Who Should Pay for Credit Ratings and How?

Anil K Kashyap
Booth School of Business, University of Chicago

Natalia Kovrijnykh
Department of Economics, Arizona State University

December 2015

Abstract

We analyze a model where investors use a credit rating to decide whether to finance a firm. The rating quality depends on unobservable effort exerted by a credit rating agency (CRA). We study optimal compensation schemes for the CRA when a planner, the firm, or investors order the rating. Rating errors are larger when the firm orders it than when investors do (and both produce larger errors than is socially optimal). Investors overuse ratings relative to the firm or planner. A trade-off in providing time-consistent incentives embedded in the optimal compensation structure makes the CRA slow to acknowledge mistakes. (*JEL:* D82, D83, D86, G24)

We have benefited from suggestions and comments by the editor, two anonymous referees, Bo Becker, Hector Chade, Simon Gilchrist, Ben Lester, Robert Lucas, Rodolfo Manuelli, Marcus Opp, Chris Phelan, Francesco Sangiorgi, Joel Shapiro, Robert Shimer, Nancy Stokey, and Joel Watson. We are also grateful for comments by seminar participants at ASU, Atlanta Fed, Philadelphia Fed, Purdue University, NYU Stern, University of Arizona, University of Chicago, University of Iowa, University of Oxford, University of Southern California, University of Wisconsin–Madison, Washington University in St. Louis, and conference participants of the LAEF Accounting for Accounting in Economics Conference in 2013, Fall 2012 NBER Corporate Finance Meeting, 2013 NBER Summer Institute, 2013 SED Meetings, and 2015 SAET Conference. Kashyap thanks the National Science Foundation, as well as the Initiative on Global Markets and the Center for Research on Security Prices at Chicago Booth for research support. For information on Kashyap's outside compensated activities, see http://faculty.chicagobooth.edu/anil.kashyap/. Send correspondence to Anil K Kashyap, Booth School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL 60637; telephone: (773) 702-7260, E-mail: anil.kashyap@chicagobooth.edu

Virtually every government inquiry into the 2008–2009 financial crisis has assigned some blame to credit rating agencies. For example, the Financial Crisis Inquiry Commission (2011, xxv) concludes that "this crisis could not have happened without the rating agencies." Likewise, the United States Senate Permanent Subcommittee on Investigations (2011, 6) states that "inaccurate AAA credit ratings introduced risk into the U.S. financial system and constituted a key cause of the financial crisis." In announcing its lawsuit against S&P, the U.S. government claimed that "S&P played an important role in helping to bring our economy to the brink of collapse."

The details of the indictments, however, differ slightly across the analyses. For instance, the Senate report points to inadequate staffing as a critical factor, while the Financial Crisis Inquiry Commission highlights the business model that had firms seeking to issue securities pay for ratings as a major contributor, and the U.S. Department of Justice lawsuit identifies the desire for increased revenue and market share as a critical factor.[1] In this paper we explore the role that these and other factors might play in creating inaccurate ratings.

We study a one-period model where a firm is seeking funding for a project from investors. The project's quality is unknown, and a credit rating agency can be hired to evaluate the project. So, the CRA creates value by generating information that can lead to more efficient financing decisions. The CRA must exert costly effort to acquire a signal about the quality of the project, and the higher the effort, the more informative the signal about the project's quality. The key friction is that the CRA's effort is unobservable, so a compensation scheme must be designed to provide incentives to the CRA to exert it. We consider three settings, where we vary who orders a rating—a planner, the firm, or potential investors.

This simple framework makes it possible to directly address the claims made in the government reports. In particular, we can ask: how do you compensate the CRA to avoid shirking? Does the issuer-pays model generate more shirking than when the investors pay for ratings? In addition, in natural extensions of the basic model we can see whether a

[1]The United States Senate Permanent Subcommittee on Investigations (2011, 7) reported that "factors responsible for the inaccurate ratings include rating models that failed to include relevant mortgage performance data, unclear and subjective criteria used to produce ratings, a failure to apply updated rating models to existing rated transactions, and a failure to provide adequate staffing to perform rating and surveillance services, despite record revenues." The Financial Crisis Inquiry Commission (2011, 212) concluded that"the business model under which firms issuing securities paid for their ratings seriously undermined the quality and integrity of those ratings; the rating agencies placed market share and profit considerations above the quality and integrity of their ratings." In the press releasing announcing that it was suing S&P, the United States Department of Justice (2013) states that because of "the desire to increase market share and profits, S&P issued inflated ratings on hundreds of billions of dollars' worth of CDOs."

battle for market share would be expected to reduce ratings quality, or whether different types of securities create different incentives to shirk.

Our model explains five observations about the ratings business that are documented in the next section, in a unified fashion. The first is that rating mistakes are in part due to insufficient effort by rating agencies. The second is that outcomes and accuracy of ratings do differ depending on who pays for a rating. Third, increases in competition between rating agencies are accompanied by a reduction in the accuracy of ratings. Fourth, ratings mistakes are more common for newer securities with shorter histories than for more established types of securities. Finally, revisions to ratings are slow to occur.

We begin our analysis by characterizing the optimal compensation scheme for the CRA. The need to provide incentives for effort requires setting compensation that is contingent on the rating and the project's performance, which can be interpreted as rewarding the CRA for establishing a reputation for accuracy.[2] Moreover, the problem of effort underprovision argues for giving the surplus from the investment project to the rating agency, so that the higher the CRA's profits, the higher the effort it exerts.

We proceed by comparing the CRA's effort and the total surplus in this model depending on who orders a rating. Generically, under the issuer-pays model, the rating is acquired less often and is less informative (i.e., the CRA exerts less effort) than in the investor-pays model (or in the second-best case, where the planner asks for a rating). However, the total surplus in the issuer-pays model may be higher or lower than in the investor-pays model, depending on the agents' prior belief about the project's quality. The ambiguity about the total surplus arises because even though investors induce the CRA to exert more effort, they ask for a rating even when the social planner would not. So the extra accuracy achieved by having investors pay is potentially dissipated by an excessive reliance on ratings.

We also extend the basic setup in three ways. First, we introduce competition among CRAs, an immediate implication of which is a tendency to reduce fees in order to win business. But with lower fees comes lower effort in evaluating projects, which reduces ratings accuracy. Next, we suppose that some types of securities are inherently more difficult to evaluate, presumably because they have a short track record. We show that there will be more mistakes for those types of securities. Finally, we allow for a second period in the model and posit that investment is needed in each period, so that there is a role for ratings in both periods. The need to elicit effort in both periods creates a dilemma. Paying the CRA if it makes a "mistake" in the initial rating (when a high rating

[2]We discuss this interpretation of the outcome-contingent fee structure in more detail in Section 3.2.

is followed by the project's failure) is detrimental for the first period's incentives. However, not paying to the CRA after a "mistake" will result in zero effort in the second period, when the rating needs to be revised. Balancing this trade-off involves the compensation in the second period after a "mistake" being too low ex post, causing the CRA to be slow to acknowledge mistakes.

While our simple model neatly explains the five observations described above using relatively few assumptions, our approach does come with several limitations. For instance, due to complexity, we do not study the problem when multiple ratings can be acquired in equilibrium. Thus we cannot address debates related to ratings shopping—a common criticism of the issuer-pays model.[3] Also, we assume that the firm has the same knowledge about the project's quality ex ante as everyone else. Without this assumption the analysis becomes much more complicated, since in addition to the moral hazard problem on the side of the CRA there is an adverse selection problem on the side of the firm. We do offer some cursory thoughts on this problem in our conclusions.

Despite these caveats, a strength of our model is in explaining all the aforementioned observations using a single friction (moral hazard);[4] in contrast, the existing literature uses different models with different frictions to explain the various phenomena. Hence, we are comfortable arguing that a full understanding of what went wrong with the CRAs will recognize that there were several problems and that moral hazard was likely one of them.

1 Motivating Facts and Literature Review

Given the intense interest in the causes of the financial crisis and the role that official accounts of the crisis ascribe to the rating agencies, it is not surprising that there has been an explosion of research on credit rating agencies.[5] We offer a new look at the recent events through the lens of a model with moral hazard created by the unobservability of CRA effort. In doing so, we are in no way intending to deny the role of other frictions, but instead merely trying to isolate the potential impact of one contributing factor.

To understand our contribution, we separate our review into two parts. We first review the papers that document various empirical regularities that our model can explain. Several

[3]See the literature review below for discussion of papers that do have ratings shopping. Notice, however, that even without ratings shopping we are able to identify some problems with the issuer-pays model.

[4]We also have an adverse selection problem arising from the fact that the CRA can misreport its signal, but this friction is not essential for our results, and we allow for it only because it seems realistic to do so.

[5]See White (2010) for a concise description of the rating industry and its role in the crisis.

of these establish evidence on the importance of CRA effort in the ratings process. We then review theoretical papers that are most closely related to ours.

1.1 Empirical studies of the rating business

The first body of research consists of the empirical studies that seek to document mistakes or perverse rating outcomes. There are so many of these papers that we cannot cover them all, but it is helpful to note that there are five observations that our analysis sheds light on. So we will point to specific contributions that document these particular facts.

First, prior work shows that who pays for a rating matters. The rating industry is currently dominated by Moody's, S&P, and Fitch Ratings, which are each compensated by issuers. So comparison of their recent performance does not speak to this issue. But Cornaggia and Cornaggia (2013) compare Moody's ratings to those of Rapid Ratings, a small CRA funded by subscription fees from investors, and find that Moody's ratings are slower to reflect bad news than those of Rapid Ratings. Jiang, Stanford, and Xie (2012) provide complementary evidence by analyzing data from the 1970s when Moody's and S&P were using different compensation models. From 1971 until June 1974 S&P was charging investors, while Moody's was charging issuers. During this period the Moody's ratings systematically exceeded those of S&P. S&P adopted the issuer-pays model in June 1974, and over the next three years their ratings essentially matched Moody's.

Second, as documented by Mason and Rosner (2007), most of the rating mistakes occurred for structured products that were primarily related to asset-backed securities. As Pagano and Volpin (2010) note, the volume of these new securities increased tenfold between 2001 and 2010. Coval, Jacob, and Stafford (2009) explain that ratings for collateralized debt obligations (CDOs) are very sensitive to the assumed correlation of defaults of the securities in the CDOs.[6] Griffin and Tang (2012) describe the ratings process for structured products and observe that "defaults are rare and irreversible events, historical data are sparse, and modeling default time is challenging as it is a point process. Consequently, deriving default correlation from fundamental default drivers can be inaccurate." In other words, inferring the correlations would require considerable effort. In contrast, rating traditional corporate bonds requires estimating only the probability of default for the firm under consideration. Even for new firms, the dominant rating agencies have long industry

[6]They also note that prior to the crisis, it was common for CDOs to be constructed with components of prior CDOs, creating what came to be called CDO^2. The accuracy of CDO^2 ratings are even more sensitive to mistakes in assessing correlation.

histories that span many recessions.[7]

Typically it might be hard to separately isolate mistakes due to shirking from those arising from incompetence. But, Griffin and Tang (2012) uncover some unusual evidence regarding structured products that clearly points to the former. In their Internet Appendix, they describe what they call "coincidental CDOs" that reek of shirking. They write: "A number of CDOs seem to use the same constant default probability criterion for each of the 19 rating scales, regardless of their maturities. . . . Not only are their default probability criteria constant and identical, their scenario default rates are identical for each of the 19 rating scales from AAA to CCC—across all 27 CDOs. This result will only be possible if they are all drawn from the same portfolio loss distribution or the CDOs refer to the same collateral asset pool. . . . It would seem extremely improbable that all 27 CDOs could have the same SDRs across all rating scales. The closing dates range from December 28, 2000 to July 19, 2007. One interesting finding is that all but one of the CDOs are rated by a group of credit analysts located in New York City and monitored by one surveillance analyst." We see this as the cleanest evidence that shirking did lead to ratings errors.

