Security Design with Ratings*

Brendan Daley†
Duke University

Brett Green‡
UC Berkeley

Victoria Vanasco§
Stanford University

July 2016

Abstract

We investigate the effect of ratings on the security design problem of a privately informed issuer. We find that the presence of ratings has important implications for the form of security designed (e.g., equity, debt, etc.), the level of seller retention, and price informativeness. The model rationalizes the issuance of securities that are informationally insensitive (standard debt) and informationally sensitive (levered-equity), depending on the informativeness of ratings. Furthermore, we show that the introduction of sufficiently informative ratings efficiently increases market liquidity by decreasing the reliance on inefficient retention to convey high quality. Perhaps counterintuitively, the presence of informative ratings actually decreases the amount of information transmitted to investors and prices become less informative.

*We thank Vladimir Asriyan, Peter DeMarzo, Darrell Duffie, Ben Hébert, and seminar participants at Stanford GSB and the EIFE Junior Conference in Finance for helpful feedback and suggestions. Green gratefully acknowledges support from the Fisher Center for Real Estate and Urban Economics.

†Duke University, The Fuqua School of Business, 100 Fuqua Drive, Durham, NC 27708, bd28@duke.edu

‡Haas School of Business, University of California Berkeley, 2200 Piedmont Avenue, Berkeley, CA 94720, greenb@berkeley.edu

§Stanford University, Graduate School of Business, 655 Knight Way, Stanford, CA 94305, vvanasco@stanford.edu

1 Introduction

Financial markets are plagued with informational asymmetries. For example, managers know more about their firm's future cash flows than do potential stock and bond holders, banks know more about the quality of the loans that they originate than do investors in asset backed securities (ABS) backed by these loans, an entrepreneur knows more about the quality of her project than do her angel investors. Understanding how these information asymmetries affect firms' access to funding and financial market liquidity has been a central topic in corporate finance and especially the security design literature. Two key results arise from the extensive literature on this topic: first, retaining exposure to underlying cash flows can be used as a costly "signal" that ameliorates informational frictions (Leland and Pyle (1977), DeMarzo and Duffie (1999)); second, because of its relative insensitivity to (private) information, debt often emerges as the form of security issued in equilibrium (Myers and Majluf (1984), Nachman and Noe (1994), DeMarzo (2005)).

The development of financial markets has also been accompanied by creation of additional institutions and mechanisms intended to alleviate informational frictions. One of the most salient examples are credit rating agencies (CRAs). CRAs originated in the United States at the beginning of the 20th century, and the industry rapidly grew after the passage of the Glass-Steagall Act of 1933.[1] A primary role of these agencies has been to provide an independent assessment (i.e., a rating) regarding the creditworthiness of both issuers (corporations, financial institutions, municipalities, sovereign nations, etc.) and issued securities (bonds, structured financial products, managed investments, etc.). According to White (2010), each of the three largest CRAs has ratings outstanding on tens of trillions of dollars of securities. Given the widespread use of ratings in modern times, it is natural to ask how the presence of ratings affects the aforementioned results on security design.

In this paper, we introduce ratings into a canonical security design problem. A liquidity-constrained issuer has existing assets that generate a random future cash flow X. To raise capital, the issuer can design and issue a security F, backed by her asset's cash flows, in a competitive market of risk-neutral investors.[2] The issuer has private information about the quality of her existing assets (high or low), which affects her ability to raise funds in the market because investors are concerned about purchasing a security backed by low-quality assets. To address this problem, the seller can choose a security that signals information to buyers. Furthermore, after the security is designed, an (imperfectly) informative rating

[1] By 1924, the antecedents of "The Big Three" rating companies: Moody's, Standard & Poor's, and Fitch Publishing Company had already been established.

[2] For example, the issuer could be a firm with profitable investment opportunities that raises capital by selling claims to cash flows generated by existing assets, or a bank selling asset-backed securities in order to make more loans.

R is publicly observed. Thus, there are two potential sources of information that investors receive about X: (i) the choice of security, F, and (ii) the rating, R. After observing F and R, investors bid competitively for the security and the market clearing price is determined.

As a benchmark, we first analyze the model without ratings in which we obtain the standard results.[3] Without ratings, an issuer with high-quality assets perfectly signals her "type" to investors by choosing to retain (i.e., not sell) some of her cash flows. Under general conditions, we show that the design chosen by the high-type issuer is debt, $F = \min\{d, X\}$, where the debt level d is determined by the minimum amount of cash flow retention needed to separate from the low type, who issues a claim to all of her cash flows, $F = X$. Because asset quality is perfectly revealed by the choice of security, prices reflect all available information. This information transmission, however, is not free, since retention of cash flows is costly for the high-type issuer.

We then analyze the model with ratings. We show there exists a unique equilibrium satisfying standard refinements and provide a full characterization of the equilibrium as it depends on the informativeness of ratings.[4] We then focus on how ratings affect: (1) the *form* of security isssued, (2) the *level* of cash flows retained by the issuer, and (3) the informativeness of prices.

How do ratings affect the form of security issued? When ratings are not very informative, the form of security issued remains standard debt and the issuer retains, therefore, a levered equity claim. Intuitively, when there is little information conveyed by R, the most credible signal of high quality is the seller's willingness to retain the most informationally sensitive portion of the cash flow (i.e., issue debt). However, once ratings are sufficiently informative, the opposite is true: the most credible signal of high value is to create exposure to "ratings risk" by issuing the most informationally sensitive portion of the cash flow. We characterize the precise condition on ratings informativeness at which the issuer switches from issuing a debt contract (and retaining levered equity) to issuing a levered equity claim (and retaining debt), and refer to this condition as *α-informativeness*.

How do ratings affect inefficient retention? When ratings are not very informative, the types separate by choice of retention levels. For sufficiently informative ratings, however, the high-type issuer starts to rely (at least in part) on the rating to convey information to investors. Doing so requires some degree of pooling—if all information is revealed by choice of security, there is nothing left for the ratings to convey. Therefore, the high-type issuer retains a smaller portion of the residual cash flows, thereby reducing her amount of

[3] See, for example, DeMarzo (2005).

[4] We refine the set of equilibria using D1 (Banks and Sobel, 1987; Cho and Kreps, 1987).

inefficient retention. On the other hand, because the low-type issuer also chooses this level of retention (with at least some probability) as opposed to selling everything, her inefficient retention increases. We precisely characterize the condition on ratings informativeness at which retention levels which from separating to (at least some degree of) pooling, and refer to this condition as *β-informativeness*.

Notably, β-informativeness is a strictly weaker condition than α-informativeness. Intuitively, when ratings are α-informative, the incentive to issue informationally sensitive securities dissipates if there is no uncertainty for the ratings to speak to (as is the case when equilibrium play is separating). Hence, the effect on form requires the effect on retention (but, as we will show, the reverse is not true).

How do ratings affect the informativeness of prices? When ratings are β-informative, some degree of pooling occurs in equilibrium and hence the rating conveys meaningful, but not complete, information to investors. However, the amount of information transmitted to investors in equilibrium is strictly less than without ratings—because, without ratings, equilibrium play is completely revealing. Thus, perhaps counterintuitively, the presence of informative ratings actually decreases the amount of information transmitted to investors and prices become less informative.

Finally, while our exposition focuses on ratings generated by CRAs, our framework can speak to the effect of public information in security design more generally. For example, R could represent mandated information disclosure from the issuer. As a specific illustration, in order to implement Section 942(b) of the Dodd-Frank act, the SEC introduced rules that require ABS issuers to provide standardized, asset-level information to potential investors prior to the offering. That said, given the motivation provided above, and for concreteness, we refer to R as a rating throughout.

1.1 Related Literature

This paper contributes to the extensive literature on security design in the presence of adverse selection due to the seller's private information initiated by Myers and Majluf (1984), and followed by Nachman and Noe (1994) and DeMarzo and Duffie (1999). These papers show that standard debt is the design optimally chosen by sellers, under general conditions, due to its low sensitivity to information. DeMarzo (2005) shows that debt continues to be optimal when the seller has multiple assets and can pool asset cash flows and sell tranches of this pool. Biais and Mariotti (2005) show that standard debt not only minimizes the adverse effects of asymmetric information, but can also be used to minimize the consequences of the buyers market power on gains from trade.

The optimality of debt is less robust when buyers have private information. This was highlighted by Axelson (2007), who studies security design in a setting where buyers (rather than the seller) have private signals about the underlying cash flows. He shows that the seller may choose to issue an informationally sensitive security (i.e., a call option) when there are sufficiently many buyers. In Axelson (2007), the seller also issues some debt, but the call option component becomes relatively more important as competition increases. In contrast, buyers are identically informed and perfectly competitive in our setting, and we focus on the role of a public signal. In our model, the seller uses an informationally sensitive security because creating exposure to the rating (and, unlike in Axelson (2007), retaining debt) is the most credible way to signal when the rating is sufficiently informative.

The use of informationally sensitive securities may also be desirable in order to extract information from buyers or induce them to acquire information. Yang (2015) and Yang and Zeng (2015) study security design in environments where the buyer can acquire private information about asset quality at a cost. Yang (2015) shows that when information is valuable for the seller, she may design an informationally sensitive security to provide incentives to the buyer to acquire information. In line with this result, Dang et al. (2010) show that debt is the optimal security to discourage investor's information acquisition in equilibrium.

Our model builds on the framework developed in Daley and Green (2014) who study how the presence of *grades* affects equilibrium behavior in signaling games, such as in the model of Spence (1973).[5] They show that the presence of informative grades leads to some degree of pooling on the costly signal. In their model, the signal space is one dimensional (i.e., a subset of the real line). We extend their results to a setting where the space of signals is instead a much larger set of functions from the realized cash flow to the payment of the security owner. We recover the result that an informative grade leads to some degree of pooling and characterize the type of securities on which this pooling occurs.

There is an extensive literature that has focused on understanding CRAs' incentives to issue unbiased, informative, ratings. Important considerations include reputation and moral hazard (Mathis et al., 2009; Bar-Isaac and Shapiro, 2013; Fulghieri et al., 2014; Kashyap and Kovrijnykh, 2016), coordination and feedback effects (Boot et al., 2006; Manso, 2013), investors' biases (Skreta and Veldkamp, 2009; Sangiorgi and Spatt, 2012; Bolton et al., 2012), and rating-contingent regulation (Opp et al., 2013; Josephson and Shapiro, 2015). Though these issues are surely important, we abstract from them here and model the rating as an informative public signal. We believe that understanding how unbiased signals impact the design of securities and liquidity in markets is a good starting point.

[5]How ratings may affect equilibrium behavior is also explored in Boot et al. (2006), who show that ratings can work as a coordination mechanism in situations where multiple equilibria would otherwise arise.

