Rating agencies in the face of regulation
Online Appendix

July 18, 2012

C.   Robustness checks

C.1.   Signal structure

C.1.1.   Symmetry of signals

In the main text, we assume a symmetric error probability, so that "type I" and "type II" errors, $\alpha(\iota)$, are identical. If we allow these to differ, such that $\alpha(\iota)$ represents the error on good types and $\beta(\iota)$ the error on bad types, the comparative statics in the full-disclosure region (see Proposition 3) change somewhat. The optimal choice of information acquisition satisfies

$$C'(\iota^*) = -\alpha'(\iota^*)\pi_g x_g(y) + \beta'(\iota^*)\pi_b x_b(y). \tag{1}$$

As a result, information acquisition is increasing in y if $\pi_g > \left(1 + \frac{\alpha'(\iota^*)}{\beta'(\iota^*)}\right)^{-1}$, independent of y for $\pi_g = \left(1 + \frac{\alpha'(\iota^*)}{\beta'(\iota^*)}\right)^{-1}$, and decreasing in y otherwise. Qualitatively, all the results are unaffected.

C.1.2.   Binary vs. continuum of signals

Although our *binary* signal structure clearly simplifies the exposition, adding multiple or even a continuum of signals would not alter the results. For each signal, such as A and B in our model, the agents need to compute the posterior belief about the underlying types. If the expected revenue contribution using these posterior beliefs is positive, public financing for this signal class generates positive revenue for the rating agency. The rating agency's optimal information acquisition trades off more precise signals (and thus better investment allocations) with the increased cost of information acquisition. We refrain from this generalization because it would require us to specify how effort translates into

1

the distribution of (all) signals for each firm type. Moreover with multiple signals, a regulatory policy should be a schedule $y(s)$ and could not be characterized by a single regulatory advantage y. The regulatory advantage between a good and a bad signal captures the important discontinuities in regulation at relevant cutoffs, such as the one between investment-grade and junk (or AAA and AA). The comparative statics of our analysis would be qualitatively unaffected.

Also, our signal structure does not ensure the fraction of favorable signals (A) coincides with the fraction of good types (g) in the population. This "bias" is a natural feature of binary test outcomes. For example, medical tests for ex-ante unlikely diseases usually produce test outcomes where the positive signal for a disease occurs more frequently than the disease in the underlying population. These signals combined with the prior provide Bayesian agents with sufficient statistics for the posterior probabilities.

C.1.3. Binary vs. continuum of types

Our model features only two population types. As the discussion about the signal structure suggests, one could extend our setup to a continuum of underlying types once a continuum of signals is assumed. It is, however, important that the worst type project has negative NPV. Otherwise, information acquisition does not improve investment allocations. The main advantage of a two-type, binary signal economy is that one can summarize outside options and regulatory distortions with one variable each, instead of specifying a functional form or working with a less tractable signal structure.

C.2. Private messages

In this section, we allow issuers to send private messages to the rating agency. Otherwise, the setup is unchanged relative to the discussion in the main text. Without loss of generality, we consider binary messages of firm type n, i.e., $m(n) \in \{b, g\}$. The rating agency may use its information acquisition technology to verify the report as a function of the report m, i.e., $\iota(m)$. As before, the rating agency can commit to any of its actions (see justification in repeated game setup). In addition, we make the following assumption:

Assumption 2. *The rating agency may not make positive net transfers to any set of firms.*

This assumption is restrictive from a mechanism design perspective, as it prohibits the rating agency to pay off firms that identify themselves as bad. We make this assumption on economic grounds, because slight perturbations in the economic environment (say the endogenous creation of bad firms) would cause negative fees to be endogenously suboptimal.

