

## Securitization, Ratings, and Credit Supply

BRENDAN DALEY, BRETT GREEN, and VICTORIA VANASCO\*

### ABSTRACT

We develop a framework to explore the effect of credit ratings on loan origination. We show that ratings endogenously shift the economy from a *signaling* equilibrium, in which banks inefficiently retain loans to signal quality, toward an *originate-to-distribute* equilibrium with zero retention and inefficiently low lending standards. Ratings increase overall efficiency, provided that the reduction in costly retention more than compensates for the origination of some negative net present value loans. We study how banks' ability to screen loans affects these predictions and use the model to analyze commonly proposed policies such as mandatory "skin in the game."

ASSET-BACKED SECURITIZATION IS AN IMPORTANT driver of credit supply (Loutskina and Strahan (2009), Shivdasani and Wang (2011)). In the United States, 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 (2013)), significantly increasing households' access to credit. The development of markets for securitized products has been facilitated in part by credit rating agencies (CRAs), which allow issuers access to a large pool of investors who would otherwise have perceived these securities as opaque and complex (Coval, Jurek, and Stafford (2009), Pagano and Volpin (2010)).

\*Brendan Daley is at Johns Hopkins University. Brett Green is at Washington University in St. Louis. Victoria Vanasco is at Centre de Recerca en Economia Internacional (CREI), UPF, and Barcelona GSE. Brett Green gratefully acknowledges support from the Fisher Center for Real Estate and Urban Economics. The authors are grateful to Philip Bond, an anonymous associate editor, and two anonymous referees for their valuable feedback and suggestions. We thank Barney Hartman-Glaser, Joel Shapiro, Anjan Thakor, Pablo Ruiz Verdú, and Alessio Piccolo for their thoughtful discussions. We are also grateful to 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, McCombs School of Business at UT Austin, Carlson School of Management at the University of Minnesota, Imperial College, London Business School, Toulouse School of Economics, HEC Paris, and Wharton School of Business; 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, NBER Corporate Finance Meeting, FIRS, Barcelona Summer Forum, and Oxford FIT for helpful feedback and suggestions. We have read *The Journal of Finance* disclosure policy and have no conflicts of interest to disclose.

DOI: 10.1111/jofi.12866

© 2019 the American Finance Association

In the aftermath of the recent financial crisis, the practice of securitization has been under intense scrutiny. The roles of originators in screening loans and of rating agencies in evaluating securitized products have come into question.<sup>1</sup> A variety of regulations have been proposed in an 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 with respect to 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 analyze the role of both public and private information. We then explore the implications for lending standards, credit supply, and welfare. Our main finding is that the availability of public information, such as credit ratings, improves the allocation of cash flow rights but reduces lending standards and can lead to an oversupply of credit. Despite the potential for an oversupply of credit, in most cases, total welfare increases with rating accuracy. We also illustrate how the effectiveness of banks’ screening technology influences the effect of ratings. With a better understanding of these forces, we investigate 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.<sup>2</sup> Each bank then decides whether to fund its pool—the *origination stage*. Following origination, banks have an incentive to reallocate the cash flow rights from their loan pool 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, 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). Banks’ ability to signal through retention is consistent with evidence in Begley and Purnanandam (2017) and Ivashina (2009), who study the markets for residential mortgage-backed

<sup>1</sup> See Dell’Ariccia, Igan, and Laeven (2012), Keys et al. (2010), Jaffee et al. (2009), Mian and Sufi (2009), and Agarwal, Chang, and Yavas (2012) for how securitization negatively affects lending standards, and Pagano and Volpin (2010) and Benmelech and Dlugosz (2010) for the role, and failures, of CRAs in the securitization process.

<sup>2</sup> The assumption that banks acquire private information about borrowers at the loan screening stage is consistent with findings in Mikkelsen and Partch (1986), Lummer and McConnell (1989), Slovin, Sushka, and Polonchek (1993), Degryse and Ongena (2005), Plantin (2009), Agarwal and Hauswald (2010), and Botsch and Vanasco (2019).

securities (RMBS) and syndicated loans, respectively.<sup>3</sup> Second, information about the pool of loans underlying each security can be conveyed to investors through a noisy public signal about the quality of the underlying collateral. We refer to this signal as a *rating*, but it can be interpreted more broadly as any form of public information. This rating is observed after the bank's retention decision but prior to the sale of the security. The primary question that we seek to answer is how ratings affect lending standards and the supply of credit.

To address this question, it is useful to describe the benchmark model without ratings. Absent ratings or the 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 of the loan they issue if they originated a good pool and sell 100% of their loan if they originated a bad pool. 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.

When informative ratings are available, banks that originate good loan pools no longer fully separate through retention. Instead, there is some degree of pooling at a lower retention level.<sup>4</sup> Since retention is inefficient, ratings improve allocative efficiency in the securitization stage.<sup>5</sup> But, because less is retained and ratings are imperfect, their introduction actually reduces banks' lending standards and may induce an oversupply of credit. In essence, when ratings are introduced, the equilibrium of the securitization stage endogenously shifts from a signaling-through-retention equilibrium to an originate-to-distribute (OTD) equilibrium, where banks forgo signaling and sell 100% of the loans they originate.

There are two potential sources of inefficiency in our model. First, cash flow retention may be inefficiently high due to asymmetric information at the securitization stage, which induces banks to engage in costly signaling. Second, lending standards may be inefficiently high or low since banks do not necessarily internalize the social value of the loans they originate. More accurate ratings reduce costly retention, but may also induce inefficiently low lending standards. Therefore, ratings increase ex-ante efficiency, provided that the benefits of reduced retention outweigh the costs of originating negative net present value (NPV) loans. We show that these benefits necessarily outweigh the costs when bank screening technology is sufficiently effective. Further, as ratings

<sup>3</sup> Adelino, Gerardi, and Hartman-Glaser (2019) find evidence of banks signaling their private information about RMBS deals by delaying trade—a different form of cash flow retention that could be studied within our context as well.

<sup>4</sup> A similar feature is present in Hartman-Glaser (2017), who shows that when sellers are able to signal through both retention and reputation (as opposed to with a public signal), the equilibrium is no longer separating.

<sup>5</sup> This result is consistent with empirical evidence 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 (2005), Sufi (2007), Derrien and Kecskés (2013)).

become perfectly informative, retention, lending standards, credit supply, and efficiency converge to first best.

We next explore how the precision of banks' screening technology (i.e., banks' private information at origination) interacts with the effect of ratings. Without ratings, as banks' screening technology becomes arbitrarily precise, only good loans are originated. With ratings, however, as banks' screening technology becomes more precise, their lending standards fall and a nonnegligible mass of bad loans is always (deliberately) originated.

We use the model to evaluate several different regulations. An intuitive and often proposed regulation is to require that banks retain a fraction of all originated loans. Proponents argue that this would provide banks incentives 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 requiring skin in the game, which goes as follows. If banks had been 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 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. We then discuss situations in which it does not. Finally, motivated by central banks' policy of easing credit constraints 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 and fewer loans being originated.

The remainder of the paper is organized as follows. In Section I, we relate the model's predictions to existing literature, and we discuss several novel testable implications of our model. In Section II, we introduce the model and our solution concept. In Section III, we present benchmarks. We analyze the equilibrium of the model in Section IV and its comparative statics in Section V. In Section VI, we explore policy implications. Section VII concludes. All proofs are relegated to the Appendix.

## **I. Related Literature**

### *A. Relation to Existing Empirical Literature*

Our model suggests that the widespread use of ratings as a source of public information for securitized products may have been an important driver of

the credit expansion and the proliferation of OTD practices observed in the years leading up to the 2008 to 2009 financial crisis. We show that when banks originate to distribute, a decrease in rating accuracy results in an expansion of credit to negative NPV borrowers (Proposition 4). This result is consistent with evidence of decreasing lending standards (Mian and Sufi (2009), Purnanandam (2010)), as rating technologies worsened due to changes in banks' screening behavior not incorporated in statistical models (Rajan, Seru, and Vig (2015)) and/or to pervasive rating shopping and manipulation practices (Ashcraft, Goldsmith-Pinkham, and Vickery (2010)). The effect of rating accuracy on lending standards, however, is nonmonotonic (see Proposition 6). When rating accuracy falls sufficiently, our model predicts that banks will cease their OTD practices and sharply contract credit, consistent with the reduction in lending and securitization during and after the financial crisis.

We obtain several cross-sectional implications, some of which are in line with empirical evidence. Consistent with our model predictions, prior empirical studies find a positive relation between retention and underlying loan quality (Ashcraft, Gooriah, and Kermani (2019)), between loan screening and underlying loan quality (Berger and Udell (2004)), and between rating accuracy and underlying loan quality (Rajan, Seru, and Vig (2015)). Furthermore, we find that excessive lending and inefficiently low lending standards are more likely to arise when banks' liquidity needs are high (Proposition 11), in line with the finding in Begley and Purnanandam (2017), the quality of screening effort fell more during the lending boom in more capital-constrained banks.

### *B. Testable Implications*

Our results first highlight the importance of studying the interaction between cash flow retention and rating accuracy for the determination of credit supply and lending standards. Our model predicts that a reduction in rating accuracy results in an expansion of credit to negative NPV loans when banks originate to distribute (e.g., credit card loans, mortgages), and a contraction of credit to positive NPV loans when banks retain some of their loan cash flows (e.g., syndicated lending).

Our results also support that the interaction between banks' screening technology and rating accuracy is important in determining the level of cash flow retention and underlying loan quality. Our model predicts that banks are more likely to OTD, and thus lend excessively, in asset classes for which the screening technology is more effective and/or more public information is available. This prediction is consistent with the observation that banks generally retain loans to small businesses, for which screening is costly and public information is relatively scarce, while they originate to distribute credit card loans and a large fraction of their mortgages, for which screening costs are low and the availability of public information is relatively high (e.g., FICO score, credit reports). We view syndicated lending as being somewhere in the middle, consistent with banks retaining a small fraction of these loans.

*C. Related Theoretical Literature*

Several papers highlight the trade-off between productive and allocative efficiency studied in this paper. Dell’Ariccia and Marquez (2006) explore how the information structure of loan markets interacts with competitive banks’ strategic decisions and impacts lending standards and the overall supply of credit. Parlour and Plantin (2008) study the effect of loan sales on banks’ origination decisions. Malherbe (2012) explores the relation between postorigination risk-taking and market discipline. Chemla and Hennessy (2014) study 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. A rich literature also focuses on optimal contracting with loan sales and moral hazard (Gorton and Pennacchi (1995), Hartman-Glaser, Piskorski, and Tchisty (2012), Vanasco (2017)). None of these papers, however, study the release of public information to investors about the assets being traded.

The theoretical approach adopted in this paper builds on Daley and Green (2014). That paper considers 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 that the distribution of the quality of assets brought to market is endogenous, similar to Vanasco (2017).

In our model, we take ratings accuracy as an exogenous parameter and abstract from strategic incentives of CRAs as well as an issuer’s ability to influence the information content of ratings, both of which are of practical relevance. Indeed, an extensive literature studies these considerations.<sup>6</sup> In an earlier working paper version of the paper and inspired by the CRA models in Skreta and Veldkamp (2009), Sangiorgi and Spatt (2016), Bolton, Freixas, and Shapiro (2012), and Opp, Opp, and Harris (2013), we considered extensions allowing for ratings shopping and rating manipulation.<sup>7</sup> In both cases, the information content of the rating is endogenously determined, which effectively reduces their accuracy. Thus, incorporating these considerations has an effect similar to a reduction in the accuracy of (exogenously generated) ratings.

<sup>6</sup> Important considerations include the role of CRA reputation and moral hazard (Mathis, McAndrews, and Rochet (2009), Bar-Isaac and Shapiro (2013), Fulghieri, Strobl, and Xia (2014), Goel and Thakor (2015), Kashyap and Kovrijnykh (2015)), feedback effects and ratings as coordination devices (Boot, Milbourn, and Schmeits (2006), Manso (2013), Goldstein and Huang (2017)), and the implications of rating-contingent regulation (Opp, Opp, and Harris (2013), Josephson and Shapiro (2019)).

<sup>7</sup> See <http://faculty.haas.berkeley.edu/bgreen/files/RatingsWP2017.pdf>. These extensions are also in line with empirical studies on ratings shopping and manipulation: Ashcraft et al. (2011), Griffin and Tang (2011), Griffin, Nickerson, and Tang (2013), Becker and Milbourn (2011), He, Qian, and Strahan (2011), Kraft (2015), and Piskorski, Seru, and Witkin (2015).

## II. 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.

### A. 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 quality of the underlying loans in the pool,  $t \in \{\text{good}, \text{bad}\}$ . We refer to  $t$  as the pool's type although there remains (an arbitrary amount of) residual uncertainty about the cash flow generated by the loan pool, which is captured by the fact that  $Y$  is still random after conditioning on  $t$ . We let  $v_t = E[Y|t]$  and  $\xi \in (0, 1)$  denote the proportion of good pools in the economy. We assume  $v_b < 1 < v_g$ , meaning that only good loan pools create value.

Prior to making origination decisions, banks acquire information about loan pools using their screening technology.<sup>8</sup> 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  $\psi_t$ . When screening results in a realization  $s$ , then the bank's *appraisal* about its loan pool, denoted by  $p$ , is given by

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

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  $[0, \infty)$ .<sup>9</sup> 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 that observes signal  $s$  satisfying (1) when it screens its loan pool.<sup>10</sup>

<sup>8</sup> See footnote 2 for references to evidence consistent with this assumption.

<sup>9</sup> This assumption holds if, for example,  $\psi_t$  is a normal density with mean  $m_t$ ,  $m_g \neq m_b$ , and variance  $\sigma^2$ .

<sup>10</sup> 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 that it satisfies Bayes plausibility (i.e.,  $\xi = \int p dH(p)$ ).

After observing the realization from the screening technology, each bank decides whether to originate the loans in its pool. If the bank chooses not to originate, it takes no further actions and earns a payoff of zero. If the bank originates its loans, it has the opportunity to securitize the cash flows from the pool as we describe next.

### *B. Securitization Stage*

Each originating bank has an incentive to raise cash through securitization of the cash flows from its loan pool. This incentive could arise for a variety of reasons (e.g., credit constraints or capital requirements). 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 one. 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.