2 Model

There are two periods. A risk-neutral seller owns an asset that generates a random period-2 cash flow X, with support $[0, \overline{x}]$, $0 < \overline{x}$. The seller has an incentive to raise cash by issuing a claim to some portion of the cash flow (e.g., due to credit constraints, capital requirements, or productive investment opportunities). To capture this incentive, we assume the seller discounts period-2 payoffs at $\delta \in (0, 1)$, while there is a competitive market of risk-neutral buyers, whose common discount factor is 1.[6]

At the outset, the seller privately observes a signal $t \in \{L, H\}$ (also referred to as her *type*), and the conditional distribution and density functions of X are Π_t and π_t, respectively, where $\pi_t(x) > 0$ for all $x \in [0, \overline{x}]$. A signal of H indicates a higher value for the asset's cash flow according to the monotone likelihood ratio property (MLRP): $\frac{\pi_H(x)}{\pi_L(x)}$ is increasing in x. There is a common prior $\mu_0 \equiv \Pr(t = H) \in (0, 1)$.

After observing her signal, the seller selects a *security*, $F = \phi(X)$, where $\phi : [0, \overline{x}] \to [0, \overline{x}]$, to offer for sale.[7] Specifically, for any realization of the cash flow x, $\phi(x)$ is the amount paid to the purchaser of the security and $x - \phi(x)$ is the amount retained by the seller. Both the amount paid and the amount retained must be nondecreasing in x. Denote the set of all such securities by Δ.

After the seller designs the security, it receives a *rating*, which is a public signal correlated with t (Section 4 provides the formal details). Based upon the security offered for sale, F, and the realized rating, r, the buyers in the market update their (common) belief to a final belief $\mu_f(F, r) \equiv \Pr(t = H|F, r)$. Since the market is competitive, the price paid for the security is then

$$P(F|\mu_f) = E^{\mu_f}[F] = \mu_f E[F|H] + (1 - \mu_f)E[F|L]. \tag{1}$$

The seller's total payoff is $U(F, P, x) \equiv P + \delta(x - \phi(x))$.

Notice that because the seller values cash today more than buyers do, the uniquely efficient outcome is to sell the entire cash flow (i.e., $F = X$). Of course, information frictions may impede this outcome, which is the purpose of the present study.

Solution Concept

To handle the common problems posed by the freedom of off-equilibrium-path beliefs in signaling games, our solution concept is perfect Bayesian equilibrium that satisfy the D1

[6]This approach is also used in DeMarzo and Duffie (1999), Biais and Mariotti (2005), DeMarzo (2005), and Holmström and Tirole (2011), among others.

[7]That is, $\phi(\cdot)$ is a function, whereas the security F is the random variable $\phi(X)$.

refinement (Banks and Sobel, 1987; Cho and Kreps, 1987), hereafter simply referred to as *equilibrium*. Roughly put, D1 requires buyers to attribute the offer of an unexpected security to the type who is "more likely" to gain from the offer compared to her equilibrium payoff (a formal description is found in the Appendix).

Debt and Levered Equity

Before beginning the analysis, it will be useful to develop notation for two particular forms of securities.

Definition 1. *A **debt** security, F_d^D, is characterized by its **face value**, $d \in [0, \overline{x}]$, as $F_d^D = \min\{d, X\}$. Let $\Delta^D \equiv (F_d^D)_{d \in [0,\overline{x}]}$ be the set of all debt securities.*

If the seller issues a debt security with face value d, she retains a *levered equity* claim: $X - F_d^D = \max\{0, X - d\}$. Conversely, if the seller issues a levered equity security, she retains a debt claim.

Definition 2. *A **levered equity** security, F_a^A, is characterized by the face value of its residual debt, $a \in [0, \overline{x}]$, as $F_a^A = \max\{0, X - a\}$. Let $\Delta^A \equiv (F_a^A)_{a \in [0,\overline{x}]}$ be the set of all levered equity securities.*

Of course, $F_{\overline{x}}^D = F_0^A = X$. That is, selling the entire cash flow is a special case of both forms of securities.

3 The No-Ratings Benchmark

As a benchmark, consider the model without ratings (or, equivalently, one in which ratings are completely uninformative). For the moment, also suppose that only debt securities are permitted to sell. Then the choice of the seller is single-dimensional—select $d \in [0, \overline{x}]$—and the model is similar to many signaling environments that have been studied. In particular, the seller indifference curves satisfy the single-crossing property and the unique equilibrium is *least-cost separating*, as we now detail.

If the low type's private information was revealed, she would efficiently sell rights to the entire cash flow, $F = X$. The resultant price and seller payoff would be $P(X|\mu_f = 0) = E[X|L]$ referred to as her full-information payoff, which is also a lower bound on her equilibrium payoff in the true game, denoted $\underline{u} \equiv E[X|L]$. On the other hand, if offering a security F would lead the market to be convinced her asset was high-value, the low type's payoff would be $E[F|H] + \delta(E[X - F|L])$. Let, $F_{d^{LC}}^D$ be the unique debt security that equates the two:

$$E[F_{d^{LC}}^D|H] + \delta(E[X - F_{d^{LC}}^D|L]) = \underline{u}.$$

That is, $F^D_{d^{LC}}$ is unique debt security such that, if selecting it would convince the market that $t = H$, the low type would be indifferent between selecting it or getting her full-information payoff. In the least-cost separating equilibrium of this no-rating debt-only model, the low type selects $F = X$ (i.e., $F^D_{\overline{x}}$), and the high type selects $F = F^D_{d^{LC}}$.

Our first result is that this is the unique equilibrium if the seller can select any security from Δ. Intuitively, since high X-realizations are more indicative of $t = H$, the high type is more willing to retain the claim that only pays off for her in high X-realizations. Hence, issuing debt is the "least costly" way to separate from the low type. A similar result is found in DeMarzo (2005).

Proposition 3.1. *Without ratings, there is a unique equilibrium. In it, the low type selects $F = X$ and the high type selects $F = F^D_{d^{LC}}$.*

Because the equilibrium is separating, the seller's information is revealed to buyers and security prices accurately reflect all information.

4 Ratings and Informativeness

Formally, the rating is a random variable R, with type-dependent density function q_t on $\mathbb{R}$.[8] The informativeness of a rating realization, r, is captured by $\beta(r) \equiv \frac{q_L(r)}{q_H(r)}$.[9] Without loss, order the ratings such that β is weakly decreasing. For convenience, we assume that q_H, q_L are continuous almost everywhere, the informativeness of ratings is bounded,[10] and, unless otherwise stated, ratings have some informativeness.[11]

While $\beta(r)$ measures the informativeness of a particular ratings realization, r, the informativeness of the rating system, $\{q_L, q_H\}$, will be the critical aspect for our results. Blackwell (1951) and Lehmann (1988) provide the two predominant notions for what it means for one system to be unambiguously more informative than another, which endow only partial orderings of systems. We will show that there are two critical measures of informativeness for our analysis, each of which is strictly weaker than the notions of Blackwell and Lehmann and endows a complete ordering over rating systems. Both measures are differences in the expectations between the types: the first is the maximal difference in expected market posteriors, $E[\mu_f|t]$, and the second is the difference between expected likelihood ratios, $E[\beta(R)|t]$.

[8]To accommodate situations with a finite or countable set of rating outcomes $\{y_1, y_2, \dots\}$—for example, *pass/fail*, letter grades A to F, etc.—with probabilities $p_t(y_n)$, let $q_t(r) = p_t(y_n)$ for $r \in [n, n+1)$ and $q_t(r) = 0$ for all other r.

[9]If $q_H(r) = q_L(r) = 0$, we adopt the convention that $\beta(r) = 1$.

[10]That is, $\inf_r \beta(r) > 0$ and $\sup_r \beta(r) < \infty$.

[11]That is, there exists $\mathcal{R} \subseteq \mathbb{R}$ such that $\beta(r) \neq 1$ for all $r \in \mathcal{R}$ and $\int_{\mathcal{R}} q_t(r)dr > 0$ for $t \in \{L, H\}$.

Consider the market belief that $t = H$ conditional on the chosen security, F, but prior to the realization of the rating. We refer to this as the *interim* belief. An arbitrary interim belief is denoted μ, and $\mu(F)$ indicates the interim belief conditional on the seller's chosen security F. For an interim belief μ, the final market belief given $R = r$ is calculated

$$\mu_f(\mu, r) = \frac{\mu q_H(r)}{\mu q_H(r) + (1-\mu)q_L(r)} = \frac{\mu}{\mu + (1-\mu)\beta(r)}. \tag{2}$$

Now, let $\alpha_t(\mu)$ be the expected final market belief from the type t's perspective:

$$\alpha_t(\mu) \equiv E_R[\mu_f(\mu, R)|t]. \tag{3}$$

Immediately, $\alpha_H(\mu) \geq \alpha_L(\mu)$ with the inequality being strict if and only if $\mu \notin \{0, 1\}$. Also, the difference between them, denoted $\alpha(\mu) \equiv \alpha_H(\mu) - \alpha_L(\mu)$, is single-peaked.

Lemma 4.1. *α is continuous and single-peaked in μ.*

Let $\hat{\mu}$ denote, then, the unique maximizer of α, and $\hat{\alpha} \equiv \alpha(\hat{\mu})$ be the maximum difference in expected final market belief between the seller types. Our first measure of a more informative rating system is a higher value for of $\hat{\alpha}$. The key will be how this measure of informativeness compares to the gains from trade generated by the seller's need for capital.

Definition 3. *Ratings are **α-informative** if $\hat{\alpha} > \delta$.*

To ease exposition, assume that we not are in the knife-edge case where $\hat{\alpha} \neq \delta$ unless otherwise stated.

The second relevant measure of informativeness compares the expected likelihood ratios of the seller types, $E[\beta(R)|t]$. Immediately,

$$E[\beta(R)|H] = \int_{\mathbb{R}} \frac{q_L(r)}{q_H(r)} q_H(r)dr = 1 < \int_{\mathbb{R}} \frac{q_L(r)}{q_H(r)} q_L(r)dr = E[\beta(R)|L].$$

Again, the key is how the difference in them compares to payoff parameters.

Definition 4. *Ratings are **β-informative at x** if $E[\beta(R)|L] - E[\beta(R)|H] > \frac{\delta}{1-\delta}\frac{\Pi_L(x)-\Pi_H(x)}{1-\Pi_H(x)}$.*[12]

Notice that this measure is parameterized by cash flow values, x. The following lemma establishes the relative strength of the informativeness measures.

Lemma 4.2. *For any ratings system,*

- *β-informativeness at x, implies β-informativeness at all $x' < x$. The reverse is not true.*
- *α-informativeness, implies β-informativeness at all x. The reverse is not true.*

[12] For the case of $x = \overline{x}$, we use L'Hospital's rule to replace the RHS with $\frac{\delta}{1-\delta}\frac{\pi_H(\overline{x})-\pi_L(\overline{x})}{\pi_H(\overline{x})}$.

5 Equilibrium Security Design with Ratings

We begin with the statements of our main results, followed by the key piece of the analysis underlying them. More detailed results follow in the subsequent section. The theorems below decompose the effect of ratings into two parts: 1) How do ratings affect the *form* of the security (e.g., equity, debt, etc.) issued by the seller? 2) How do ratings affect the seller's *retention* level?

Theorem 1 (Effect on Security Form)**.**

(a) If ratings are <u>*not*</u> *α-informative, then in the unique equilibrium both types issue debt securities.*

(b) If ratings are α-informative, then in the unique equilibrium both types issue levered-equity securities.