Given this augmented setup, we are looking for a mechanism in which both firm types

report their true type, i.e.,

$$m(n) = n. \tag{2}$$

For these reports to be incentive compatible, bad types must not have a (strict) incentive to misreport themselves as good, i.e., reporting $m(b) = g$. (Trivially, good types will always report truthfully). Let $P_b(m)$ denote the payoff to the bad firm type using message m. We thus require:

$$P_b(b) \geq P_b(g). \tag{3}$$

Since $P_b(g) > 0$, the rating agency needs to offer bad firm types with rents, i.e., $P_b(b) > 0$. First, note that $\iota(b) = 0$, since the rating agency does not need to verify that the reported bad type is indeed bad. Following Assumption 2, the only way to provide bad types with (expected) rents is to publicly rate a fraction η of those bad firms as A (assuming that A-rated firms get financing). Let $\tilde{P}_b(A) = (1 - d_b)(R - \bar{N})$ represent the payoff to the bad firm of being rated A, then:

$$P_b(b) = \eta \tilde{P}_b(A). \tag{4}$$

Now consider the choice of $\iota(g)$. It is clear that $\iota(g) > 0$ (in the full disclosure region) because otherwise bad firms would simply report $m(b) = g$. Similar to the discussion in the main text, the efficient use of acquired information implies that if the rating agency signal confirms the report, i.e., $s = A$, then the issuer will be rated A. If the signal is negative, i.e., $s = B$, the firm will be rated as B. This implies that the payoff of a bad firm to report g is determined by the erroneous classification as A:

$$P_b(g) = \alpha(\iota(g)) \tilde{P}_b(A). \tag{5}$$

For the rating agency, it is optimal to cause the incentive constraint for self-reporting to bind (see Eq. 3), i.e.,

$$\eta = \alpha(\iota(g)). \tag{6}$$

For good types, the equilibrium implications are virtually unaffected compared to the setup discussed in the main text. Since all good types report themselves as good, they will always be screened and identified as good with probability $1 - a(\iota(g))$. Moreover, the revenue contribution of both types will still be $x_n(y)$. Thus, profits of the rating agency with type-dependent messages are now given by:

$$\Pi = (1 - \alpha(\iota(g))) \pi_g x_g(y) + \alpha(\iota(g)) \pi_b x_b(y) - \pi_g C(\iota_g) \tag{7}$$

Compared to the previous setup, private messages allow the rating agency save the cost of information acquisition on bad types. Mathematically, the additional feature of private messages can simply be interpreted as a proportional adjustment to the cost function (by $\pi_g < 1$).

C.3. Regulation-induced entry

In the benchmark model of the paper, we considered the case in which the cost of information acquisition is sufficiently low. Now consider the case in which any positive level of information acquisition would yield negative profits in the absence of a preferential regulatory treatment of A-rated securities, that is, $(1 - \alpha(\iota^*))\pi_g x_g(0) + \alpha(\iota^*)\pi_b x_b(0) - C(\iota^*) < 0$, where ι^* is such that $C'(\iota^*) = \pi_g x_g(0) - \pi_b x_b(0)$. In this case, the high cost of information acquisition prevents the rating agency from operating without regulatory distortions. However, if y is sufficiently large, $y > |x_g \pi_g(0) + x_b \pi_b(0)|$, the business of regulatory arbitrage becomes profitable. The rating agency would still not acquire information. Nonetheless, due to the regulatory advantage, investors would be willing to pay for ratings. Thus, if regulatory accreditation is associated with sufficiently large benefits, a rating agency finds it profitable to enter lines of business that are unprofitable in the absence of regulation.

C.4. Disclosure rule

In the main text, we restrict the rating agency to upward biases in the disclosure rule, i.e., the probability that a firm with an A signal is labeled B, is set to 0. We now allow for this possibility. Let ε_{ij} refer to the conditional probability of labeling a firm with an i signal as j. We reconsider the proof of Proposition 1.

Given the optimal fee level $f^*(\iota, \varepsilon)$, revenue $S(\iota, \varepsilon)$ is just a function of information acquisition and the disclosure rule ε. Full-disclosure revenue can be written as

$$S(\iota, \mathbf{0}) = (1 - \alpha(\iota))\pi_g x_g(0) + \alpha(\iota)\pi_b x_b(0). \tag{8}$$

For an arbitrary disclosure rule, revenue can be decomposed into the full-disclosure revenue and the deviation from full disclosure:

$$S(\iota, \varepsilon) = \underbrace{S(\iota, \mathbf{0})}_{\text{Full-Disclosure Revenue}} + \underbrace{[\pi_g x_g(0)\alpha(\iota) + \pi_b x_b(0)(1 - \alpha(\iota))]\varepsilon_{BA} - S(\iota, \mathbf{0})\varepsilon_{AB}}_{\text{Revenue from Full-Disclosure Deviation}}. \tag{9}$$