After origination but prior to securitizing a loan pool, banks uncover additional information about the quality of their loan pools, which we capture as the bank learning the loan pool type  $t$ .<sup>11</sup> The timing assumption, while not crucial to our findings, is motivated by the fact that there is a lag between origination and securitization during which the bank can observe loan performance and conduct additional analysis.<sup>12</sup>

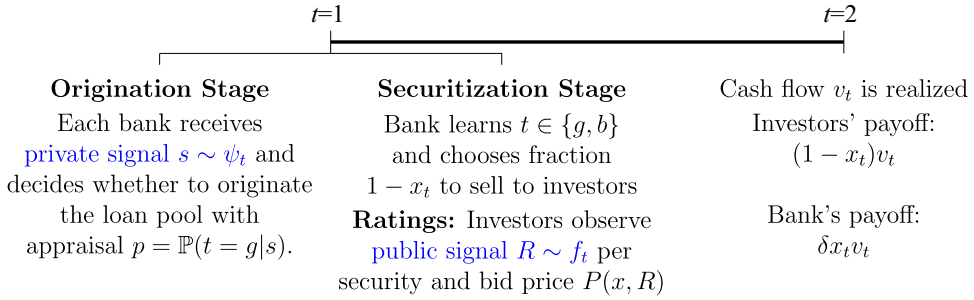
**REMARK 1:** *In Section V.C, we demonstrate that our main results are robust to an alternative specification in which each bank observes a single binary signal prior to origination and no additional information thereafter.*

For convenience, we focus on a simple securitization structure whereby 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 2:** *In principle, each bank could design and sell a security that is an arbitrary function of its cash flow. In Daley, Green, and Vanasco (2016), we study the relevant security design game with ratings. Using the results therein, we demonstrate that the main insights of this paper remain unchanged when*

<sup>11</sup> This assumption conveniently reduces the type space to binary, which allows for a sharp characterization of equilibrium play at the securitization stage.

<sup>12</sup> Adelino, Gerardi, and Hartman-Glaser (2019) document that for the period 2002 to 2007, the average loan seasoning is 3.3 months for private-label MBS. Aiello (2016) shows that borrowers who make payments at least a day prior to the due date early in the life of the mortgage—information that is privately observed by the issuer prior to securitization—are significantly less likely to become delinquent.



**Figure 1. Time line of the game.** (Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com))

we allow banks to design and sell arbitrary securities (see the Internet Appendix Section I).<sup>13</sup>

### C. Ratings

In addition to observing the level of costly retention  $x$ , we consider a second channel through which information may be conveyed to investors, which we refer to as a *rating*. We model 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  $\mathbb{R}$ .<sup>14</sup>

The accuracy of a rating realization,  $r$ , is captured by the likelihood ratio:  $\Gamma(r) \equiv \frac{f_b(r)}{f_g(r)}$ .<sup>15</sup> Without loss of generality, order the ratings such that  $\Gamma$  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 accuracy, we 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 a higher  $\gamma$  corresponds to more accurate ratings.

A time line summarizing the sequence of events is presented in Figure 1. Though stylized, our model captures the timing of the securitization and rating process in practice, which we describe in more detail in the Internet Appendix Section II.

### D. 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., securitization) stage and work backward.

<sup>13</sup> The Internet Appendix is available in the online version of the article on *The Journal of Finance* website.

<sup>14</sup> To encompass 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  $r \in \mathbb{R}$ .

<sup>15</sup> If  $f_g(r) = f_b(r) = 0$ , we adopt the convention that  $\Gamma(r) = 1$ .

At the beginning of the securitization stage, investors have a (common) prior belief  $\mu_0$  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 banks' retention strategy.<sup>16</sup>

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

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

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 value of the cash flows generated by the security given  $\mu_f$ :

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

Given a schedule of interim beliefs  $\mu(\cdot)$ , the expected payoff of a bank that originates 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 the continuation payoffs and a bank's 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,  $\mu_0$ , 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 requires  $\mu_0 = E[p|p \in O^*]$ .

#### D.1. 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, such that 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 to guarantee that the lending standard is always interior, we make the following assumption.

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

<sup>16</sup> 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.

In effect, 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*

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

An immediate corollary is that investors' prior belief 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 as follows:*

- (1) *In the securitization stage: Given  $\mu_0^*$ , 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 that is more likely to gain from this deviation. See the Appendix for a formal definition. Hereafter, we use the term *equilibrium* to refer to a PBE that satisfies D1 in the securitization stage.

### III. Benchmarks

#### A. Full Information / First Best

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 would also be achieved as loan pools would be 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  $Q(p^{FB}) = 1 - H(p^{FB})$ .

### B. Strategic Model without Ratings

Consider now the model as described in Section II, but with completely uninformative ratings (i.e.,  $\Gamma(r) = 1$  for all  $r \in \mathbb{R}$ ). 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

$$\underbrace{u_b(0, 0)}_{v_b} = \underbrace{u_b(\bar{x}, 1)}_{(1-\bar{x})v_g + \delta\bar{x}v_b}. \quad (5)$$

That is, the originator of a  $b$ -pool is indifferent between efficiently selling all of its cash flow rights at the full-information price,  $v_b$ , and retaining fraction  $\bar{x}$  if doing so convinces investors that its pool is good, which leads to a price of  $(1 - \bar{x})v_g$  for the security it sells. Therefore,  $\bar{x}$  is the minimum amount the  $g$ -type must retain to separate from the  $b$ -type in the securitization stage.

Without ratings, the securitization stage is a standard signaling game in which indifference curves over  $(x, \mu)$  pairs satisfy the single-crossing property (i.e., the  $g$ -type's indifference curve is flatter than that of the  $b$ -type) because it is less costly for a  $g$ -type to retain cash flows. As a result, D1 selects this "least-cost separating" equilibrium (LCSE).

**PROPOSITION 1:** *Without informative ratings, 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 by  $p^{NR}$ , 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, that is,  $p^{NR} > p^{FB}$ .*

## IV. 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, which we model as the random variable  $R$ . We first characterize the equilibrium of the securitization stage for any investor belief  $\mu_0$  (Section A). We then characterize banks' lending standard in the origination stage along with the consistent investor belief (Section B). We conclude this section with one of our main results (Proposition 3), which characterizes when the equilibrium involves an over- or undersupply of credit.

### A. Securitization Stage

The analysis of this stage closely follows that of Daley and Green (2014).

#### A.1. The Maximization Problem

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

$$\max_{x, \mu} u_g(x, \mu) \quad \text{s.t.} \quad u_b(x, \mu) = v_b. \quad (6)$$

That is, among all  $(x, \mu)$  pairs that deliver the  $b$ -type, its full information expected payoff  $v_b$ , 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 by  $(\tilde{x}, \tilde{\mu})$ . As formalized in Result 1, the solution is a critical part of the equilibrium characterization (and where the D1 refinement plays its role). We first provide some intuition for how ratings affect the solution.

Recall that without ratings, the  $g$ -type's lower cost of retention leads to the single-crossing property, and thus in this case  $(\tilde{x}, \tilde{\mu}) = (\tilde{x}, 1)$ . That is, if there are no ratings to convey information to investors, the  $g$ -type uses the LCSE retention level to perfectly distinguish the superior quality of its cash flows. Consider now the effect of introducing ratings. For any interior interim belief, the  $g$ -type expects a higher final belief than does a  $b$ -type. Therefore, the  $g$ -type now has two advantages: a lower cost of retention and a superior distribution of the rating. At the solution to (6),  $(\tilde{x}, \tilde{\mu})$ , the  $g$ -type makes optimal use of the two advantages.

The rating advantage is largest for intermediate levels of the interim belief  $\mu$ : the effect of new information is greatest when Bayesians are most uncertain. Therefore, the addition of ratings pushes the optimal interim belief away from the corner solution of  $\mu = 1$  and toward the interim belief at which the rating is most impactful. When the solution  $\tilde{\mu}$  is interior, it represents the belief at which the  $g$ -type's benefit of a marginal increase in the belief is exactly offset by its loss from decreasing the significance of the rating. Hence, more accurate ratings lead to a lower value of  $\tilde{\mu}$  (see Lemma 3). The following proposition captures how informative the rating must be to alter the solution from the no-ratings case.

LEMMA 2: *In the solution to (6),  $(\tilde{x}, \tilde{\mu}) < (\bar{x}, 1)$  if and only if*

$$E[\Gamma(R)|b] > \frac{v_g - \delta v_b}{(1 - \delta)v_g}. \quad (7)$$

The accuracy 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 accuracy of the *rating system*,

$\{f_g, f_b\}$ .<sup>17</sup> The right-hand side of (7) measures the relative cost advantage of the  $g$ -type in retaining cash flows. Thus, the solution to (6) 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. The exact form of (7) arises from checking when ratings are sufficiently informative to reverse the original single-crossing property at  $(x, \mu) = (\bar{x}, 1)$ .<sup>18</sup>

Given Lemma 2, it is perhaps not surprising that if (7) does not hold, then ratings are simply too noisy to alter the prediction from the no-ratings benchmark studied in Section III.B. In the remainder of the paper, we analyze the model in which ratings are informative enough to affect the equilibrium outcome.

ASSUMPTION 2: *Henceforth, we assume that (7) holds unless otherwise stated.*

### A.2. Equilibrium Securitization

While our model does not satisfy the primitive assumptions of Daley and Green (2014), Lemmas 2 and A.2 establish properties that are sufficient to apply proposition 3.8 of Daley and Green (2014). With these properties established, we obtain the following characterization of the equilibrium at the securitization stage.

RESULT 1 (Daley and Green (2014), proposition 3.8): *For any  $\mu_0 \neq \tilde{\mu}$ , there is a unique equilibrium of the securitization stage. The equilibrium is characterized as follows:*

- (1) *If  $\mu_0 < \tilde{\mu}$ , there is partial pooling at  $\tilde{x} < \bar{x}$ . That is, all banks with  $g$ -pools retain  $\tilde{x}$ , a fraction  $\frac{\mu_0(1-\tilde{\mu})}{(1-\mu_0)\tilde{\mu}}$  of banks with  $b$ -type pools retains  $\tilde{x}$ , and a fraction  $\frac{\tilde{\mu}-\mu_0}{(1-\mu_0)\tilde{\mu}}$  retains zero. Hence, the interim belief for  $x = \tilde{x}$  is  $\mu(\tilde{x}) = \tilde{\mu}$ .*
- (2) *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  $x \in [0, \bar{x}]$ .*

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

<sup>17</sup> The more informative the rating system, the higher the value of  $E[\Gamma(r)|b]$ . This measure is consistent with the notion of accuracy 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.

<sup>18</sup> That is, under (7), the  $g$ -type's indifference curve is now steeper than the  $b$ -type's at  $(x, \mu) = (\bar{x}, 1)$ . Hence, in the LCSE profile, D1 now prescribes that investors interpret a downward deviation in retention from  $\bar{x}$  as a  $g$ -type bank demonstrating its confidence that its rating will realize favorably, which breaks the LCSE.

$\mu_0 < \tilde{\mu}$ , 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. Of course, banks' behavior must be consistent with that interim belief. Therefore, the fraction of banks with  $b$ -type pools that retain  $\tilde{x}$  is such that the Bayesian consistent interim belief conditional on observing  $\tilde{x}$  is exactly  $\tilde{\mu}$ .

The switch from full to partial pooling happens at  $\tilde{\mu}$  precisely because it is the solution to (6). It represents the level of interim belief at which *further* increases through costly retention are no longer beneficial for the  $g$ -type. When the prior,  $\mu_0$ , is below  $\tilde{\mu}$ , the originators of good pools signal through retention to increase the belief to  $\tilde{\mu}$ , but no higher. Alternatively, if  $\mu_0$  is already above  $\tilde{\mu}$ , there is no reason to signal through retention.

### B. Origination Stage

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

#### B.1. 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, if  $p \geq \frac{1-u_b^*}{u_g^*-u_b^*}$ . From Result 1,  $u_g^*$  and  $u_b^*$  vary with investors' belief  $\mu_0$  when ratings are informative—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 that a bank is willing to originate (i.e., the lending standard) given investors' beliefs  $\mu_0$ .

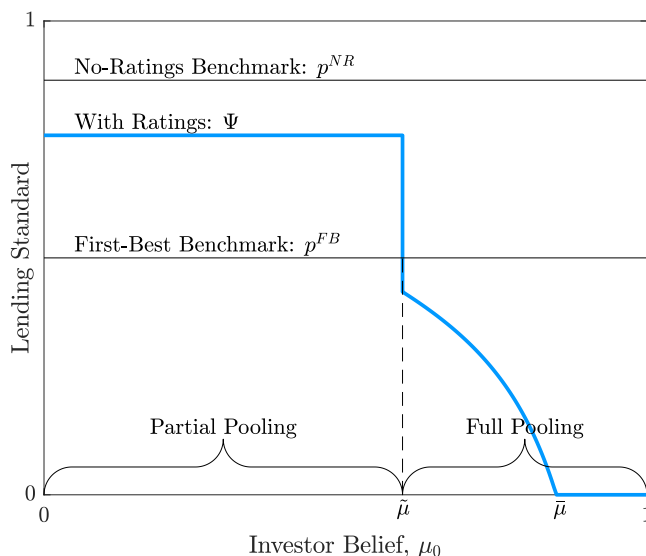
DEFINITION 1:  $\Psi(\mu_0) \equiv \{\max\{\frac{1-u_b^*}{u_g^*-u_b^*}, 0\} | u_g^*, u_b^* \text{ are equilibrium payoffs given } \mu_0\}$ .

The max operator in  $\Psi$  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 zero. From Result 1,  $\Psi$  is single-valued for all  $\mu_0 \neq \tilde{\mu}$ .

