{\mu}$. If $\mu_0^* = \tilde{\mu}$, then $p^* = A^{-1}(\tilde{\mu}) < A^{-1}(\mu_1) = p^{FB}$. If, instead, $\mu_0^* > \tilde{\mu}$, then by from Corollary 3, $\Psi(\mu_0^*) = \tilde{\Psi}(\mu_0^*) < p^{FB}$, where the inequality follows from Lemma A.1(1)-(2).

From Lemma 3, $\tilde{\mu}_\gamma$ is decreasing in γ. Hence, it is sufficient to show that $\lim_{\gamma\to 1}\tilde{\mu}_\gamma \leq p^{FB}$ if $v_g v_b \leq 1$. This is a matter of direct calculation. With binary-symmetric ratings, one can use (38) to obtain the closed-form expression:

$$\tilde{\mu} = \frac{\frac{\sqrt{(\delta-1)(1-2\gamma)^2(\delta(\gamma v_b+(\gamma-1)v_g)((\gamma-1)v_b+\gamma v_g)-(1-2\gamma)^2 v_b v_g)}}{(\delta-1)(1-2\gamma)^2}+v_b}{v_b-v_g} \underbrace{\to \frac{v_b}{v_b+\sqrt{v_b v_g}}}_{\gamma\to 1} \tag{44}$$

Immediately, the limit value is no greater than $p^{FB}=\frac{1-v_b}{v_g-v_b}$ if $v_g v_b \leq 1$. ■

Proof of Proposition 8. Restating (39):

$$\tilde{x} = \frac{\alpha_b(\tilde{\mu})(v_g-v_b)}{\alpha_b(\tilde{\mu})(v_g-v_b)+(1-\delta)v_b}$$

From (41), as $\gamma\to 1$, $\alpha_b(\mu)\to 0$ for all $\mu\in(0,1)$. Hence, $\tilde{x}\to 0$ as $\gamma\to 0$.

In addition, $u_t(0,\mu_0) = \mathbb{E}_R[P(0,R)|t] = v_b + \mathbb{E}_R[\mu_f(x,r)|t](v_g-v_b)$. Finally, as $\gamma\to 1$, $\mathbb{E}_R[\mu_f(x,r)|t]\to \mathbb{I}_{t=g}$, and $u_t(0,\mu_0)\to v_t$. Hence, for all μ_0, $u_t^*(\mu_0)\to v_t$ and $\Psi(\mu_0)\to p^{FB}$, as $\gamma\to 1$. Since (μ_0^*,p^*) is the unique intersection of Ψ and A^{-1}, $p^*\to p^{FB}$. ■

A.5 Proofs for Section 5

Proof of Proposition 9. The skin-in-the-game rule requires all securitizers to retain at least a fraction x_s. As a result, $\{\tilde{x},\tilde{\mu}\}$ from Proposition 3 are now given by the solution to:

$$\begin{aligned} &\max_{\mu,x}(1-x)(\alpha_g(\mu)(v_g-v_b)+v_b)+\delta x v_g \\ &s.t.(1-x)(\alpha_g(\mu)(v_g-v_b)+v_b)+\delta x v_g = (1-x_s+\delta x_s)\,v_b \end{aligned} \tag{45}$$

where the only adjustment has been a change in the outside option (full information payoff) of the banks with b-type pools in the constraint. From Proposition 3, we know that the solution to (45) fully characterizes the PBE of the securitization stage with skin-in-the-game.

From the constraint, it follows that $\tilde{x}\geq x_s$. Therefore, equilibrium retention levels satisfy: $x_t\in[x_s,\tilde{x}]$. We have shown that $\{\tilde{x},\tilde{\mu}\}$ are given by the problem's FOC. In particular, $\tilde{\mu}$ continues to be determined by condition (32), while $\tilde{x}$ is determined by the new constraint (45). As a result, $\tilde{\mu}$ is unaffected by the skin-in-the-game rule. In contrast, from the constraint, we know $\tilde{x}$ responds to the retention rule as follows. For $\mu>\tilde{\mu}$, equilibrium retention increases from 0 to x_s. For $\mu=\tilde{\mu}$, retentions in the range $x\in[\tilde{x},x_s]$ can be D1-equilibria. Finally, for $\mu<\tilde{\mu}$, there is partial pooling at the new (higher) $\tilde{x}$, where banks with g-type pools retain $\tilde{x}$

and those with b-type pools mix between $\{\tilde{x}, x_s\}$ as described in Proposition 3. Since payoffs at the securitization stage, u_t^*, decrease with the skin-in-the-game rule, the lending standard weakly increases, credit supply weakly decreases, and so does ex-ante efficiency.