Thus, for a fixed ι, the revenue (and thus profits) of the rating agency is linear in ε_{AB} and ε_{BA}. The coefficient on ε_{BA} is given by

$$\frac{dS}{d\varepsilon_{BA}} = \pi_g x_g(0)\alpha(\iota) + \pi_b x_b(0)(1 - \alpha(\iota)) \tag{10}$$

$$< (1 - \alpha(\iota))(\pi_g x_g(0) + \pi_b x_b(0)) \tag{11}$$

$$< (1 - \alpha(\iota))(\pi_g V_g + \pi_b V_b) < 0. \tag{12}$$

The first relation follows because $\alpha(\iota) < \frac{1}{2}$ and $x_g(0) > 0$. The second one follows from $x_n(0) < V_n$. The third one follows from the assumption that the average project is not worth financing. Thus, for any ι, revenue is decreasing in ε_{BA}. Hence, choosing $\varepsilon_{BA} = 0$ must be optimal.

4

Now, consider ε_{AB}. The coefficient on ε_{AB} is given by

$$\frac{dS}{d\varepsilon_{AB}} = -S(\iota, \mathbf{0}). \tag{13}$$

The revenue under full disclosure $S(\iota, \mathbf{0})$ must be nonnegative in equilibrium. For suppose $S(\iota, \mathbf{0}) < 0$; then $\varepsilon^* = (0, 1)$ and $S(\iota, \varepsilon^*) = S(\iota, \mathbf{0}) - S(\iota, \mathbf{0}) = 0$ for any ι, which is a contradiction.

Using full disclosure, the profit of the rating agency conditional on any level of information acquisition ι satisfies

$$\Pi = \mu_A(\iota) f^*(\iota) - C(\iota) = (1 - \alpha(\iota)) \pi_g x_g(0) + \alpha(\iota) \pi_b x_b(0) - C(\iota). \tag{14}$$

The optimal level of information acquisition must solve the first-order condition,

$$\pi_g x_g(0) - \pi_b x_b(0) - C'(\iota^*) = 0. \tag{15}$$

The second-order condition is satisfied since $C''(\iota)$ is positive. The restrictions on the cost function ensure that a unique interior level of information acquisition $0 < \iota^* < 1$ exists. The remaining parts of the Proposition follow directly.

D. Generalized parameter region

We now consider the case in which $y > |V_b|$. In this parameter region, the purchasing decision, $p(A) = (0, 1)$, may occur in equilibrium.

Lemma 4. *For $p(A) = (0, 1)$ to be incentive compatible for bad types, the maximum fee that the rating agency can charge for an A-rating is given by:*

$$f_{\max}(\gamma, y) = \begin{cases} (1 - d_b) N_U(\gamma) + y - 1 & \text{for} \quad \gamma < \gamma^*, \\ (1 - d_b) R + y - 1 = V_b + y & \text{for} \quad \gamma \geq \gamma^*, \end{cases} \tag{16}$$

with

$$N_U(\gamma) = \frac{1}{1 - d_U(\gamma)}, \tag{17}$$

$$d_U(\gamma) = \frac{\pi_g}{\pi_g + \pi_b \gamma} d_g + \frac{\pi_b \gamma}{\pi_g + \pi_b \gamma} d_b, \tag{18}$$

$$\gamma^* = \frac{x_g(0)}{-x_b(0)} \frac{\pi_g}{\pi_b}, \tag{19}$$

where $N_U \in \left[\frac{1}{1 - d_g}, \bar{N}\right)$ is the face value of unrated bonds and γ the fraction of bad firms that are assigned a B-rating.