Interestingly, the Dodd-Frank Act in the United States also presumes that shirking was a problem during the crisis and takes several steps to try to correct it. First, section 936 of the Act requires the Securities and Exchanges Commission to take steps to guarantee that any person employed by a nationally recognized statistical rating organization "(1) meets standards of training, experience, and competence necessary to produce accurate ratings for the categories of issuers whose securities the person rates; and (2) that employees are tested for knowledge of the credit rating process." The law also requires the agencies to identify and then notify the public and other users of ratings which five assumptions would have the largest impact on their ratings in the event that they were incorrect.

Fourth, revisions to ratings are typically slow to occur. This issue attracted considerable attention in the early 2000s when the rating agencies were slow to identify problems at Worldcom and Enron ahead of their bankruptcies. But, Covitz and Harrison (2003) show that 75% of the price adjustment of a typical corporate bond in the wake of a downgrade occurs prior to the announcement of the downgrade. So these delays are pervasive.

Finally, it appears that competition among rating agencies reduces the accuracy of ratings. Direct evidence on this comes from Becker and Milbourn (2011), who study how the

[7] See also Morgan (2002) who argues that banks (and insurance companies) are inherently more opaque than other firms, and this opaqueness explains his finding that Moody's and S&P differ more in their ratings for these intermediaries than for non-banks.

rise in market share by Fitch influenced ratings by Moody's and S&P (who had historically dominated the industry). Prior to its merger with IBCA in 1997, Fitch had a very low market share in terms of ratings. Thanks to that merger, and several subsequent acquisitions over the next five years, Fitch substantially raised its market share, so that by 2007 it was rating around 1/4 of all the bonds in a typical industry. Becker and Milbourn (2011) exploit the cross-industry differences in Fitch's penetration to study competitive effects. They find an unusual pattern. Any given individual bond is more likely to be rated by Fitch when the ratings from the other two big firms are relatively low.[8] Yet, in the sectors where Fitch issues more ratings, the overall ratings for the sector tend to be higher.

This pattern is not easily explained by the usual kind of catering that the rating agencies have been accused of. If Fitch were merely inflating its ratings to gain business with the poorly performing firms, the Fitch intensive sectors would be ones with more ratings for these underperforming firms and hence lower overall ratings. This general increase in ratings suggests instead a broader deterioration in the quality of the ratings, which would be expected if Fitch's competitors saw their rents declining; consistent with this view, the forecasting power of the ratings for defaults also declined.

Griffin, Nickerson, and Tang (2013) do find patterns consistent with competitive forces leading to catering in the ratings of CDOs. They show that when Moody's and S&P both rated CDOs, the AAA tranches were more likely to default than when only one of them granted a rating (even though investors accepted lower interest rates on dual-rated deals relative to solo-rated ones). In particular, they demonstrate that each of the firms would judgmentally adjust their ratings upward to match the other one whenever a pure model-based rating by one of the firms was lower than the other.

1.2 Theoretical models of the rating business

Next, we compare our paper with the many theoretical studies on rating agencies that have been proposed to explain these and other facts.[9] However, we believe our paper is the only

---

[8] Bongaerts, Cremers, and Goetzmann (2012) identify another interesting competitive effect. If two CRAs disagree about whether a security qualifies as an investment grade, then it does not qualify as an investment grade. But if a third rating is sought and an investment grade rating is given, then the security does qualify. Since Moody's and S&P rate virtually every security, this potential of tiebreaking creates an incentive for an issuer to seek an opinion from Fitch when the other two disagree. The authors find exactly this pattern: Fitch ratings are more likely to be sought precisely when Moody's and S&P disagree about whether a security is of investment-grade quality.

[9] While not applied to rating agencies, there are a number of theoretical papers on delegated information acquisition—see, for example, Chade and Kovrijnykh (forthcoming), Gromb and Martimort (2007), and

one that simultaneously accounts for the five observations described above.

The paper by Bongaerts (2014) is closest to ours in that it studies an environment where a CRA's effort that determines rating precision is unobservable, and like us, he is interested (among other things) in comparing issuer- and investor-pays models. Unlike us, he assumes that projects produce private benefits for the owner of the technology, which create incentives for owners to fund bad projects. Also, he allows for heterogeneous competition, where issuer- and investor-paid CRAs coexist and compete for business. Finally, while his model is dynamic, he analyzes stationary rather than Pareto-efficient equilibria.

Opp, Opp, and Harris (2013) also have a model where a CRA's effort affects rating precision, but unlike us, they assume that it is observable, and they do not study optimal contracts. They find that introducing rating-contingent regulation leads the rating agency to rate more firms highly, although it may increase or decrease rating informativeness.[10]

Bolton, Freixas, and Shapiro (2012) study a model where a CRA receives a signal about a firm's quality, and can misreport it (although investors learn about a lie ex post). Some investors are naive, which creates incentives for the CRA—which is paid by the issuer—to inflate ratings. The authors show that the CRA is more likely to inflate ratings in booms, when there are more naive investors, and/or when the risks of failure, which could damage the CRA's reputation are lower. Unlike in our model, in theirs both the rating precision and reputation costs are exogenous. Similar to us, the authors predict that competition among CRAs may reduce market efficiency, but for a very different reason than we do: the issuer has more opportunities to shop for ratings and to take advantage of naive investors by only purchasing the best ratings. In contrast, we assume rational expectations, and predict that larger rating errors occur because of insufficient effort.

Our result that competition reduces surplus is also reminiscent of the result in Strausz (2005) that certification constitutes a natural monopoly. In Strausz this result obtains because honest certification is easier to sustain when certification is concentrated with one party. In contrast, in our model the ability to charge a higher price increases rating accuracy even when the CRA cannot lie.

Skreta and Veldkamp (2009) analyze a model where the naïveté of investors gives issuers incentives to shop for ratings by approaching several rating agencies and publishing only favorable ratings. They show that a systematic bias in disclosed ratings is more likely to

references therein.

[10] See also a recent paper by Cole and Cooley (2014), who argue that distorted ratings during the financial crisis were more likely caused by regulatory reliance on ratings rather than by the issuer-pays model.

occur for more complex securities—a finding that resembles our result that rating errors are larger for new securities. Similar to our findings, in their model, competition also worsens the problem. They also show that switching to the investor-pays model alleviates the bias, but as in our setup the free-rider problem can eliminate the ratings market completely.

Sangiorgi and Spatt (2015) generate ratings shopping in a model with rational investors. In equilibrium, investors cannot distinguish between issuers who only asked for one rating, which turned out to be high, and issuers who asked for two ratings and only disclosed the second high rating but not the first low one. They show that too many ratings are produced, and while there is ratings bias, there is no bias in asset pricing as investors understand the structure of the equilibrium. While we conjecture that a similar result would hold in our model, the analysis of the case where multiple ratings are acquired in equilibrium is hard since, unlike in Sangiorgi and Spatt (2015), the rating quality is endogenous in our setup.

Similar to us, Faure-Grimaud, Peyrache, and Quesada (2009) study optimal contracts between a rating agency and a firm, but their focus is on providing incentives to the firm to reveal its information, while we focus on providing incentives to the CRA to exert effort. Goel and Thakor (2011) have a model where the CRA's effort is unobservable, but they do not analyze optimal contracts; instead, they are interested in the impact of legal liability for "misrating" on the CRA's behavior.

As we later discuss, the structure of our optimal contracts can be interpreted as a reduced form of the optimal reputational mechanism that would arise in a fully dynamic model. Our mechanism differs, however, from the one in the well-known paper by Mathis, McAndrews, and Rochet (2009).[11] The friction in their model is adverse selection (the CRA's type is unobservable), while the main friction in ours is moral hazard (the CRA's effort is unobservable). They use a different concept of reputation than we do. In their model the CRA can be one of two possible types—committed to tell the truth or opportunistic—and reputation is the investors' belief that the CRA is committed. In our model, reputation captures how the CRA's future profits change based on how project performance matches the announced ratings.[12] Mathis, McAndrews, and Rochet (2009) show that when the CRA perfectly observes the project's quality, an opportunistic CRA lies (i.e., gives a good rating to a bad security) with some probability if the fraction of the CRA's income that comes from rating the complex products is large enough. If reputation is high enough, then

[11] Other papers that model reputational concerns of rating agencies include, for example, Bar-Isaac and Shapiro (2013) and Fulghieri, Strobl, and Xia (2014).

[12] Also, we consider equilibria that depend on the whole history of events, while Mathis, McAndrews, and Rochet (2009) look at Markov equilibria.

the opportunistic CRA lies with probability one. Importantly, if the CRA's signal about the project's quality is imperfect, then the incentive provision collapses completely and the opportunistic CRA will always lie. In our model, the optimal fee structure is designed so that the CRA does not lie. Also, and perhaps more importantly, incentives (for effort and truthful reporting) are provided even though the CRA does not observe the project's quality with certainty.

1.3 Summary

Our literature review is intended to make three points. First, there is substantial evidence suggesting that shirking by rating agencies is a genuine issue. We are not saying that it is the only issue that is relevant for CRAs, but it seems very difficult to deny that it is present. Second, there now are many facts about the types of problems in the rating business. We believe that the more a single model can explain, the better, and that is one of our goals in what follows. Accounting for moral hazard helps us to simultaneously explain several facts, which to us strengthens the presumption that it matters in the ratings business.

Finally, the theoretical approach we take is very different from past approaches. Very few other papers look at optimal contracts between the CRA and its clients. Of the few that do, none explore how contracts differ depending on who pays for the ratings. Instead, the majority of the literature takes certain features of the ratings process as given and tries to understand their implications. Relative to these papers, our framework is valuable because it allows us to separate the fundamental problems that come from one business model or institutional arrangement versus another, from those that arise because of a badly designed compensation scheme (that could perhaps be eliminated with better contracting).

2 The Model

We consider a one-period model with one firm, a number ($n \geq 2$) of investors, and one credit rating agency. All agents are risk neutral and maximize expected profits.

The firm (the issuer of a security) is endowed with a project that requires one unit of investment (in terms of the consumption good) and generates the end-of-period return, which equals y units of the consumption good in the event of success and 0 in the event of failure. The likelihood of success depends on the quality of the project, q. The project's quality can be good or bad, $q \in \{g, b\}$, and is unobservable to everyone. A project of

	High signal ($\theta = h$)	Low signal ($\theta = \ell$)
Good quality ($q = g$)	$\alpha + \beta_h e$	$1 - \alpha - \beta_h e$
Bad quality ($q = b$)	$\alpha - \beta_\ell e$	$1 - \alpha + \beta_\ell e$

Table 1
Information structure

quality q succeeds with probability p_q, where $0 < p_b < p_g < 1$. We assume that $-1 + p_b y < 0 < -1 + p_g y$, so that it is profitable to finance a high-quality project but not a low-quality one. The prior belief that the project is of high quality is denoted by γ, where $0 < \gamma < 1$.

The CRA can acquire information about the quality of the project. It observes a signal $\theta \in \{h, \ell\}$ that is correlated with the project's quality. The informativeness of the signal about the project's quality depends on the level of effort $e \geq 0$ that the CRA privately exerts. Specifically, $\Pr\{\theta = h | q = g, e\} = \alpha + \beta_h e$ and $\Pr\{\theta = h | q = b, e\} = \alpha - \beta_\ell e$.[13] Table 1 shows the full matrix of probabilities of observing a particular signal realization conditional on the project's quality. We assume that $0 < \alpha < 1$, $\beta_i \geq 0$, $i = h, \ell$, and $\beta_\ell + \beta_h > 0$. Also, to ensure that the probabilities are between zero and one, we require $e \leq \bar{e}$, where $\bar{e} = \min\{(1 - \alpha)/\beta_h, \alpha/\beta_\ell\}$.