A marginal increase in the retention rule around $x_s = 0$ increases overall efficiency (given by 9) if

$$\left[-\left(p^*(u_g^*+s_g)+(1-p^*)(u_b^*+s_b)-1\right)h(p^*)+\int_{p^*}^1\left(p\frac{\partial u_g^*}{\partial \mu_0^*}+(1-p)\frac{\partial u_b^*}{\partial \mu_0^*}\right)A'(p^*)dH(p)\right]\frac{\partial p^*}{\partial x_s}|_{x_s=0} > -\int_{p^*}^1\left(p\frac{\partial u_g^*}{\partial x_s}+(1-p)\frac{\partial u_b^*}{\partial x_s}\right)|_{x_s=0}dH(p) \quad (46)$$

$$\iff \left[-\left(p^*s_g+(1-p^*)s_b\right)h(p^*)+\left(\mu_0^*\frac{\partial u_g^*}{\partial \mu_0^*}+(1-\mu_0^*)\frac{\partial u_b^*}{\partial \mu_0^*}\right)A'(p^*)Q(p^*)\right]\frac{\partial p^*}{\partial x_s}|_{x_s=0} > -\left(\mu_0^*\frac{\partial u_g^*}{\partial x_s}+(1-\mu_0^*)\frac{\partial u_b^*}{\partial x_s}\right)|_{x_s=0}Q(p^*) \quad (47)$$

which is obtained by differentiating (9) with respect to x_s around $x_s = 0$, and where we have used the fact that in any equilibrium, lending standards are chosen so that the marginal originated loan has zero NPV: $p^*u_g^* + (1-p^*)u_b^* = 1$.

In an OTD equilibrium, (47) becomes (12). To see this, we first compute the marginal gain from increasing the lending standard by differentiating our efficiency measure in an OTD equilibrium with respect to p^*. Note that (10) can be written as:

$$\int_{p^*}^1\left(p(v_g+s_g)+(1-p)(v_b+s_b)-1\right)dH(p) \quad (48)$$

Which implies that the marginal gain from increasing p^* in an OTD equilibrium is:

$$-\left(p^*(v_g+s_g)+(1-p^*)(v_b+s_b)\right)h(p^*) \quad (49)$$

On the other hand, a marginal increase in retention levels would decrease efficiency by the expected increase in the cost of retention: $(1-\delta)(\mu_0^*\mu_g^*+(1-\mu_0^*)\mu_b^*)Q(p^*) = (1-\delta)(\mu_0^*v_g + (1-\mu_0^*)v_b)Q(p^*)$. Thus, there exists an x_s that increases efficiency in the OTD equilibrium if the marginal gain from increasing retention levels more than compensates for the cost of retention:

$$-\left(p^*(v_g+s_g)+(1-p^*)(v_b+s_b)\right)h(p^*)\frac{\partial p^*}{\partial x_s}|_{x_s=0} > (1-\delta)(\mu_0^*v_g+(1-\mu_0^*)v_b)Q(p^*) \quad (50)$$

In a signaling equilibrium, since u_t^* are independent of μ_0^*, condition (47) becomes:

$$-\left(p^*s_g+(1-p^*)s_b\right)h(p^*)\frac{\partial p^*}{\partial x_s}|_{x_s=0} > -\left(\mu_0^*\frac{\partial u_g^*}{\partial x_s}+(1-\mu_0^*)\frac{\partial u_b^*}{\partial x_s}\right)Q(p^*) \quad (51)$$

Since retention is costly, and we have shown that x_s increases retention levels type-by-type in a signaling equilibrium, the RHS of condition (51) is positive.