COROLLARY 3: *Taking investors' belief  $\mu_0$  as given, banks' lending standard 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 2 depicts  $\Psi$ , 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

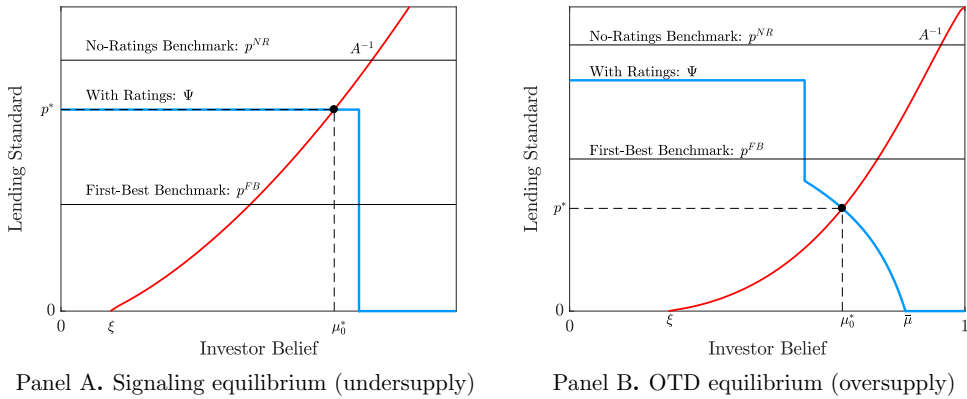


**Figure 2. The lending standard, as it depends on investor belief.** This figure illustrates banks' lending as a function of investor beliefs ( $\Psi(\mu_0)$ ), as well as the lending standard in the first-best ( $p^{FB}$ ) and no-ratings ( $p^{NR}$ ) benchmarks. Note that for  $\mu_0 > \tilde{\mu}$ , banks choose to sell 100% of originated loan pools regardless of  $t$ . (Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com))

benchmarks, payoffs in the securitization stage do not depend on investors' prior beliefs, so the lending standards are independent of  $\mu_0$ . Furthermore,  $p^{FB} < p^{NR}$ , as documented in Corollary 2.

With ratings, the lending standard adopted by banks depends on investors' 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  $\mu_0 < \tilde{\mu}$ ). Notice that the lending standard is independent of  $\mu_0$  in this region because investors' belief conditional on the retention level is independent of  $\mu_0$  for all  $\mu_0 \leq \tilde{\mu}$ . As a result, the bank payoffs conditional on loan type ( $u_b^*, u_g^*$ ) are also constant in this region.

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  $\mu_0 = \tilde{\mu}$ . The lending standard continues to decrease as  $\mu_0$  increases further, as a higher investor belief translates directly into higher security prices for both types. Eventually,  $u_b^*$  reaches one, the cost of origination. We denote this belief level as  $\bar{\mu}$ . 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 earn a positive return. Consequently,  $\Psi(\mu_0) = 0$  for all  $\mu_0 \geq \bar{\mu}$ , as seen in Figure 2.



**Figure 3. The equilibrium lending standard and investor belief.** This figure illustrates how the lending standard and investors' belief are jointly determined in equilibrium. Panel (A) illustrates an example of a signaling equilibrium in which the lending standard is above  $p^{FB}$ , and hence, there is an undersupply of credit. Panel (B) illustrates an example of an OTD equilibrium in which the lending standard is below  $p^{FB}$ , and hence, there is an oversupply of credit. (Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com))

### B.2. Investor Belief Consistency

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

**PROPOSITION 2:** *There exists a unique equilibrium. The equilibrium lending standard is given by the (unique)  $p^*$  satisfying  $p^* = A^{-1}(\mu_0) \in \Psi(\mu_0)$ .*

### C. Equilibrium Properties

Figure 3 illustrates how the bank-origination-optimality and investor-belief-consistency conditions pin down the equilibrium lending standard,  $p^*$ , and investor beliefs,  $\mu_0^*$ , by the unique intersection of  $A^{-1}$  and  $\Psi$ . Panel (A) depicts an equilibrium with  $\mu_0^* < \tilde{\mu}$ , where the lending standard,  $p^*$ , is above first-best (i.e., there is an undersupply of credit). We refer to this as a *Signaling equilibrium* because information is conveyed to investors by a bank's retention level. In a signaling equilibrium,  $u_b^* = v_b$  but  $u_g^* < v_g$ , which implies that

$$p^* = \frac{1 - u_b^*}{u_g^* - u_b^*} = \frac{1 - v_b}{u_g^* - v_b} > \frac{1 - v_b}{v_g - v_b} = p^{FB}.$$

Panel (B) depicts an equilibrium with  $\mu_0^* > \tilde{\mu}$ , in which case the securitization stage involves full pooling at zero retention. We refer to this as an *OTD equilibrium* because banks originate loans with no intention of retaining them. In an OTD equilibrium, the lending standard is inefficiently low (i.e., there is

an oversupply of credit). Intuitively, perfectly informative ratings would return us to the first-best benchmark, in which case there is no retention, a bank's payoff is equal to the value of the pool it originates, and the first-best lending standard is attained ( $p^* = p^{FB}$ ). When there is no retention, and the rating only *imperfectly* distinguishes good pools from bad ones, a bank's payoff is no longer equal to the value of the pool they originate as there is some subsidization from good to bad loan pools. Hence, banks originate too many loans compared to the first best.

The third possibility (not depicted) is that  $\mu_0^* = \tilde{\mu}$ , in which case the lending standard may be above or below first best. We refer to this as a *retention equilibrium*, since there is full pooling at positive retention. The following corollary fully characterizes when each of these three types of equilibria obtains.

COROLLARY 4: *The unique equilibrium is*

- (1) *a signaling equilibrium if and only if  $\tilde{\mu} > A(\Psi^+(\tilde{\mu}))$ ;*
- (2) *a retention equilibrium if and only if  $\tilde{\mu} \in (A(\Psi^-(\tilde{\mu})), A(\Psi^+(\tilde{\mu})))$ ; and*
- (3) *an OTD equilibrium if and only if  $\tilde{\mu} \leq A(\Psi^-(\tilde{\mu}))$ ;*

where  $\Psi^+(\tilde{\mu})$  and  $\Psi^-(\tilde{\mu})$  denote the maximum and minimum values of  $\Psi$  at  $\tilde{\mu}$ .

Perhaps, the most surprising implication of our equilibrium analysis is that the introduction of ratings can lead to lending standards that are too lax and an oversupply of credit. The following proposition characterizes when this occurs.

PROPOSITION 3: *The equilibrium lending standard is strictly below the first-best level if and only if*

$$\tilde{\mu} < A(p^{FB}). \quad (8)$$

Fixing the payoff parameters (i.e.,  $\delta, v_\theta$ ), the accuracy of ratings determines  $\tilde{\mu}$  and has no effect on  $A$ , while the precision of the screening technology determines  $A(p^{FB})$  and has no effect on  $\tilde{\mu}$ . As we will show in the next section,  $\tilde{\mu}$  is decreasing with rating accuracy, while  $A(p^{FB})$  is increasing with screening precision. Therefore, all else equal, (8) is more likely to hold when ratings are more informative or when the screening technology is more effective.

## V. Determinants of Credit Supply

Using Proposition 3 to guide our analysis, in this section, we study how changes in rating accuracy and banks' screening technology affect loan origination and overall efficiency. To do so, notice that because investors break even, the total surplus generated by the banking sector—our measure of overall efficiency—is given by

$$\int_{p^*}^1 \mathbb{E}[u_t^* - 1|p] dH(p) = [\mu_0^* u_g^* + (1 - \mu_0^*) u_b^* - 1] Q(p^*), \quad (9)$$

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.

It is important to highlight the type of inefficiencies that may arise in a given equilibrium. In signaling equilibria, retention is inefficient,  $x > 0$ , while the lending standard is efficient (given equilibrium retention), resulting in  $p^* > p^{FB}$ . This is because in such equilibria, each bank fully internalizes the social value of the loans it originates. In contrast, in OTD equilibria, retention is efficient but the lending standard is inefficient, since  $p^* < p^{FB}$ . This is because in such equilibria, each bank does not fully internalize the social value of the loans it originates as ratings are imperfect and affect the loan pool's pricing.

Finally, in retention equilibria, both sources of inefficiency are generically present.

### A. Ratings Accuracy

How does the accuracy of ratings affect origination, securitization, and overall efficiency? To answer this question, we focus on the binary-symmetric rating system,  $P(R = G|g) = P(R = B|b) = \gamma \in (\frac{1}{2}, 1)$ , where higher  $\gamma$  implies more informative ratings. To begin, we examine how an increase in rating accuracy affects the securitization stage and, in turn, banks' reaction function for origination,  $\Psi$ .

LEMMA 3: *As the accuracy of ratings ( $\gamma$ ) increases:*

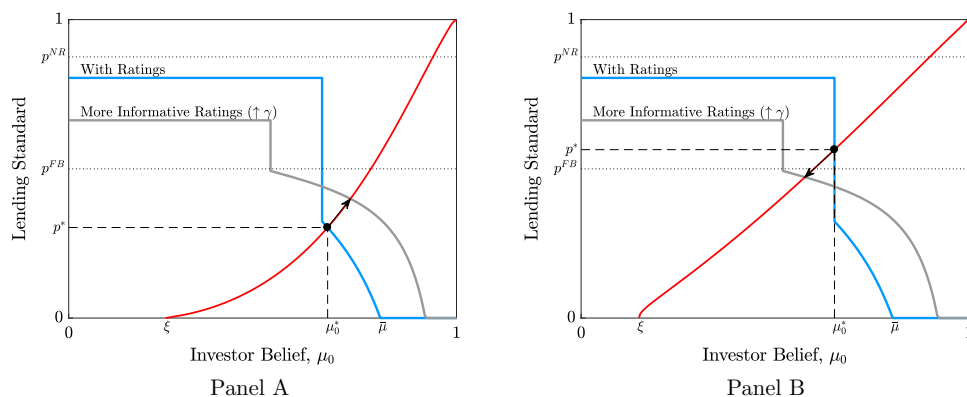
- (1) *Both  $\tilde{\mu}$  and  $\tilde{x}$  decrease.*
- (2)  *$\Psi(\mu_0)$  decreases if  $\Psi(\mu_0) > p^{FB}$  and increases if  $\Psi(\mu_0) < p^{FB}$ .*

The first statement formalizes the intuitive claim in Section IV.A: more accurate ratings lead to less reliance on retention. The interpretation of the second statement is that as rating accuracy improves, banks' incentive to originate tends toward their incentive to originate in the first-best benchmark.

The first statement implies that the condition for an oversupply of credit (i.e., (8)) is more likely to be satisfied when ratings are more informative. However, this does not imply that  $p^*$  monotonically decreases with  $\gamma$  nor does it imply that more accurate ratings harm overall efficiency. Indeed, the second statement suggests (and Figure 4 confirms) that more accurate ratings can increase or decrease the lending standard depending on which type of equilibrium is being played.

PROPOSITION 4: *A marginal increase in rating accuracy ( $\gamma$ ) has the following implications:*

- (1) *Retention decreases when positive.*
- (2) *The lending standard and average loan quality decrease in a signaling or retention equilibrium and increase in an OTD equilibrium.*
- (3) *The quantity of loans originated increases in a signaling or retention equilibrium and decreases in an OTD equilibrium.*



**Figure 4. Ratings accuracy, the lending standard, and investor belief.** This figure illustrates how the accuracy of the rating technology ( $\gamma$ ) affects the equilibrium lending standard and investor beliefs. In Panel (A), an increase in rating accuracy leads to a higher lending standard, whereas in Panel (B), the lending standard decreases. (Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com))

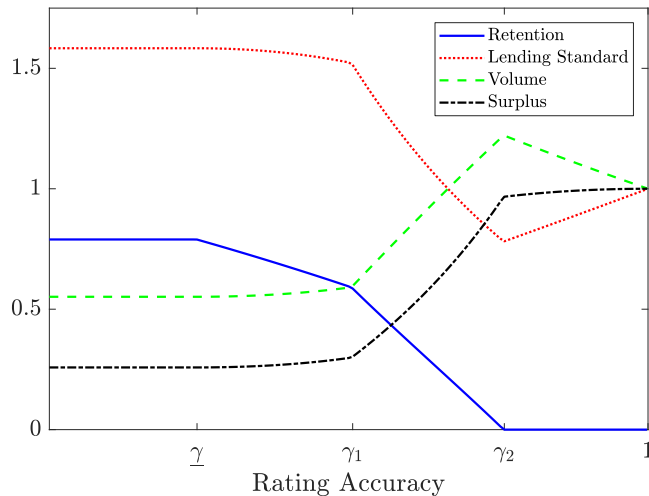
(4) *Efficiency increases in a signaling or OTD equilibrium, but may increase or decrease in a retention equilibrium.*

The overall effect of  $\gamma$  is summarized in Figure 5. In this example and as described below, four types of equilibria occur for a positive measure of  $\gamma$ . Starting from uninformative ratings (i.e.,  $\gamma < \underline{\gamma}$ ), where  $p^* = p^{NR} > p^{FB}$ , the equilibrium is fully separating and increasing  $\gamma$  has no effect on the economy until  $\gamma = \underline{\gamma}$ , where (7) holds with equality. At this point, the economy enters a signaling equilibrium, where increasing  $\gamma$  decreases both the lending standard and retention. Because the only source of inefficiency is retention in a signaling equilibrium, more accurate ratings lead to higher total surplus.

At  $\gamma_1$ , the economy transitions to a retention equilibrium, where further increases in  $\gamma$  continue to lower retention and the lending standard. Since in retention equilibria, both inefficiencies are at play, increasing rating accuracy improves total surplus if the gains from lower retention outweigh the costs of lower lending standards.<sup>19</sup>

At  $\gamma_2$ , retention hits zero, at which point the equilibrium is OTD and further increases in  $\gamma$  lead to higher lending standards. Because the only source of inefficiency in OTD equilibria is inefficiently low lending standards, more accurate ratings increase total surplus. Finally, as summarized in the next proposition, the equilibrium converges to the first-best benchmark as  $\gamma \rightarrow 1$ .

<sup>19</sup> In Figure 5, total surplus increases with  $\gamma$  in the retention region, but it is possible to construct examples in which surplus decreases for some  $\gamma$  in this region. It is therefore also possible to construct examples in which surplus can decrease if a discrete increase in rating accuracy moves the economy from a signaling to an OTD equilibrium.