Theorem 2 (Effect on Retention)**.**

(a) If ratings are <u>*not*</u> *β-informative at $x = d^{LC}$, then the unique equilibrium is least-cost separating using debt securities as in Proposition 3.1.*

(b) If ratings are β-informative at $x = d^{LC}$, then in the unique equilibrium the high type retains less than is required for separation. Hence, there is at least some degree of pooling on the chosen security.

Recall that in the no-ratings benchmark, the seller issues debt because it is the most informationally insensitive security. In addition, equilibrium play completely reveals the seller's type to the market since the high type (inefficiently) retains enough to accomplish separation. Theorems 1 and 2 demonstrate that both of these features are upended with sufficiently informative ratings.

The idea behind Theorem 1 is as follows. With no or relatively uninformative ratings, the most credible signal of high value is the seller's willingness to retain the most information sensitive portion of the cash flow (i.e., issue debt). However, if ratings are sufficiently informative, the opposite is true: the most credible signal of high value is by creating exposure to "ratings risk" by offering up for sale the most informationally sensitive portion of the cash flow (i.e., issue levered equity).

Critically, however, for the ratings to have any effect, there cannot be separation in equilibrium. Hence, ratings must also have an effect on retention—specifically, they must lower the retention of the high type and lead to at least some degree of pooling, as described in Theorem 2. In particular, the retention effect kicks in at a strictly lower level of ratings informativeness than does the effect on security form (see Lemma 4.2). Perhaps counterintuitive at first pass, then, is that ratings decrease the total information transmitted to buyers. Consequently, if ratings alter retention, securities are mispriced in that they do not reflect the total information available to market participants, unlike in the no-ratings environment.

Sketch of the Argument

To arrive at these results, first recall that the final update from the interim to final belief is a straightforward application of Bayes rule. We can use (1)-(3) to write the seller's expected payoff given any security F and interim belief μ as

$$\begin{aligned} u_t(F,\mu) &= E[P(F)|t,\mu] + \delta\left(E[X-F|t]\right) \\ &= \alpha_t(\mu)E[F|H] + (1-\alpha_t(\mu))E[F|L] + \delta\left(E[X-F|t]\right). \end{aligned} \tag{4}$$

A key to understanding Theorems 1 and 2 is the following maximization problem. Fix a candidate equilibrium expected payoff level for the low type $u_L = k$, and consider the problem

$$\begin{aligned} &\max_{F,\mu} u_H(F,\mu) \\ &s.t.\ u_L(F,\mu) = k. \end{aligned} \tag{M(k)}$$

This problem is important because, if the low type's payoff is $u_L = k$, then equilibrium requires that the high type select a security $F^*(k)$ that leads to an the interim belief $\mu(F^*(k))$ such that $(F^*(k), \mu(F^*(k)))$ is a solution to M(k) (see the proof in the Appendix).[13]

Because $u_L = k$ by the constraint in M(k) and subtracting a constant from the objective does not meaningfully alter the maximization problem, we can replace the objective in M(k) with $u_H(F,\mu) - u_L(F,\mu)$ without changing the set of solutions. Next, recalling the expression for $u_t(F,\mu)$ from (4), we have that

$$u_H(F,\mu) - u_L(F,\mu) = (\alpha(\mu)-\delta)\left(E[F|H]-E[F|L]\right) + \delta\left(E[X|H]-E[X|L]\right).$$

Finally, $\delta\left(E[X|H]-E[X|L]\right)$ is a constant unaffected by F and μ. So, the solutions to M(k)

[13]Intuitively, if $u_L = k$ but the high type does not select a security that solves M(k), then by D1, the off-path issuance of a security that does solve the problem will be attributed to the high type since she stands to gain more so than does the low type. This attribution makes the deviation profitable, breaking the equilibrium.

are identical to the solutions to,

$$\max_{F,\mu} \left(\alpha(\mu) - \delta\right) \left(E[F|H] - E[F|L]\right) \qquad (\mathrm{M}'(k))$$
$$s.t.\ u_L(F,\mu) = k.$$

This restatement sheds light on importance of α-informativeness for the form of security chosen (Theorem 1). Notice that $E[F|H] - E[F|L] \geq 0$ for all $F \in \Delta$. Therefore, if ratings are not α-informative, the objective in $\mathrm{M}'(k)$ is negative, and the high type seeks to minimize the sensitivity of the security's expected payment to her private information about the quality of the underlying asset. This is accomplished by issuing debt, which is said to be minimally *information sensitive* among security forms in Δ.

Inversely, if ratings are α-informative, the high type wants to select a security F that instead maximizes the sensitivity to the private information, so long as it leads to a belief $\mu(F)$ such that $\alpha(\mu(F)) > \delta$.[14] Intuitively, if ratings are sufficiently informative, the way for the high type to maximize her payoff is by designing a security that is sensitive to the true quality of the underlying asset, which the rating is likely to authenticate.

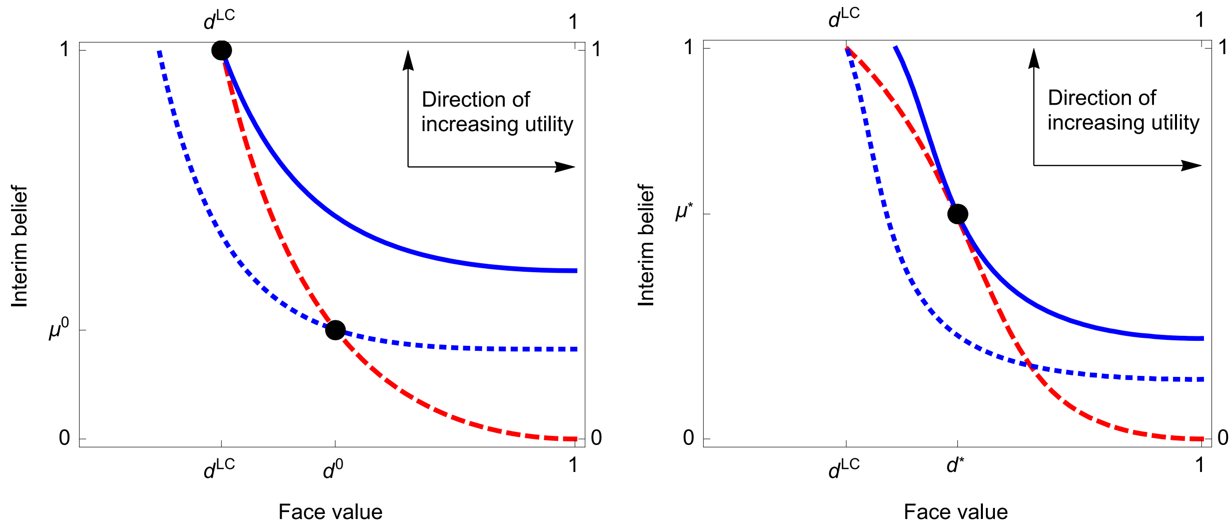
Turing to the effect on retention, consider first the case where ratings are not α-informative, meaning the solution to $\mathrm{M}(k)$ involves issuing debt. Hence, we can restate the problem as

$$\max_{d,\mu} \left(\alpha(\mu) - \delta\right) \left(E[\min\{d,X\}|H] - E[\min\{d,X\}|L]\right)$$
$$s.t.\ u_L(F^D_d,\mu) = k.$$

Now that the security-design component is single-dimensional, the solution can be illustrated graphically. Let us examine how the solution to $\mathrm{M}(\underline{u})$ depends on rating informativeness. Starting with the case of no ratings, Figure 1(a) shows the low type's indifference curve for $u_L = \underline{u}$ (in dashed-red). Therefore, the depicted (d^0, μ^0) satisfies the constraint, but it is does not solve $\mathrm{M}(\underline{u})$ since points with higher μ-values on the low type's indifference curve strictly increase u_H. In fact, without ratings, the indifference curves for the two types satisfy the single-crossing property, meaning the unique solution to $\mathrm{M}(\underline{u})$ is the boundary solution: $(d,\mu) = (d^{LC}, 1)$. This is the property underlying Proposition 3.1: without ratings, $u_L = \underline{u}$, and the high type separates by choosing $F^D_{d^{LC}}$ which leads to the separating interim belief $\mu(F^D_{d^{LC}}) = 1$.

For any (d,μ), the addition of ratings decreases u_L and increases u_H. This is depicted in

[14] This will not be feasible if k is sufficiently high. We will see that, in this case, both types sell their entire cash flows (see Proposition 6.3 below).

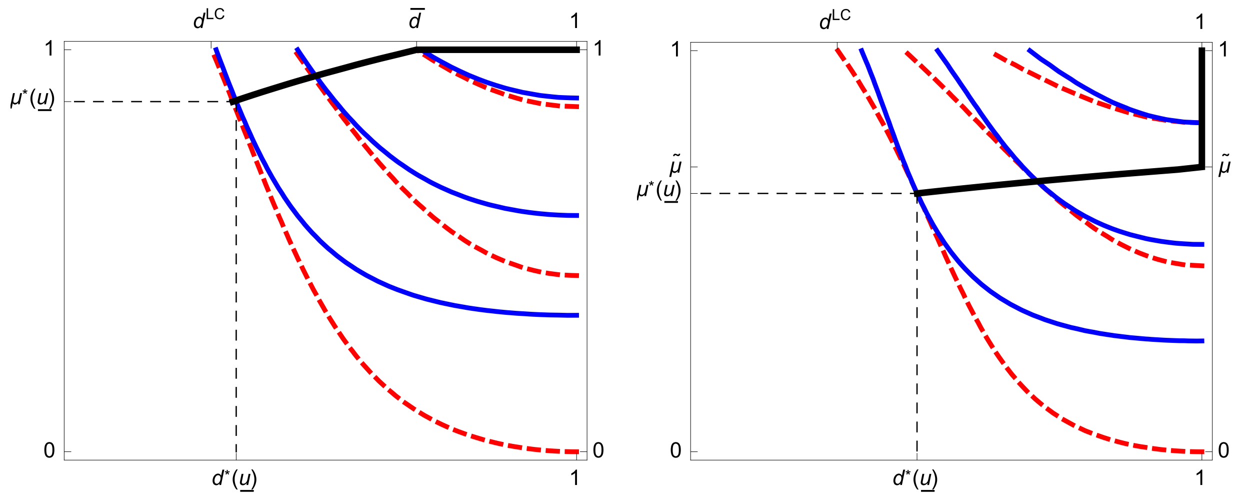


(a) No Ratings (b) Ratings β-informative at $x = d^{LC}$

FIGURE 1: Solving M($\underline{u}$). In both panels, the dashed-red indifference curve is for the low-type $u_L(F_d^D, \mu) = \underline{u}$, the solid-blue indifference curve is for the high type at the optimal value of M($\underline{u}$), and the dotted-blue indifference curve is for the high type at a suboptimal value.

Figure 1(b). The low type's indifference curve is higher than without ratings (i.e., a higher interim belief is need to offset the negative impact of the rating and keep $u_L = \underline{u}$), whereas the high type's indifference curve is lower than without ratings. Whether or not the solution to M($\underline{u}$) is altered, then, is determined by whether the high type's curve falls below the low's at $(d^{LC}, 1)$. This is why Theorem 2 hinges on β-informativeness at $x = d^{LC}$, which is the precise condition that determines whether ratings are informative enough to take the solution to M($\underline{u}$) away from the boundary. Finally, if $(d^{LC}, 1)$ does not solve M($\underline{u}$), then separation is not possible in equilibrium since (1) separation implies $u_L = \underline{u}$, and, recall, (2) if $u_L = k$, D1 requires that the high type select a security $F^*(k)$ that leads to an the interim belief $\mu(F^*(k))$ such that $(F^*(k), \mu(F^*(k)))$ is a solution to M(k).