($\Leftarrow$) It is straightforward that if externalities at the margin, $p^* s_g + (1-p^*)s_b$, are sufficiently negative, condition (51) holds, and thus there exists a skin-in-the-game rule that can increase overall efficiency. ($\Rightarrow$) If there is no retention rule x_s that increases efficiency, it must be that the cost of increased retention is always higher than the gains from increasing the lending standard. This means there is a bound (B) on the gains from increasing the lending standard from p^* to the retention level $p^*_{x_s}$ implied by retention rule x_s:

$$-\int_{p^*}^{p^*_{x_s}} \left(ps_g + (1-p)s_b\right) dH(p) < B_{x_s}, \quad \forall x_s \in (0,1]$$

Which implies a bound for how negative externalities can be at the margin, p^*, since $h(p) > 0$ for all p and $\left(ps_g + (1-p)s_b\right) h(p)$ is continuous in p. ■

Proof of Proposition 10. First, consider the economy in an *OTD equilibrium*: there is full pooling at zero retention and $\mu_0^* \geq \tilde{\mu}$. Total surplus is given by

$$TS = \int_{p^*}^{1} \left(pv_g + (1-p)\, v_b - 1\right) dH(p).$$

A marginal increase in informativeness, γ, has no effect on retention, so

$$\frac{dTS}{d\gamma} = \frac{dTS}{dp^*}\frac{dp^*}{d\gamma} = -\left(p^* v_g + (1-p^*)\, v_b - 1\right) h\left(p^*\right) \frac{dp^*}{d\gamma}.$$

To establish that $\frac{dTS}{d\gamma} > 0$, we have (i) $p^* < p^{FB}$ in an *OTD equilibrium* (Proposition 1), which implies $p^* v_g + (1-p^*)\, v_b - 1 < 0$ and (ii) $h(p^*) > 0$. Finally, $\frac{dp^*}{d\gamma} > 0$ in an *OTD equilibrium* since Lemma 3 establishes that the relevant range of Ψ increases with γ and A^{-1} is an increasing function.

Second, consider the economy in a *signaling equilibrium*: there is partial pooling at retention level $\tilde{x}$ and $\mu_0^* \leq \tilde{\mu}$. Total surplus is given by

$$TS = \int_{p^*}^{1} \left(pu_g\left(\tilde{x}, \tilde{\mu}\right) + (1-p)\, v_b - 1\right) dH(p)$$

A marginal increase in informativeness, γ, affects both retention (and therefore u_g) and the lending standard, so

$$\frac{dTS}{d\gamma} = \frac{dTS}{dp^*}\frac{dp^*}{d\gamma} + \frac{dTS}{du_g}\frac{du_g\left(\tilde{x}, \tilde{\mu}\right)}{d\gamma}.$$

The first term,

$$\frac{dTS}{dp^*}\frac{dp^*}{d\gamma} = -\underbrace{\left(p^* u_g\left(\tilde{x},\tilde{\mu}\right) + \left(1-p^*\right) v_b - 1\right)}_{=0} h\left(p^*\right)\frac{dp^*}{d\gamma} = 0,$$

since p^* is optimal for banks, given retention. The second term,

$$\frac{dTS}{du_g}\frac{du_g\left(\tilde{x},\tilde{\mu}\right)}{d\gamma} = \int_{p^*}^{1} p\frac{du_g\left(\tilde{x},\tilde{\mu}\right)}{d\gamma} dH(p).$$

Since $(\tilde{x},\tilde{\mu})$ is the maximizer of (7), $\frac{du_g(\tilde{x},\tilde{\mu})}{d\gamma} = \frac{\partial u_g(\tilde{x},\tilde{\mu})}{\partial\gamma} > 0$, as the g-type gains from a more informative rating for fixed (x,μ). Hence, $\frac{dTS}{d\gamma} > 0$. ■