**Figure 5. Comparative statics on rating accuracy illustrates the effect of increasing rating accuracy.** The lending standard, volume, and total surplus are all plotted as a fraction of the first-best benchmark. (Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com))

**PROPOSITION 5:** *As ratings become perfectly informative ( $\gamma \rightarrow 1$ ), the equilibrium converges to the first-best outcome.*

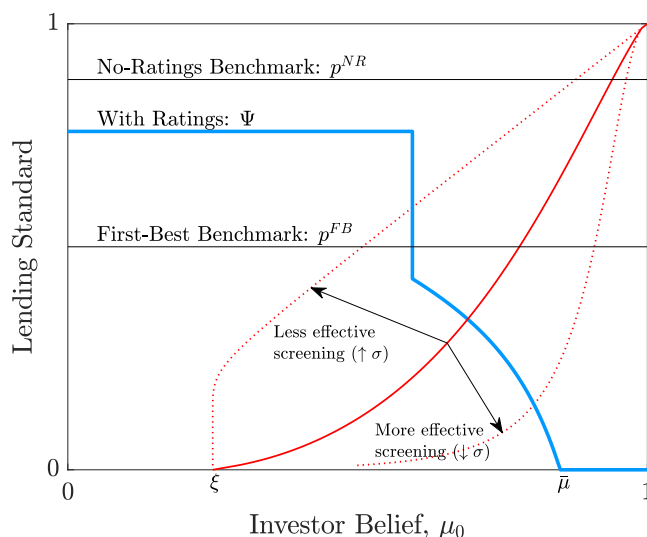
In general, there need not exist a positive measure of  $\gamma$  over which the retention and OTD equilibria exist. Consider, for instance,  $\gamma_1 = \gamma_2 = 1$ , in which case  $p^*$  converges monotonically to  $p^{FB}$ . As Proposition 3 suggests, such an example necessarily requires a relatively ineffective screening technology. Monotonic convergence of the equilibrium lending standard to  $p^{FB}$  also requires that  $\tilde{\mu}_1 \equiv \lim_{\gamma \rightarrow 1} \tilde{\mu} > p^{FB}$ . Calculation yields that  $\tilde{\mu}_1 = \frac{\sqrt{v_g v_b} - v_b}{v_g - v_b}$ , which provides a sufficient condition for an oversupply of credit as stated in the following result.

**PROPOSITION 6:** *If  $v_g v_b < 1$ , then there is an oversupply of credit for all  $\gamma$  large enough, regardless of the screening technology. Moreover:*

- (i) *The equilibrium is OTD for  $\gamma$  close to one (i.e.,  $\gamma_2 < 1$ ).*
- (ii) *The lending standard is U-shaped in  $\gamma$ .*
- (iii) *The quantity of loans originated is hump-shaped in  $\gamma$ .*

Graphically, as the rating becomes more informative,  $\Psi$  converges pointwise to  $p^{FB}$ , but it does so from above (below) to the left (right) of  $\tilde{\mu}$ . The condition  $v_g v_b < 1$  implies that  $\tilde{\mu} < p^{FB}$  for large enough  $\gamma$ , which ensures that any intersection with  $A^{-1}$  must occur at a lending standard below  $p^{FB}$ .<sup>20</sup> By contrast,  $v_g v_b > 1$  implies  $\tilde{\mu}_1 > p^{FB}$ , in which case statements (i) to (iii) of

<sup>20</sup> 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).



**Figure 6. Comparative statics on the screening technology.** This figure illustrates how the effectiveness of the screening technology affects the equilibrium lending standard. In this example, the screening technology,  $\{\psi_g, \psi_b\}$ , consists of normal density functions with means  $m_g > m_b$  and common standard deviation  $\sigma$ . (Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com))

Proposition 6 hold if and only if the screening technology is sufficiently effective (see Figure 7).

### B. Screening Technology

We now study how changes in banks' screening technology affect equilibrium outcomes and efficiency. The following definition will be useful to conduct our analysis.

**DEFINITION 2:** We say that screening technology  $\{\psi_b^i, \psi_g^i\}$  is more effective than screening technology  $\{\psi_b^j, \psi_g^j\}$  if  $A^i(p) \geq A^j(p)$  for all  $p$ .

A better screening technology means that, overall, banks become more certain of 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 zero or one, which then has implications for the  $A(\cdot)$  function that is used to pin down the equilibrium lending standard. As the screening technology becomes more effective,  $A^{-1}(\mu_0)$  decreases for all  $\mu_0 \in (\xi, 1)$  because, for any  $p \in (0, 1)$ , if the loan pool is good (bad), it is more likely that it would have generated an appraisal above (below)  $p$ .

Figure 6 illustrates how a change in banks' screening technology affects equilibrium outcomes. In this example, the screening technology,  $\{\psi_g, \psi_b\}$ , consists of normal density functions with means  $m_g > m_b$  and common standard deviation  $\sigma$ . As  $\sigma$  decreases, the screening technology becomes more effective.

Consequently, the lending standard falls, and, from Proposition 3, the equilibrium is more likely to feature an oversupply of credit relative to first-best (since  $A(p^{FB})$  increases). However, because the screening technology is more effective, the average quality of loans originated actually increases, and therefore so does the average payoff to banks. The quantity of loans originated and hence total surplus in the economy may increase or decrease. These findings are summarized in the next proposition.

PROPOSITION 7: *As the screening technology becomes more effective:*

- (1) *retention weakly decreases;*
- (2) *the lending standard falls but average loan quality increases;*
- (3) *the total quantity of originated loans may increase or decrease; and*
- (4) *if total quantity increases, then efficiency increases; otherwise, efficiency may increase or decrease.*

To further understand the impact of a more effective screening technology, it is useful to analyze the limit case in which banks observe the loan pool type prior to origination.

### C. Perfect Screening: A Limiting Case

Suppose that banks' screening technology reveals the loan type prior to origination. Although there can remain an arbitrary amount of residual uncertainty about the cash flows that the loan pool will generate, we will refer to this as the case of *perfect screening*. Note that with perfect screening, the bank acquires no additional information between origination and securitization. Technically, perfect screening is not a special case of our model as it does not satisfy our assumption that the distribution over appraisals is continuous. It is not difficult, however, to extend our analysis to address this case.<sup>21</sup> First, with perfect screening, only good loans are originated in the first best. Second, the analysis of the securitization stage remains unchanged.

Figure 6 suggests that ever-increasing the effectiveness of screening will send  $\mu_0^*$ , the proportion of good pools among all those originated, to  $\bar{\mu}$ . Notably,  $\bar{\mu} < 1$ , implying that originated loan pools are not exclusively good. Indeed, this result holds in the perfect screening case. To see why, consider the incentive to originate a loan pool, given that the type is known by the bank prior to the origination decision.

As before, given expected payoffs in the securitization stage of  $u_g^*$ ,  $u_b^*$ , a bank (weakly) prefers to originate if and only if  $u_t^* - 1 \geq 0$ . From Result 1 and Assumption 1, it is immediate that  $u_g^* > 1$  for any  $\mu_0$ . Hence, good pools are always originated when screening is perfect. What about bad pools? If no bad pools are originated, then  $\mu_0^* = 1 > \bar{\mu}$ , and the equilibrium is full pooling with zero retention, leading to  $u_b^* = v_g > 1$ . Thus, any bank with a bad

<sup>21</sup> Furthermore, it can be shown that there is continuity in the limit. That is, the equilibrium with perfect screening (as described in Proposition (8)) corresponds to the limiting equilibrium of our model as the screening technology becomes arbitrarily effective.

pool should originate, which is a contradiction. At the other extreme, if all bad pools are originated, then  $\mu_0^* = \xi$  and  $u_b^* \leq \xi v_g + (1 - \xi)v_b < 1$ , which is also a contradiction.<sup>22</sup> Therefore, equilibrium requires an interior fraction of bad pools to be originated. These findings are summarized in the following proposition.

**PROPOSITION 8:** *With perfect screening, there exists a unique equilibrium, which is characterized as follows:*

- (1) *All good loan pools are originated.*
- (2) *The average quality of originated loan pools is  $\bar{\mu} \in [\bar{\mu}, 1)$ .*
- (3) *The measure of bad loans originated is  $\frac{\xi(1-\bar{\mu})}{\bar{\mu}} \in (0, 1)$ .*
- (4) *The securitization stage involves full pooling at retention level  $x_m \in [0, \tilde{x})$ , where  $x_m$  is such that  $u_b(x_m, \bar{\mu}) = 1$ .*

Hence, when banks are very good at appraising loan opportunities prior to origination, they fund all good loan pools as well as a strictly positive amount of loan pools that they are certain are bad. This is because there is an incentive to originate until the average quality is driven down to  $\bar{\mu}$ —the level of investor beliefs at which origination of a bad pool is expected to exactly break even (i.e., when  $u_b^* = 1$ ). If investor beliefs were any higher (lower), then banks would strictly prefer to (not to) originate bad loans.

To relate these findings back to Proposition 3, notice that with perfect screening,  $A(p^{FB}) = 1$ . It follows that (8) necessarily holds for any informative rating system. Thus, with perfect screening and informative ratings, there is *always* an oversupply of credit. It is worth noting that informative ratings are critical for this result. With no-ratings and perfect screening, only good loans are originated.

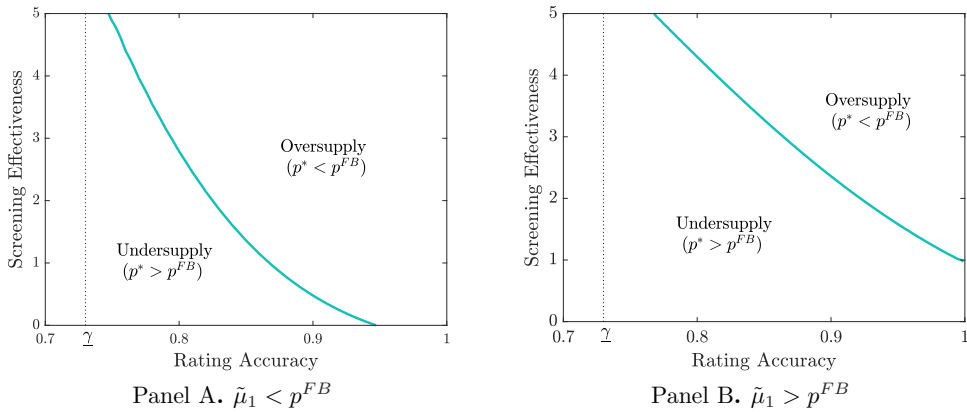
Our comparative static results from Section V.A can be extended to the case of perfect screening. In particular, Proposition 4(i) to (iii) and Proposition 5 remain true. Moreover, Proposition 6(i) to (iii) holds for any  $v_g v_b$ . Two aspects of the perfect screening case are somewhat different and worth mentioning. First, with perfect screening, a signaling equilibrium cannot exist; the equilibrium is either retention or OTD. Second, the effect of ratings accuracy on overall efficiency is now unambiguous.

**PROPOSITION 9:** *With perfect screening, overall efficiency is strictly increasing in  $\gamma$ .*

#### D. Ratings and Screening

Having separately analyzed the effects of ratings accuracy and screening effectiveness, Figure 7 illustrates the two considered together. Essentially, the figure provides an illustration of Proposition 3, where the heavy gray line corresponds to pairs of  $(\gamma, \sigma)$  such that (8) holds with equality. Recall from Lemma 2 that a minimum level of ratings accuracy, labeled  $\underline{\gamma}$  in the figure, is required

<sup>22</sup> For any  $\mu_0$ , a  $b$ -type can earn no more in the securitization stage than it does under full pooling on zero retention:  $u_b^* \leq u_b(0, \mu_0) < \mu_0 v_g + (1 - \mu_0)v_b$ .



**Figure 7. Ratings accuracy, screening effectiveness, and the supply of credit.** This figure illustrates the interaction between ratings accuracy and screening effectiveness on the equilibrium lending standard in relation to the first-best benchmark. (Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com))

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 technology. By contrast, for ratings accuracy above  $\underline{\gamma}$ , there is a strictly decreasing threshold of screening effectiveness above which the lending standard is below first-best. As ratings become more informative, a less effective screening technology is required for the equilibrium to exhibit oversupply. In Panel (A),  $\tilde{\mu}_1 < p^{FB}$  (i.e.,  $v_g v_b < 1$ ) and thus oversupply obtains for any screening precision when  $\gamma$  is large enough (see Proposition 6), whereas in Panel (B),  $\tilde{\mu}_1 > p^{FB}$  and thus for low enough screening precision, the lending standard monotonically converges to  $p^{FB}$  as  $\gamma \rightarrow 1$ .

## VI. Efficiency and Policy Analysis

After the recent financial crisis, both the United States 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.<sup>23</sup> In addition, central banks intervened 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 VI.A), disclosure requirements and regulation of CRAs (Section VI.B), and liquidity provision policies (Section VI.C).

The effect of the policies under consideration depends on the type of equilibrium in which it is introduced. Recall from Section V that the nature of origination and securitization depends on both the screening technology and

<sup>23</sup> See Schwarcz (2015) for an analysis of the regulatory changes in securitization in response to the financial crisis, both in the United States and in Europe.

the accuracy of ratings, which determine whether the equilibrium is signaling, retention, or OTD.

### A. Skin-in-the-Game Requirements

In October 2014, as part of the Dodd-Frank Wall Street Reform and Consumer Protection Act, the United States 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% of the credit risk. 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 originators and investors. A similar rule has been imposed in Europe via the Capital Requirements Regulation (CRR).

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, or proxies for the quality of the underlying cash flows.<sup>24</sup> 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 formalizes these findings.

**PROPOSITION 10:** *Imposing a retention requirement of  $x_s > 0$  increases  $\tilde{x}$  (and  $\tilde{x} > x_s$ ), but does not affect  $\bar{\mu}$ . As a result, the lending standard increases and aggregate credit supply falls. Furthermore:*

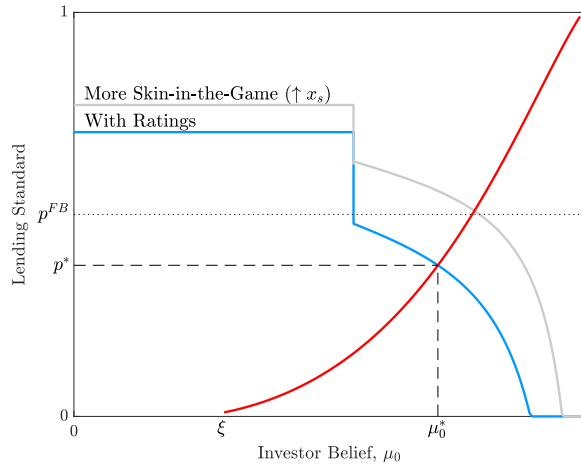
- (1) *In an OTD equilibrium, there exists a retention requirement  $x_s > 0$  that increases overall efficiency if*

$$\underbrace{(p^* v_g + (1 - p^*) v_b - 1)}_{\text{Social value of marginal loan}} \underbrace{\left( -h(p^*) \frac{dp^*}{dx_s} \right) \Big|_{x_s=0}}_{\substack{\text{Reduction in credit} \\ \text{supply for marginal } \uparrow x_s}} > \underbrace{(1 - \delta)(\mu_0^* v_g + (1 - \mu_0^*) v_b) Q(p^*)}_{\text{Social cost of marginal } \uparrow x_s}. \quad (10)$$

- (2) *In a signaling equilibrium, retention requirements worsen overall efficiency.*

Forced cash-flow retention reduces 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 to signal the same information, and thus  $\tilde{x}$  increases to levels above  $x_s$ . In turn, the reduction in

<sup>24</sup> The present regulation does make exceptions for particular asset classes. However, for a given asset class, retention rules are equal for all asset qualities.