Intuitively, as ratings become increasingly informative, the high type wishes to rely at least partially on ratings in equilibrium. But reliance on ratings requires some degree of pooling—if the types separate by choice of security, there is no useful information left for the ratings to convey. Simultaneously, reliance on the rating allows the high type to reduce her degree of inefficient retention by selecting a face value of debt that is strictly higher than the d^{LC}, as we precisely characterize in the subsequent section.



(a) β-informative at $x = d^{LC}$, but not at $x = \overline{x}$.

(b) β-informative at $x = \overline{x}$.

FIGURE 2: The effect of rating informativeness on $(d^*(\cdot), \mu^*(\cdot))$, depicted as heavy-black curve. Low type indifference curves in dashed-red, high's in solid-blue.

6 Detailed Equilibrium Characterization

We now present the full equilibrium characterization, in order of increasing ratings informativeness. Theorem 2(a) describes the equilibrium when the ratings are sufficiently uninformative (or nonexistent, subsuming Proposition 3.1 from the benchmark case).

As ratings become more informative, separation no longer holds in equilibrium. To find the equilibria, we solve $\mathrm{M}(k)$ for all $k \in [\underline{u}, \overline{u})$, where $\overline{u} \equiv E[X|H]$ is a strict upper bound on the payoff the low type can achieve in equilibrium. The properties of the solution are recorded in the following lemma and illustrated in Figure 2. Let $\underline{d}(k)$ be the unique solution to $u_L(F^D_{\underline{d}(k)}, 1) = k$, i.e, the face value of debt required for $u_L = k$ given that it engenders a belief that $t = H$ with probability one.[15]

Lemma 6.1. *Suppose ratings are not α-informative. Then,*

(a) The solution to $\mathrm{M}(k)$, denoted $(F^*(k), \mu^*(k))$, is unique for all $k \in [\underline{u}, \overline{u})$.

(b) $F^*(k)$ is a debt security, with face value $d^*(k)$, for all $k \in [\underline{u}, \overline{u})$.

(c) $\mu^*(k) < 1$ if and only if ratings are β-informative at $\underline{d}(k)$.

(d) d^* and μ^* are continuous and strictly increasing in k (modulo boundary conditions).

The final component is connecting the solutions of $\mathrm{M}(k)$ to equilibrium. For each prior belief μ_0, only a single value for the low type's payoff is consistent with equilibrium.

[15] For example, $\underline{d}(\underline{u}) = d^{LC}$.

Proposition 6.1. *Suppose ratings are β-informative at $x = d^{LC}$, but not α-informative. Then, there is unique equilibrium. In it, both types select debt securities, and*

(a) if $\mu_0 < \mu^*(\underline{u})$, the high type selects face value $d^*(\underline{u})$ and the low type mixes between face value $d^*(\underline{u})$ (i.e., retaining the same amount as the high type) and face value $\overline{x}$ (i.e., selling the entire cash flow) with probability $\frac{\mu_0(1-\mu^*(\underline{u}))}{\mu^*(\underline{u})(1-\mu_0)} \in (0,1)$ and the complementary probability, respectively.

(b) if $\mu_0 \geq \mu^*(\underline{u})$, both types select the unique face value $d^*(k)$ such that $\mu^*(k) = \mu_0$. Hence, the face value increases with μ_0.

Informative ratings decrease the reliance on inefficient retention to convey high value, and more so when the market's prior belief is more favorable. It is natural to ask then if ratings can eliminate signaling via retention, which is answered in the next proposition.

Proposition 6.2. *Suppose ratings are β-informative at $x = \overline{x}$, but not α-informative. Then, in addition to the characterization in Proposition 6.1, there exists $\tilde{\mu} \in (\hat{\mu}, 1)$ such that both types efficiently sell their entire cash flow (i.e., select $d = \overline{x}$, equivalently $F = X$) for all $\mu_0 \geq \tilde{\mu}$.*

Panel (a) of Figure 2 depicts a case in which ratings are informative enough to lead to pooling, but with positive levels of retention for all priors. That is, for all μ_0, the high type issues debt with face value $d < \overline{d}$. Panel (b) of Figure 2 increases the informativeness of ratings such that the seller efficiently sells her entire cash flow for high priors.

As seen in Theorem 1, if we further increase the informativeness of ratings, the seller no longer issues debt, but instead its complement: levered equity. We begin again by characterizing the solutions to $\mathrm{M}(k)$.

Lemma 6.2. *Suppose ratings are α-informative. Then,*

(a) The solution to $\mathrm{M}(k)$, denoted $(F^*(k), \mu^*(k))$, is unique for all $k \in [\underline{u}, \overline{u})$.

(b) $F^*(k)$ is a levered equity security with retention level $a^*(k)$, for all $k \in [\underline{u}, \overline{u})$.

(c) $\mu^*(k) < 1$ for all $k \in [\underline{u}, \overline{u})$.

(d) a^* is strictly decreasing, and μ^* strictly increasing, in k (modulo boundary conditions).

Again, for each prior belief μ_0, only a single value for the low type's payoff is consistent with equilibrium.

Proposition 6.3. *If ratings is α-informative, then there is unique equilibrium. In it, both types select levered securities, and*

(a) if $\mu_0 < \mu^*(\underline{u})$, the high type selects $a^*(\underline{u})$ and the low type mixes between $a^*(\underline{u})$ (i.e., retaining the same amount as the high type) and $a = 0$ (i.e., selling the entire cash flow) with probability $\frac{\mu_0(1-\mu^*(\underline{u}))}{\mu^*(\underline{u})(1-\mu_0)}$ and the complementary probability, respectively.

(b) if $\mu_0 \geq \mu^*(\underline{u})$, both types select the unique $a^*(k)$ such that $\mu^*(k) = \mu_0$. Hence, retention decreases with μ_0.

Since α-informativeness is stronger than β-informativeness at $x = \overline{x}$ (Lemma 4.2), we maintain the efficient-retention result of Proposition 6.2. In fact, we can also put an upper-bound on the threshold prior needed for efficiency.

Proposition 6.4. *Suppose ratings are α-informative. Then, in addition to the characterization in Proposition 6.3, there exists $\tilde{\mu} < \hat{\mu}$ such that both types efficiently sell their entire cash flow (i.e., select $a = 0$, equivalently $F = X$) for all $\mu_0 \geq \tilde{\mu}$.*

7 Conclusion

We have analyzed the effect of ratings (or other publicly disclosed information) on a general security design problem. Sufficiently informative ratings incent high-value sellers to issue informationally sensitive securities and to decrease their inefficient retention (which is completely eliminated when the initial market belief about types is favorable). Consequently, low-value sellers are induced to pool with high-value ones with positive probability, and prices are less informative than in the (separating) equilibrium of the no-ratings environment.

In this paper, the distribution of underlying cash flow is an exogenous aspect of the model. In a companion paper, we build on this analysis to ask questions about bank-loan origination and credit supply by endogenizing the distribution of cash flows. That is, we study how ratings affect which loans are originated, in addition to their securitization and retention. As ratings become more informative, the bank relies on them more and on retention less for conveying information to investors. Since retention is costly and inefficient, the rating improves efficiency in the securitization stage, but since less is being retained, the presence of informative ratings can actually reduce the incentives to issue good loans and lead to an over-supply of credit; as was witnessed in the years leading up the the recent financial crisis (Mian and Sufi, 2009).

References

Axelson, U. (2007): "Security design with investor private information," *The Journal of Finance*, 62, 2587–2632.

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

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

Biais, B. and T. Mariotti (2005): "Strategic liquidity supply and security design," *The Review of Economic Studies*, 72, 615–649.

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

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

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

Cho, I. K. and D. Kreps (1987): "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102, 179–221.

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

Dang, T. V., G. Gorton, and B. Holmström (2010): "Financial Crises and the Optimality of Debt for Liquidity," *Working paper.*

DeMarzo, P. M. (2005): "The pooling and tranching of securities: A model of informed intermediation," *Review of Financial Studies*, 18, 1–35.

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

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

Holmström, B. and J. Tirole (2011): *Inside and outside liquidity*, MIT press.

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

Karlin, S. (1968): *Total Positivity*, vol. 1, Stanford University Press.

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

Lehmann, E. (1988): "Comparing Location Experiments," *Annals of Statistics*, 16, 521–523.

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

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

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

Mian, A. and A. Sufi (2009): "The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis," *The Quarterly Journal of Economics*, 124, 1449–1496.

Myers, S. C. and N. S. Majluf (1984): "Corporate Financing and Investment Decisions when Firms Have Information that Investors do not Have," *Journal of Financial Economics*, 13, 187–221.

Nachman, D. and T. Noe (1994): "Optimal Design of Securities under Asymmetric Information," *The Review of Financial Studies*, 7, 1–44.

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

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

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

White, L. J. (2010): "The Credit Rating Agencies," *Journal of Economic Perspectives*, 24, 211–226.

Yang, M. (2015): "Optimality of debt under flexible information acquisition," *Available at SSRN 2103971.*

Yang, M. and Y. Zeng (2015): "Financing Entrepreneurial Production: Security Design with Flexible Information Acquisition," *Available at SSRN 2194194.*

A Proofs

A.1 Preliminaries and Definitions

Fact A.1. *For any $t \in \{L, H\}$ and $F \in \Delta \backslash \{0\}$,*

1. *$\alpha_t(\cdot)$ is strictly increasing.*

2. *$E[F|H] > E[F|L]$.*

3. *$u_t(F, \mu)$ is strictly increasing in μ.*

4. *There exists unique $d, a \in [0, \overline{x}]$ such that $E[F|t] = E[F_d^D|t] = E[F_a^A|t]$.*

5. *For $\gamma \in [0,1]$, let $F^\gamma \equiv (1-\gamma)F + \gamma X$. Then $F^\gamma \in \Delta$, and if $F \neq X$, then $E[F^\gamma|t]$ and $u_t(F^\gamma, 1)$ are strictly increasing in γ.*

Fact A.2. *In any PBE, $u_t \in [\underline{u}, \overline{u})$ for any $t \in \{L, H\}$.*

The D1 Refinement

Fix $k \in [\underline{u}, \overline{u})$ and $F \in \Delta$, and consider the equation $u_t(F, \mu) = k$. By Fact A.1(3), there is at most one solution for μ. If it exists, denote it by $b_t(F, k)$—that is, $u_t(F, b_t(F, k)) = k$. Next, let $B_t(F, k) \equiv \{\mu : u_t(F, \mu) > k\}$. From Fact A.1(1), the connection between b_t and B_t is immediate: if $b_t(F, k)$ exists, then $B_t(F, k) = (b_t(F, k), 1]$. If $b_t(F, k)$ fails to exist, then either $B_t(F, k) = [0, 1]$ or $B_t(F, k) = \emptyset$.