Note that if effort is zero, the conditional distribution of the signal is the same regardless of the project's quality (the high signal is observed with probability α and the low one with $1 - \alpha$), and thus the signal is uninformative. Conditional on the project being of a certain quality, the probability of observing a signal consistent with that quality is increasing in effort. So higher effort makes the signal more informative in Blackwell's sense.[14]

The assumed information structures nests the extreme cases $\beta_h = 0$ or $\beta_\ell = 0$, as well as the symmetric case with $\beta_h = \beta_\ell$. When $\beta_h = 0$, the CRA's effort only affects the distribution of the signal if the project's quality is low, so the CRA's effort matters only in detecting bad projects. The situation is reversed if instead $\beta_\ell = 0$. And when $\beta_h = \beta_\ell$, the CRA's effort increases the likelihood of observing a signal consistent with the project's quality by the same amount in both states.

Exerting effort e entails a cost of $\psi(e)$ to the CRA. The function ψ satisfies $\psi(0) = 0$, $\psi'(e) > 0$, $\psi''(e) > 0$, $\psi'''(e) > 0$ for all $e > 0$, and $\lim_{e \to \bar{e}} \psi(e) = +\infty$. The assumptions on the second and third derivatives of ψ guarantee that the CRA's and planner's problems, respectively, are strictly concave in effort, so that the first-order conditions describe the

[13] The information structure follows Chade and Kovrijnykh (forthcoming).
[14] See Blackwell and Girshick (1954), chapter 12.

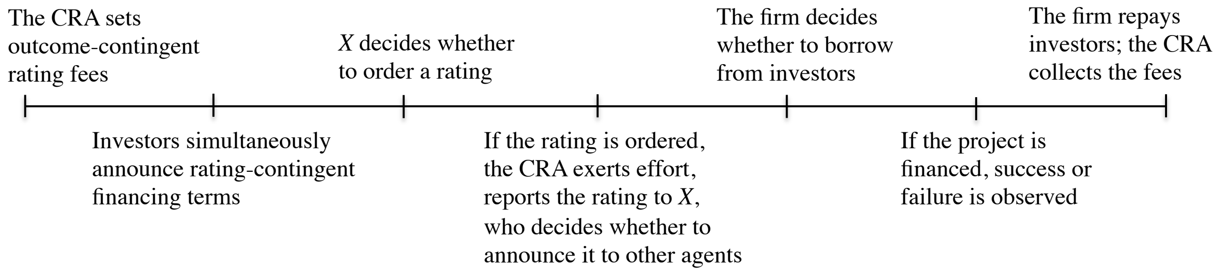

Figure 1
Timing

global optimum.[15] We also assume that $\psi'(0) = 0$ and $\psi''(0) = 0$, which guarantee an interior solution for effort in the CRA's and planner's problems, respectively.

We assume that the signal realization is the CRA's private information so that the CRA can potentially misreport it. Thus, in addition to the moral hazard problem due to effort unobservability, there is also an adverse selection problem due to the signal unobservability. While allowing for misreporting affects the form of the optimal CRA compensation, it does not fundamentally alter other key implications of the model. In other words, we allow for misreporting mostly as an appeal to realism, and it is neither needed for, nor changes, any important results. Finally, we also assume that the CRA is protected by limited liability, so that all payments that it receives must be non-negative.

The firm has no internal funds, and hence needs investors to finance the project. Investors have funds, behave competitively, and will make zero profits in equilibrium.

We will consider three scenarios depending on who decides whether a rating is ordered—the social planner, the issuer, or each of the investors. Let X refer to the identity of the player ordering a rating. The timing of events, illustrated in Figure 1, is as follows.

At the beginning of each period, the CRA posts a compensation schedule—the fees to be paid at the end of the period, conditional on the outcome.[16] Each investor announces

[15] Convexity of the marginal disutility of effort ψ' ensures that the planner's marginal cost of implementing effort under moral hazard is increasing in e. This is a common assumption in principal-agent problems—see, e.g., Jewitt, Kadan, and Swinkels (2008). Technically, since the planner's problem includes the first-order condition with respect to effort from the CRA's problem as an incentive constraint, we impose (sufficient) conditions not only on the second but also on the third derivatives to guarantee that local second-order conditions are satisfied globally.

[16] When X is the firm, it might not be able to pay for a rating if the compensation structure requires payments when no output is generated. Thus we assume that in this case the firm can borrow from investors in order to pay to the CRA, and that the firm repays the loan out of generated output in the event of the project's success. Since this stage is not important in our analysis, for simplicity of exposition we exclude it from the timeline depicted in Figure 1.

project financing terms (interest rates) that are contingent on a rating or the absence of one. Then X decides whether to ask for a rating, and chooses whether to reveal to the public that a rating has been ordered.[17] If a rating is ordered, the CRA exerts effort, observes a signal realization, and reports a rating to X, who then decides whether it should be published (and hence made known to other agents). The firm decides whether to borrow from investors in order to finance the project. If the project is financed, its success or failure is observed. The firm repays investors, and the CRA collects its compensation. (We assume that X can commit to paying the fees due to the CRA, and that the firm can commit to paying investors.)

We are interested in analyzing perfect Bayesian equilibria with the highest total surplus. The rationale for considering total surplus comes from thinking about a hypothetical consumer who owns both the firm and CRA, in which case it would be natural for the social planner to maximize the consumer's utility.

3 Analysis and Results

Before deriving any results, it will be convenient to introduce some notation. First, let π_1 denote the ex-ante probability of success (before observing a rating), so $\pi_1 = p_g\gamma + p_b(1-\gamma)$. Then the ex-ante probability of failure is $\pi_0 = 1 - \pi_1$. Next, let $\pi_h(e)$ be the probability of observing a high rating given effort e, that is, $\pi_h(e) = (\alpha + \beta_h e)\gamma + (\alpha - \beta_\ell e)(1-\gamma)$. The probability of observing a low rating given effort e is then $\pi_\ell(e) = 1 - \pi_h(e)$. Also, let $\pi_{h1}(e)$ and $\pi_{h0}(e)$ denote the probabilities of observing a high rating followed by the project's success/failure given effort e: $\pi_{h1}(e) = p_g(\alpha + \beta_h e)\gamma + p_b(\alpha - \beta_\ell e)(1-\gamma)$ and $\pi_{h0}(e) = (1-p_g)(\alpha + \beta_h e)\gamma + (1-p_b)(\alpha - \beta_\ell e)(1-\gamma)$. Similarly, the probabilities of observing a low rating followed by success/failure given e are $\pi_{\ell 1}(e) = p_g(1-\alpha-\beta_h e)\gamma + p_b(1-\alpha+\beta_\ell e)(1-\gamma)$ and $\pi_{\ell 0}(e) = (1-p_g)(1-\alpha-\beta_h e)\gamma + (1-p_b)(1-\alpha+\beta_\ell e)(1-\gamma)$.

The probability of observing a high rating bears directly on the earlier discussion of the possibility that rating agencies issue inflated ratings for securities that eventually fail. In our model, when the CRA puts insufficient effort, its ratings will be unreliable. Thus, for bad projects, the underprovision of effort will make it more likely to incorrectly assign high ratings. Moreover, the information structure given by Table 1 implies that unconditionally

[17] Equivalently, we could instead assume that everyone automatically observes whether or not a rating has been ordered, but they do not learn the rating unless X reveals it. What matters is that when X is the firm and it decides not to order a rating, then it must be able to credibly announce this fact to investors. We discuss this issue further in Section 3.3.

the high rating is produced more often when less effort is put in if $\pi_h'(e) < 0$, that is, for $\gamma < \beta_\ell/(\beta_\ell + \beta_h)$. Notice that the lower the ratio β_h/β_ℓ—that is, the more important the CRA's effort in detecting bad projects relative to recognizing good ones,—the higher this cutoff. In particular, in the extreme case $\beta_h = 0$, the cutoff is equal to one, and lower effort will always lead to more ratings inflation. As we will show later, it will not be optimal to acquire ratings when γ is close enough to either zero or one. Thus, even if β_h/β_ℓ is not zero but is small enough, lower effort by the CRA will lead to more (unconditional) ratings inflation for all priors in equilibrium. It seems plausible to assume that β_h/β_ℓ is low, so that detecting bad securities takes more effort than is needed to identify good ones.

3.1 First best

As a benchmark, we begin by characterizing the first-best case, where the CRA's effort is observable, and the social planner decides whether to order a rating.[18] Given a rating (or the absence of one), the project is financed if and only if it has a positive net present value (NPV). Thus, the total surplus in the first-best case is

$$S^{FB} = \max_e -\psi(e) + \pi_h(e) \max\left\{0, -1 + \frac{\pi_{h1}(e)}{\pi_h(e)} y\right\} + \pi_\ell(e) \max\left\{0, -1 + \frac{\pi_{\ell 1}(e)}{\pi_\ell(e)} y\right\},$$

where $\pi_{i1}(e)/\pi_i(e)$ is the conditional probability of success after a rating $i \in \{h, \ell\}$ given the level of effort e. Notice that since $\pi_{h1}(e)/\pi_h(e) \geq \pi_{\ell 1}(e)/\pi_\ell(e)$, with strict inequality if $e > 0$, the project will never be financed after the low rating if it is not financed after the high rating. So only the following three cases can occur: (i) the project is financed after both ratings, (ii) the project is not financed after both ratings, and (iii) the project is financed after the high rating but not after the low rating. It immediately follows that in cases (i) and (ii) the optimal effort choice is zero: it is never efficient to expend effort if the information it produces is not used. In case (iii), the optimal effort, e^*, is strictly positive and (given our assumptions) uniquely solves $\max_e -\psi(e) - \pi_h(e) + \pi_{h1}(e)y$. Thus,

$$S^{FB} = \max\{0, -1 + \pi_1 y, \max_e -\psi(e) - \pi_h(e) + \pi_{h1}(e)y\}.$$

Letting e^{FB} denote first-best effort, the following lemma describes how the prior γ determines which alternative the planner picks.

[18] It is easy to check that with observable effort, the total surplus does not depend on who orders a rating.

Lemma 1. There exist thresholds $\underline{\gamma}^{FB}$ and $\bar{\gamma}^{FB}$ satisfying $0 < \underline{\gamma}^{FB} < \bar{\gamma}^{FB} < 1$, such that

(i) $e^{FB} = 0$ for $\gamma \in [0, \underline{\gamma}^{FB}]$, and the project is never financed;

(ii) $e^{FB} = 0$ for $\gamma \in [\bar{\gamma}^{FB}, 1]$, and the project is always financed;

(iii) $e^{FB} > 0$ for $\gamma \in (\underline{\gamma}^{FB}, \bar{\gamma}^{FB})$, and the project is only financed after the high rating.

The intuition behind this result is quite simple. If the prior about the project's quality is close to either zero or one, so that investment opportunities are thought to be either very good or very bad, then it does not pay off to acquire information about the project.

We now turn to the analysis of the more interesting case when the CRA's effort is unobservable, the CRA can misreport its signal, and payments are subject to limited liability.

3.2 Second best: The social planner orders a rating

It will be convenient to first analyze the case where the planner gets to decide whether to order a rating and in doing so sets the CRA's compensation structure. This construct allows us to write a standard optimal contracting problem and characterize the constrained Pareto frontier. We identify an equilibrium on the frontier where the total surplus is maximized and demonstrate that it is the same one that prevails when the CRA chooses the fees (which is the actual assumption in our model).

Just as in the first-best case, there are three options: do not acquire a rating and do not finance the project, do not acquire a rating and finance the project, and acquire a rating and finance the project only if the rating is high. In the first two cases the CRA exerts no effort, so only in the third case is there a nontrivial problem of finding the optimal compensation structure. To allow for the richest possible contract space, the compensation must be contingent on all possible outcomes. When the project is financed only after the high rating, there are three possible outcomes: the rating is high and the project succeeds, the rating is high and the project fails, and the rating is low (in which case the project is not financed). Let f_{h1}, f_{h0}, and f_ℓ denote the payments to the CRA in each case.