Proof. If all bad firms obtain an A-rating, i.e., $\gamma = 0$, then all unrated firms are good firms, i.e. $d_U = d_g$. In this case, good firm types obtain financing from the public bond market with a face value of $N_g = \frac{1}{1-d_g} < \bar{N} < R$. Bad types would only have an incentive to buy an A-rating if the face value N_A is lower than or equal to $N_U = N_g$. This generates a maximum fee $f_{\max}(0, y) = (1 - d_b) N_g + y - 1$. As γ is increased, the quality of the unrated pool is decreased, so that N_U and the fee $f_{\max}(\gamma, y)$ are increased. At some point, $\gamma = \gamma^*$, the pool is so diluted that good types are indifferent between public financing with $N = N_U(\gamma^*)$ and their outside option with $N = \bar{N}$, that is,[1]

$$N_U(\gamma^*) = \bar{N}. \tag{20}$$

Solving for γ^* yields

$$\gamma^* = \frac{\bar{N}(1 - d_g) - 1}{1 - \bar{N}(1 - d_b)} \frac{\pi_g}{\pi_b} = \frac{x_g(0)}{-x_b(0)} \frac{\pi_g}{\pi_b}. \tag{21}$$

∎

Lemma 5. *Inducing only bad types to purchase an A-rating, $p(A) = (0, 1)$, and setting $0 < \gamma < \gamma^*$ implies smaller rating agency profits than rating inflation with participation of all types.*

Proof. If $p(A) = (0, 1)$ and $\gamma < \gamma^*$, the maximum fee satisfies $f_{\max}(\gamma, y) < x_b(y)$ as $N_U < \bar{N}$. Rating inflation with participation of all types yields rating agency profits of $\pi_g x_g(y) + \pi_b x_b(y) > x_b(y)$. Rating agency profits with rating inflation and participation of only bad types yields profits of $\pi_b(1 - \gamma) f_{\max}(\gamma, y) < f_{\max}(\gamma, y) < x_b(y)$. This proof only covers the relevant case in which $f_{\max}(\gamma, y) > 0$, i.e., $y \gg x_b(0)$. Otherwise, rating inflation of this kind is strictly dominated by full disclosure (as it would imply negative profits). ∎

At γ^*, the maximum fee that the rating agency may charge, increases discontinuously from $x_b(y) = (1 - d_b) \bar{N} + y - 1$ to $y + V_b$ as good types move to their outside option and unrated firms can no longer obtain financing. Thus, if $y > |V_b|$, positive profits can be obtained setting $p(A) = (0, 1)$ and $\gamma \geq \gamma^*$.

Lemma 6. *If $y > |V_b|$ and $p(A) = (0, 1)$, achievable rating agency profits for $\gamma = \gamma^*$ are greater than for any $\gamma > \gamma^*$.*

Proof. Since the maximum fee does not depend on γ if $\gamma > \gamma^*$, profits are given by $(1 - \gamma) \pi_b (y + V_b)$. Since $y > |V_b|$ by assumption, rating agency profits are maximized by minimizing γ, subject to $\gamma \geq \gamma^*$. ∎

Thus, for $y > |V_b|$, profit maximization of the rating agency ensures that equilibrium profits, Π^*, are given by the maximum of three relevant candidates: full disclosure profits, $\Pi_{FD}(y) = \Pi(\iota^*(y), 0, f^*(\iota^*(y), 0, y), y)$, rating inflation with participation of all firms,

---

[1]We assume that good types choose their outside option if indifferent.

$\Pi_{RI}(y) = \pi_g x_g(y) + \pi_b x_b(y)$, and rating inflation in which only bad firms are A-rated, $\tilde{\Pi}_{RI}(y) = (1 - \gamma^*) \pi_b (V_b + y)$. Formally,

$$\Pi^*(y) = \max \left\{ \Pi_{FD}(y), \Pi_{RI}(y), \tilde{\Pi}_{RI}(y) \right\}. \tag{22}$$

It is useful to write $\Pi_{RI}(y)$ as

$$\Pi_{RI}(y) = \pi_g x_g(0) + \pi_b x_b(0) + y. \tag{23}$$

Since all firms receive the regulatory advantage y under the rating inflation regime, the corresponding rating agency profits $\Pi_{RI}(y)$ are most sensitive to changes in regulatory advantages.

$$\frac{d\Pi_{RI}(y)}{dy} = 1 > \max \left\{ \frac{d\Pi_{FD}(y)}{dy}, \frac{d\tilde{\Pi}_{RI}(y)}{dy} \right\}. \tag{24}$$