Proof of Proposition 11. We analyze how $\{\tilde{x},\tilde{\mu}\}$ from Proposition 3 change with δ. We know that $\tilde{\mu}$ satisfies condition (38). Let $RHS\left(\tilde{\mu},\delta\right)$ denote the right-hand side of this condition. We have shown in the proof of Lemma 3 that $\frac{\partial RHS}{\partial\tilde{\mu}}$ takes the sign of $\delta - \alpha\left(\tilde{\mu}\right)$. In addition, we have

$$\frac{\partial RHS}{\partial\delta} = \frac{1}{[...]^2}\frac{1}{\alpha'\left(\tilde{\mu}\right)}\left[\left(1-\alpha_g\left(\tilde{\mu}\right)\right)\alpha_b'\left(\tilde{\mu}\right) + \alpha_g'\left(\tilde{\mu}\right)\alpha_b\left(\tilde{\mu}\right)\right]$$

which takes the sign of $\alpha'\left(\tilde{\mu}\right)$. Since the left-hand-side of the constraint is positive, so has to be the the RHS, which requires $\alpha'\left(\tilde{\mu}\right)\times\left(\alpha\left(\tilde{\mu}\right)-\delta\right)\geq 0$. Thus, we have that $\frac{\partial\tilde{\mu}}{\partial\delta}\geq 0$. To see the effect on retention, we do a total differentiation of the constraint:

$$\begin{aligned}\left[-\alpha_b\left(\tilde{\mu}\right)\left(v_g-v_b\right) - \left(1-\delta\right)v_b\right]d\tilde{x} + \left[\left(1-x\right)\alpha_b'\left(\tilde{\mu}\right)\left(v_g-v_b\right)\frac{d\tilde{\mu}}{d\delta} + \tilde{x}v_b\right]d\delta &= 0\\ \frac{d\tilde{x}}{d\delta} = \frac{\left[\left(1-x\right)\alpha_b'\left(\tilde{\mu}\right)\left(v_g-v_b\right)\frac{d\tilde{\mu}}{d\delta} + \tilde{x}v_b\right]}{\left[\alpha_b\left(\tilde{\mu}\right)\left(v_g-v_b\right) + \left(1-\delta\right)v_b\right]} &> 0\end{aligned}$$

Finally, to study the effect of a change in δ on the lending standard and credit supply, we characterize the changes in Ψ. We know that $\tilde{\mu}$ has increased, and since δ does not affect payoffs when retention is zero, $\tilde{\Psi}$ remains unaffected. For $\mu_0 < \tilde{\mu}$, we need to analyze the effect of δ on $u_g\left(\tilde{x},\tilde{\mu}\right)$:

$$\begin{aligned}u_g^* \equiv u_g\left(\tilde{x},\tilde{\mu}\right) = \max_{\mu,x}\left(1-x\right)\left[\alpha_g\left(\mu\right)\left(v_g-v_b\right)+v_b\right] + \delta x v_g\\ s.t.\left(1-x\right)\left[\alpha_b\left(\mu\right)\left(v_g-v_b\right)+v_b\right] + \delta x v_b = v_b\end{aligned}$$

We have that

$$\frac{\partial u_g^*}{\partial\delta} = v_g - \frac{\alpha_g'(\tilde{\mu})}{\alpha_b'(\tilde{\mu})}v_b \tag{52}$$

since $\frac{\alpha_g'(\tilde{\mu})}{\alpha_b'(\tilde{\mu})}$ is the Lagrange multiplier of the constraint in this problem at $\{\tilde{\mu}, \tilde{x}\}$.

Therefore, to complete the proof it suffices to show that $\frac{\alpha_g'(\tilde{\mu})}{\alpha_b'(\tilde{\mu})} < \frac{v_g}{v_b}$. To see this, first rewrite the FOC of $\mathcal{M}(v_b)$ as

$$\frac{\alpha_g'(\tilde{\mu})}{\alpha_b'(\tilde{\mu})} = \frac{\alpha_g(\tilde{\mu})v_g + (1-\alpha_g(\tilde{\mu}))v_b - \delta v_g}{\alpha_b(\tilde{\mu})v_g + (1-\alpha_b(\tilde{\mu}))v_b - \delta v_b} \tag{53}$$

and observe that the numerator on the RHS is increasing in α_g, while the denominator is increasing in α_b, therefore

$$\frac{\alpha_g(\tilde{\mu})v_g + (1-\alpha_g(\tilde{\mu}))v_b - \delta v_g}{\alpha_b(\tilde{\mu})v_g + (1-\alpha_b(\tilde{\mu}))v_b - \delta v_b} < \frac{\alpha_g(1)v_g + (1-\alpha_g(1))v_b - \delta v_g}{\alpha_b(0)v_g + (1-\alpha_b(0))v_b - \delta v_b} = \frac{v_g}{v_b}.$$

Thus, the lending standard increases in the signaling equilibrium. In what follows, we analyze the effect of a marginal increase in δ on overall efficiency.