On the Pareto frontier, the payoff to one party is maximized subject to delivering at least certain payoffs to other parties. Since investors earn zero profits, we can maximize the value to the firm subject to delivering at least a certain value to the CRA. Let $u(v)$ denote the value to the firm given that the value to the CRA is at least v, and the project is financed only after the high rating. It can be written as

$$u(v) = \max_{e, f_{h1}, f_{h0}, f_\ell} -\pi_h(e) + \pi_{h1}(e)y - \pi_{h1}(e)f_{h1} - \pi_{h0}(e)f_{h0} - \pi_\ell(e)f_\ell \tag{1}$$

$$\text{subject to } -\psi(e) + \pi_{h1}(e)f_{h1} + \pi_{h0}(e)f_{h0} + \pi_\ell(e)f_\ell \geq v, \tag{2}$$

$$\psi'(e) = \pi'_{h1}(e)f_{h1} + \pi'_{h0}(e)f_{h0} + \pi'_\ell(e)f_\ell, \tag{3}$$

$$-\psi(e) + \pi_{h1}(e)f_{h1} + \pi_{h0}(e)f_{h0} + \pi_\ell(e)f_\ell \geq \max\{\pi_1 f_{h1} + \pi_0 f_{h0}, f_\ell\}, \tag{4}$$

$$e \geq 0,\ f_{h1} \geq 0,\ f_{h0} \geq 0,\ f_\ell \geq 0. \tag{5}$$

Constraint (2) ensures that the CRA's profits are at least v. Constraint (3) is the incentive constraint, which reflects the fact that the CRA chooses its effort privately, and is obtained by maximizing the left-hand side of (2) with respect to e. Constraint (4) precludes the CRA from misreporting the rating. Given that we have both the moral hazard and adverse selection problems, we need to worry about double deviations. However, it is easy to show that whenever the CRA plans to misreport a signal, it optimally exerts zero effort. The left-hand side of (4) is the CRA's payoff if it exerts effort e and truthfully reports the acquired signal. The right-hand side is the payoff from exerting no effort and always reporting the rating that delivers the highest expected compensation. Notice that constraint (4) is equivalent to imposing the following pair of constraints:

$$-\psi(e) + \pi_{h1}(e)f_{h1} + \pi_{h0}(e)f_{h0} + \pi_\ell(e)f_\ell \geq \pi_1 f_{h1} + \pi_0 f_{h0}, \tag{6}$$

$$-\psi(e) + \pi_{h1}(e)f_{h1} + \pi_{h0}(e)f_{h0} + \pi_\ell(e)f_\ell \geq f_\ell. \tag{7}$$

The constraints in (5) reflect limited liability and the nonnegativity of effort. Finally, we assume that the firm can choose not to operate at all, so its profits must be nonnegative, that is, $u(v) \geq 0$, which restricts the values of v that can be delivered to the CRA.

Our first main result demonstrates how the optimal compensation must be structured in order to provide incentives to the CRA to exert effort and to report the rating truthfully.

Proposition 1 (**Optimal Compensation Structure**). Suppose the project is financed only after the high rating, and the implemented effort is below the first-best level e^*. Then $f_{h1} > 0$, $f_\ell > 0$, and $f_{h0} = 0$.[19] Furthermore, there is a threshold $\hat{\gamma} \in [0, 1]$ such that (6) will bind for $\gamma > \hat{\gamma}$ and (7) will bind for $\gamma < \hat{\gamma}$.

[19] If e^* is implemented, then $f_{h1} > f_{h0}$, $f_\ell > f_{h0}$, and $f_{h0} \geq 0$. This is equivalent to paying an upfront fee equal to f_{h0} and rewarding the CRA with $f_{h1} - f_{h0}$ and $f_\ell - f_{h0}$ after the high rating followed by the project's success and after the low rating, respectively.

The proposition states that the CRA should be rewarded in only two cases: if it announces the high rating and the project succeeds or if it announces the low rating. Quite intuitively, the CRA is never paid for announcing the high rating if it is followed by the project's failure.[20]

The CRA's ability to misreport the rating is crucial for the result that both f_{h1} and f_ℓ must be positive. In the absence of (4) we would have $f_{h1} > 0 = f_\ell = f_{h0}$ if $\gamma > \hat{\gamma}$, and $f_\ell > 0 = f_{h1} = f_{h0}$ otherwise.[21] Given this, the incentive to always report the high (low) rating constrains the compensation scheme when $\gamma > \hat{\gamma}$ ($\gamma < \hat{\gamma}$), as Proposition 1 states.

Our presumption that the compensation structure is contingent on the rating and the project's performance might appear unrealistic at first. Instead, one might prefer to analyze a setup where fees are paid up front. But, in any static model an upfront fee will never provide the CRA with incentives to exert effort—the CRA will take the money and shirk. So it is necessary to introduce some sort of reward for accuracy to prevent shirking.

Many papers in this literature impose exogenous penalties and rewards that influence CRA behavior; for example, Bolton, Freixas, and Shapiro (2012) (see also references therein) introduce exogenous reputation costs. They essentially assume that investors can punish the CRA by withholding business and thus the value of future profits when the CRA is not caught lying serves as a disciplining device. In our paper, the CRA's outcome-contingent payoff can be interpreted in precisely this way, except that the reputation costs are endogeneous because the compensation structure is endogenous.

In the online Appendix, we present a repeated infinite-horizon model that mimics the key features of our static model, to formally analyze the optimal reputation structure. There, we allow for an infinitely-lived CRA, infinitely-lived investors, and a sequence of firms (with i.i.d. projects) who operate for a single period, but who are informed of all previous play and correctly form expectations about all future play when choosing their actions. In each period the CRA charges an upfront (flat) fee, but the fee can vary over time. Formally, in the recursive formulation the CRA's "continuation values" (future present discounted profits) depend on histories. Thus even if the fees are restricted to be paid up front in each period, the CRA will be motivated to exert effort by the prospect of higher

[20] The stark result that the CRA's limited liability constraint binds after the $h0$ outcome so that the CRA receives nothing if it makes a "mistake" is an artifact of the one-period setup. An analog of this result in an infinitely-repeated version of the model that we discuss in the online Appendix is that the punishment for a "mistake" involves a fall in the present discounted value of CRA's future profits.

[21] The proof relies on the standard maximum likelihood ratio argument: the CRA should be rewarded for the event whose occurrence is the most consistent with its exerting effort, which in turn depends on the prior—see the proof of Proposition 1 in the Appendix.

future profits—via the ability to charge higher future fees—that follow from developing a "reputation" by correctly predicting the firms' performance. Market participants will be willing to pay those higher fees because they will rationally anticipate that the CRA will be motivated to produce high-quality ratings when it is appropriately compensated for its effort. So, unlike in our static model, the outcome-contingent compensation structure is not simply the CRA's choice, but is tied to future strategies of all market participants.

The dynamic model is not only much more complicated to analyze, but also yields no new important insights. At the same time, the only way to approximate the critical role that reputation plays in the dynamic setting in our static model is to allow the compensation to depend on outcomes. So outcome-contingent compensation should not be interpreted literally, but instead should be recognized as a simplification to bring reputational considerations into the analysis in a tractable way. Conversely, if we ruled out this kind of compensation, it would be impossible to provide incentives that are needed to elicit effort, and the static model would have very different properties than the dynamic one.

The next proposition derives several properties of the Pareto frontier that will be important for our subsequent analysis.

Proposition 2 (**Pareto Frontier**)**.** Suppose the project is financed only after the high rating.

(i) There exists v^* such that for all $v \geq v^*$ $e(v) = e^*$. Moreover, $u(v) < 0$ for $v \geq v^*$ if $\gamma \geq \hat{\gamma}$, and $u(v) < -1 + \pi_1 y$ for $v \geq v^*$ if $\gamma < \hat{\gamma}$.

(ii) There exists $v_0 > 0$ such that (2) is slack for $v < v_0$ and binds for $v \geq v_0$, so that u is strictly decreasing in v for $v \geq v_0$. Moreover, $e(v_0) > 0$.

(iii) Effort and total surplus are increasing in v, strictly increasing for $v \in (v_0, v^*)$.

Part (i) says that there is a threshold value, v^*, above which the first-best effort is implemented. Notice, however, that if $\gamma \geq \hat{\gamma}$ or $-1 + \pi_1 y \leq 0$, that is, if $\gamma \leq \gamma_0 \equiv (1/y - p_b)/(p_g - p_b)$, the resulting profit to the firm is strictly negative, violating individual rationality, and so this arrangement cannot be sustained in equilibrium. Obtaining the first best requires $\gamma \in (\gamma_0, \hat{\gamma})$, and this set could be empty.[22]

There is an interesting economic reason why implementing the first-best effort requires the firm's profits to be negative when, for example, $\gamma \geq \hat{\gamma}$. To convey the intuition, suppose

[22]Notice that this condition on γ is only necessary but not sufficient for $u(v^*)$ to be non-negative. In fact, e.g., in the symmetric case with $\beta_h = \beta_\ell$ and $\alpha = 1/2$, it can be shown that $u(v^*) < 0$ for all γ.

the CRA cannot misreport the rating.[23] In this case, when $\gamma \geq \hat{\gamma}$, the CRA is only paid after outcome $h1$. With observable effort, the problem is equivalent to one where the firm acquires information itself: $\max_e -\psi(e) + \pi_{h1}(e)(y - R(e))$, where $R(e)$ is the break-even interest rate, $-\pi_h(e) + \pi_{h1}(e)R(e) = 0$. So when the firm chooses effort, it accounts for two effects. One is that higher effort increases the probability that a surplus is generated, $\pi_{h1}(e)$. The other is that more effort delivers a more accurate rating, which investors reward by lowering the interest rate $R(e)$. The lower interest rate increases the size of the surplus.

When the CRA's effort is unobservable, the CRA internalizes the fact that more effort generates a higher probability of the fee being paid. However, its fees cannot be contingent on effort. Formally, the CRA solves $\max_e -\psi(e) + \pi_{h1}(e)f_{h1}$, where f_{h1} does not depend on e. Thus, the only way to implement the first-best level effort is to set f_{h1} above $y - R(e)$, which leaves the firm with negative profits, $\pi_{h1}(e)(y - R(e) - f_{h1}) < 0$.

It will be handy to denote the highest value that can be delivered to the CRA without leaving the firm with negative profits by $\bar{v} \equiv \max\{v | u(v) = 0\}$.

Part (ii) of Proposition 2 identifies the lowest value that can be delivered to the CRA on the Pareto frontier. This value, denoted by v_0, is strictly positive. So the rating agency will still be making profits and will exert positive effort. For $v \leq v_0$ $u(v) = u(v_0)$, while for $v \geq v_0$ constraint (2) binds, and thus $u(v)$ is strictly decreasing in v.

Finally, part (iii) shows that the higher the CRA's profits, the higher the total surplus, and the higher the effort. This is an important result, and will be crucial for our further analysis. Intuitively, when effort is unobservable (and there is limited liability), higher fees are required to give incentives to the CRA to exert more effort.[24] To implement the highest possible effort, one needs to extract all surplus from the firm and give it to the CRA. However, as part (i) implies, implementing the first-best level of effort often results in negative profits to the firm. Combining (i) and (iii) tells us that the level of effort that can be implemented is strictly smaller than the first-best one whenever $\gamma \geq \hat{\gamma}$ or $\gamma \leq \gamma_0$.

Importantly, while a higher payoff to the CRA increases the total surplus, it makes the firm worse off. The firm's payoff is maximized at v_0, which is the lowest payoff to the CRA

[23] Without misreporting the payoff to the firm is higher for a fixed payoff to the CRA. So if the firm's payoff is negative without misreporting, it will only be more so when misreporting is allowed.

[24] Clearly, our assumption of limited liability plays an important role in these results. Without it, it would be possible to punish the CRA in some states and achieve the first best for all v. In particular, selling the project to the CRA and making it an investor would provide it with incentives to exert the first-best level of effort. However, forcing rating agencies to co-invest does not appear to be a practical policy option, as it would require them to have implausibly large levels of wealth, given that they rate trillions of dollars' worth of securities each year.