As a result, the function $\Pi_{RI}(y)$ crosses the functions $\Pi_{FD}(y)$ (at $\bar{y}$) and $\tilde{\Pi}_{RI}(y)$ (at $\tilde{y}$) exactly once where $\tilde{y}$ satisfies

$$\tilde{y} = \frac{\kappa \bar{U}_g}{\pi_g} \frac{|\pi_g x_g(0) + \pi_b x_b(0)|}{x_g(0) - x_b(0)} > 0, \tag{25}$$

and $\kappa = \frac{1-d_b}{1-d_g}$. It follows immediately that $\Pi_{RI}(y) > \tilde{\Pi}_{RI}(y)$ for $y > \tilde{y}$, and $\Pi_{RI}(y) < \tilde{\Pi}_{RI}(y)$ for $y < \tilde{y}$. Due to the convexity of $\Pi_{FD}(y)$, the linear function $\tilde{\Pi}_{RI}(y)$ (with slope less than one) may intersect $\Pi_{FD}(y)$ twice. Let us define $\breve{y}_i$ such that:

$$\Pi_{FD}(\breve{y}_i) = \tilde{\Pi}_{RI}(\breve{y}_i), \tag{26}$$

where we define $\breve{y}_i$ such that $\breve{y}_2 > \breve{y}_1$ if two solutions exist and define $\breve{y}_2 = \infty$ if only one solution exists.[2] If the functions never cross in the relevant region, then $\Pi_{FD}(y) > \tilde{\Pi}_{RI}(y)$ for all y (this case obviously rules out rating inflation with only bad types). If they cross once, then $\Pi_{FD}(y) > \tilde{\Pi}_{RI}(y)$ for $y < \breve{y}$ and $\tilde{\Pi}_{RI}(y) > \Pi_{FD}(y)$ for $y > \breve{y}$. This is because $\Pi_{FD}(0) > 0$ (by assumption) and $\tilde{\Pi}_{RI}(0) < 0$ (as $V_b < 0$). If they cross twice, then $\Pi_{FD}(y) > \tilde{\Pi}_{RI}(y)$ for $y < \breve{y}_1$, $\tilde{\Pi}_{RI}(y) > \Pi_{FD}(y)$ for $\breve{y}_1 < y < \breve{y}_2$, and $\Pi_{FD}(y) > \tilde{\Pi}_{RI}(y)$ for $y > \breve{y}_2$ (see Fig. 4).

Proposition 5. *Equilibrium Characterization:*
If $\bar{y} \leq \breve{y}_1$, the equilibrium outcome is full disclosure for all $y \leq \bar{y}$ and rating inflation with participation of all types for all $y > \bar{y}$.
If $\breve{y}_1 < \bar{y} \leq \breve{y}_2$, then the equilibrium outcome is full disclosure for all $y \leq \breve{y}_1$, rating inflation with only bad types for all $\breve{y}_1 < y \leq \tilde{y}$, and rating inflation with participation of all types for $y > \tilde{y}$.
If $\bar{y} > \breve{y}_2$ (only relevant if $\breve{y}_2$ finite), the equilibrium outcome is full disclosure for all $y \leq \breve{y}_1$, rating inflation with only bad types for all $\breve{y}_1 < y < \breve{y}_2$, full disclosure for $\breve{y}_2 \leq y \leq \bar{y}$, and rating inflation with participation of all types for $y > \bar{y}$.