In an OTD equilibrium with $\mu_0^* > \tilde{\mu}$, after a marginal increase in δ we continue to have $\mu_{new}^* > \tilde{\mu}$, by the continuity of the RHS of condition (38) in μ_0. Thus, the economy continues to be in an OTD equilibrium. From (10), it follows that changes in δ do not affect efficiency, since retention is zero in such an equilibrium.

In a signaling equilibrium, with the same argument, after a marginal increase in δ, the economy moves to a new signaling equilibrium, i.e., $\mu_{new}^* < \tilde{\mu}$. To see that efficiency has increased, note that (i) the lending standard has increased, (ii) u_g^* has increased, as shown above, and (iii) $u_b^* = v_b$ remains unchanged. From (11) it follows that, absent externalities, efficiency has increased. ■

A.6 Proofs for Section 6

Proof of Proposition 12. Consider a candidate equilibrium in which $\beta = 0$. If $m < \Delta$, then it is strictly profitable for b-type banks to manipulate, a contradiction. Now suppose that $\beta = 1$. In this case, recognizing that ratings will be uninformative, g-type banks will choose $x_g = \bar{x}$ in the securitization stage (see Proposition 2). Therefore, the equilibrium payoff to b-banks who manipulate is no greater than $v_b - m$. In this case, a b-type can profitably deviate by not manipulating and setting $x_b = 0$ in the securitization stage. ■

Proof of Proposition 13. Fix any belief assigned to an unrated issuance, μ_n. Define $\bar{r}$ be such that

$$\frac{\mu}{\mu + (1-\mu)\Gamma(\bar{r})} - \mu_n = \phi,$$

where if $\bar{r}$ exists it is unique since the left-hand side is strictly increasing in r for $\mu \in (0,1)$. If ϕ is large enough so that an intersection does not occur, then we say $\bar{r} = \inf r$. Suppose

$r < \bar{r}$ and $r \in \mathcal{R}$. Since $\mu_f(\mu, r; \phi)$ is increasing in r, $u_t(\mu, x, r) < u_t^n(\mu, n, r)$, which violates bank's optimality with respect to publishing the rating. Now suppose there exists $r > \bar{r}$ such that $r \notin \mathcal{R}$. Then, if a bank receives rating r and chooses to pay and report it, investors assign beliefs $\mu_f(\mu, r; \phi)$ and the bank profits from this deviation, a contradiction. Thus, it must be that $\mathcal{R} = \{r : r \geq \bar{r}\}$.

Then, in any equilibrium, not being rated indicates that $r < \bar{r}(\mu; \phi)$. As a result:

$$\mu_n(\mu; \phi) = \frac{\mu}{\mu + (1-\mu)\frac{F_b(\bar{r}(\mu;\phi))}{F_g(\bar{r}(\mu;\phi)))}}.$$

It remains to show that the set $\mathcal{R}$ differs from the full set of ratings iff $\phi > 0$. ($\Leftarrow$) Let the set $\{r : \ r < \bar{r}\}$ have zero measure. Then, $\bar{r} = \inf r$ and $\mu_f(\mu; \phi) - \mu_n(\mu; \phi) = 0$ since by L'Hopital,

$$\lim_{r \to \inf r} \frac{F_b(r)}{F_g(r)} = \frac{f_b(r)}{f_g(r)}.$$

This is only consistent with $\phi = 0$. ($\Rightarrow$) Now assume that $\phi = 0$. Then, $\bar{r}$ is given by: $\mu_f(\mu, \bar{r}; \phi) - \mu_n(\mu; \phi) = 0$, which implies $\bar{r} = \inf r$ since $\mu_f(\mu, r; \phi) - \mu_n(\mu; \phi) > 0$ for $\mu \in (0, 1)$ and all $r > \inf r$. Thus, $\{r : \ r < \bar{r}(\mu; \phi)\}$ has zero measure. ■

B Security Design

Let the underlying cash flow be a continuous random variable Y, with type-dependent density functions π_H, π_L satisfying the monotone likelihood ratio property (i.e., $\pi_H(y)/\pi_L(y)$ is increasing in y.). Thus far, we have studied how much banks will retain taking the "class" of securities, $F = (1-x)Y$, as given. In this section, we demonstrate that the main results of the paper remain unchanged when banks can choose the design of the security. Our demonstration relies heavily on Daley et al. (2016) (henceforth, DGV16) in which we study the optimal security design in the presence of public information (e.g., ratings).