**Figure 8. Mandatory skin in the game.** This figure illustrates the effect of imposing a retention requirement in an OTD equilibrium. (Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com))

expected payoffs from securitization due to forced retention reduces the profitability of origination. As a result,  $\Psi$  shifts upward as illustrated in Figure 8. In OTD equilibria, retention policies can increase efficiency if the marginal gain from increasing the lending standard (which was otherwise inefficiently low) more than compensates for the marginal increase in the cost of retention, as stated in (10). In signaling equilibria, however, the lending standard is optimal given equilibrium retention levels. As a result, an increase in retention levels can only worsen efficiency. It follows that such policies also increase efficiency in a retention equilibrium if and only if the efficiency gains from higher lending standards dominate the costs of increased retention.

### B. Disclosure Requirements and CRA Regulation

Rules have been adopted in both the United States 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 this information. The regulations 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, and conflicts of interest, quality of information disclosure, and so on.<sup>25</sup> Similarly, the European

<sup>25</sup> 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.

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 increasing the quality of public information available to investors. In our model, the “rating” represents any release of public information. Hence, an increase in the quality of public information corresponds to an increase in the accuracy of the rating. We analyze the effect of ratings accuracy on origination and securitization decisions and overall efficiency in Propositions 4 to 6 and 9. As we show, increasing ratings accuracy generally improves overall efficiency.<sup>26</sup> However, more informative public information leads banks to reduce retention and can lead, in turn, to lending standards that are too lax. Naturally, if originating bad loans has significant negative externalities that are not internalized by the banks, then better public information has more potential to reduce overall welfare.

In our analysis, we have taken the ratings accuracy to be an exogenous parameter. An extensive literature studies the strategic nature of CRAs and the strength of their incentives to provide unbiased information.<sup>27</sup> Inspired by the CRA models in Skreta and Veldkamp (2009), Sangiorgi and Spatt (2016), Bolton, Freixas, and Shapiro (2012), and Opp, Opp, and Harris (2013), in an earlier version of our paper, we consider extensions of our model that allow for ratings shopping and rating manipulation.<sup>28</sup> In both cases, the accuracy of ratings is endogenously determined. Provided investors are fully rational (i.e., they understand that banks can shop/manipulate ratings), allowing for either possibility effectively reduces the accuracy of the published rating. Incorporating these considerations therefore has an effect that is qualitatively similar to a reduction in the accuracy of (exogenously generated) ratings. As a result, oversight of CRAs that limits the scope for such behavior would also serve to increase the quality of publicly available information.

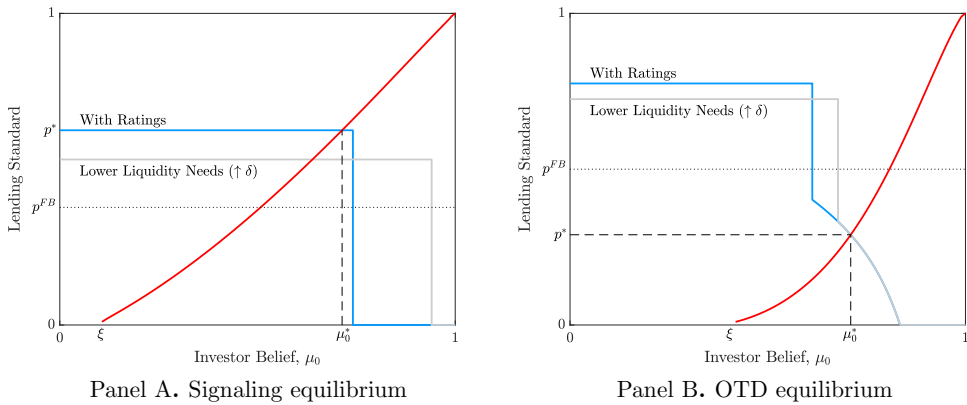
### *C. Liquidity Needs*

In an attempt to stimulate the economy by inducing banks to lend more, central banks often implement policies aimed at easing credit constraints

<sup>26</sup> The only caveat is if the screening technology is noisy and the economy is in a retention equilibrium, then it is possible to construct examples in which efficiency decreases with rating accuracy for some  $\gamma$ . A necessary condition for such an example is that the density of loans at  $p^*$  (i.e.,  $h(p^*)$ ) be sufficiently large.

<sup>27</sup> Important considerations include the role of CRA reputation and moral hazard (Mathis, McAndrews, and Rochet (2009), Bar-Isaac and Shapiro (2013), Fulghieri, Strobl, and Xia (2014), Goel and Thakor (2015), Kashyap and Kovrijnykh (2015)), feedback effects and ratings as coordination devices (Boot, Milbourn, and Schmeits (2006), Manso (2013), Goldstein and Huang (2017)), and the implications of rating-contingent regulation (Opp, Opp, and Harris (2013), Josephson and Shapiro (2019)).

<sup>28</sup> See <http://faculty.haas.berkeley.edu/bgreen/files/RatingsWP2017.pdf>.



**Figure 9. The effect of reducing liquidity needs.** (Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com))

of distressed financial institutions.<sup>29</sup> Within our model, such policies can be interpreted as increasing  $\delta$ .

**PROPOSITION 11:** *A marginal increase in  $\delta$  (i.e., reduction in liquidity needs) leads to an increase in both  $\tilde{\mu}$  and  $\tilde{x}$ . Furthermore:*

- (1) *In an OTD equilibrium, such policy has no effect.*
- (2) *In a signaling equilibrium, lending standards fall while credit supply and overall efficiency increase.*

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 the cost of retention 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 first-best (see Figure 9A). In an OTD equilibrium, a small change in  $\delta$  has no effect because banks do not retain anything in the first place (see Figure 9B). However, large enough increases in  $\delta$  can cause the economy to shift to an equilibrium with positive retention levels, which reduces the value of origination and leads 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 (10)). It follows that such policies can increase efficiency in retention equilibria with high retention levels.

<sup>29</sup> For example, in March 2008, the Federal Reserve announced the Term Securities Lending Facility, which 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.

## VII. Conclusion

We study the effect of both ratings (i.e., public information) and screening (i.e., private information) on loan origination, securitization decisions, and overall welfare. Without ratings, banks use retention as a way to signal quality to investors in the asset-backed securities market. With informative ratings, banks eschew inefficient retention in favor of relying on public information. This improves market liquidity and allocative efficiency, and typically leads to higher welfare. Yet, the introduction of public information also reduces lending standards and can lead to an oversupply of credit. This oversupply of credit is more likely to occur when banks' screening technology is more effective; as screening becomes perfect, the lending standard goes to zero and a positive fraction of bad loans are (intentionally) originated by banks. We use our model to explore the implications of policies such as mandatory skin in the game, disclosure requirements for CRAs, and liquidity provision. In doing so, we identify conditions under which such policies are welfare improving.

Initial submission: October 17, 2017; Accepted: January 15, 2019  
 Editors: Stefan Nagel, Philip Bond, Amit Seru, and Wei Xiong

## Appendix: Proofs

### A. 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  $\mu$ . 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  $\mu$  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)  $\frac{\partial}{\partial \mu}(\frac{\alpha'_g(\mu)}{\alpha'_b(\mu)}) < 0$  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  $\mu$  for any  $x \in (0, 1]$ ; and
- (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  $t \in \{b, g\}$ .

### A.1. The D1 Refinement

DEFINITION A.1: We denote by  $b_t(x, v)$  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  $\mu(x) = 0$ .

### B. Proofs for Section II

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  $p_d$  satisfy  $p_d \delta v_g + (1 - p_d) \delta v_b = 1$ . Hence, any bank with loan opportunity  $p \in (p_d, 1]$  originates in any PBE (and  $p_d < 1$  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, we have  $u_g^* > u_b^*$ . To see this, 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 nondegenerate because 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,  $p^* \leq p_d < 1$ . 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,  $u_b^* < \xi u_g^* + (1 - \xi)u_b^* \leq \xi v_g + (1 - \xi)v_b < 1$  because  $\xi v_g + (1 - \xi)v_b$  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 zero earns negative expected profit by originating, contradicting  $p^* = 0$ .  $\square$

### C. Proofs for Section III

DEFINITION A.2: Let  $\tilde{\Psi}(\mu_0) \equiv \max\{\frac{1 - u_b(0, \mu_0)}{u_g(0, \mu_0) - u_b(0, \mu_0)}, 0\}$ .

LEMMA A.1: Under any rating system,

- (1) if  $\tilde{\Psi}(\mu_0) \in (0, 1)$ , then  $\tilde{\Psi}'(\mu_0) < 0$ ;

- (2)  $\tilde{\Psi}(p^{FB}) = p^{FB}$ ; and  
 (3) with binary-symmetric ratings, if  $\tilde{\Psi}(\mu_0) > 0$ , then  $\text{sign}(\frac{d\tilde{\Psi}(\mu_0)}{d\gamma}) = \text{sign}(\mu_0 - p^{FB})$ .

PROOF: For part (1), note that  $0 < u_b(0, \mu_0) < u_g(0, \mu_0)$  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 observation, 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  $\mu_0$  for  $t = \{b, g\}$ .

For part (2), because investors earn zero expected profit, banks' expected gross return in the securitization stage must equal the expected value of their offering. Therefore, for any  $\mu_0$ ,

$$\begin{aligned}\mu_0 u_g(0, \mu_0) + (1 - \mu_0) u_b(0, \mu_0) &= \mu_0 v_g + (1 - \mu_0) v_b \\ u_g(0, \mu_0) &= v_g + \frac{1 - \mu_0}{\mu_0} (v_b - u_b(0, \mu_0)).\end{aligned}\quad (\text{A1})$$

Substituting (A1) into the definition for  $\tilde{\Psi}$  gives

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

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

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

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

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

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

PROOF OF PROPOSITION 1: 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 incentive compatibility (IC) for the  $b$ -type binds. Second, a binding IC for the  $b$ -type implies a slack IC for the  $g$ -type because  $v_g > v_b$  and thus  $v_b < x_g v_g + \delta(1 - x_g) v_g$ . The 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  $g$ -type and a  $b$ -type loan pool are  $u_g^{LC} = (1 - x_g)v_g + \delta x_g v_g > 1$  and  $u_b^{LC} = v_b < 1$ , respectively. Since there is a continuum of banks and lending standards are not observable by investors, deviations in individual bank lending decisions do not impact securitization-stage payoffs.  $\square$

#### D. Proofs for Section IV

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

$$\max_{\mu \in [0,1], x \in [0,1]} u_g(x, \mu) \quad \text{s.t.} \quad u_b(x, \mu) = k$$

denoted by  $\{\mu(k), x(k)\}$  is unique and characterized by the problem's first-order conditions (FOCs). In addition, there exists constant  $\bar{k} > v_b$  such that  $\mu(k)$  is constant for all  $k \leq \bar{k}$  and increasing otherwise, and  $x(k)$  is decreasing in  $k$ .

PROOF: We write expected securitization-stage payoffs as a function of the retention level and investors' beliefs as

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

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  $\mu$ . It will be useful to restate the  $\mathcal{M}(k)$  problem as follows:

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

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

$$\begin{aligned} & \max_{\mu, x} (1 - x)(\alpha(\mu) - \delta)(v_g - v_b) + \delta(v_g - v_b) \\ & \text{s.t.} \quad (1 - x)(\alpha_b(\mu)(v_g - v_b) + v_b) + \delta x v_b = k. \end{aligned} \quad (\text{A2})$$

Let  $\{\mu(k), x(k)\}$  satisfy the problem's FOCs of the unconstrained problem (since constraints  $x \in [0, 1]$  and  $\mu \in [0, 1]$  are not taken into account):

$$\alpha(\mu) - \delta - \frac{\alpha'(\mu)}{\alpha'_b(\mu)} \alpha_b(\mu) = \frac{\alpha'(\mu)}{\alpha'_b(\mu)} \frac{(1 - \delta)v_b}{v_g - v_b}, \quad (\text{A3})$$

$$(1 - x)(\alpha_b(\mu)(v_g - v_b) + v_b) + \delta x v_b = k. \quad (\text{A4})$$

Let  $\tilde{\mu}$  denote the solution to (A3) and let  $\bar{k} \equiv u_b(\tilde{\mu}, 0)$ . If  $k \leq \bar{k}$ ,  $\{\mu(k), x(k)\}$  are given by the above FOC and thus  $\mu(k) = \tilde{\mu}$  and is independent of  $k$ . To

see that  $\tilde{\mu}$  is unique, we analyze the left-hand side (LHS) and right-hand side (RHS) of (A3) separately. From Facts A.1.3 and A.1.2, we have that the RHS is strictly decreasing in  $\mu$ , 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  $\mu$ , and negative ( $-\delta$ ) for  $\mu = 0$ . Therefore, if a solution to (A3) exists, it is unique. Otherwise, the solution is given by the corner  $\tilde{\mu} = 1$  (which by (A4) implies a retention level of  $x^{LC}$ ). Finally, note that  $x(k)$  is strictly decreasing in  $k$  and given by (A4). If  $k > \bar{k}$ ,  $x(k) = 0$  and  $\mu(k)$  is given by the binding constraint (A4), and thus  $x(k)$  is constant and  $\mu(k)$  increasing in this scenario.