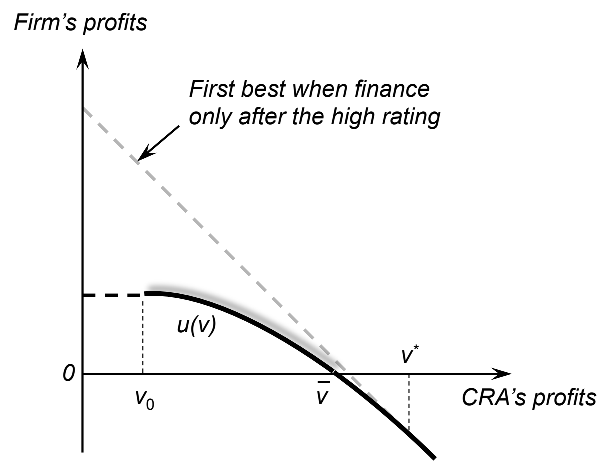

Figure 2
The function $u(v)$
The Pareto frontier when the project is financed only after the high rating is the shaded area of the $u(v)$ curve. The case where $u(v^*) < 0$ is depicted.

on the frontier. Thus, while the planner wants a more precise rating, the firm actually prefers a less precise rating (but still an informative one, as effort is positive at v_0).

The function $u(v)$ is graphed in Figure 2. Recall that the set $\{(v, u(v))|v \geq 0, u(v) \geq 0\}$ is the Pareto frontier conditional on the project being financed only after the high rating. Specifically, conditional on such a strategy being optimal, each point on this frontier corresponds to an equilibrium where, given that the payoff to the CRA is at least v, the payoff to the firm is maximized by optimally choosing the compensation structure. The corresponding total surplus and implemented effort are $v + u(v)$ and $e(v)$, respectively. In addition, there are two other cases to consider. If the solution to problem (1)$-$(5) involves an effort level such that the NPV is positive (negative) after both ratings, then the planner would choose not to acquire a rating and investors will always (never) finance. Combining the three cases, the total surplus is $\max\{0, -1 + \pi_1 y, v + u(v)\}$.

Recall that we are considering equilibria where the total surplus is maximized. It immediately follows from Proposition 2 that if the project is financed only after the high rating, then the planner will choose the point $(\bar{v}, u(\bar{v}))$ on the frontier. This corresponds to maximum feasible CRA profits and effort, and zero profits for the firm.[25] The implemented effort, which we denote by e^{SB} (where SB stands for the second best), is smaller than e^{FB}, and is strictly smaller at least for some priors (for which $u(v^*) < 0$).

[25]If $\bar{v} > v^*$, then choosing any $v \in [v^*, \bar{v}]$ is feasible for the planner and yields the same total surplus and effort.

To close the loop, let us return to the issue of what happens if the CRA rather than the planner sets the fees. When the planner does it (and the project is financed only after the high rating), all the surplus goes to the CRA. Thus, the CRA will select the same fees as the planner.

We summarize our results in the following proposition.

Proposition 3 (X = **Planner**)**.** If the planner decides whether to order a rating, then

(i) The maximum total surplus in equilibrium is $S^{SB} = \max\{0, -1 + \pi_1 y, \bar{v} + u(\bar{v})\}$;

(ii) $e^{SB} \leq e^{FB}$ and $S^{SB} \leq S^{FB}$, with strict inequalities if $e^{FB} > 0$ and either $\gamma > \hat{\gamma}$ or $-1 + \pi_1 y \leq 0$.

It is straightforward to show that rating acquisition will take place for a smaller set of priors under the second best than under the first best. Formally, one can prove an analog of Lemma 1, and show that the bounds for the prior belief between which the rating is acquired in the second-best case—call them $\underline{\gamma}^{SB}$ and $\bar{\gamma}^{SB}$—lie inside the $[\underline{\gamma}^{FB}, \bar{\gamma}^{FB}]$ interval.

Next, we will analyze how the maximum total surplus and the corresponding effort in cases where the issuer or investors order ratings, compare those in the second-best case. We will ask ourselves: does it matter who orders ratings? We will see that the answer depends on the prior belief about the project's quality. The identity of who orders a rating will not matter in "bad times," when the average project has negative NPV (i.e., $-1 + \pi_1 y \leq 0$), but it will matter in "good times," when the expected NPV is positive ($-1 + \pi_1 y > 0$).

3.3 The issuer orders a rating

Consider the case where the firm decides whether to order a rating. Recall that in setting its fees the CRA picks the highest ones that the firm is willing to pay. The firm's willingness to pay is pinned down by its profit if it chooses not to order a rating. Without a rating, investors finance the firm's project if and only if $-1 + \pi_1 y > 0$. Since investors break even, the firm's profit in this case is $\underline{u} \equiv \max\{0, -1 + \pi_1 y\}$.[26] Thus, if a rating is acquired

[26] This argument relies on the assumption that the firm can credibly announce that it did not get rated. Without this assumption, the issuer's payoff is still strictly positive when $-1+\pi_1 y > 0$, although it is lower than $-1 + \pi_1 y$; see Claim 1 in the online Appendix. Furthermore, if instead of the CRA posting fees we assumed that X sets them, then in the issuer-pays model the firm's payoff would be maximized and equal to $u(v_0)$, where v_0 is defined in Proposition 1. Then rating precision under the issuer-pays model would be even lower (as effort is the lowest at v_0). Moreover, the equilibrium in the issuer-pays model would not depend on whether or not the issuer can credibly announce that it did not get rated, as the firm's outside option no longer plays a role in pinning down the equilibrium surplus and effort.

in equilibrium, the firm receives $\underline{u}$, and the corresponding value to the CRA is $v^{iss} \equiv \max\{v|u(v) = \underline{u}\} \leq \bar{v}$ with strict inequality if $-1 + \pi_1 y > 0$.

Proposition 2 can be used to compare v^{iss} with various benchmarks. Part (iii) of the proposition tells us that the total surplus and effort are strictly increasing in v for $v < v^*$. Also, part (i) tells us that $u(v^*)$ is strictly smaller than either 0 or $-1 + \pi_1 y$. Therefore, it is strictly smaller than the maximum of the two, $u(v^*) < \max\{0, -1 + \pi_1 y\} = u(v^{iss})$, implying $v^{iss} < v^*$. This means that the total surplus and effort in the issuer-pays model are strictly lower than in the first best, and also strictly lower than in the second best whenever $-1 + \pi_1 y > 0$.

Denote the total surplus and effort in the issuer-pays case by S^{iss} and e^{iss}, respectively. The preceding logic implies the following results:

Proposition 4 (X = **Issuer**)**.** Suppose the firm decides whether to order a rating. Then

(i) The maximum total surplus in equilibrium is $S^{iss} = \max\{0, -1 + \pi_1 y, v^{iss} + u(v^{iss})\}$;

(ii) If $-1 + \pi_1 y \leq 0$, then $S^{iss} = S^{SB}$ and $e^{iss} = e^{SB}$. If $-1 + \pi_1 y > 0$, then $S^{iss} \leq S^{SB}$ and $e^{iss} \leq e^{SB}$, with strict inequalities so long as $e^{SB} > 0$;

(iii) $S^{iss} \leq S^{FB}$ and $e^{iss} \leq e^{FB}$, with strict inequalities so long as $e^{FB} > 0$.

Importantly, the mechanism that delivers these results is not driven by the CRA trying to cater to firms to gain business. In our model, the CRA must be paid more in order to have incentives to produce a more accurate rating. More accurate ratings raise the total surplus, but they can leave less for the firm. In particular, if the rating precision is above $e(v_0)$, the increase in fees for a more precise rating dominates the decrease in the interest rate. So the firm prefers cheaper and less precise ratings than what is socially optimal. If it was up to the firm, it would choose fees that result in the payoff to the CRA equal to v_0. But since the CRA sets the fees, it extracts everything up to the firm's willingness to pay, but still yielding less accurate ratings than is socially optimal. So this result is not due to the CRA making mistakes to gain business from issuers.

As usual, the firm will decide not to ask for a rating if the prior belief γ is sufficiently close to zero or one. Moreover, since the implemented effort with the firm choosing whether to request a rating is lower relative to when the planner picks, rating acquisition will occur on a smaller set of priors in the former case than in the latter.[27]

[27] Formally, the bounds for the prior belief between which the rating is acquired, $\underline{\gamma}^{iss}$ and $\bar{\gamma}^{iss}$, satisfy $\underline{\gamma}^{iss} = \underline{\gamma}^{SB}$ and $\bar{\gamma}^{iss} < \bar{\gamma}^{SB}$. The equality follows from the fact that $v^{iss} + u(v^{iss}) = \bar{v} + u(\bar{v})$ when $-1 + \pi_1 y \leq 0$, and the inequality follows from $v^{iss} + u(v^{iss}) < \bar{v} + u(\bar{v})$ for $-1 + \pi_1 y > 0$.

3.4 Investors order a rating

Consider finally the case when each investor decides whether to order a rating. It is helpful to recall the timing of events in this case. The period starts with the CRA announcing the fee structure. The investors simultaneously announce rating-contingent interest rates. Then they simultaneously choose whether to commit to pay the fees and order a rating. If a rating is ordered by at least one investor, the CRA exerts effort, and reports the rating to those investors who ordered it. The firm borrows from investors, production takes place, and payments between the agents are made conditional on the outcome.

The assumption that investors who do not pay for a rating can be excluded from learning it is critical. If the spread of information cannot be precluded, investors will want to free-ride on others paying for a rating. As a result, no rating will be acquired in equilibrium, and investors will make their financing decisions solely based on the prior. Until the mid-1970s, the investor-pays model was widely used. However, the rise of photocopying made protecting the sort of information described above increasingly impractical, which arguably resulted in the switch to the issuer-pays model. Recently, though, the investor-pays model has made a comeback with the emergence of several new CRAs who fund themselves with subscriptions from investors (see Cornaggia and Cornaggia, 2013). Hence, it seems relevant to examine what would happen if the subscription-based model, or some other variant of the investor-pays model, were to once again become prevalent.

When X is the social planner or the firm, investors play only a passive role by pricing the loans competitively and earning zero profits. Thus, looking for equilibrium with the highest total surplus was equivalent to solving a fairly standard optimal contracting problem. Now, when each investor makes a decision of whether to ask for a rating, the problem is no longer standard, and solving for the equilibrium becomes much trickier. In fact, for some values of the prior, payoffs in the equilibrium with the highest total surplus actually do not lie on the constrained Pareto frontier described in Section 3.2.

Lemma 2 describes this important inefficiency of the investor-pays model. Recall that for γ close enough to one, it is efficient not to ask for a rating and always finance the project, so that S^{SB} (and S^{iss}) equal $-1+\pi_1 y$. However, financing without a rating never happens in the investor-pays case; investors always ask for a rating, even when it is inefficient.

Lemma 2. Suppose that $-1+\pi_1 y > 0$. Then there is no equilibrium where investors do not ask for a rating and always finance the project. That is, in (any) equilibrium $e^{inv} > 0$.

The intuition is as follows. If the project is financed without a rating, then all surplus from the production, $-1+\pi_1 y$, goes to the firm and the CRA earns nothing. The CRA can try to sell a rating; the planner would not want one unless the generated surplus is at least $-1+\pi_1 y$, nor would the firm if it were deciding, unless its profit were at least that amount. However, when investors decide whether to ask for a rating, they are not concerned with either the total or the firm's surplus. They always make zero profits, and pass along the rating fees to the firm by charging higher interest rates.

Why then do investors choose to order a rating if they earn zero profits either way? Suppose no one asks for a rating regardless of what the fees are. Then if fees are low enough, one investor could generate profits by ordering a rating, hiding it from other investors, only offering financing conditional on the high rating and charging a slightly lower interest rate than do other investors who do not know the rating. Recognizing this, the CRA can set fees low enough to entice someone to ask for a rating.