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

A.2 Proofs of Lemmas

Proof of Lemma 4.1. First, note that

$$\alpha(\mu) = \int \frac{\mu}{\mu + (1-\mu)\beta(r)} (q_H(r) - q_L(r)) dr$$

is bounded, twice continuously differentiable and meets the criteria for exchanging the order of integration and differentiation by the functional form of the integrand, which has bounded first and second partial derivatives with respect to μ. Second, it is straightforward that $\alpha(0) = \alpha(1) = 0$, and since $\alpha(\mu) > 0$ for all $\mu \in (0, 1)$ because the ratings are informative, it must be that $\alpha'(0) > 0$

and that $\alpha'(1) < 0$. Finally, note that:

$$\begin{aligned}
\alpha'(\mu) &= \int \frac{(1-\beta(r))}{(\mu+(1-\mu)\beta(r))^2} q_L(r)dr \\
\alpha''(\mu) &= -2\int \frac{(1-\beta(r))^2}{(\mu+(1-\mu)\beta(r))^3} q_L(r)dr
\end{aligned}$$

and thus $\alpha''(\mu) < 0$ for all $\mu \in (0,1)$. Thus, $\alpha(\mu)$ is single-peaked in μ. □

It is easiest to prove Lemma 4.2 after Lemmas 6.1 and 6.2.

Proof of Lemma 6.1. Recall that the solutions to M(k) are identical to the solutions to M$'(k)$, which we show are characterized by (a)-(d).

Starting with (b), fix $k \in [\underline{u}, \overline{u})$ and let $\{F^*, \mu^*\}$ be a solution to M$'(k)$, where $F^* = \phi^*(X)$ and $F^* \notin \Delta^D$. By Fact A.1(4), let d be the unique solution to $E\left[F^*|L\right] = E[F_d^D|L]$, and $\phi_d(x) = \min\{d, x\}$. Since $F^* \in \Delta$, ϕ^* is non-decreasing and $\phi^*(x) \leq x$ for all x. Thus, there exists an $\tilde{x} \in (0, \overline{x})$ such that $\phi^*(x) \leq \phi_d(x)$ for all $x \leq \tilde{x}$ with strict inequality for a positive measure set of $x < \tilde{x}$, and $\phi^*(x) \geq \phi_d(x)$ for all $x \geq \tilde{x}$ with strict inequality for a positive measure set of $x > \tilde{x}$. Next,

$$\begin{aligned}
&E[F^* - F_d^D|H] - E[F^* - F_d^D|L] \\
&\quad= \int_0^{\overline{x}} (\phi^*(x) - \phi_d(x))(\pi_H(x) - \pi_L(x))dx \\
&\quad= \int_0^{\tilde{x}} (\phi^*(x) - \phi_d(x))(\pi_H(x) - \pi_L(x))dx + \int_{\tilde{x}}^{\overline{x}} (\phi^*(x) - \phi_d(x))(\pi_H(x) - \pi_L(x))dx \\
&\quad= \int_0^{\tilde{x}} (\phi^*(x) - \phi_d(x))\left(\frac{\pi_H(x)}{\pi_L(x)} - 1\right)\pi_L(x)dx + \int_{\tilde{x}}^{\overline{x}} (\phi^*(x) - \phi_d(x))\left(\frac{\pi_H(x)}{\pi_L(x)} - 1\right)\pi_L(x)dx \\
&\quad= \int_0^{\tilde{x}} (\phi^*(x) - \phi_d(x))\left(\frac{\pi_H(x)}{\pi_L(x)}\right)\pi_L(x)dx + \int_{\tilde{x}}^{\overline{x}} (\phi^*(x) - \phi_d(x))\left(\frac{\pi_H(x)}{\pi_L(x)}\right)\pi_L(x)dx \\
&\quad> \int_0^{\tilde{x}} (\phi^*(x) - \phi_d(x))\left(\frac{\pi_H(\tilde{x})}{\pi_L(\tilde{x})}\right)\pi_L(x)dx + \int_{\tilde{x}}^{\overline{x}} (\phi^*(x) - \phi_d(x))\left(\frac{\pi_H(\tilde{x})}{\pi_L(\tilde{x})}\right)\pi_L(x)dx \\
&\quad= \left(\frac{\pi_H(\tilde{x})}{\pi_L(\tilde{x})}\right)\int (\phi^*(x) - \phi_d(x))\pi_L(x)dx = \left(\frac{\pi_H(\tilde{x})}{\pi_L(\tilde{x})}\right)\left(E[F^*|L] - E[F_d^D|L]\right) = 0,
\end{aligned}$$

where the last inequality follows from MLRP: $\frac{\pi_H(\tilde{x})}{\pi_L(\tilde{x})} = \max_{x \leq \tilde{x}} \frac{\pi_H(x)}{\pi_L(x)}$ and that $\frac{\pi_H(\tilde{x})}{\pi_L(\tilde{x})} = \min_{x \geq \tilde{x}} \frac{\pi_H(x)}{\pi_L(x)}$. Thus, the last inequality results from maximizing the weights assigned to the negative points and minimize the weights assigned to the positive ones.

It follows that $E\left[F_d^D|H\right] - E\left[F_d^D|L\right] < E\left[F^*|H\right] - E\left[F^*|L\right]$, and

$$u_L(F_d^D, \mu^*) = \alpha_L(\mu^*)\left(E[F_d^D|H] - E[F_d^D|L]\right) + (1-\delta)E[F_d^D|L] + \delta E[X|L] < k.$$

Since $u_L(F_d^D, \mu^*)$ is continuous and increasing in d, there exists security $F_{d'}^D \in \Delta^D$, with $d' > d$, such that $u_L(F^D_{d'}, \mu^*) = k$, satisfying the constraint for M$'(k)$. Further, because $E[F^D_{d'}|L] > E[F^D_d|L] = E[F^*|L]$ and $u_L(F^D_{d'}, \mu^*) = k = u_L(F^*, \mu^*)$, it must be that $E[F^D_{d'}|H] - E[F^D_{d'}|L] < E[F^*|H] - E[F^*|L]$. But then the objective in M$'(k)$ attains a higher value at $\{F^D_{d'}, \mu^*\}$ than at $\{F^*, \mu^*\}$ since $\alpha(\mu^*) - \delta < 0$. This contradicts that $\{F^*, \mu^*\}$ solves M$'(k)$. Hence, any solution to M$'(k)$ must be a debt security.

To establish (a), we show that there is a unique solution to the following re-statement of M$'(k)$:

$$\max_{d,\mu} \; v^D(d,\mu) \quad s.t. \;\; h^D(d,\mu) = k$$

where

$$\begin{aligned} v^D(d,\mu) &\equiv (\alpha(\mu) - \delta)\,(E[\min\{d,X\}|H] - E[\min\{d,X\}|L]) \\ h^D(d,\mu) &\equiv \alpha_L(\mu)\,(E[\min\{d,X\}|H] - E[\min\{d,X\}|L]) + (1-\delta)E[\min\{d,X\}|L] + \delta E[X|L]. \end{aligned}$$

Define $\mu_\ell(k)$ to be the unique solution to $u_L(X, \mu_\ell(k)) = k$. Since $X = F^D_{\overline{x}}$ and $u_L(F^D_d, \mu)$ is increasing in both d and μ, any $\{d,\mu\}$ that satisfies the constraint in M$'(k)$ must have $\mu \in [\mu_\ell(k), 1]$.

Let us look first for a solution, $\{d^*, \mu^*, \gamma^*\}$, with interior $\mu^* \in (\mu_\ell(k), 1)$, where γ denotes the multiplier on the constraint. Such a solution is characterized by the first-order conditions (second-order conditions are verified at the end of this proof):

$$\left(\alpha'(\mu) - \gamma\alpha_L'(\mu)\right)(E[\min\{d,X\}|H] - E[\min\{d,X\}|L]) = 0 \qquad \text{(FOC-}\mu\text{)}$$

$$\left(\alpha(\mu) - \delta - \gamma\alpha_L(\mu)\right)\int_d^{\overline{x}} (\pi_H(x) - \pi_L(x))\,dx - \gamma(1-\delta)\int_d^{\overline{x}} \pi_L(x)\,dx = 0 \qquad \text{(FOC-}d\text{)}$$

In any solution, $d^* > 0$, as $u_L(0,\mu) = \delta E[X|L] < \underline{u} \le k$, in violation of the constraint. From Fact A.1(2) then, the second term in the LHS of FOC-μ is positive, and $\gamma^* = \frac{\alpha'(\mu^*)}{\alpha_L'(\mu^*)}$. The second condition, FOC-d, requires $\gamma^* < 0$, since $\alpha(\mu) - \delta < 0$ for all μ. Thus, $\alpha'(\mu^*) < 0$ which implies $\mu^* \in [\hat{\mu}, 1]$. Combining these two conditions, we obtain the system of equations that determines $\{d^*, \mu^*\}$:

$$\frac{1 - \Pi_H(d^*)}{1 - \Pi_L(d^*)} - 1 = \frac{1-\delta}{(\alpha(\mu^*) - \delta)\frac{\alpha_L'(\mu^*)}{\alpha'(\mu^*)} - \alpha_L(\mu^*)} \qquad (5)$$

$$h^D(d^*, \mu^*) = k. \qquad (6)$$

Let $d_1 : \mu \mapsto d_1(\mu)$ denote the mapping from beliefs to debt levels implied by (5) and $d_2^k : \mu \mapsto d_2^k(\mu)$ denote the mapping implied by (6). First, note that $d_1(\cdot)$ is continuous and strictly increasing in μ since the LHS of (5) is strictly increasing in d due to hazard rate dominance (implied by MLRP):

$$\frac{d}{dd}LHS = \left[\frac{\pi_L(d)}{1 - \Pi_L(d)} - \frac{\pi_H(d)}{1 - \Pi_H(d)}\right]\frac{1 - \Pi_H(d)}{1 - \Pi_L(d)} > 0,$$

and the RHS of (5) is strictly increasing in μ when $\alpha(\mu) - \delta < 0$:

$$\frac{d}{d\mu} RHS = \frac{1-\delta}{\left((\alpha(\mu) - \delta) \frac{\alpha_L'(\mu)}{\alpha'(\mu)} - \alpha_L(\mu) \right)^2} \left(\frac{\alpha(\mu) - \delta}{\alpha'(\mu)^2} \left(\alpha_H''(\mu) \alpha_L'(\mu) - \alpha_H'(\mu) \alpha_L''(\mu) \right) \right) > 0,$$

since $\frac{d}{d\mu} \frac{\alpha_H'(\mu)}{\alpha_L'(\mu)} < 0$ (see Daley and Green (2014, Lemma A.1) and Karlin (1968, Chapter 3, Proposition 5.1)).