It remains to verify the second-order conditions (SOCs) of the unconstrained problem. 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  $L(x, \mu) = u_g(x, \mu) - \lambda(u_b(x, \mu) - k)$  where  $\lambda$  is the Lagrange multiplier.

$$L_{xx} = 0,$$

$$L_{\mu\mu} = (\alpha'_g(\mu) - \lambda\alpha''_b(\mu))(1 - x)(v_g - v_b),$$

$$L_{x\mu} = L_{\mu x} = -(v_g - v_b)(\alpha'_g(\mu) - \lambda\alpha'_b(\mu)) = 0.$$

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  $|BH_2| = -(\frac{\partial u_b(x, \mu)}{\partial x})^2 L_{\mu\mu} > 0$ , since  $L_{\mu\mu}|_{\{x(k), \mu(k), \lambda(k)\}} < 0$  from  $\frac{\partial}{\partial \mu}(\frac{\alpha'_g(\mu)}{\alpha'_b(\mu)}) < 0$ . Thus, the SOC's are satisfied.  $\square$

PROOF OF LEMMA 2: From Lemma A.2 and its proof, the solution to (A2) is interior with  $\tilde{\mu} < 1$  (and thus  $\tilde{x} < \bar{x}$ ) if and only if condition (A3) holds for an interior  $\mu$ , which requires that the RHS and the LHS of (A3) 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  $\mu$  for all  $\mu \in (0, 1)$ . Thus, a necessary and sufficient condition for  $\tilde{\mu} < 1$  is that LHS > RHS at  $\mu = 1$ :

$$\begin{aligned} \alpha(1) - \delta - \frac{\alpha'(1)}{\alpha'_b(1)}\alpha_b(1) &> \frac{\alpha'(1)}{\alpha'_b(1)} \frac{(1 - \delta)v_b}{v_g - v_b} \\ \iff \frac{\alpha'_b(1)}{\alpha'_g(1)} &> \frac{v_g - \delta v_b}{(1 - \delta)v_g}. \end{aligned}$$

Because  $\alpha'_b(1) = E[\Gamma(R)|b]$  and  $\alpha'_g(1) = E[\Gamma(R)|g] = 1$ , the result follows.  $\square$

PROOF OF RESULT 1: Lemma A.2 establishes that the solution locus,  $(x(k), \mu(k))_{k \in [v_b, v_g]}$ , in the present model fits the characterization of the solution locus in lemma A.5 of Daley and Green (2014): the solution is unique for all  $k$  and the locus is “L-shaped.” Lemma 2 establishes that  $\tilde{\mu} \equiv \mu(v_b) < 1$ . With these properties established, the equilibrium characterization follows from proposition 3.8 of Daley and Green (2014).  $\square$

PROOF OF PROPOSITION 2: From Result 1 and Corollaries 1 and 3, an equilibrium is pinned down by  $\mu_0^*$  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, note that  $\Psi(\mu_0)$  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  $\mu_0 > \tilde{\mu}$ ,  $\Psi(\mu_0) = \tilde{\Psi}(\mu_0) = \max\{\frac{1-u_b(0, \mu_0)}{u_g(0, \mu_0)-u_b(0, \mu_0)}, 0\}$  (Definition A.2). The function  $\tilde{\Psi}$  is continuous and strictly decreasing at all  $\mu_0 \in (0, 1)$  (Lemma A.1), which guarantees existence and uniqueness of  $\mu_0^*$  (and therefore equilibrium).  $\square$

PROOF OF PROPOSITION 3: Let  $\mu_1$  be the unique solution to  $\mu_1 = A(p^{FB})$ . By definition,  $\mu_0 < A(\mu_0)$  for all  $\mu_0 \in [\xi, 1)$ , meaning  $\mu_1 > p^{FB}$ .

First suppose that  $\tilde{\mu} < \mu_1$ . We claim that  $p^* < p^{FB}$  is implied. To see this, recall that  $(\mu_0^*, p^*)$  is the unique intersection of  $A^{-1}$  and  $\Psi$ . From Corollary 3 and the fact that  $A^{-1}$  is strictly increasing,  $\Psi(\mu_0) > p^{FB} > A^{-1}(\mu_0)$  for all  $\mu_0 < \tilde{\mu}$ . 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 Corollary 3,  $\Psi(\mu_0^*) = \tilde{\Psi}(\mu_0^*) < p^{FB}$ , where the inequality follows from Lemma A.1, parts (1) and (2).

For the other direction, if  $\mu_1 \leq \tilde{\mu}$ , then either  $\mu_1 < \tilde{\mu}$  or  $\mu_1 = \tilde{\mu}$ . In the first case (i.e.,  $\mu_1 < \tilde{\mu}$ ),  $\Psi(\mu_1)$  is unique and strictly above  $p^{FB}$ , meaning that  $\mu_0^* > \mu_1$ , which implies that  $A^{-1}(\mu_0^*) > p^{FB}$  and hence  $p^* > p^{FB}$ . If  $\mu_1 = \tilde{\mu}$ , then we have both (i)  $\Psi(\mu_0) > A^{-1}(\mu_0)$  for  $\mu_0 < \tilde{\mu}$  and (ii)  $\Psi(\mu_0) < A^{-1}(\mu_0)$  for  $\mu_0 > \tilde{\mu}$ , meaning that  $\mu_0^* = \tilde{\mu}$  and therefore  $p^* = A(\tilde{\mu}) = p^{FB}$ .  $\square$

### E. Proofs for Section V

PROOF OF LEMMA 3: For statement (i) of the lemma, we analyze how  $\{\tilde{x}, \tilde{\mu}\}$  change with ratings accuracy  $\gamma$ . After some algebra, we have that  $\{\tilde{x}, \tilde{\mu}\}$  solve

$$\frac{(1-\delta)v_b}{v_g - v_b} = (\alpha(\tilde{\mu}) - \delta) \frac{\alpha'_b(\tilde{\mu})}{\alpha'(\tilde{\mu})} - \alpha_b(\tilde{\mu}), \quad (\text{A5})$$

$$\tilde{x} = \frac{\alpha_b(\tilde{\mu})(v_g - v_b)}{\alpha_b(\tilde{\mu})(v_g - v_b) + (1-\delta)v_b}. \quad (\text{A6})$$

We proceed to characterize how this solution changes with  $\gamma$ . From the RHS of condition (A5), we have that

$$\begin{aligned}\frac{\partial RHS}{\partial \gamma}|_{\mu=\tilde{\mu}} &= \frac{\delta}{(2\gamma - 1)^3(1 - 2\tilde{\mu})}, \\ \frac{\partial RHS}{\partial \mu}|_{\mu=\tilde{\mu}} &= -\frac{\alpha(\tilde{\mu}) - \delta}{\alpha(\tilde{\mu})^2} \frac{\partial}{\partial \mu} \left( \frac{\alpha'_g(\tilde{\mu})}{\alpha'_b(\tilde{\mu})} \right).\end{aligned}$$

If  $\alpha(\tilde{\mu}) - \delta < 0$ , then from (A5), 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  $\frac{\partial RHS}{\partial \mu} < 0$ . 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  $\frac{\partial RHS}{\partial \mu} > 0$ . It follows that as  $\gamma$  increases,  $\tilde{\mu}$  has to decrease.

We have established that  $\tilde{\mu}$  decreases in  $\gamma$ . It remains to characterize how  $\tilde{x}$  changes in  $\gamma$ . Let  $RHS_c$  denote the RHS of the constraint (A6). We then have that

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

where the inequality 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)  $RHS_c$  increasing in  $\alpha_b(\mu)$  for all  $\mu \in (0, 1)$ .

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

**PROOF OF PROPOSITION 4:** We prove the proposition by analyzing the effect of a marginal increase in  $\gamma$  from a signaling, retention, and OTD equilibrium. To reduce notation, we use  $\Psi(x, \mu)$  to refer to the continuous and differentiable function  $\frac{1-u_b(x, \mu)}{u_g(x, \mu)-u_b(x, \mu)}$  when needed.

*Case 1: Signaling equilibrium.* Average quality is  $\mu_0^* = A(p^*) < \tilde{\mu}$ , where the lending standard is  $p^* = \frac{1-v_b}{u_g(\tilde{x}, \tilde{\mu})-v_b}$ . Thus, total surplus is

$$\int_{p^*}^1 (pu_g(\tilde{x}, \tilde{\mu}) + (1-p)v_b - 1)dH(p). \quad (A7)$$

First, retention falls because  $\tilde{x}$  decreases with  $\gamma$  (Lemma 3). Second, note that

$$\begin{aligned} u_g(\tilde{\mu}, \tilde{x}) &= \max_{\mu, x} (1-x)(\alpha_g(\mu)(v_g - v_b) + v_b) + \delta x v_g \\ \text{s.t. } & (1-x)(\alpha_b(\mu)(v_g - v_b) + v_b) + \delta x v_b = v_b. \end{aligned}$$

Using the envelope theorem, as  $\gamma$  increases, the payoff to  $g$ -type pools increases as well:

$$\frac{du_g^*}{d\gamma} = \frac{\partial \alpha_g(\mu^*)}{\partial \gamma} - \frac{\alpha'_g(\mu^*)}{\alpha'_b(\mu^*)} \frac{\partial \alpha_b(\mu^*)}{\partial \gamma} > 0.$$

Thus, as  $\gamma$  increases,  $u_g^*$  increases and  $u_b^* = v_b$  remains constant, which implies a decrease in the lending standard,  $p^*$ , and in average quality,  $\mu_0^*$ . Because the lending standard falls and the distribution of appraisals does not change, the quantity of loans originated,  $Q(p^*)$ , must increase. Note that the marginal pool,  $p^*$ , is zero NPV (for the bank and the economy), and thus by inspection, total surplus as given by (A7) increases.

*Case 2: Retention equilibrium.* Average quality is  $\mu_0^* = \tilde{\mu}$ , the lending standard is  $p^* = A^{-1}(\tilde{\mu})$  and retention is determined by the equilibrium condition  $\Psi(x, \tilde{\mu}) = A^{-1}(\tilde{\mu})$ . Finally, total surplus is

$$\int_{p^*}^1 (p u_g(x, \tilde{\mu}) + (1-p)u_b(x, \tilde{\mu}) - 1) dH(p).$$

After a marginal change in  $\gamma$ , the economy continues to be at a retention equilibrium. Thus, because  $\tilde{\mu}$  decreases in  $\gamma$  (Lemma 3), so does the lending standard and average loan quality, and quantity originated increases. The resulting change in retention is obtained from implicit differentiation of the equilibrium condition above to obtain

$$\frac{dx}{d\gamma} = \frac{1 - A'(\Psi(x, \tilde{\mu}))\Psi_\mu(x, \mu)}{A'(\Psi(x, \tilde{\mu}))\Psi_x(x, \mu)} \frac{d\tilde{\mu}}{d\gamma} < 0,$$

where the inequality follows from  $\Psi_x > 0$ ,  $\Psi_\mu < 0$ ,  $A' > 0$ , and  $\frac{d\tilde{\mu}}{d\gamma} < 0$ . Thus, retention decreases with ratings accuracy. Finally, we analyze the effect of a marginal increase in  $\gamma$  on total surplus

$$\begin{aligned} & \int_{p^*}^1 \left( p \left[ \frac{du_g}{dx} \frac{dx}{d\gamma} + \frac{du_g}{d\tilde{\mu}} \frac{d\tilde{\mu}}{d\gamma} \right] + (1-p) \left[ \frac{du_b}{dx} \frac{dx}{d\gamma} + \frac{du_g}{d\tilde{\mu}} \frac{d\tilde{\mu}}{d\gamma} \right] \right) dH(p) \\ & + [p^* u_g(x, \tilde{\mu}) + (1-p^*)u_b(x, \tilde{\mu}) - 1] h(p^*) \frac{dp^*}{d\gamma}. \end{aligned}$$

Because  $p^*$  is the zero NPV pool for a bank, the last term is zero. Thus, we are left to analyze the first term, which can be stated as

$$\left( -(1-\delta)(\tilde{\mu}(v_g - v_b) + v_b) \frac{dx}{d\gamma} + (1-x)(\tilde{\mu}\alpha'_g(\tilde{\mu}) + (1-\tilde{\mu})\alpha'_b(\tilde{\mu}))(v_g - v_b) \frac{d\tilde{\mu}}{d\gamma} \right) Q(p^*),$$

where the expression is obtained by computing the respective derivatives, integrating the expressions, and using the law of iterated expectations, where  $Q(p^*) = 1 - H(p^*)$ . We now plug in our expression for  $\frac{dx}{d\gamma}$  and obtain that efficiency decreases (increases) in  $\gamma$  if and only if

$$\frac{(1-x)(\tilde{\mu}\alpha'_g(\tilde{\mu}) + (1-\tilde{\mu})\alpha'_b(\tilde{\mu}))(v_g - v_b)}{(1-\delta)(\tilde{\mu}(v_g - v_b) + v_b)} > (<) \frac{1 - A'(p^*)\Psi_\mu(x, \tilde{\mu})}{A'(p^*)\Psi_x(x, \tilde{\mu})}.$$

Note that as  $A'(p^*) \rightarrow 0$ , the RHS goes to  $+\infty$ , while the LHS remains constant (and finite). Hence, efficiency clearly increases in  $\gamma$  for  $A'(p^*)$  sufficiently small. To see that the inequality can also go the other way, let  $A'(p^*)$  be arbitrarily large. It then suffices to show that the LHS of (A8)  $+\Psi_\mu/\Psi_x$  is strictly positive some for admissible parameters. Such parameter configurations are not difficult to construct. For example, taking  $\gamma \rightarrow 1$ ,  $\tilde{\mu} \rightarrow \tilde{\mu}_1$ , and  $\tilde{x} \rightarrow 0$ , we get that

$$\text{LHS of (A8)} + \Psi_\mu/\Psi_x \rightarrow (1-\delta)^2(v_g - v_b)\sqrt{v_g v_b} > 0.$$

*Case 3: OTD equilibrium.* We now have  $\mu_0^* = A(p^*) > \tilde{\mu}$ , where the lending standard is  $p^* = \frac{1-u_b(0, \mu_0^*)}{u_g(0, \mu_0^*)-u_b(0, \mu_0^*)}$  and retention is zero. Thus, total surplus is

$$\int_{p^*}^1 (p u_g(0, \mu_0^*) + (1-p)u_b(0, \mu_0^*) - 1)dH(p) = \int_{p^*}^1 (p v_g + (1-p)v_b - 1)dH(p). \quad (\text{A9})$$

First, it is immediate that retention continues to be zero since the economy stays in an OTD equilibrium in response to a marginal change in  $\gamma$ . Second, from Lemma 3, we know that in an OTD equilibrium,  $\mu_0^* \geq \hat{\mu} = \max\{p^{FB}, \tilde{\mu}\}$  and  $\Psi(\mu)$  (weakly) increases in  $\gamma$  for all  $\mu \in (\hat{\mu}, 1)$ . Since  $A^{-1}(\mu)$  monotonically increases in  $\mu$ , the new intersection must occur at a higher lending standard and average loan quality. Hence, quantity decreases. It is easy to check that total surplus increases because there are fewer loans with negative social value that are originated due to the increase in  $p^*$ , that is, those with  $p \in (p^*, p^{FB})$ .  $\square$

PROOF OF PROPOSITION 5: Recall from (A6) that

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

As  $\gamma \rightarrow 1$ ,  $\alpha_b(\mu) \rightarrow 0$  for all  $\mu \in (0, 1)$ . Hence,  $\tilde{x} \rightarrow 0$  as  $\gamma \rightarrow 0$ . 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  $\gamma \rightarrow 1$ ,  $\mathbb{E}_R[\mu_f(x, r)|t] \rightarrow \mathbb{I}_{t=g}$  and  $u_t(0, \mu_0) \rightarrow v_t$ . Hence, for all  $\mu_0$ ,  $u_t^*(\mu_0) \rightarrow v_t$  and  $\Psi(\mu_0) \rightarrow p^{FB}$  as  $\gamma \rightarrow 1$ . Because  $(\mu_0^*, p^*)$  is the unique intersection of  $\Psi$  and  $A^{-1}$ ,  $p^* \rightarrow p^{FB}$ .  $\square$

PROOF OF PROPOSITION 6: We first show that  $\lim_{\gamma \rightarrow 1} \tilde{\mu} = \tilde{\mu}_1$  and  $\tilde{\mu}_1 < p^{FB} \iff v_g v_b < 1$ . This is a matter of direct calculation. Using (A5) with a binary-symmetric rating, we get that

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

As  $\gamma \rightarrow 1$ , the first term in the numerator goes to  $\sqrt{v_g v_b}$ ; therefore,  $\lim_{\gamma \rightarrow 1} \tilde{\mu} = \frac{\sqrt{v_g v_b} - v_b}{v_g - v_b} = \tilde{\mu}_1$ , which, by inspection, is strictly less than  $p^{FB}$  if and only if  $v_g v_b < 1$ .