Proposition 5 further describes the equilibrium properties of the investor-pays model, and compares them with those in the second-best and issuer-pays cases.

Proposition 5 ($X =$ **Investors**)**.** Suppose investors decide whether to order a rating.

(i) If $-1+\pi_1 y \leq 0$ (i.e., $\gamma \leq \gamma_0$), then $S^{inv} = S^{iss} = S^{SB}$ and $e^{inv} = e^{iss} = e^{SB}$.

(ii) Suppose that $-1+\pi_1 y > 0$ (i.e., $\gamma > \gamma_0$).

  (a) Suppose both the planner and investors ask for a rating and finance after the high rating only. Then the rating precision is lower when investors order ratings: $e^{inv} \leq e^{SB}$, with strict inequality unless the first best is achieved, that is, $e^{inv} = e^{SB} = e^*$. When the planner would not ask for a rating (i.e., for $\gamma \geq \bar{\gamma}^{SB}$), investors still do, and so $e^{inv} > e^{SB} = 0$. For all $\gamma > \gamma_0$, the total surplus is lower under the investor-pays regime than under the planner: $S^{inv} \leq S^{SB}$, with strict inequality unless $e^{inv} = e^{SB} = e^*$.

  (b) The investor-pays model produces ratings with higher precision than the issuer-pays regime, $e^{inv} > e^{iss}$. However, the comparison of the total surpluses is ambiguous. If $S^{inv} > -1+\pi_1 y$ so that the investor-pays model generates more surplus than financing without a rating, then $S^{inv} > S^{iss}$. Otherwise, when $S^{inv} < -1+\pi_1 y$, $S^{inv} < S^{iss}$.

In the region $(0, \gamma_0]$ where the project is not optimal to finance ex ante, the investor-pays model delivers the same total surplus and effort as when the planner or the firm orders a rating. Important differences arise only when the project is ex-ante profitable, that is, when $\gamma \in (\gamma_0, 1)$. As Lemma 2 shows, investors always ask for a rating in this case. But, we know that for γ high enough this is not socially optimal, so that the total surplus is below $-1 + \pi_1 y$, which is what would be obtained if the project is not rated and always financed. This means that when the prior is high enough, the total surplus in the investor-pays case is lower than that in the issuer-pays or the second-best cases. Also, if γ is high enough, then e^{SB} and e^{iss} are zero (for $\gamma > \bar{\gamma}^{SB}$ and $\gamma > \bar{\gamma}^{iss}$, respectively), while $e^{inv} > 0$, so that e^{inv} exceeds both e^{SB} and e^{iss}.

Moreover, for high enough γ the project's NPV will be positive even after the low rating, meaning that investors fund the firm after both ratings, and waste the rating resources. Formally, when $-\pi_\ell(e^{inv}) + \pi_{\ell 1}(e^{inv})y > 0$, the total surplus in the investor-pays case is $S^{inv} = -1 + \pi_1 y - \psi(e^{inv})$.

For less favorable priors, the project's NPV is negative after the low rating ($-\pi_\ell(e^{inv}) + \pi_{\ell 1}(e^{inv})y \leq 0$), and investors will only finance after the high rating. In this case, the equilibrium solves the optimal contracting problem (1)−(5) for some specific payoff to the CRA that we denote by v^{inv}. Accordingly, $S^{inv} = u(v^{inv}) + v^{inv}$. Proposition 2 tells us that comparisons of the total surplus and effort for when investors order ratings relative to when the firm or the planner does will depends on how v^{inv} compares to v^{iss} and $\bar{v}$. As we explain below, v^{inv} must lie strictly between v^{iss} and $\bar{v}$. Thus, the investor-pays model generates a lower rating precision than under the second best ($e^{inv} < e^{SB}$), but a higher precision than in the issuer-pays case ($e^{inv} > e^{iss}$).

Competitive considerations explain why investors choose less accurate ratings than a planner would. Recall that in the second best, the value to the CRA is the highest possible ($\bar{v}$) so that the firm earns no profits, that is, the equilibrium interest rate equals y. But competition among investors constrains the interest rates that they will charge. One investor can choose not to ask for a rating and undercut other investors by offering an interest rate $R < y$ and earn $-1 + \pi_1 R$, which is positive so long as $R > 1/\pi_1$. Thus $1/\pi_1$ is the highest interest rate that informed investors can charge in equilibrium.[28] This leaves the firm with $u(v^{inv}) = \pi_{h1}(e^{inv})(y - 1/\pi_1) > 0 = u(\bar{v})$. So v^{inv} must be strictly less than $\bar{v}$.

A different consideration matters in the contrast with the issuer-pays model. In the

[28]Note that although the probability that the project succeeds increases to π_{h1}/π_h after the high rating compared with π_1 without a rating, the informed investors still charge the higher interest rate $1/\pi_1 > \pi_h/\pi_{h1}$. Thus, ignoring the rating fees, the informed investors make rents (after the high rating). In equilibrium, the expected rents exactly equal to the expected cost of the rating fees.

issuer-pays model, the firm's payoff is the same as in the outside option of being financed without a rating, where investors charge $1/\pi_1$ and finance with probability one. In the investor-pays case, the interest rate is the same, but the firm receives financing less often—only when the rating is high. Formally, $u(v^{iss}) = -1 + \pi_1 y = \pi_1(y - 1/\pi_1) > \pi_{h1}(e^{inv})(y - 1/\pi_1) = u(v^{inv})$. Thus the issuer receives a higher payoff when it orders ratings than when investors do. So v^{inv} must be strictly lower than v^{iss}.

Intuitively, since the issuer prefers a less informative rating, the CRA cannot charge as much when the issuer decides whether to ask for a rating as when investors do. Put differently, when the firm orders ratings, its outside option is to not order a rating and receive financing without one. When the investors order a rating, the only thing that the firm can do is to refuse financing, which leaves it with a zero payoff. Competition between the investors keeps them from raising interest rates all the way to where the firm has zero profits, but they do extract more from the firm than when it is in charge of ordering ratings. Since the total surplus is split between the firm and CRA, a lower payoff to the firm means a higher payoff to the CRA, which it turn means higher rating precision under the investor-pays than under the issuer-pays model.

To consolidate and reiterate the results for the different cases, Figure 3 shows a numerical example. The parameters are listed in the note to the figure, and they are not calibrated in any way other than to permit comparisons across different cases.

The left panel depicts the CRA effort in the different cases. For $\gamma \le \underline{\gamma}^{SB} = \underline{\gamma}^{iss} = \underline{\gamma}^{inv}$, regardless of the identity of X, the project is not financed because the expected profitability is simply too low. For $\gamma \in (\underline{\gamma}^{SB}, \gamma_0]$, the project is promising enough to incur the cost of a rating to assess its profitability. In this region, the planner, the firm, and the investors all ask for a rating and finance only if it is high.

To the right of γ_0, the project becomes ex-ante profitable. For $\gamma \in (\gamma_0, \bar{\gamma}^{iss})$, in all the three models effort is positive, and the effort level for the investor-pays model lies in between those for the issuer-pays model and the second best. In this region, the total surplus in the issuer-pays model is lower than that in the investor-pays model, as the right panel shows. This difference arises because the CRA can squeeze more surplus out of the firm when the investors order a rating than when the firm does (because the firm's willingness to pay is lower). Notice that as the prior rises, the firm's outside option improves and thus the compensation that it is prepared to offer the CRA declines, so that effort in the issuer-pays model falls with γ to the right of γ_0.

In this region, the effort is lower when the investors pay than in the second best because

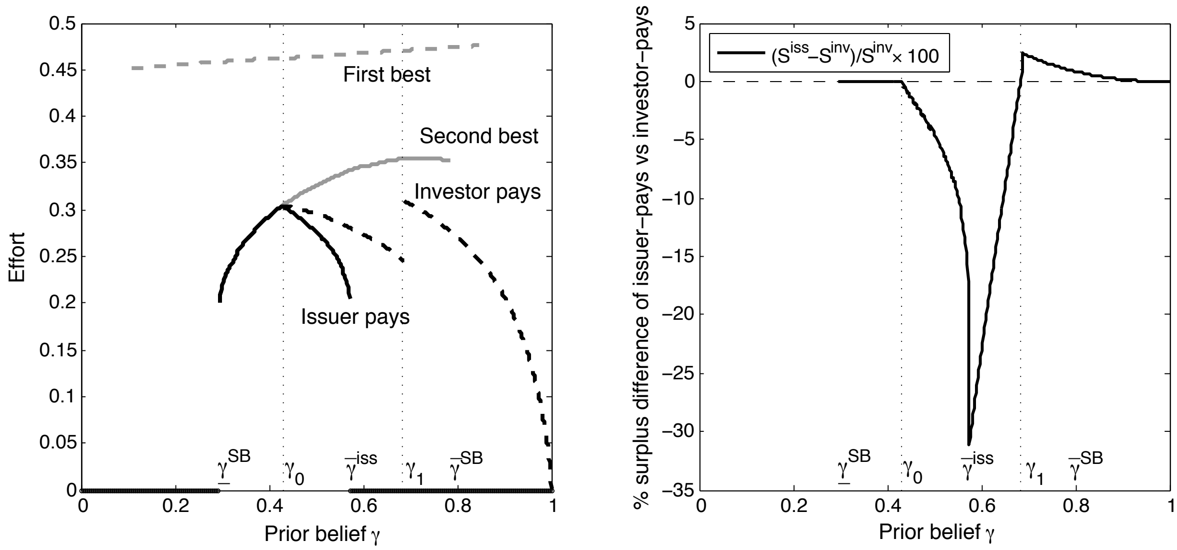

Figure 3
Comparison of different models: A numerical example
Effort (left) and the percentage surplus difference between issuer-pays and investor-pays, $\frac{S^{iss}-S^{inv}}{S^{inv}} \times 100$, (right) are displayed as functions of the prior belief γ. The parameters and the cost function used in the example are: $\alpha = 1/2$, $\beta_h = \beta_\ell = 1$, $y = 2$, $p_g = .9$, $p_b = .2$, $\psi(e) = 3e^5$.

of the competitive effects described earlier. The planner is giving all the surplus to the CRA and getting the most accurate ratings possible. The option of offering funding without a rating caps the level of CRA compensation that the investors are willing to pay. Hence, when the investors pay, ratings are less accurate than in the second-best case.

At $\bar{\gamma}^{iss}$, e^{iss} drops to zero, so that for $\gamma \in [\bar{\gamma}^{iss}, \bar{\gamma}^{SB})$ the firm is no longer willing to pay for a rating (because the option of getting funded without a rating becomes more attractive relative to paying for a noisy rating), while the planner and investors still do. Inside this region the investors are still asking for a rating, but as the pool of projects improves, the total surplus from financing without a rating keeps rising. At some point, γ_1, the investor-pays model ceases to generate higher total surplus than that alternative. Since by the time we reach γ_1 the issuer had already stopped ordering ratings, to the right of γ_1 the total surplus from the issuer-pays model exceeds that of when the investors pay.

In addition, there is a discontinuous jump in the investor-pays effort level just to the right of γ_1. The jump occurs because for a high enough γ the project's NPV is positive even after the low rating, and hence the investors finance it regardless of the rating. As a result, the project's success/failure is now observed after both high and low ratings. These extra potential outcomes mean that the CRA's compensation can be conditioned on the outcome

$\ell 0$ instead of just ℓ. The richer contracting opportunities make incentive provision more effective (in other words, less costly), which leads to a higher implemented effort relative to when the project is not funded after the low rating.

Finally, for $\gamma \in (\bar{\gamma}^{SB}, 1)$, the planner no longer asks for a rating, but the investors still do. The effort level they implement, e^{inv}, goes to zero as γ goes to one, and thus the total surplus $S^{inv} = -\psi(e^{inv}) - 1 + \pi_1 y$ converges to $S^{iss} = -1 + \pi_1 y$ (see the right panel).