Second, $d_2^k(\cdot)$ is continuous and strictly decreasing in μ for all k since $\alpha_L'(\cdot) > 0$. Therefore, there is at most one pair $\{d^*(k), \mu^*(k)\}$ such that $d_1(\mu^*(k)) = d_2^k(\mu^*(k)) = d^*(k)$ (i.e., solves (5) and (6)). If this pair exits, it is then the unique solution to $\mathrm{M}'(k)$. If it fails to exists, the unique solution to $\mathrm{M}'(k)$ is a boundary solution. In this case: (i) if $d_1(1) < d_2^k(1)$ the unique solution is $\{d^*(k), \mu^*(k)\} = \{d_2^k(1) = \underline{d}(k), 1\}$; and (ii) if instead $d_1(1) > d_2^k(1)$, then given there is no intersection, $d_1(\mu_\ell(k)) > d_2^k(\mu_\ell(k))$ as well, and the unique solution is $\{d^*(k), \mu^*(k)\} = \{\overline{x}, \mu_\ell(k)\}$.

Next, (c) is a matter of direct calculation. For any $k \in [\underline{u}, \overline{u})$, $\mu^*(k) < 1$ if and only if $d_1(1) > d_2^k(1) = \underline{d}(k)$. This holds if and only if

$$\frac{1 - \Pi_H(\underline{d}(k))}{1 - \Pi_L(\underline{d}(k))} - 1 < \frac{\left(\frac{\alpha'(1)}{\alpha_L'(1)} \right) (1-\delta)}{\alpha(1) - \delta - \left(\frac{\alpha'(1)}{\alpha_L'(1)} \right) \alpha_L(1)}. \tag{7}$$

By straightforward calculations $\alpha(1) = 0$, $\alpha_L(1) = 1$, $\alpha_t'(1) = E[\beta(R)|t]$, and $E[\beta(R)|H] = 1$. So (7) becomes

$$E[\beta(R)|L] > \frac{(1 - \Pi_H(\underline{d}(k))) - \delta(1 - \Pi_L(\underline{d}(k)))}{1 - \Pi_H(\underline{d}(k)) - \delta(1 - \Pi_H(\underline{d}(k)))} = \frac{\delta(\Pi_L(\underline{d}(k)) - \Pi_H(\underline{d}(k)))}{(1-\delta)(1 - \Pi_H(\underline{d}(k)))} + 1,$$

which is the definition of ratings being β-informative at $\underline{d}(k)$.

Finally, for (d), note that changes in k do not impact the mapping $d_1(\cdot)$. For any two $k, k' > 0$ such that $k' > k$ we have that $d_2^{k'}(\mu^*(k)) > d_2^k(\mu^*(k)) = d_1(\mu^*(k))$. Since we have shown that $d_1(\cdot)$ is strictly increasing, it must be that (modulo boundary solutions) $\mu^*(k') > \mu^*(k)$ and thus $d^*(k') > d^*(k)$. Finally, both $\mu^*(k)$ and $d^*(k)$ are continuous in k since d_1 and d_2^k are continuous in μ and k.

Verifying Second-Order Conditions. We now verify that the solution given by the first-order conditions (5)-(6) is, in fact, a solution to $\mathrm{M}'(k)$. We verify that the determinant of the Bordered Hessian is negative at our interior critical point:

$$BH = \begin{bmatrix} 0 & h_d^D & h_\mu^D \\ h_d^D & L_{dd} & L_{d\mu} \\ h_\mu^D & L_{\mu d} & L_{\mu\mu} \end{bmatrix}$$

where $L(d,\mu) = v^D(d,\mu) - \gamma\left(h^D(d,\mu) - k\right)$.

$$\begin{aligned}
h_d^D &= \alpha_L(\mu^*)\int_{d^*}^{\overline{x}} (\pi_H(x) - \pi_L(x))\,dx + (1-\delta)\int_{d^*}^{\overline{x}} \pi_L(x)\,dx && > 0\\
d_\mu^D &= \alpha_L'(\mu^*)\left[E[\min\{d^*, X\}|H] - E[\min\{d^*, X\}|L]\right] && > 0\\
L_{dd} &= -[\left(\alpha(\mu^*) - \delta - \gamma^*\alpha_L(\mu^*)\right)(\pi_H(d^*) - \pi_L(d^*)) - \gamma^*(1-\delta)\pi_L(d^*)] && < 0\\
L_{\mu\mu} &= \left(\alpha''(\mu^*) - \gamma^*\alpha_L''(\mu^*)\right)\left[E[\min\{d^*, X\}|H] - E[\min\{d^*, X\}|L]\right] && < 0\\
L_{d\mu} = L_{\mu d} &= \left(\alpha'(\mu^*) - \gamma^*\alpha_L'(\mu^*)\right)\int_{d^*}^{\overline{x}} (\pi_H(x) - \pi_L(x))\,dx && = 0
\end{aligned}$$

Where $L_{dd} < 0$ since hazard rate dominance implies $\frac{1-\Pi_H(d)}{1-\Pi_L(d)} > \frac{\pi_H(d)}{\pi_L(d)}$ which combined with the FOC implies:

$$\frac{\pi_H(d^*)}{\pi_L(d^*)} - 1 < \frac{\gamma^*(1-\delta)}{\alpha(\mu^*) - \delta - \gamma^*\alpha_L(\mu^*)}$$

$$(\alpha(\mu^*) - \delta - \gamma^*\alpha_L(\mu^*))(\pi_H(d^*) - \pi_L(d^*)) - \pi_L(d^*)\gamma^*(1-\delta) > 0$$

where the inequality changes since since $(\alpha(\mu^*) - \delta - \gamma^*\alpha_L(\mu^*)) < 0$. Finally, $L_{\mu\mu}(\mu^*, d^*, \gamma^*) < 0$ since $\frac{d}{d\mu}\left(\frac{\alpha_H'(\mu)}{\alpha_L'(\mu)}\right) < 0$. A sufficient condition for our solution to be a local maximum is that the bordered Hessian is negative definite. That is, $|BH_1| < 0$ and $|BH_2| > 0$. It is easy to see that $|BH_1| = -(h_d^D)^2 < 0$ and that $|BH_2| = -\left(h_d^D\right)^2 L_{\mu\mu} - \left(h_\mu^D\right)^2 L_{dd} > 0$. □

Proof of Lemma 6.2. Recall that the solutions to M(k) are identical to the solutions to M$'(k)$, which we show are characterized by (a)-(d). Lemma 4.1 and α-informativeness imply that there are exactly two solutions to $\alpha(\mu) = \delta$, which we denote $\underline{\mu}, \overline{\mu}$, and that $\underline{\mu} < \hat{\mu} < \overline{\mu}$. For all $\mu \in (\underline{\mu}, \overline{\mu})$, $\alpha(\mu) > \delta$, and for all $\mu \notin [\underline{\mu}, \overline{\mu}]$, $\alpha(\mu) < \delta$. As in the proof of Lemma 6.1, define $\mu_\ell(k)$ to be the unique solution to $u_L(X, \mu_\ell(k)) = k$. Because, for any μ and $F \neq X$, $u_L(X,\mu) > u_L(F,\mu)$, and u_L is increasing μ, any $\{F, \mu\}$ that satisfies the constraint in M$'(k)$ must have $\mu \in [\mu_\ell(k), 1]$.

<u>Case 1</u>: $\mu_\ell(k) \in [0, \overline{\mu})$.
First, in any solution it must be that $\mu^*(k) \in (\underline{\mu}, \overline{\mu})$. To see this, recall that $E[F|H] - E[F|L] \geq 0$ for any $F \in \Delta$ (Fact A.1(2)). Hence, if $\mu \notin (\underline{\mu}, \overline{\mu})$, then the objective in M$'(k)$ is weakly negative. However, the objective can attain positive value. For example, select arbitrary $\mu \in (\mu_\ell(k), \overline{\mu})$ and let $\gamma > 0$ solve $u_L(\gamma X, \mu) = k$ (it is straightforward to show such a γ always exists, and is positive). Then,

$$\underbrace{(\alpha(\mu) - \delta)}_{>0}\underbrace{(E[(\gamma X|H] - E[\gamma X|L])}_{>0} > 0.$$

Because any solution must do at least this well, $\mu^*(k) \in (\underline{\mu}, \overline{\mu})$, and $\alpha(\mu^*(k)) - \delta > 0$. Notice that this establishes claim (c) of the lemma.

To establish (b), fix $k \in [\underline{u}, \overline{u})$ and let $\{F^*, \mu^*\}$ be a solution to M$'(k)$, where $F^* = \phi^*(X)$ and $F^* \notin \Delta^A$. By Fact A.1(4), let a be the unique solution to $E[F^*|L] = E[F_a^A|L]$, and $\phi_a(x) =$

$\max\{0, x-a\}$. Since $F^* \in \Delta$, $\phi^*(x) - x$ is non-decreasing and $\phi^*(x) \geq 0$ for all x. Thus, there exists an $\tilde{x} \in (0, \overline{x})$ such that $\phi^*(x) \geq \phi_a(x)$ for all $x \leq \tilde{x}$ with strict inequality for a positive measure set of $x < \tilde{x}$, and $\phi^*(x) \leq \phi_a(x)$ for all $x \geq \tilde{x}$ with strict inequality for a positive measure set of $x > \tilde{x}$. From here the calculations run analogously to those in the proof of Lemma 6.1(b), to show that $E\left[F_a^A|H\right] - E\left[F_a^A|L\right] > E\left[F^*|H\right] - E\left[F^*|L\right]$, and

$$u_L(F_a^A, \mu^*) = \alpha_L(\mu^*)\left(E[F_a^A|H] - E[F_a^A|L]\right) + (1-\delta)E[F_a^A|L] + \delta E[X|L] > k.$$

Since $u_L(F_a^A, \mu^*)$ is continuous and decreasing in a, there exists $F_{a'}^A \in \Delta^A$, with $a' > a$, such that $u_L(F_{a'}^A, \mu^*) = k$, satisfying the constraint for $\mathrm{M}'(k)$. Further, because $E[F_{a'}^A|L] < E[F_a^A|L] = E[F^*|L]$ and $u_L(F_{a'}^A, \mu^*) = k = u_L(F^*, \mu^*)$, it must be that $E[F_{a'}^A|H] - E[F_{a'}^A|L] > E[F^*|H] - E[F^*|L]$. But then the objective in $\mathrm{M}'(k)$ attains a higher value at $\{F_{a'}^A, \mu^*\}$ than at $\{F^*, \mu^*\}$ since $\alpha(\mu^*) - \delta > 0$. This contradicts that $\{F^*, \mu^*\}$ solves $\mathrm{M}'(k)$. Hence, any solution to $\mathrm{M}'(k)$ must be a levered equity security.

To establish (a), we show that there is a unique solution to the following re-statement of $\mathrm{M}'(k)$:

$$\max_{a,\mu}\ v^A(a,\mu) \quad s.t.\ \ h^A(a,\mu) = k,$$

where

$$\begin{aligned} v^A(a,\mu) &\equiv (\alpha(\mu) - \delta)\left(E[\max\{0, X-a\}|H] - E[\max\{0, X-a\}|L]\right), \\ h^A(a,\mu) &\equiv \alpha_L(\mu)\left(E[\max\{0, X-a\}|H] - E[\max\{0, X-a\}|L]\right) \\ &\qquad\qquad\qquad + (1-\delta)E[\max\{0, X-a\}|L] + \delta E[X|L]. \end{aligned}$$

Again, the constraint implies that in any solution $\mu^* \geq \mu_\ell(k)$, and we have already established that $\mu^* \in (\underline{\mu}, \overline{\mu})$.