B.1 Summary

In DGV16, we characterize the equilibrium of the securitization stage where the securitizer can choose any *security*, $F = \psi(Y)$, to offer for sale. Specifically, for any realization of the cash flow y, $\psi(y)$ is the amount paid to the purchaser of the security and $y - \psi(y)$ is the amount retained by the securitizer, where $0 \leq \psi(y) \leq y$ for all y. As in much of the security design literature, we focus on securities for which both the amount paid and the amount retained must be nondecreasing in y. We retain the assumption that each pool of loans is either good or bad

(i.e., $t \in \{b, g\}$). Further, a type-t pool delivers a cash flow distributed according to the cdf and pdf denoted by Π_t and π_t respectively on a common support $[0, \bar{y}]$, where $\frac{\pi_g(y)}{\pi_b(y)}$ is weakly increasing (i.e., MLRP holds). We refer to this setting as the *Security Design game*.

We show that the form of the security that emerges in the equilibrium of the Security Design game depends on a new measure of rating informativeness (denoted RI) and the cost of retention, δ.[31] In particular, if $RI < \delta$, then the securitizer issues debt and retains a levered-equity claim, while if $RI > \delta$, then the securitizer issues a levered-equity claim and retains debt (see DGV16, Theorem 1).[32] We also show that a result analogous to Proposition 3 holds in the Security Design game. That is, when ratings are informative enough the unique equilibrium involves some degree of pooling (either partial or full) whereas when ratings are not sufficiently informative the unique equilibrium is separating (see DGV16, Theorem 2).

Below, we characterize the equilibrium payoffs of the Security Design game as a function of the prior belief, μ_0, about the type of the pool. Denote these payoffs by $u_t^{SDG}(\mu_0)$. Importantly, these payoff functions share similar characteristics to the ones derived earlier, where the bank is restricted to issuing (and retaining) equity.

Fact B.1. *There exists a unique equilibrium of the Security Design game. Moreover,*

(i) $u_t^{SDG}(\mu_0)$ *is continuous and weakly increasing in* μ_0.

(ii) There exists $\mu_1 \in (0, 1]$ *such that for all* $\mu_0 \leq \mu_1$, $u_b^{SDG}(\mu_0) = v_b$ *and* $u_g^{SDG}(\mu_0) \in (u_g(\tilde{x}, \tilde{\mu}), v_g)$.

Proof. See DGV16. ■

B.2 Security Design: Lending Standards and Credit Supply

Let us now turn to the implications for lending standards and the supply of credit when the Securitization stage is replaced by the Security Design game. Analogous to Corollary 3, the lending standard is given by

$$p^{SDG} \in \Psi^{SDG}(\mu_0) \equiv \max\left\{\frac{1 - u_b^{SDG}(\mu_0)}{u_g^{SDG}(\mu_0) - u_b^{SDG}(\mu_0)}, 0\right\} \tag{54}$$

From Fact B.1, we know that $\Psi^{SDG}(\mu_0)$ is decreasing and continuous in the relevant range. Hence, there is a unique solution to (54) and a unique level of credit supply that is consistent

[31] The measure is defined as $RI \equiv \max_\mu \alpha_g(\mu) - \alpha_b(\mu)$.

[32] In the knife-edge case of $RI = \delta$, the form of security designed is not unique in equilibrium, though (for each μ_0) the equilibrium payoffs u_g^{SDG}, u_b^{SDG} are unique. Debt, equity, or levered equity, among other possibilities, can be used so long as the proper quantity is issued. Note that since the unique equilibrium payoffs can be obtained by issuing equity when $RI = \delta$, origination and credit supply are unaltered by expanding the set of available securities and the analysis from the body of the paper holds in this case. Therefore, we omit this non-generic case for the remainder of this Appendix.