For (i), because  $\tilde{\mu}$  decreases with  $\gamma$  (Lemma 3), if  $v_g v_b < 1$ , then there exists  $\hat{\gamma} < 1$  such that  $\tilde{\mu} < p^{FB}$  for  $\gamma \in (\hat{\gamma}, 1)$ . By Corollary 4, it suffices to show that  $A^{-1}(\tilde{\mu}) \leq \Psi^{-}(\tilde{\mu})$  for  $\gamma \in (\hat{\gamma}, 1)$ . To this end, fix any  $\gamma \in (\hat{\gamma}, 1)$ . Recall that  $A^{-1}(\mu_0) \leq \mu_0$  for all  $\mu_0$  and thus  $A^{-1}(\tilde{\mu}) \leq \tilde{\mu} < p^{FB}$ . Further,  $\Psi^{-}(\tilde{\mu}) = \tilde{\Psi}(\tilde{\mu})$  and  $\tilde{\Psi}$  is decreasing with  $\tilde{\Psi}(p^{FB}) = p^{FB}$  (Lemma A.1). Hence,  $\Psi^{-}(\tilde{\mu}) \geq \tilde{\Psi}(p^{FB}) = p^{FB}$ . Combining the previous two inequalities yields  $A^{-1}(\tilde{\mu}) < p^{FB} < \Psi^{-}(\tilde{\mu})$  as desired.

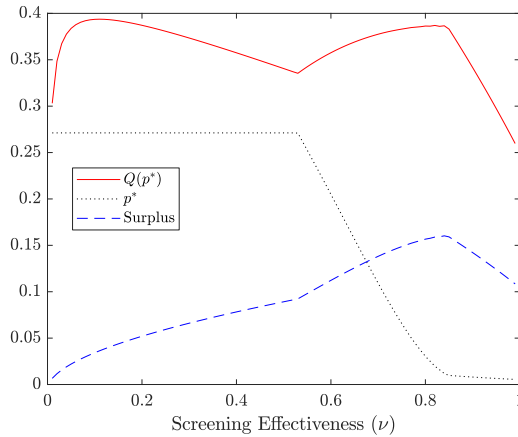
For (ii), from Proposition 4, we know that the lending standard is decreasing in  $\gamma$  when the equilibrium is signaling or retention, and is increasing when the equilibrium is OTD. Thus, it suffices to argue that equilibrium is OTD for all  $\gamma \in (\gamma_2, 1)$ . This follows from the fact that (I) the lending standard is increasing in  $\gamma$ , and hence so too is  $\mu_0^*$ , (II)  $\tilde{\mu}$  is decreasing in  $\gamma$ , and (III) the equilibrium is OTD when  $\mu_0^* > \tilde{\mu}$ .

Statement (iii) of the proposition follows immediately from (ii) because in  $Q(p^*) = 1 - H(p^*)$ , a decrease (increase) in  $p^*$  implies an increase (decrease) in the quantity of loans originated.  $\square$

PROOF OF PROPOSITION 7: Note that increasing the effectiveness of the screening technology has no impact on  $\Psi$ . For (i) and (ii), take any two screening technologies  $i$  and  $j$  such that  $i$  is more effective than  $j$ . Hence,  $A_i^{-1} \leq A_j^{-1}$ . Suppose that  $p_i^* > p_j^*$  and therefore  $A_i^{-1}(\mu_i^*) > A_j^{-1}(\mu_j^*)$ . Because  $A^{-1}$  is increasing for any screening technology and  $i$  is more effective than  $j$ , it must be that  $\mu_i^* > \mu_j^*$ . But then any element of  $\Psi(\mu_i^*)$  must be weakly less than any element of  $\Psi(\mu_j^*)$ , which contradicts  $p_i^* > p_j^*$ . Hence,  $p_i^* \leq p_j^*$ . Moreover, because  $\Psi$  is a decreasing correspondence, that  $p_i^* \leq p_j^*$  implies  $\mu_i^* \geq \mu_j^*$ . Retention being lower under screening technology  $i$  then follows immediately.

For (iii), that the total quantity of loans originated may increase or decrease can be demonstrated with an example. Let  $v_g = 2$ ,  $v_b = 0.8$ ,  $\delta = 0.7$ ,  $\xi = 0.25$ , and  $\gamma = 0.85$ . For the screening technology, or equivalently, the distribution of appraisals, let  $H$  be a beta distribution with parameters  $\alpha = \frac{\xi(1-\nu)}{\nu}$  and  $\beta = \frac{(1-\xi)(1-\nu)}{\nu}$ , where  $\nu \in (0, 1)$  and higher  $\nu$  corresponds to a more effective screening technology. Figure A1 plots total quantity as a function of  $\nu$ .

For (iv), notice that the term in brackets on the RHS of (9) (i.e., the average payoff) is higher under screening technology  $i$  because  $\mu_i^* \geq \mu_j^*$  and  $u_i^*$  is increasing in  $\mu_0^*$  and decreasing in retention. Thus, if  $Q_i(p^*) \geq Q_j(p^*)$ , then so too



**Figure A1.** Illustration of Proposition 7. (Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com))

must be the product of average payoff times quantity. That surplus can increase or decrease if quantity decreases is also shown (by example) in Figure A1.  $\square$

**PROOF OF PROPOSITION 8:** Given expected payoffs in the securitization stage of  $u_g^*, u_b^*$ , a bank (weakly) prefers to originate if and only if  $u_t^* - 1 \geq 0$ . From Result 1 and Assumption 1, it is immediate that  $u_g^* > 1$  for any  $\mu_0$ . Hence, good pools are always originated (result (i)). On the other hand, if no bad pools are originated,  $\mu_0 = 1 > \tilde{\mu}$  and the equilibrium is full pooling with zero retention, leading to  $u_b^* = v_g > 1$ ; thus, all bad pools should be originated, which is a contradiction. If all bad pools are originated,  $\mu_0 = \xi$  and  $u_b^* \leq \xi v_g + (1 - \xi)v_b < 1$ , which is another contradiction, where we use the fact that given any  $\mu_0$ , the best a  $b$ -type can do is full pooling on zero retention:  $u_b^* \leq u_b(0, \mu_0) < \mu_0 v_g + (1 - \mu_0)v_b$  (where the inequality is strict since the  $b$ -type is more likely to have a lower rating than the  $g$ -type). Thus, it must be the case that in equilibrium, an interior fraction,  $m \in (0, 1)$ , of bad pools is originated, where  $m$  is such that banks are indifferent between originating a bad pool or not:  $u_b(\frac{\psi}{\psi + m(1 - \psi)}) = 1$ . Because  $\tilde{\mu}$  is defined as the belief at which  $u_b(\tilde{\mu}) = 1$ , it follows that  $m = \frac{\psi}{1 - \psi} \frac{1 - \tilde{\mu}}{\tilde{\mu}}$  (results (ii) and (iii)).

To see that an equilibrium exists and is unique, the results from Result 1, which continue to hold with perfect screening, are useful to study how  $u_b$  changes with  $\mu$ :

$$u_b(\mu) = \begin{cases} v_b & \mu < \tilde{\mu} \\ \{(1 - x)(\alpha_b(\mu)(v_g - v_b) + v_b) + \delta x v_b | x \in [0, \tilde{x}]\} & \mu = \tilde{\mu} \\ \alpha_b(\mu)(v_g - v_b) + v_b & \mu > \tilde{\mu}. \end{cases}$$

First, for  $\mu < \tilde{\mu}$ ,  $u_b(\mu) = v_b < 1$ . Second, for  $\mu = \tilde{\mu}$ ,  $u_b$  strictly increases in  $x$  going from  $v_b < 1$  when  $x = \tilde{x}$  to  $\alpha_b(\tilde{\mu})(v_g - v_b) + v_b$  when  $x = 0$ . Finally, for  $\mu > \tilde{\mu}$ ,  $u_b$  strictly increases in  $\mu$  until  $u_b(1) = v_g > 1$ . Thus, it follows that there

exists a unique pair  $\{\bar{\mu}, x_m\}$  such that  $u_b(x_m, \bar{\mu}) = 1$ , with  $\bar{\mu} \in [\tilde{\mu}, 1]$  and  $x_m \in (\tilde{x}, 0]$  (result (iv)).  $\square$

PROOF OF PROPOSITION 9: In any equilibrium  $u_b^* = 1$ , that is, there is no loss/surplus associated with originating  $b$ -type loans. Because all  $g$ -type loans are originated, total welfare varies only with changes in  $u_g$ . In a retention equilibrium, we have that

$$u_g^* \equiv u_g(x^*, \mu^*) = \max_{\mu, x} (1-x) [\alpha_g(\mu)(v_g - v_b) + v_b] + \delta x v_g$$

$$\text{s.t. } (1-x) [\alpha_b(\mu)(v_g - v_b) + v_b] + \delta x v_b = 1.$$

Using the envelope theorem, we obtain

$$\frac{\partial u_g^*}{\partial \gamma} = \frac{\partial \alpha_g(\mu^*)}{\partial \gamma} - \frac{\alpha'_g(\mu^*)}{\alpha'_b(\mu^*)} \frac{\partial \alpha_b(\mu^*)}{\partial \gamma} > 0$$

because  $\frac{\alpha'_g(\mu^*)}{\alpha'_b(\mu^*)} > 0$ , the Lagrange multiplier on the constraint at  $\{\mu^*, x^*\}$ , and  $\frac{\partial \alpha_g}{\partial \gamma} > 0$ , while  $\frac{\partial \alpha_b}{\partial \gamma} < 0$ .

In an OTD equilibrium,  $x^* = 0$  and  $\mu^*$  is given by the solution to

$$\alpha_b(\mu^*)(v_g - v_b) + v_b = 1.$$

It is easy to check that  $\frac{\partial \alpha_b(\cdot)}{\partial \gamma} < 0$  and  $\alpha'_b(\mu) > 0$  for all  $\mu \in (0, 1)$ . It follows that  $\mu^*$  has to increase in response to an increase in  $\gamma$ . Thus,  $u_g^*$  increases due to both the direct effect of an increase in  $\gamma$  (i.e.,  $\frac{\partial \alpha_g(\cdot)}{\partial \gamma} > 0$ ) and the indirect effect of an increase in  $\mu^*$  (i.e.,  $\alpha'_g(\mu) > 0$  for all  $\mu \in (0, 1)$ ).  $\square$

## F. Proofs for Section VI

PROOF OF PROPOSITION 10: 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 Result 1 are now given by the solution to:

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

$$\text{s.t. } (1-x) (\alpha_g(\mu)(v_g - v_b) + v_b) + \delta x v_g = (1-x_s + \delta x_s) v_b, \quad (\text{A10})$$

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 Result 1, we know that the solution to (A10) 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 (A3), and thus is not affected by the skin-in-the-game rule, while  $\tilde{x}$  is now determined by the new constraint (A10). Equilibrium retention level changes in follows. For

$\mu > \tilde{\mu}$ , 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 Result 1. Since payoffs at the securitization stage,  $u_t^*$ , decrease with the skin-in-the-game rule, the lending standard weakly increases and credit supply weakly decreases.