Let us look first for a solution, $\{a^*, \mu^*, \gamma^*\}$, with interior $\mu^* \in (\mu_\ell(k), 1)$, where γ denotes the multiplier on the constraint. Such a solution is characterized by the following first-order conditions (second-order conditions are verified at the end of this proof):

$$\left(\alpha'(\mu) - \gamma\alpha_L'(\mu)\right)\left(E[\max\{0, X-a\}|H] - E[\max\{0, X-a\}|L]\right) = 0 \qquad \text{(FOC-}\mu\text{)}$$

$$\left(\alpha(\mu) - \delta - \gamma\alpha_L(\mu)\right)\int_a^{\overline{x}} \left(\pi_L(x) - \pi_H(x)\right)dx + \gamma(1-\delta)\int_a^{\overline{x}} \pi_L(x)\,dx = 0. \qquad \text{(FOC-}a\text{)}$$

In any solution, $a^* < \overline{x}$, as $u_L(0,\mu) = \delta E[X|L] < \underline{u} \leq k$, in violation of the constraint. From Fact A.1(2) then, the second term in the LHS of FOC-μ is positive, and $\gamma^* = \frac{\alpha'(\mu^*)}{\alpha_L'(\mu^*)}$. The second condition, FOC-a, requires $\gamma^* > 0$ because we have already established that $\mu^* \in (\underline{\mu}, \overline{\mu})$, meaning $\alpha(\mu^*) - \delta > 0$. Thus, $\alpha'(\mu^*) > 0$ which implies $\mu^* < \hat{\mu}$. Combining the two conditions, we obtain the following system of equations that determines $\{a^*, \mu^*\}$:

$$\frac{1-\Pi_H(a^*)}{1-\Pi_L(a^*)}-1=\frac{1-\delta}{(\alpha(\mu^*)-\delta)\frac{\alpha_L'(\mu^*)}{\alpha'(\mu^*)}-\alpha_L(\mu^*)} \tag{8}$$

$$h^A(a^*,\mu^*)=k. \tag{9}$$

Let $a_1 : \mu \mapsto a_1(\mu)$ denote the mapping from beliefs to residual debt levels implied by (8) and $a_2^k : \mu \mapsto a_2^k(\mu)$ denote the mapping implied by (9). First, note that $a_1(\cdot)$ is continuous and strictly decreasing in $\mu \in (\underline{\mu}, \overline{\mu})$ since the LHS of (9) is strictly increasing in a due to hazard rate dominance (implied by MLRP), whereas the RHS of (9) is strictly decreasing in μ when $\alpha(\mu)-\delta > 0$, as shown in the proof of Lemma 6.1(a). Second, $a_2^k(\cdot)$ is strictly increasing in μ for all k since $\alpha_L'(\cdot) > 0$. Therefore, there is at most one pair $\{a^*(k), \mu^*(k)\}$ such that $a_1(\mu^*(k)) = a_2^k(\mu^*(k)) = a^*(k)$ (i.e., solves (8) and (9)). If this pair exits, it is then the unique solution to $\mathrm{M}'(k)$. If it fails to exists, the unique solution to $\mathrm{M}'(k)$ is a boundary solution: $\mu^*(k) \in \{\mu_\ell(k), 1\}$. Since we established at the outset that $\mu^*(k) \in (\underline{\mu}, \overline{\mu})$, if the solution is boundary it must be that $\mu^*(k) = \mu_\ell(k)$ and (by definition of $\mu_\ell(k)$) $a^*(k) = 0$ (i.e., $F^*(k) = X$).

Finally, The argument for (d) is analogous to that provided for Lemma 6.1(d).

Case 2: $\mu_\ell(k) \in [\overline{\mu}, 1)$.
To begin, let $\mu_\ell(k) = \overline{\mu}$. We claim that $\{F^*, \mu^*\} = \{X, \mu_\ell(k)\}$ is the unique solution to $\mathrm{M}'(k)$. To see this, note that it is feasible (by definition of $\mu_\ell(k)$) and produces a value of 0 for the objective since $\alpha(\overline{\mu}) = \delta$. Consider now any other candidate $\{F, \mu\}$. First, if $\mu = \mu_\ell(k)$ but $F \neq X$, then $u_L(F, \mu_\ell(k)) < u_L(X, \mu_\ell(k)) = k$, in violation of the problem's constraint. Second, if $F = 0$, then for any μ, $u_L(0, \mu) = \delta E[X|L] < \underline{u} \leq k$, also in violation of the constraint. The only remaining possibility is that $F \neq 0$ and $\mu \neq \mu_\ell(k)$. In order to satisfy the constraint, it must be that $\mu \in (\mu_\ell(k), 1] = (\overline{\mu}, 1]$. But then the objective attains a negative value, establishing the claim. Notice that $X \in \Delta^A \cap \Delta^D$.

Now let $\mu_\ell(k)$ be arbitrary in $[\overline{\mu}, 1]$, and consider the restricted version of $\mathrm{M}(k)$ in which only *debt* securities can be offered:

$$\max_{d,\mu} u_H(F_d^D, \mu) \quad s.t.\ u_L(F_d^D, \mu) = k \tag{$\mathrm{M}_d(k)$}$$

For this problem, claims (a), (c), and (d) of Lemma 6.1 remain true (whereas (b) is simply assumed). Because $X = F_{\overline{x}}^D \in \Delta^D$, for $\tilde{k}$ such that $\mu_\ell(\tilde{k}) = \overline{\mu}$ the unrestricted solution is feasible in the restricted problem, so it must remain the solution in the restricted problem. Lemma 6.1(d) then implies that, in $\mathrm{M}_d(k)$, $d^*(k) = \overline{x}$ for all $k > \tilde{k}$ as well. Note that $k > \tilde{k}$ is equivalent to $\mu_\ell(k) > \overline{\mu}$. Thus we have that if $\mu_\ell(k) > \overline{\mu}$, the optimal debt security to offer has face value $d^* = \overline{x}$.

Turing back now to the unrestricted problem $\mathrm{M}'(k)$, for any $\mu_\ell(k) > \overline{\mu}$, since any feasible μ is in $[\mu_\ell(k), 1]$, $\alpha(\mu) - \delta < 0$ for all feasible μ. The same argument given for Lemma 6.1 implies that any solution must be a debt security. So restricting to debt securities is without loss, and the solution to $\mathrm{M}'(k)$ is the same as the solution to $\mathrm{M}_d(k)$, which is $\{F^*(k), \mu^*(k)\} = \{X, \mu_\ell(k)\}$. Claims (a)-(d) follow immediately.

Verifying Second-Order Conditions. We now verify that the solution given by the first-order conditions (8)-(9) is, in fact, a solution to $\mathrm{M}'(k)$. We verify that the determinant of the Bordered Hessian is negative at our interior critical point:

$$BH = \begin{bmatrix} 0 & h_a^A & h_\mu^A \\ h_a^A & L_{aa} & L_{a\mu} \\ h_\mu^A & L_{\mu a} & L_{\mu\mu} \end{bmatrix}$$

where $L(a,\mu) = v^A(a,\mu) - \gamma\left(h^A(a,\mu) - k\right)$.

$$\begin{aligned}
h_a^A &= -\alpha_L(\mu^*)\int_{a^*}^{\infty}(\pi_H(x) - \pi_L(x))\,dx - (1-\delta)\int_{a^*}^{\infty}\pi_L(x)\,dx && < 0 \\
h_\mu^A &= \alpha_L'(\mu^*)\left[E[\max\{0, X-a^*\}\,|H] - E[\max\{0, X-a^*\}\,|L]\right] && > 0 \\
L_{aa} &= \big(\alpha(\mu^*) - \delta - \gamma^*\alpha_L(\mu^*)\big)(\pi_H(a^*) - \pi_L(a^*)) + \gamma^*(1-\delta)\pi_L(a^*) && < 0 \\
L_{\mu\mu} &= \left(\alpha''(\mu^*) - \gamma^*\alpha_L''(\mu^*)\right)\left[E[[\max\{0, X-a^*\}\,|H] - E[[\max\{0, X-a^*\}\,|L]\right] && < 0 \\
L_{a\mu} &= L_{\mu a} = -\left(\alpha'(\mu^*) - \gamma^*\alpha_L'(\mu^*)\right)\int_{a^*}^{\infty}(\pi_H(x) - \pi_L(x))\,dx && = 0
\end{aligned}$$

Where $L_{aa} < 0$ since hazard rate dominance implies $\frac{1-\Pi_H(d)}{1-\Pi_L(d)} > \frac{\pi_H(a)}{\pi_L(a)}$ which combined with the FOC implies:

$$\frac{\pi_H(a^*)}{\pi_L(a^*)} - 1 < \frac{\gamma^*(1-\delta)}{\alpha(\mu^*) - \delta - \gamma^*\alpha_L(\mu^*)}$$

$$(\alpha(\mu^*) - \delta - \gamma^*\alpha_L(\mu^*))(\pi_H(a^*) - \pi_L(a^*)) - \pi_L(d^*)\gamma^*(1-\delta) < 0$$

Finally, $L_{\mu\mu}(\mu^*, a^*, \gamma^*) < 0$ since $\frac{d}{d\mu}\left(\frac{\alpha_H'(\mu)}{\alpha_L'(\mu)}\right) < 0$. A sufficient condition for our solution to be a local maximum is that the bordered Hessian is negative definite. That is, $|BH_1| < 0$ and $|BH_2| > 0$. It is easy to see that $|BH_1| = -(h_a^A)^2 < 0$ and that $|BH_2| = -\left(h_a^A\right)^2 L_{\mu\mu} - \left(h_\mu^A\right)^2 L_{aa} > 0$.

□

Proof of Lemma 4.2. For the first claim, it is sufficient to show that $\frac{\Pi_L(x)-\Pi_H(x)}{1-\Pi_H(x)}$ is nondecreasing in x. Taking the derivative yields

$$\begin{aligned}
&\frac{(1-\Pi_H(x))\pi_L(x) - (1-\Pi_L(x))\pi_H(x)}{(\Pi_H(x)-1)^2} \geq 0 \\
\iff\ & (1-\Pi_H(x))\pi_L(x) - (1-\Pi_L(x))\pi_H(x) \geq 0 \\
\iff\ & \frac{\pi_L(x)}{1-\Pi_L(x)} \geq \frac{\pi_H(x)}{1-\Pi_H(x)},
\end{aligned}$$

where the last inequality is the definition of hazard rate dominance, which holds due to MLRP.

For the second claim, first suppose that ratings are α-informative. Then, from the proof of Lemma 6.2, there exists $\tilde{k}$ such that $F^*(k) = X$ and $\mu^*(k) = \mu_\ell(k) < 1$ for all $k \geq \tilde{k}$. Consider now the restricted version of $\mathrm{M}(k)$ in which only debt securities can be offered:

$$\begin{aligned} &\max_{d,\mu} u_H(F_d^D, \mu) \\ &s.t.\ u_L(F_d^D, \mu) = k. \end{aligned} \tag{$\mathrm{M}_d(k)$}$$

For this problem, claims (a), (c), and (d) of Lemma 6.1 remain true (whereas (b) is simply assumed). Because $X \in \Delta^D$, for all $k \geq \tilde{k}$ the unrestricted solution is feasible in the restricted problem, so it must remain the solution in the restricted problem. From Lemma 6.1(d), then, in $\mathrm{M}_d(k)$, $\mu^*(k) \leq \mu_\ell(\tilde{k}) < 1$ for all $k \leq \tilde{k}$. Lemma 6.1(d) then implies that ratings are β-informative all $x \in [0, \overline{x}]$.