Overall, the various cases can be understood as arising from the interplay of four forces. First, the planner prefers not to compensate the CRA when the pool of projects is relatively homogeneous, so that the chance of a mistaken financing decision is small. Second, regardless of who is deciding whether to order a rating, when the projects are heterogeneous enough, it pays to get a signal from the CRA and fund only when the signal is favorable. Third, when the issuer is paying for ratings, its willingness to compensate the CRA is influenced by its option to receive financing without a rating. Finally, when the investors pay for ratings, they pass along the costs of the ratings to the firm through higher interest rates and tend to use them excessively. We believe that these insights are broader than the particular model that we have analyzed and would be present in other setups where compensation for the CRA is optimally designed to provide incentives for effort.

4 Extensions

We now consider three variants of the baseline model that allow us to study the effects of competition, differences in the complexity of securities on ratings accuracy, and the speed of downgrades. The first two have been analyzed by others in the literature, so that our contribution is more to bring a slightly different perspective on these issues. The extension to address downgrades is novel; prior analyses do not connect downgrade decisions to the optimal form of CRA compensation.

4.1 Multiple CRAs

Suppose there are multiple CRAs. If multiple ratings are acquired in equilibrium, the problem becomes quite complicated. In particular, contracts will depend on the CRAs' relative performance (i.e., a CRA's compensation would in part depend on other CRAs' ratings).[29] In fact, it may be advantageous to order an extra rating only to fine-tune

[29]An example of a paper that considers relative performance incentives is Che and Yoo (2001).

the contract, while planning to ignore that rating for the purpose of the financing decision. Also, because different CRAs rely on models and data that have common features, it would seem doubtful that the signals from the various CRAs would be conditionally independent. This adds further modeling complications, but also implies that the benefits of having more information will be smaller if the signals are more correlated. Finally, if ratings are acquired sequentially and are published only at the end, in the issuer-pays model the firm's decision whether to acquire the second rating will depend on its first rating. Since this rating is the firm's private information, it introduces an adverse selection problem. For all these reasons, the analysis of this problem is complicated enough that we leave it for future research.

Instead, as a first step, we restrict our attention to the case when, even though there are multiple CRAs, only one rating is acquired in equilibrium. (Of course, this may or may not happen in equilibrium, so we simply operate under the assumption that it does.)

We modify the timing of our original model as follows. The game starts with the CRAs simultaneously posting fees. The issuer then chooses which CRA to ask for a rating. Under these assumptions the problem becomes very simple to analyze. The CRAs compete in fees, which leads to maximizing the issuer's profits. Recall from Proposition 2 that the firm's profits are maximized at v_0.[30] Hence, the total surplus in this case, denoted by S^{iss}_{many}, equals $\max\{0, -1+\pi_1 y, v_0 + u(v_0)\}$. Let e^{iss}_{many} denote the corresponding level of effort. Since $v_0 < v^{iss}$, it immediately follows from part (iii) of Proposition 2 that $e^{iss}_{many} \leq e^{iss}$ and $S^{iss}_{many} \leq S^{iss}$, with strict inequalities if $e^{iss} > 0$.

This extension suggests that a battle for market share and desire to win business will lead to lower fees, which means less accurate ratings and lower total surplus. However, the firm's surplus is higher despite the lower overall surplus. Also note that despite Bertrand competition, the CRAs still make positive profits, because $v_0 > 0$.

Interestingly, if instead the planner were ordering the rating, he would want the most precise one possible. This will prevent the CRAs from attempting to undercut each others' fees, because doing so will not gain them any business. Therefore, the optimal level of effort in this case will be the same as with one CRA. Hence the problem of increased rating errors associated with competition is specific to the issuer-pays model.[31]

[30] The value v_0 now corresponds to the payoff of the CRA whose rating has been ordered in equilibrium.

[31] We do not explore the effects of competition in the investor-pays model because it is impossible to do so without checking investors' deviations that involve the acquisition of multiple ratings.

4.2 New securities

Suppose some types of investment projects are inherently more difficult for the CRA to evaluate—presumably because they have a short track record that makes comparisons difficult. We proceed by assuming that the cost of effort is given by $\psi(e) = A\varphi(e)$, with $A > 0$; new, or more complex securities, are those with higher values of A. A higher value of A means that it is more costly for a CRA to obtain a rating of the same quality for a new security, or, alternatively, the same level of effort will lead to a less accurate rating.

Suppose that A increases to A'. We consider two scenarios. First, suppose the increase in A is unanticipated, and hence the fee structure remains unchanged. Claim 2 in the online Appendix shows that in this case constraint (4) with A' instead of A will be violated (recall from Proposition 1 that it was binding with A). Thus, when the CRA realizes that the cost of evaluating the security is higher than expected, its optimal response is to exert zero effort and always report either h or ℓ, depending on the prior. If the prior is high enough (above $\hat{\gamma}$), the CRA always reports the high rating.

This possibility could be important if the optimal reputation system that our static fee structure is meant to capture evolves slowly when new securities appear. In other words, suppose it takes some time for the market to gauge the difficulty in judging new, complex securities (e.g., the correlation risk of CDOs) and to settle on the appropriate compensation structure. Then our model predicts that the CRA might shirk and lie when rating the new complex products (such as CDOs), but report accurately on standard, established securities (such as corporate bonds), for which the reputation system is tailored.

Next, consider the second scenario where the shift in A is anticipated, and thus rating fees change appropriately. Claim 3 in the online Appendix shows that it is optimal to implement lower effort with A' than with A, which results in more rating inaccuracies. Intuitively, since the marginal cost of information acquisition is higher, it is optimal to implement a lower level of effort.[32] Thus, our model predicts that under both scenarios the quality of ratings deteriorates for new securities.

[32] The result that information acquisition is decreasing in the cost parameter is also obtained in Opp, Opp, and Harris (2013). However, in their case this result is obvious since the CRA can commit to any level of effort, and will choose less effort if its marginal cost is higher. Our result is less straightforward since fees are optimally chosen, but nonetheless the new optimal fee structure results in lower effort.

4.3 Delays in downgrading

Finally, suppose that there are two periods. The firm is endowed with the same project in each period (i.e., the quality is the same in both periods), which requires investment in both periods. The CRA exerts effort in each period to rate the project. In the optimal contract, all payments to the CRA will be made at the end of the second period, conditional on the outcome. Denote these payments by $f_{i,j}$, where $i, j \in \{h1, h0, \ell\}$, whenever positive effort is exerted in both periods.

To illustrate the idea, assume for simplicity that the CRA cannot misreport ratings (a very similar type of argument can be used when misreporting is allowed). Suppose the CRA announced a high rating in period 1, which was followed by the project's failure. We call this outcome a "mistake," because the project's performance did not match the rating. By the same argument as in the original model, to provide incentives for effort in period 2 at this point, the CRA should be paid either $f_{h0,h1} > 0$ or $f_{h0,\ell} > 0$ (only one of the two values needs to be positive since the CRA cannot misreport the rating). We are interested in the case where the mistake leads the market to expect the CRA to downgrade the security, that is, announce the low rating in period 2. This occurs when after $h0$ the optimal incentives require rewarding the CRA for announcing the low rating (i.e., to pay $f_{h0,\ell} > 0$).

Now consider how offering this payment affects incentives for effort in period 1. As we show in the online Appendix, paying $f_{h0,\ell}$ is never the best way to provide incentives for effort in period 1. In fact, unless the prior is very high, paying $f_{h0,\ell}$ actually reduces effort in the first period. So, from the point of view of incentive provision for the initial rating, the contract should never reward the CRA for changing its rating after a mistake.

This means that there is a trade-off between providing incentives for effort in period 1 (the initial rating) and effort in period 2 after a mistake. The optimal contract is designed to balance this trade-off. The desire to support effort in period 1 makes the fee structure after a mistake ex-post suboptimal. The fee $f_{h0,\ell}$ is reduced relative to what is optimal ex post (after the initial effort has been exerted), or could even be set to zero and replaced with $f_{h0,h1}$ if paying such a fee increases effort after $h0$. That is, either the reward for downgrading the security after a mistake is too low, or the CRA is being paid for sticking with the high rating after the project has failed instead of being rewarded for downgrading.

As fees after a mistake are ex-post suboptimal, the effort level in period 2 after a mistake is too low ex post. This means that if the agents were to renegotiate fees after the CRA has initially evaluated the security, they would set them to implement a higher level of effort after a mistake.[33] (Of course, ex ante it is optimal to commit not to renegotiate fees.) As a result of the low effort ex post, the probability of not downgrading conditional on the project's quality being bad is too high ex post. Hence, the CRA will appear too slow to acknowledge mistakes. Remarkably, this inertia seems to be a very general property of an optimal compensation scheme. We want to stress that such delays in downgrading are not inefficient—quite the opposite, they arise as part of an optimal arrangement.

Note that the result relies on the fact that the same CRA rates the project in both periods. This suggests a remedy to this problem. If the initial rating and the rating revision were assigned to different CRAs, then they could each be paid differently depending on the outcome, and thus the conflict between the incentive provision in the two periods would be eliminated.

5 Conclusions

We develop a parsimonious optimal contracting model that addresses multiple issues regarding ratings performance. Unobservability of the CRA's effort leads to its underprovision. Giving all surplus to the CRA maximizes rating accuracy and total surplus.

Regarding the question of pros and cons of the issuer- and investor-pays model, we find that in the issuer-pays model the rating is less accurate than in the second-best case. The reason is that the option to finance without a rating puts a bound on the firm's willingness to pay for one. The investor-pays model generates a more precise rating than the issuer-pays model, although still not as precise as what the planner could attain. However, investors tend to ask for a rating even when it is socially inefficient—in particular, when the prior about the project's quality is sufficiently high. In addition, the investor-pays model suffers from a potential free-riding problem, which can collapse security rating all together.

We show under certain conditions that battling for market share by competing CRAs leads to less accurate ratings, which yields higher profits to the firm. We also find that rating errors tend to be larger for new securities. Finally, we demonstrate that optimal provision of incentives for initial rating and revision naturally generates delays in downgrading.

Although we view the mileage that is possible with our very parsimonious framework as impressive, there are many ways in which the model can be extended. One would be to allow the firm to have superior information about its project relative to other agents.

[33]See the online Appendix for the formal analysis.

While a general analysis of moral hazard combined with adverse selection is typically quite complicated, there are a few things we can say in some special cases.

First, suppose that the firm knows the quality of its project perfectly. Then if a separating equilibrium exists, the bad type must receive no financing, since investors know that the bad project has a negative NPV. If the firm has no initial wealth, as in our original model, there is no way to separate the two types of firms in equilibrium. The reason is that the only (net) payment that the firm can possibly make occurs when the project succeeds, and either both types will want to make such a payment, or neither one will. Thus only a pooling equilibrium exists, and the analysis is essentially the same as in our original model. By continuity, the same will be true if the initial wealth is positive but sufficiently small.

If the firm has sufficient internal funds (but not enough to fund the project), then even in the absence of a rating agency, investors can separate firms with different information about their projects. They could do so by requiring the issuer to make an upfront payment in addition to a payment in the event of success (or, equivalently, requiring the issuer to invest its own funds into the project).

A more interesting but also a more complicated problem is when the firm has some private information about the project's quality, but does not know it perfectly. In this case, even in the absence of internal funds it might be possible to use the CRA to separate different types of firms by inducing them to choose different compensation schemes for the CRA and thus produce ratings of different accuracy. In particular, suppose there are two types of firms, one being more optimistic about its project than the other, and there are no internal funds. Then one can show that in a separating equilibrium where both types get rated, the firm that has a lower prior about its quality must receive a more precise rating.

Notice that different rating precision means that the same signal for different types will lead to different posterior beliefs about the project's quality. The different posteriors can be interpreted as reflecting different ratings. That is, with two signals there can be effectively four different ratings in equilibrium, associated with four different posteriors.

We leave a more complete treatment of this problem for future work.