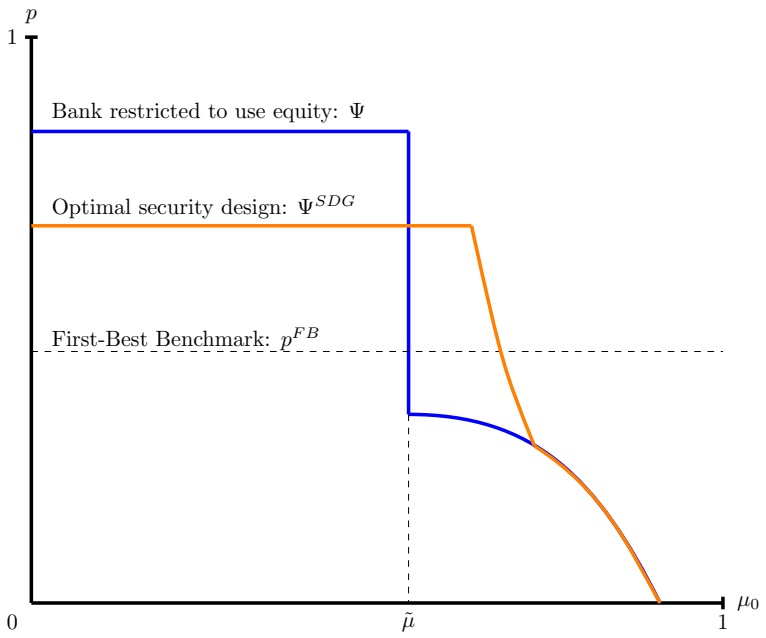


FIGURE 7: Security Design and Lending Standards.

with an equilibrium. We also know that $\Psi^{SDG}(\mu_0) \in \left(\frac{1-v_b}{v_g-v_b}, \Psi(\mu_0)\right)$ for $\mu_0 \leq \tilde{\mu}$. In this region, a bank with a bad pool of loans gets the full-information value, while the bank with a good pool of loans does strictly better by being able to choose the security design, which eases banks lending standards.

For convenience, we say that ratings are Γ-*informative* if inequality (8) in Lemma 2 of the present paper is satisfied. Provided the rating is Γ-informative, one can also show that Ψ^{SDG} lies weakly above Ψ for all priors above a threshold (and strictly above for at least some priors).[33] In the region where the inequality is strict, banks use retention to signal quality when they can design the security, but rely purely on ratings when they are restricted to equity. These properties are summarized in Figure 7 and formally stated in Lemma B.1.

Lemma B.1. *If the rating is* Γ*-informative, then the following statements are true.*

(i) $\Psi^{SDG}(\mu_0) \in \left(\frac{1-v_b}{v_g-v_b}, \Psi(\mu_0)\right)$ *for all* $\mu_0 < \tilde{\mu}$.

(ii) $\Psi^{SDG}(\mu_0) \geq \Psi(\mu_0)$ *for all* $\mu_0 > \tilde{\mu}$, *where the inequality holds strictly for at least some* μ_0.

If the rating is not Γ*-informative then* $\Psi^{SDG} < p^{NR}$ *for all* μ_0.

The next proposition summarizes the implications of security design on the equilibrium lending standard.

[33] Under some conditions (e.g., if the rating is not β-informative at $\bar{y}$, as defined in DGV16), Ψ^{SDG} lies strictly above Ψ for *all* priors above a threshold.

Proposition B.1. *If each bank can optimally design the security that they issue (i.e., if the Securitization stage is replaced with the Security Design game) then:*

- *If $A^{-1}(\tilde{\mu}) > \Psi^{SDG}(\tilde{\mu})$: the lending standard decreases toward p^{FB}, though an undersupply of credit persists.*
- *If $A^{-1}(\tilde{\mu}) < \Psi^{SDG}(\tilde{\mu})$: the lending standard increases, which may be toward or away from p^{FB}.*
- *If $A^{-1}(\mu_s) < \Psi^{SDG}(\mu_s)$, where μ_s is defined implicitly by $\Psi^{SDG}(\mu_s) = p^{FB}$: the lending standard increases toward p^{FB}, though oversupply of credit persists.*

In essence, Proposition B.1 says that the ability to design the security improves productive efficiency when the precision of banks' screening technology is sufficiently high or low. For intermediate levels of precision, optimal security design results in tighter lending standards, which may or may not improve productive efficiency.