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

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

which is equivalent to

$$\left[ \left( \mu_0^* \frac{\partial u_g^*}{\partial \mu_0^*} + (1 - \mu_0^*) \frac{\partial u_b^*}{\partial \mu_0^*} \right) A'(p^*) Q(p^*) \right] \frac{\partial p^*}{\partial x_s} \Big|_{x_s=0} > - \left( \mu_0^* \frac{\partial u_g^*}{\partial x_s} + (1 - \mu_0^*) \frac{\partial u_b^*}{\partial x_s} \right) \Big|_{x_s=0} \times Q(p^*), \quad (\text{A11})$$

which is obtained by differentiating (9) with respect to  $x_s$  around  $x_s = 0$ , and where we use 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, (A11) becomes (10). 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  $p^*$ . We obtain that the marginal gain from increasing  $p^*$  is (i.e., the LHS of (A11))

$$-(p^*v_g + (1 - p^*)v_b)h(p^*).$$

On the other hand, a marginal increase in retention levels decreases efficiency by the expected increase in the cost of retention (i.e., the RHS of (A11)),  $(1 - \delta)(\mu_0^*v_g + (1 - \mu_0^*)v_b)Q(p^*)$ . 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:

$$-(p^*v_g + (1 - p^*)v_b)h(p^*) \frac{\partial p^*}{\partial x_s} \Big|_{x_s=0} > (1 - \delta)(\mu_0^*v_g + (1 - \mu_0^*)v_b)Q(p^*).$$

In a signaling equilibrium, since  $u_t^*$  are independent of  $\mu_0^*$ , condition (A11) becomes

$$0 > - \left( \mu_0^* \frac{\partial u_g^*}{\partial x_s} + (1 - \mu_0^*) \frac{\partial u_b^*}{\partial x_s} \right) Q(p^*). \quad (\text{A12})$$

Because  $u_t^*$  decreases in retention levels, which increase for all types, the RHS of condition (A12) is positive. Thus, efficiency never increases in a signaling equilibrium.  $\square$

PROOF OF PROPOSITION 11: We analyze how  $\{\tilde{x}, \tilde{\mu}\}$  from Fact 1 change with  $\delta$ . We know that  $\tilde{\mu}$  satisfies condition (A5). Let  $RHS(\tilde{\mu}, \delta)$  denote the RHS of this condition. We show in the proof of Lemma 3 that  $\frac{\partial RHS}{\partial \tilde{\mu}}$  takes the sign of  $\delta - \alpha(\tilde{\mu})$ . In addition, we have that

$$\frac{\partial RHS}{\partial \delta} = \frac{1}{[\dots]^2} \frac{1}{\alpha'(\tilde{\mu})} [(1 - \alpha_g(\tilde{\mu}))\alpha'_b(\tilde{\mu}) + \alpha'_g(\tilde{\mu})\alpha_b(\tilde{\mu})],$$

which takes the sign of  $\alpha'(\tilde{\mu})$ . Because the LHS of the constraint is positive, the RHS must be as well, which requires that  $\alpha'(\tilde{\mu}) \times (\alpha(\tilde{\mu}) - \delta) \geq 0$ . Thus,  $\frac{\partial \tilde{\mu}}{\partial \delta} \geq 0$ . To see the effect on retention, we run a total differentiation of the constraint:

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

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

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

We have that

$$\frac{\partial u_g^*}{\partial \delta} = v_g - \frac{\alpha'_g(\tilde{\mu})}{\alpha'_b(\tilde{\mu})} v_b,$$

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

$$\frac{\alpha'_g(\tilde{\mu})}{\alpha'_b(\tilde{\mu})} = \frac{\alpha_g(\tilde{\mu})v_g + (1 - \alpha_g(\tilde{\mu}))v_b - \delta v_g}{\alpha_b(\tilde{\mu})v_g + (1 - \alpha_b(\tilde{\mu}))v_b - \delta v_b},$$

and observe that the numerator on the RHS is increasing in  $\alpha_g$ , while the denominator is increasing in  $\alpha_b$ . 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, since  $u_b = v_b$  and  $u_g^*$  increases, it follows that the lending standard falls in the signaling equilibrium. In what follows, we analyze the effect of a marginal increase in  $\delta$  on overall efficiency.

In an OTD equilibrium with  $\mu_0^* > \tilde{\mu}$ , after a marginal increase in  $\delta$ , we continue to have  $\mu_0^* > \tilde{\mu}_{\text{new}}$ , by the continuity of the RHS of condition (A5) in  $\mu$ . Thus, the economy continues to be in an OTD equilibrium. From (A9), it follows that changes in  $\delta$  do not affect efficiency, since retention is zero in such an equilibrium.

In a signaling equilibrium, following the same argument, after a marginal increase in  $\delta$ , the economy moves to a new signaling equilibrium with a lower lending standard, that is,  $\mu_{\text{new}}^* < \tilde{\mu}$ . To see that efficiency has increased, note that  $u_g^*$  has increased,  $u_b^* = v_b$  remains unchanged, and the new originated loans are all positive NPV.  $\square$

## REFERENCES

- Adelino, Manuel, Kristopher Gerardi, and Barney Hartman-Glaser, 2019, Are lemons sold first? Dynamic signaling in the mortgage market, *Journal of Financial Economics* 132, 1–25.
- Agarwal, Sumit, Yan Chang, and Abdullah Yavas, 2012, Adverse selection in mortgage securitization, *Journal of Financial Economics* 105, 640–660.
- Agarwal, Sumit, and Robert Hauswald, 2010, Distance and private information in lending, *Review of Financial Studies* 23, 2757–2788.
- Aiello, Darren, 2016, Information exploitation? A precrisis RMBS issuer's private information, UCLA Ziman Center for Real Estate Working Paper Series.
- Ashcraft, Adam, Paul Goldsmith-Pinkham, Peter Hull, and James Vickery, 2011, Credit ratings and security prices in the subprime MBS market, *American Economic Review* 101, 115–119.
- Ashcraft, Adam B., Paul Goldsmith-Pinkham, and James I. Vickery, 2010, *MBS ratings and the mortgage credit boom*, New York Federal Reserve Staff Report no. 449.
- B Ashcraft, Adam, Kunal Gooriah, and Amir Kermani, 2019, Does skin-in-the-game affect security performance?, *Journal of Financial Economics* 134, 333–354.
- Banks, Jeffrey, and Joel Sobel, 1987, Equilibrium selection in signaling games, *Econometrica* 55, 647–662.
- Bar-Isaac, Heski, and Joel Shapiro, 2013, Ratings quality over the business cycle, *Journal of Financial Economics* 108, 62–78.
- Becker, Bo, and Todd Milbourn, 2011, How did increased competition affect credit ratings?, *Journal of Financial Economics* 101, 493–514.
- Begley, Taylor A., and Amiyatosh Purnanandam, 2017, Design of financial securities: Empirical evidence from private-label RMBS deals, *Review of Financial Studies* 30, 120–161.
- Benmelech, Efraim, and Jennifer Dlugosz, 2010, The credit rating crisis, *NBER Macroeconomics Annual* 24, 161–208.
- Berger, Allen N., and Gregory F. Udell, 2004, The institutional memory hypothesis and the procyclicality of bank lending behavior, *Journal of Financial Intermediation* 13, 458–495.
- Blackwell, David, 1951, Comparison of experiments, Proceedings of the Second Berkeley Symposium on Mathematics, Statistics and Probability.
- Bolton, Patrick, Xavier Freixas, and Joel Shapiro, 2012, The credit ratings game, *Journal of Finance* 67, 85–111.
- Boot, Arnoud W.A., Todd T. Milbourn, and Anjolein Schmeits, 2006, Credit ratings as coordination mechanisms, *Review of Financial Studies* 19, 81–118.
- Botsch, Matthew, and Victoria Vanasco, 2019, Learning by lending, *Journal of Financial Intermediation* 37, 1–14.
- Chemla, Gilles, and Christopher A Hennessy, 2014, Skin in the game and moral hazard, *The Journal of Finance* 69, 1597–1641.
- Cho, In-Koo, and David M. Kreps, 1987, Signaling games and stable equilibria, *Quarterly Journal of Economics* 102, 179–221.
- Coval, Joshua, Jakub Jurek, and Erik Stafford, 2009, The economics of structured finance, *Journal of Economic Perspectives* 23, 3–26.

- Daley, Brendan, and Brett Green, 2014, Market signaling with grades, *Journal of Economic Theory* 151, 114–145.
- Daley, Brendan, Brett Green, and Victoria Vanasco, 2016, Security design with ratings, Working paper at SSRN 2940791.
- Degryse, Hans, and Steven Ongena, 2005, Distance, lending relationships, and competition, *Journal of Finance* 60, 231–266.
- Dell'Ariccia, Giovanni, Deniz Igan, and Luc A. Laeven, 2012, Credit booms and lending standards: Evidence from the subprime mortgage market, *Journal of Money, Credit and Banking* 44, 367–384.
- Dell'Ariccia, Giovanni, and Robert Marquez, 2006, Lending booms and lending standards, *Journal of Finance* 61, 2511–2546.
- DeMarzo, Peter M., 2005, The pooling and tranching of securities: A model of informed intermediation, *Review of Financial Studies* 18, 1–35.
- DeMarzo, Peter M., and Darrell Duffie, 1999, A liquidity-based model of security design, *Econometrica* 67, 65–99.
- Derrien, François, and Ambrus Kecskés, 2013, The real effects of financial shocks: Evidence from exogenous changes in analyst coverage, *Journal of Finance* 68, 1407–1440.
- Faulkender, Michael, and Mitchell A. Petersen, 2005, Does the source of capital affect capital structure?, *Review of Financial Studies* 19, 45–79.
- Fulghieri, Paulo, Günter Strobl, and Han Xia, 2014, The economics of solicited and unsolicited credit ratings, *Review of Financial Studies* 27, 484–518.
- Goel, Anand M., and Anjan V. Thakor, 2015, Information reliability and welfare: A theory of coarse credit ratings, *Journal of Financial Economics* 115, 541–557.
- Goldstein, Itay, and Chong Huang, 2017, Credit rating inflation and firms' investments, Working paper at SSRN 3082428.
- Gorton, Gary, and Andrew Metrick, 2013, Securitization, in George M. Constantinides, Milton Harris, and René M. Stulz, eds.: *Handbook of the Economics of Finance*, volume 2, 1–70 (Elsevier, Boston, MA).
- Gorton, Gary B., and George G. Pennacchi, 1995, Banks and loan sales marketing nonmarketable assets, *Journal of Monetary Economics* 35, 389–411.
- Griffin, John M., Jordan Nickerson, and Dragon Yongjun 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, John M., and Dragon Yongjun Tang, 2011, Did credit rating agencies make unbiased assumptions on CDOs?, *American Economic Review* 101, 125–130.
- Hartman-Glaser, Barney, 2017, Reputation and signaling in asset sales, *Journal of Financial Economics* 125, 245–265.
- Hartman-Glaser, Barney, Tomasz Piskorski, and Alexei Tchisty, 2012, Optimal securitization with moral hazard, *Journal of Financial Economics* 104, 186–202.
- He, Jie, Jun Qian, and Philip E. Strahan, 2011, Credit ratings and the evolution of the mortgage-backed securities market, *American Economic Review* 101, 131–135.
- Ivashina, Victoria, 2009, Asymmetric information effects on loan spreads, *Journal of Financial Economics* 92, 300–319.
- Jaffee, Dwight, Anthony Lynch, Matthew Richardson, and Stijn Van Nieuwerburgh, 2009, Mortgage origination and securitization in the financial crisis, in Viral Acharya and Matthew Richardson, eds.: *Restoring Financial Stability*, chapter 1 (John Wiley & Sons, Hoboken, NJ).
- Josephson, Jens, and Joel Shapiro, 2019, Credit ratings and structured finance, *Journal of Financial Intermediation*, (forthcoming).
- Kamenica, Emir, and Matthew Gentzkow, 2011, Bayesian persuasion, *American Economic Review* 101, 2590–2615.
- Kashyap, Anil K., and Natalia Kovrijnykh, 2015, Who should pay for credit ratings and how?, *The Review of Financial Studies* 29, 420–456.
- Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, 2010, Did securitization lead to lax screening? Evidence from subprime loans, *Quarterly Journal of Economics* 125, 307–362.
- Kraft, Pepa, 2015, Do rating agencies cater? Evidence from rating-based contracts, *Journal of Accounting and Economics* 59, 264–283.

- Leland, Hayne E., and David H. Pyle, 1977, Informational asymmetries, financial structure, and financial intermediation, *Journal of Finance* 32, 371–387.
- Loutskina, Elena, and Philip E. Strahan, 2009, Securitization and the declining impact of bank finance on loan supply: Evidence from mortgage originations, *Journal of Finance* 64, 861–889.
- Lummer, Scott L., and John J. McConnell, 1989, Further evidence on the bank lending process and the capital-market response to bank loan agreements, *Journal of Financial Economics* 25, 99–122.
- Malherbe, Frederic, 2012, Market discipline and securitization, Working paper London Business School.
- Manso, Gustavo, 2013, The feedback effects of credit ratings, *Journal of Financial Economics* 109, 535–548.
- Mathis, Jerome, James McAndrews, and Jean-Charles Rochet, 2009, Rating the raters: Are reputation concerns powerful enough to discipline rating agencies?, *Journal of monetary economics* 56, 657–674.
- Mian, Atif, and Amir Sufi, 2009, The consequences of mortgage credit expansion: Evidence from the us mortgage default crisis, *Quarterly Journal of Economics* 124, 1449–1496.
- Mikkelsen, Wayne H., and M. Megan Partch, 1986, Valuation effects of security offerings and the issuance process, *Journal of Financial Economics* 15, 31–60.
- Opp, Christian C., Marcus M. Opp, and Milton Harris, 2013, Rating agencies in the face of regulation, *Journal of Financial Economics* 108, 46–61.
- Pagano, Marco, and Paolo Volpin, 2010, Credit ratings failures and policy options, *Economic Policy* 25, 401–431.
- Parlour, Christine A., and Guillaume Plantin, 2008, Loan sales and relationship banking, *Journal of Finance* 63, 1291–1314.
- Piskorski, Tomasz, Amit Seru, and James Witkin, 2015, Asset quality misrepresentation by financial intermediaries: Evidence from the RMBS market, *Journal of Finance* 70, 2635–2678.
- Plantin, Guillaume, 2009, Learning by holding and liquidity, *Review of Economic Studies* 76, 395–412.
- Purnanandam, Amiyatosh, 2010, Originate-to-distribute model and the subprime mortgage crisis, *Review of Financial Studies* 24, 1881–1915.
- Rajan, Uday, Amit Seru, and Vikrant Vig, 2015, The failure of models that predict failure: Distance, incentives, and defaults, *Journal of Financial Economics* 115, 237–260.
- Sangiorgi, Francesco, and Chester Spatt, 2016, Opacity, credit rating shopping, and bias, *Management Science* 63, 4016–4036.
- Schwarcz, Steven L., 2015, Securitization and post-crisis financial regulation, *Cornell Law Review Online* 102, 115–138.
- Shivdasani, Anil, and Yihui Wang, 2011, Did structured credit fuel the LBO boom?, *Journal of Finance* 66, 1291–1328.
- Skreta, Vasiliki, and Laura Veldkamp, 2009, Ratings shopping and asset complexity: A theory of ratings inflation, *Journal of Monetary Economics* 56, 678–695.
- Slovin, Myron B., Marie E. Sushka, and John A. Polonchek, 1993, The value of bank durability: Borrowers as bank stakeholders, *Journal of Finance* 48, 247–266.
- Sufi, Amir, 2007, Information asymmetry and financing arrangements: Evidence from syndicated loans, *Journal of Finance* 62, 629–668.
- Vanasco, Victoria, 2017, The downside of asset screening for market liquidity, *Journal of Finance* 72, 1937–1982.

### Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

**Appendix S1:** Internet Appendix.  
**Replication code.**