---

[2] $\Pi_{FD}(y)$ may not intersect with $\tilde{\Pi}_{RI}(y)$ more than twice since $\Pi_{FD}(y)$ is convex in y and $\tilde{\Pi}_{RI}(y)$ is linear.
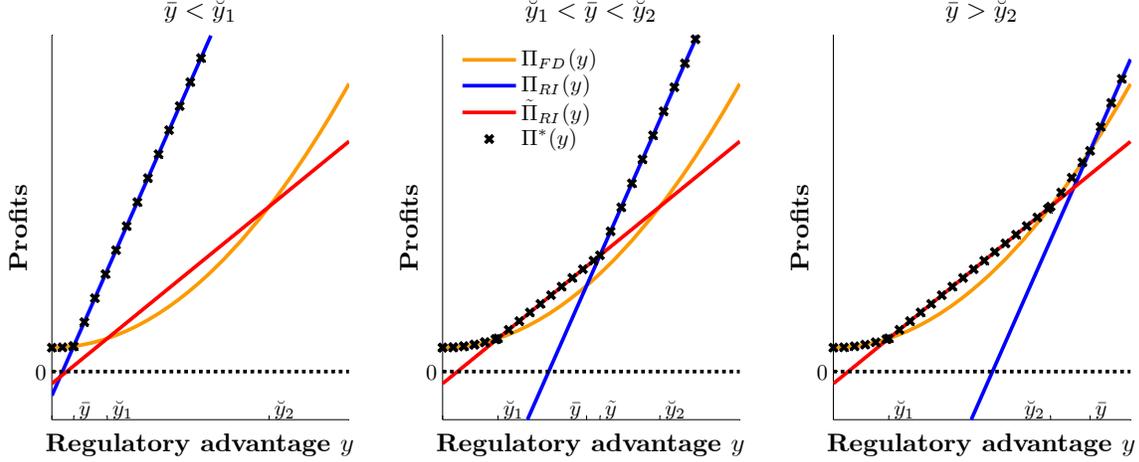
Fig. 4. The graph plots profits for full disclosure $\Pi_{FD}(y)$, rating inflation with participation of all types $\Pi_{RI}(y)$, and rating inflation with participation of only bad types $\tilde{\Pi}_{RI}(y)$. The marked line indicates the maximum achievable profits of the rating agency $\Pi^*(y)$. The left panel plots the case in which $\bar{y} < \breve{y}_1$. Here, rating inflation with participation of only bad types is always strictly dominated. The middle panel plots the case $\breve{y}_1 < \bar{y} < \breve{y}_2$. The right panel plots the case of $\bar{y} > \breve{y}_2$. In both of these cases, rating inflation with participation of only bad types will occur in equilibrium for some region of y.

Proof. As discussed above, we can ignore the case in which $\tilde{\Pi}_{RI}(y)$ and $\Pi_{FD}(y)$ do not intersect. For the purpose of the proof, we can focus on the (most complicated) case in which $\tilde{\Pi}_{RI}(y)$ and $\Pi_{FD}(y)$ intersect twice. The proof will rely heavily on Fig. 4. The single-crossing case can be obtained by simply ignoring all restrictions based on $\breve{y}_2$. If $\bar{y} < \breve{y}_1$, then $\Pi_{RI}(y) \lesseqgtr \Pi_{FD}(y)$ for all $y \lesseqgtr \bar{y}$ (by definition of $\bar{y}$). Moreover, $\Pi_{RI}(y) > \tilde{\Pi}_{RI}(y)$ for all $y \geq \bar{y}$ since $\Pi_{RI}(\bar{y}) > \tilde{\Pi}_{RI}(\bar{y})$ and $\frac{d\Pi_{RI}}{dy} > \frac{d\tilde{\Pi}_{RI}}{dy}$. As a result, $\Pi_{FD}(y) > \max\left\{\Pi_{RI}(y), \tilde{\Pi}_{RI}(y)\right\}$ for all $y < \bar{y}$ and $\Pi_{RI}(y) > \max\left\{\Pi_{FD}(y), \tilde{\Pi}_{RI}(y)\right\}$ for all $y > \bar{y}$. If $\breve{y}_1 < \bar{y} < \breve{y}_2$, full disclosure will obtain for all $y < \breve{y}_1$ and then switch to rating inflation of bad types (as $\breve{y}_1 < \bar{y}$). Since $\bar{y} < \breve{y}_2$, $\tilde{\Pi}_{RI}(y)$ will intersect $\Pi_{RI}(y)$ at $\tilde{y}$ before it intersects (again) with full disclosure profits $\Pi_{FD}(y)$. Thus, for $y > \tilde{y}$ rating inflation with participation of all types will obtain. If $\bar{y} > \breve{y}_2$, one change occurs relative to the previous case. Since the second intersection of $\tilde{\Pi}_{RI}(y)$ with full disclosure profits occurs before it hits $\Pi_{RI}(y)$, there will be an additional region between $\breve{y}_2$ and $\bar{y}$ where full disclosure obtains. ∎