That the reverse does not hold requires a only a counterexample in which, using the first claim, ratings are β-informative at $\overline{x}$ but are not α-informative. Let ratings be binary and symmetric[16]: $R \in \{l, h\}$ and $\Pr(R = h|t = H) = \Pr(R = l|t = L) \equiv p \in (\frac{1}{2}, 1)$. Since

$$\frac{\delta}{1-\delta} \geq \frac{\delta}{1-\delta} \frac{\pi_H(\overline{x}) - \pi_L(\overline{x})}{\pi_H(\overline{x})},$$

a sufficient condition for β-informativeness at $\overline{x}$ is

$$E[\beta(R)|L] - \underbrace{E[\beta(R)|H]}_{=1} > \frac{\delta}{1-\delta}, \tag{10}$$

which is equivalent to

$$\frac{E[\beta(R)|L] - 1}{E[\beta(R)|L]} = \frac{(1-2p)^2}{1-3p(1-p)} > \delta.$$

Next, α-informativeness requires $\alpha(\hat{\mu}) > \delta$. For binary-symmetric ratings, $\hat{\mu} = \frac{1}{2}$ for all p, and the requirement is $\alpha(\frac{1}{2}) = (1-2p)^2 > \delta$. Since, for all $p \in (0, 1)$,

$$0 < (1-2p)^2 < \frac{(1-2p)^2}{1-3p(1-p)} < 1,$$

β-informativeness holds for all x, while α-informativeness fails, when $\delta \in \left((1-2p)^2, \frac{(1-2p)^2}{1-3p(1-p)}\right)$, producing the counterexample. $\square$

[16]See footnote 8.

A.3 Proofs of Propositions

Proof of Proposition 3.1. Having no ratings means that ratings are not β-informative at any x. Hence, by Lemma 6.1(c), $\mu^*(k) = 1$ and $d^*(k) = \underline{d}(k)$ for all $k \in [\underline{u}, \overline{u})$. The proposition then follows directly from Proposition 6.1. $\square$

Proof of Proposition 6.1. From Lemma 6.1 we have that $F^*(k)$ and $\mu^*(k)$ are unique for all $k \in [\underline{u}, \overline{u})$. Let S_t be the support of the type t's strategy. In the proposed unique equilibrium, the high type plays a pure strategy, denoted it F_H, so $S_H = \{F_H\}$, and $S_L \subseteq \{X, F_H\}$. For completeness, we must specify the off-path beliefs: $\mu(F) = 0$ for all $F \neq F_H$.

Verifying that the proposed profile is a PBE is straightforward. To see that it satisfies D1, fix a μ_0 and consider the proposition's unique equilibrium candidate. Denote the high type's equilibrium payoff $\hat{u}_H$ and low type's equilibrium payoff k, so $F_H = F^*(k)$. Let F be an arbitrary security in Δ such that $F \neq F^*(k)$. First, if $B_L(F, k) = [0, 1]$, then the low type could deviate to F and obtain a payoff strictly greater than k, regardless of $\mu(F)$, breaking the PBE. Hence, either $b_L(F, k) \in [0, 1]$ exits or $u_L(F, 1) < k$. If $b_L(F, k)$ exits, then since $\{F^*(k), \mu^*(k)\}$ is the unique solution to M(k), $u_H(F, b_L(F, k)) < u_H(F^*(k), \mu^*(k)) = \hat{u}_H$. By Fact A.1(3) then, $b_H(F, \hat{u}_H) > b_L(F, k)$ (or $B_H(F, \hat{u}_H) = \emptyset$) implying $B_H(F, \hat{u}_H) \subseteq B_L(F, k)$. So, $\mu(F) = 0$ is consistent with D1. If instead $u_L(F, 1) < k$ (so $B_L(F, k) = \emptyset$), then there exists unique $\gamma \in (0, 1)$ such that $u_L(F^\gamma, 1) = k$. Since $\{F^*(k), \mu^*(k)\}$ solves M(k), $u_H(F^*(k), \mu^*(k)) \geq u_H(F^\gamma, 1) > u_H(F, 1)$. Hence, $B_H(F, \hat{u}_H) = \emptyset$ as well, and D1 places no restriction on $\mu(F)$.

We now establish uniqueness. Fix an equilibrium with $u_H = \hat{u}_H$ and $u_L = k$. Since the low type has the option to choose the same security as the high $u_L(F, \mu(F)) \leq k$ for all $F \in S_H$. Fix now $F \in S_H$ and suppose that $u_L(F, \mu(F)) < k$. Then $F \notin S_L$, so $\mu(F) = 1 = b_H(F, \hat{u}_H)$ and $B_L(F, k) = \emptyset$. Further, it must be that $F \neq X$ since $u_L(X, 1) = \overline{u} > k$. Then for $\gamma \in (0, 1)$ small enough $b_H(F^\gamma, \hat{u}_H) \in (0, 1)$ and $B_L(F^\gamma, k) = \emptyset$. Therefore, $F^\gamma \notin S_L$ and $\mu(F^\gamma) = 1$ (by belief consistency if $F^\gamma \in S_H$, by D1 if not). Since $u_H(F^\gamma, 1) > u_H(F, 1) = \hat{u}_H$, the high type would gain by deviating to F^γ, breaking the equilibrium. Therefore, $u_L(F, \mu(F)) = k$, or equivalently $\mu(F) = b_L(F, k)$, for all $F \in S_H$.

Suppose now there exists $F \in S_H$ such that $F \neq F^*(k)$. Then

$$u_H(F, \mu(F)) = u_H(F, b_L(F, k)) < u_H(F^*(k), \mu^*(k)) = u_H(F^*(k), b_L(F^*(k), k)),$$

and thus $b_H(F^*(k), \hat{u}_H) < \mu^*(k) = b_L(F^*(k), k)$. D1 then implies that $\mu(F^*(k)) = 1$, meaning that deviating to $F^*(k)$ is profitable for the high type and breaking the equilibrium. Hence, if the low type's equilibrium payoff is k, then $S_H = \{F^*(k)\}$ and $\mu(F^*(k)) = \mu^*(k)$.

Further, if the low type selects $F \notin S_H \cup \{X\}$, then $\mu(F) = 0$, and $u_L(F, 0) < u_L(X, 0) \leq u_L(X, \mu(X))$ for any value of $\mu(X)$. She could therefore profitably deviate to X. Hence, $S_L \subseteq S_H \cup \{X\}$.

The final step is to characterize which values of $u_L = k$ are consistent with equilibrium, which depends on the prior, μ_0. Recall from Lemma 6.1(d) that μ^* is continuous and strictly increasing in k. First, let $\mu_0 < \mu^*(\underline{u})$, and let $u_L = k$. Therefore, $S_H = \{F^*(k)\}$ and $\mu(F^*(k)) = \mu^*(k) > \mu_0$. For this belief to be consistent with seller strategies, $S_L \neq \{F^*(k)\}$. Hence, $S_L = \{F^*(k), X\}$ and $k = \underline{u}$. The precise mixing probabilities given in the proposition are required for the Bayesian consistency: $\mu(F^*(\underline{u})) = \mu^*(\underline{u})$.

Second, let $\mu_0 \geq \mu^*(\underline{u})$. Hence, there exists unique $k_0 \in [\underline{u}, \overline{u})$ such that $\mu^*(k_0) = \mu_0$. Suppose that $u_L = k > k_0$. Then $S_H = \{F^*(k)\}$ and $\mu(F^*(k)) = \mu^*(k) > \mu^*(k_0) = \mu_0$. But then for this belief to be consistent with seller strategies, $S_L \neq \{F^*(k)\}$. Hence, $S_L = \{F^*(k), X\}$ and $k = \underline{u}$, which contradicts $k > k_0$. Suppose instead that $u_L = k < k_0$. Then $S_H = \{F^*(k)\}$ and $\mu(F^*(k)) = \mu^*(k) < \mu^*(k_0) = \mu_0$. But then $\mu(F) < \mu_0$ for all F on the equilibrium path, which violates belief consistency. Hence, $k = k_0$, and $S_H = S_L = \{F^*(k_0)\}$, exactly as given in the proposition. $\square$

Proof of Proposition 6.2. Using the proof of Lemma 6.1(a) and the equilibrium characterization in Proposition 6.1, it is sufficient to show that if ratings are β-informative at $\overline{x}$, then $d_1(\mu_\ell(k)) > d_2^k(\mu_\ell(k))$ for all k such that $\mu_\ell(k)$ is sufficiently close to 1. Recalling that $d_2^k(\mu_\ell(k))$ satisfies $u_L(F^D_{d_2^k(\mu_\ell(k))}, \mu_\ell(k)) = k$, it follows that $d_2^k(\mu_\ell(k)) = \overline{x}$ by definition of $\mu_\ell(k)$. For any $k \in [\underline{u}, \overline{u})$, $d_1(\mu_\ell(k)) > \overline{x}$ if and only if

$$\frac{1 - \Pi_H(\overline{x}))}{1 - \Pi_L(\overline{x}))} - 1 < \frac{\left(\frac{\alpha'(\mu_\ell(k))}{\alpha'_L(\mu_\ell(k))}\right)(1-\delta)}{\alpha(\mu_\ell(k)) - \delta - \left(\frac{\alpha'(\mu_\ell(k))}{\alpha'_L(\mu_\ell(k))}\right)\alpha_L(\mu_\ell(k))}.$$

Because the RHS is continuous in μ, we can take the limit as $\mu_\ell(k) \to 1$, at which point the calculations are analogous to those in the proof of Lemma 6.1(c), establishing the result. $\square$

Proof of Proposition 6.3. The proof of Proposition 6.1 applies verbatim, as it only uses that the solution to M(k) is unique for all $k \in [\underline{u}, \overline{u})$ and that μ^* is continuous and strictly increasing k, all properties established in Lemma 6.2 when ratings are α-informative. $\square$

Proof of Proposition 6.4. First, for any $k \in [\underline{u}, \overline{u})$, in order to satisfy the constraint in M(k), $a^*(k) = 0$ if and only if $\mu^*(k) = \mu_\ell(k)$. Second, the proof of Lemma 6.2 establishes that, for all $k \in [\underline{u}, \overline{u})$, if $\mu^*(k) \neq \mu_\ell(k)$ then $\mu^*(k) < \hat{\mu}$ and that a^* is continuous and decreasing in k. Hence, there exists $\tilde{k}$ such that $\{a^*(k), \mu^*(k)\} = \{0, \mu_\ell(k)\}$ for all $k \geq \tilde{k}$, and that $\mu_\ell(\tilde{k}) < \hat{\mu}$. The proposition then follows from the equilibrium characterization in Proposition 6.3. $\square$

A.4 Proofs of Theorems

Both Theorem 1 and 2 are direct implications of Lemmas 6.1 and 6.2 and Propositions 6.1 and 6.3.