Proofs for Section B.2

Proof of Lemma B.1. The proof of the case in which the rating is not Γ-informative is trivial. Regardless of the prior, in any equilibrium of the Security Design game, the b-payoff is weakly greater than v_b and the g-payoff is strictly greater than u_g^{LC}. Noting the $\frac{1-u_b}{u_g-u_g}$ is decreasing in both u_b and u_g (whenever it is non-negative) yields the result.

When the rating is Γ-informative. The statement in (i) follows immediately from Fact B.1. To prove (ii), it will be useful to break the proof into three cases.

Case 1: Rating is not β-informative at $\bar{y}$.[34] In this case, the equilibrium of the Security Design game does not converge to full-pooling with zero retention as $\mu_0 \to 1$. In particular, g retains a non-trivial levered equity claim and the low-type either pools or fully separates with zero retention. In either case, $u_t^{SDG}(\mu_0) < u_t^*(\mu_0) = u_t(0, \mu_0)$ for all $\mu_0 > \tilde{\mu}$, which implies the result.

Case 2: Rating is β-informative at $\bar{y}$, but not α-informative. In this case, there is full-pooling with zero retention for μ_0 large enough. Let μ_2 denote the smallest prior belief at which the zero-retention, full-pooling outcome obtains in the SDG. It suffices to show that $\mu_2 > \tilde{\mu}$. The FOC characterizing μ_2 is

$$g(\mu_2) = \frac{\pi_g(\bar{y})}{\pi_b(\bar{y})} - 1, \tag{55}$$

where $g(\mu) \equiv \frac{1-\delta}{(\alpha(\mu)-\delta)\frac{\alpha_b'(\mu)}{\alpha'(\mu)} - \alpha_b(\mu)}$. The FOC characterizing $\tilde{\mu}$ (see (38)) can be written as $g(\tilde{\mu}) = \frac{v_g}{v_b} - 1$. Note that $g(\mu) \leq 0$ for all $\mu < \mu_{\max} \equiv \arg\max_\mu \alpha(\mu)$, whereas the RHS of the FOC is

[34] As defined in DGV16, the rating is β-informative at $\bar{y}$ if $\mathbb{E}[\Gamma(R)|L] > 1 + \frac{\delta}{1+\delta}\left(1 - \frac{\pi_b(\bar{y})}{\pi_g(\bar{y})}\right)$.

strictly positive in both cases. Hence, it must be that both μ_2 and $\tilde{\mu}$ are above $\mu_{\max}$. Further, g is positive, strictly increasing, continuous, and tends to $+\infty$ on the interval $(\mu_{\max}, \bar{m})$ and g is negative above $\bar{m}$, where $\bar{m}$ is such that the denominator of g is zero. Finally, by MLRP $\frac{\pi_g(\bar{y})}{\pi_b(\bar{y})} > \frac{v_g}{v_b}$. Therefore $g(\mu_2) > g(\tilde{\mu})$ and thus $\mu_2 > \tilde{\mu}$.

Case 3: Rating is α-informative. The proof for this case is similar to Case 2. Again, there is full-pooling with zero retention for μ_0 large enough in the SDG. Let μ_2 denote the smallest prior belief at which the zero-retention, full-pooling outcome obtains in the SDG. We will show that $\mu_2 > \tilde{\mu}$. The FOC characterizing μ_2 in this case is

$$g(\mu_2) = \frac{1 - \Pi_g(\underline{y})}{1 - \Pi_b(\underline{y})} - 1 = 0 \tag{56}$$

where $g(\mu)$ is defined in Case 2. When the rating is α-informative, g is strictly decreasing over the relevant domain and equal to zero only at $\mu_{\max}$.[35] Therefore, $\mu_2 = \mu_{\max}$, whereas $g(\tilde{\mu}) = \frac{v_g}{v_b} - 1 > 0 \implies \tilde{\mu} < \mu_{\max}$, which implies the desired result. ■

Proof of Proposition B.1. Follows immediately from Lemma B.1. ■

[35] The relevant domain includes all $\mu \in (\underline{\mu}, \bar{\mu})$, where any solution to either FOC must lie.