This proposition reveals that for small levels of y, the equilibrium will always be characterized by full disclosure and for sufficiently large levels, i.e., $y > \max(\tilde{y}, \bar{y})$, there will always be rating inflation with participation of all types. However, depending on the parameter constellation, there may exist a region in between in which the equilibrium is characterized by rating inflation with participation of only bad types. The purpose of this proposition is to show the theoretical possibility of such an outcome, not to argue that this is generally the case. In fact, rating inflation of this kind can always be ruled out if rating agency profits from rating inflation with all types are positive at $y = |V_b|$, since $\tilde{\Pi}_{RI}(|V_b|) = 0$ and $\frac{d\Pi_{RI}(y)}{dy} > \frac{d\tilde{\Pi}_{RI}(y)}{dy}$ (see left panel of Fig. 4). This results in the

following sufficient condition.

Proposition 6. *If the outside option of good types, $\bar{U}_g$, is sufficiently small, e.g., $\frac{\bar{U}_g}{V_g} \leq \frac{\pi_g}{\pi_b}(1 + V_g)$, rating inflation with participation of only bad types is never optimal.*

Proof. We simply have to check whether $\tilde{y} \leq |V_b|$. Put differently, we require that $\Pi_{RI}(|V_b|) = \pi_g x_g(0) + \pi_b x_b(0) + |V_b| \geq 0 = \tilde{\Pi}_{RI}(|V_b|)$, i.e.,

$$\pi_g\left(V_g - \bar{U}_g\right) + \pi_b\left(V_b - \kappa\bar{U}_g\right) - V_b \geq 0, \tag{27}$$

or

$$\frac{\bar{U}_g}{V_g} \leq \pi_g \frac{1 - \frac{V_b}{V_g}}{\kappa\pi_b + \pi_g}. \tag{28}$$

Since the average project has negative NPV, i.e. $-\frac{V_b}{V_g} \geq \frac{\pi_g}{\pi_b}$, a stronger condition is

$$\frac{\bar{U}_g}{V_g} \leq \frac{\pi_g}{\pi_b} \frac{1}{\kappa\pi_b + \pi_g}. \tag{29}$$

Moreover, since $\kappa = \frac{1-d_b}{1-d_g} = \frac{1+V_b}{1+V_g} \leq \frac{1 - \frac{\pi_g}{\pi_b}V_g}{1+V_g}$, an even stronger condition is

$$\frac{\bar{U}_g}{V_g} \leq \frac{\pi_g}{\pi_b}(1 + V_g). \tag{30}$$

∎

Note, that if $\pi_g \geq \frac{1}{2}$ this condition is always satisfied since $\frac{\bar{U}_g}{V_g}$ is bounded above by 1.

E.   Regulation and the value-maximizing level of information production

An interesting and important issue for the design of regulation based on ratings is the extent to which it distorts the production of information by rating agencies relative to a socially optimal level. Here we take as our criterion for socially optimal the level of information production that would maximize the aggregate NPV, *net of any regulatory advantages of A-rated securities*, of firms seeking financing, minus the cost of producing the information. We refer to this criterion as the "social objective function" and to the level of information production that maximizes the social objective as the "socially optimal" level.[3] We show, in this section, that ratings-based regulation may actually *improve* the allocation of resources to information production in this sense.

---

[3]We do not attempt here to address the more general question of what an optimal regulatory policy for financial institutions is, taking account of the effect of the regulation on the macroeconomic implications of failures of these institutions as well as its effect on information production. Instead we focus on the impact of regulation on the latter. That is, the "socially optimal" level of information production we consider here is only optimal in the sense that it results in an optimal allocation of resources to information production and projects, given the cost of information production.

The socially optimal level of information production equates the marginal cost of information, $C'(\iota)$, to the marginal *social* benefit in the full disclosure region ($y \leq \bar{y}$). In this region, an increase in information production increases the probability that good firms are financed and bad firms are not financed. For each additional good firm that is financed, aggregate NPV increases by the good firm's NPV, net of the NPV it could obtain using its outside option, i.e., by $V_g - \bar{U}_g$.[4] For each additional bad firm not financed, aggregate NPV increases by $-V_b$. Since there are π_g good firms and π_b bad firms in the economy, the marginal social benefit of an increase in information production is $\pi_g (V_g - \bar{U}_g) - \pi_b V_b$.

By comparison, in the full disclosure region, the rating agency's marginal benefit of additional information consists of the additional revenue it generates from assigning A ratings to more good firms and the additional losses it avoids by assigning B ratings to more bad firms. For each additional good firm financed, the rating agency extracts the additional NPV the firm produces relative to its outside option, as well as the regulatory advantage, i.e., $V_g - \bar{U}_g + y$. For each bad firm not financed, the rating agency avoids losing the NPV of the bad firm and the rents bad firms extract by pooling with good ones, increasing revenue by $-V_b + \kappa \bar{U}_g$. In addition, for each bad firm not financed, the rating agency loses the regulatory advantage, y. Thus the rating agency gains a total of $-V_b + \kappa \bar{U}_g - y$ for each additional bad firm it identifies. Therefore the rating agency's marginal benefit of additional information is $\pi_g (V_g - \bar{U}_g + y) + \pi_b (-V_b + \kappa \bar{U}_g - y)$.

As is clear from the above discussion, in the absence of any regulatory advantage ($y = 0$), the rating agency faces the social marginal benefit of identifying good firms, but it gains *more* than the social marginal benefit of screening out bad firms. Indeed, comparing the rating agency's marginal benefit absent regulatory advantage with the social marginal benefit, we see that the former exceeds the latter by $\pi_b \kappa \bar{U}_g$ and is, therefore, increasing in the good firms' outside option, $\bar{U}_g$. Since the private and social marginal costs are the same, with no regulatory advantage, the rating agency will overinvest in information acquisition by an amount that is increasing in the outside option $\bar{U}_g$.

Now suppose the regulatory advantage is positive. The effect on information production, compared to the case in which the regulatory advantage is zero, depends on how large the advantage is and on the fraction of firms that are good. As shown in Proposition 3, a positive advantage for A ratings induces the rating agency to increase the volume of A ratings. If good firms are in the majority, the rating agency accomplishes this by increasing information production. Since it produces too much information in the absence of a regulatory advantage, a positive advantage exacerbates the distortion. If bad firms are in the majority, the rating agency increases the volume of A ratings by reducing information production. Obviously, this reduces the distortion as long as the advantage is not too large. If the advantage is sufficiently large, however, the rating agency will reduce its production of information so much that the level with the advantage is farther *below* the social optimum than the level without the advantage was above it, thus increasing the distortion.

---

[4]As stated previously, we assume the cost of the bypass channel, $V_g - \bar{U}_g$, is socially wasteful. If it were not, one would need to incorporate the non-waste component as well.

We summarize our results on how the distortion depends on the regulatory advantage and the distribution of firm types in the following proposition.

Proposition 7. *In the absence of a regulatory advantage, the rating agency will overinvest in information acquisition, and the distortion is increasing in the outside option $\bar{U}_g$. For $0 < y < \bar{y}$, an increase in the regulatory advantage, y,*
- *increases the distortion in information production if a majority of firms is good,*
- *has no effect on the distortion if the proportion of good firms is exactly one-half, and*
- *reduces the distortion if a majority of firms is bad, provided the advantage is not too large, and increases it otherwise.*

These non-trivial effects should be taken into account when designing ratings-based regulation. Of course, adopting a regulation whose advantage for highly rated securities results in ratings hyperinflation is *never* optimal, since, in that case, no information is produced (which is below the optimum) and the regulation is completely ineffective in preventing financial institutions from taking excessive risks, because all securities are highly rated. Thus any regulation that relies on ratings as informative risk metrics must ensure the inflation regime is avoided. Since the inflation threshold $\bar{y}$ is a function of the complexity of the security class, our analysis suggests regulators should be wary of the undifferentiated treatment of ratings across different security classes. This result is independent of the precise objective function of the regulator.