A Appendix

Proof of Lemma 1. See the online Appendix. $\square$

Before we proceed to the proof of Proposition 1, we establish some intermediate results. First, we analyze the optimal compensation structure in the problem without misreporting, that is, without imposing constraint (4). Let λ and μ denote the Lagrange multipliers on constraints (2) and (3), respectively. The first-order condition with respect to f_i, $i \in \{h1, h0, \ell\}$ is

$$(-1+\lambda)\pi_i(e) + \mu\pi_i'(e) \leq 0, \qquad f_i \geq 0, \tag{A.1}$$

with complementary slackness. Summing over i and using $\sum_i \pi_i = 1$ and $\sum_i \pi_i' = 0$, obtain $-1+\lambda \leq 0$. Consider outcomes $i \in \{h1, h0, \ell\}$ for which $f_i > 0$. From equation (3), for at least one such i we must have $\pi_i' > 0$. Then from the first-order condition (A.1) for this f_i holding at equality it follows that $\mu \geq 0$. Suppose $\mu > 0$. Dividing equation (A.1) by $\pi_i(e)$ and comparing across i, one can see that the first-order condition that holds with equality (resulting in the strictly positive corresponding fee) is the one that corresponds to the highest likelihood ratio, $\pi_i'(e)/\pi_i(e)$. (The other fees will be zero so long as $\mu > 0$.) It is easy to show that $\pi_{h1}'(e)/\pi_{h1}(e) > \pi_{h0}'(e)/\pi_{h0}(e)$ for all e and all γ. As a result, $f_{h0} \geq 0$, with strict equality unless $\mu = 0$. Moreover, tedious algebra reveals that for all e, $\pi_{h1}'(e)/\pi_{h1}(e) \gtreqless \pi_\ell'(e)/\pi_\ell(e)$ if and only if $\gamma \gtreqless \hat{\gamma}$, where $\hat{\gamma}$ is the positive solution of a quadratic equation, given by

$$\hat{\gamma} = \frac{-\left[x\left(1+\dfrac{\beta_h p_g}{\beta_\ell p_b}\right)-1\right] + \sqrt{\left[x\left(1+\dfrac{\beta_h p_g}{\beta_\ell p_b}\right)-1\right]^2 + 4x}}{2},$$

where $x = \left[\alpha\left(1+\beta_h/\beta_\ell\right)\left(p_g/p_b - 1\right)\right]^{-1}$. Thus in the absence of misreporting, assuming $\mu > 0$ we have $f_{h1} > 0 = f_\ell = f_{h0}$ if $\gamma > \hat{\gamma}$ and $f_\ell > 0 = f_{h1} = f_{h0}$ otherwise. (If $\mu = 0$ so that $e = e^*$, then $f_{h1} > f_\ell = f_{h0} \geq 0$ if $\gamma > \hat{\gamma}$ and $f_\ell > f_{h1} = f_{h0} \geq 0$ otherwise.)

Next, we return to the problem with misreporting and show that constraint (6) binds when $\gamma > \hat{\gamma}$ and constraint (7) binds when $\gamma < \hat{\gamma}$. Define $\tilde{f}_{h1} \equiv f_{h1} - f_{h0}$, $\tilde{f}_\ell \equiv f_\ell - f_{h0}$, and $f \equiv f_{h0}$.[34] Suppose that $\gamma > \hat{\gamma}$ (the other case is analogous), and constraint (6) does not bind. Consider solving for the optimal compensation also ignoring (7). Then as shown above, $\tilde{f}_{h1} > 0 = \tilde{f}_\ell$ is optimal. Using this, constraint (7) becomes $-\psi(e) + \pi_{h1}(e)\tilde{f}_{h1} \geq$

[34] One can interpret f as a flat fee, and $\tilde{f}_i$'s as additional payments after outcomes $i \in \{h1, \ell\}$.

0. Using equation (3), the left-hand side is $-\psi(e)+\psi'(e)\pi_{h1}(e)/\pi'_{h1}(e)$, which is strictly increasing in e and equals 0 at $e=0$. Hence constraint (7) holds automatically. However, constraint (6) can be written as $-\psi(e)-\pi_{\ell 1}(e)\tilde{f}_{h1}\geq 0$. But the left-hand side is strictly negative, a contradiction. At $\gamma=\hat{\gamma}$, in the problem without misreporting, incentives are provided equally well with f_{h1} and f_ℓ, and thus constraints (6) and (7) can be satisfied without cost. Without loss of generality, we can assume that at $\gamma=\hat{\gamma}$, constraint (6) holds with equality.

Next, we establish some properties of the cost of implementing a particular effort level e. Consider the payoff to the CRA net of f, namely, $-\psi(e)+\pi_{h1}(e)\tilde{f}_{h1}+\pi_\ell(e)\tilde{f}_\ell$. By the previous analysis of whether constraint (6) or (7) binds, this payoff equals $\pi_1\tilde{f}_{h1}$ if $\gamma\geq\hat{\gamma}$ and $\tilde{f}_\ell$ if $\gamma<\hat{\gamma}$. In the case of $\gamma\geq\hat{\gamma}$, constraint (6) holding with equality implies $\pi_\ell\tilde{f}_\ell=\psi+\pi_{\ell 1}\tilde{f}_{h1}$. Substituting this into the incentive constraint $\psi'=\pi'_{h1}\tilde{f}_{h1}+\pi'_\ell\tilde{f}_\ell$, obtain $\psi'=\pi'_{h1}\tilde{f}_{h1}+(\psi+\pi_{\ell 1}\tilde{f}_{h1})\pi'_\ell/\pi_\ell$ or $\psi'-\psi\pi'_\ell/\pi_\ell=[\pi'_{h1}+\pi_{\ell 1}\pi'_\ell/\pi_\ell]\tilde{f}_{h1}=[-\pi'_{\ell 1}+\pi_{\ell 1}\pi'_\ell/\pi_\ell]\tilde{f}_{h1}$. We can then express $\tilde{f}_{h1}$ and substitute it into the payoff to the CRA (net of f) to express the latter as a function of effort only. Similarly, for $\gamma<\hat{\gamma}$, constraint (7) holding with equality implies $\pi_{h1}\tilde{f}_{h1}=\psi+\pi_h\tilde{f}_\ell$. Substituting into the incentive constraint, obtain $\psi'-\psi\pi'_{h1}/\pi_{h1}=\tilde{f}_\ell[\pi'_\ell+\pi_h\pi'_{h1}/\pi_{h1}]=\tilde{f}_\ell[-\pi'_h+\pi_h\pi'_{h1}/\pi_{h1}]$. This leads to the following expression for the payoff to the CRA (net of f) as a function of effort only, which we denote by $V(e)$:

$$V(e)\equiv\begin{cases}\pi_1\left[\dfrac{\psi'(e)-\psi(e)\pi'_\ell(e)/\pi_\ell(e)}{\pi_{\ell 1}(e)\pi'_\ell(e)/\pi_\ell(e)-\pi'_{\ell 1}(e)}\right], & \text{if } \gamma\geq\hat{\gamma},\\ \dfrac{\psi'(e)-\psi(e)\pi'_{h1}(e)/\pi_{h1}(e)}{\pi_h(e)\pi'_{h1}(e)/\pi_{h1}(e)-\pi'_h(e)}, & \text{if } \gamma<\hat{\gamma}.\end{cases} \tag{A.2}$$

Also denote

$$C(e)\equiv\psi(e)+V(e), \tag{A.3}$$

the expected cost (net of f) of implementing effort e. The first term is the direct cost to the CRA of exerting e, while the second term reflects the agency costs of incentive provision.

We use the following result in several places in the proofs that follow.

Lemma 3. The functions C and V have the following properties: $V'(e)>0$ and $C'(e)>0$ for all $e>0$, and $V(0)=V'(0)=C(0)=C'(0)=0$.

Proof. See the online Appendix. $\square$

We are now ready to prove Proposition 1.

Proof of Proposition 1. Let ξ_h and ξ_ℓ denote the Lagrange multipliers on constraints (6) and (7), respectively. The first-order conditions with respect to f_i and e are

$$-1+\lambda+\xi_h+\xi_\ell-\xi_h\frac{\pi_1}{\pi_{h1}(e)}+\mu\frac{\pi'_{h1}(e)}{\pi_{h1}(e)}\leq 0, \qquad f_{h1}\geq 0, \tag{A.4}$$

$$-1+\lambda+\xi_h+\xi_\ell-\xi_h\frac{\pi_0}{\pi_{h0}(e)}+\mu\frac{\pi'_{h0}(e)}{\pi_{h0}(e)}\leq 0, \qquad f_{h0}\geq 0, \tag{A.5}$$

$$-1+\lambda+\xi_h+\xi_\ell-\xi_\ell\frac{1}{\pi_\ell(e)}+\mu\frac{\pi'_\ell(e)}{\pi_\ell(e)}\leq 0, \qquad f_\ell\geq 0, \tag{A.6}$$

$$-\psi'(e)-\pi'_h(e)+\pi'_{h1}(e)y-\mu\psi''(e)\leq 0, \qquad e\geq 0, \tag{A.7}$$

all with complementary slackness, where we substituted equation (3) to obtain condition (A.7). It follows from condition (A.7) that $\mu\geq 0$ so long as $e\leq e^*$ (where e^* solves $-\psi'(e)-\pi'_h(e)+\pi'_{h1}(e)y=0$), and $\mu>0$ if $e<e^*$. Also, $e>e^*$ is never optimal. Suppose, to the contrary, that it is optimal to implement $e>e^*$ for some v . Then $V(e)=v-f(\geq 0)$ for some $f\geq 0$. Since V is continuous and $V'>0$, there is some $v'<v-f$ for which $V(e^*)=v'$. But then at v it is possible to implement e^* with the same fee structure as at v' plus a flat fee equal to $v-f-v'$. Since the total surplus with e^* is higher than that with $e>e^*$ (and the CRA's profits are the same, and equal to v'), the payoff to the firm must also be higher, which is a contradiction. Thus $\mu\geq 0$.

As we showed earlier, $\pi'_{h0}(e)/\pi_{h0}(e)<\pi'_{h1}(e)/\pi_{h1}(e)$ for all e and γ. Moreover, straightforward algebra reveals that $\pi_0/\pi_{h0}(e)>\pi_1/\pi_{h1}(e)$ for all e and γ. Then using $\mu\geq 0$ and $\xi_h\geq 0$, the left-hand side of the first-order condition (A.5) is always smaller than the left-hand side of the first-order condition (A.4) (strictly so unless $\mu=\xi_h=0$ in which case $e=e^*$). Thus $f\equiv f_{h0}\geq 0$, with equality unless $e=e^*$.

To show that both $\tilde{f}_{h1}\equiv f_{h1}-f_{h0}$ and $\tilde{f}_\ell\equiv f_\ell-f_{h0}$ must be strictly positive, suppose, for example, that $\tilde{f}_\ell=0$. Then from constraint (6), $-\psi(e)-\pi_{\ell 1}(e)\tilde{f}_{h1}\geq 0$. But the left-hand side is strictly negative since $e>0$, a contradiction. A similar argument supposing $\tilde{f}_{h1}=0$ and using constraint (7) also generates a contradiction. Finally, which of constraints (6) or (7) binds depending on γ was already shown above. □

Proof of Proposition 2. (i) Let $\tilde{u}$ denote the value function in the problem without misreporting, that is, where constraint (4) is omitted. Define $f_i^*=\psi'(e^*)/\pi'_i(e^*)$, where $i=h1$ if $\gamma\geq\hat{\gamma}$ and $i=\ell$ otherwise—the fee that implements e^*—and let $\tilde{v}^*=-\psi(e^*)+\pi_i(e^*)f_i^*$. Thus by construction e^* can be implemented at $v=\tilde{v}^*$. For $v>\tilde{v}^*$, it can be implemented by paying the fee f_i^* plus a flat fee equal to $v-\tilde{v}^*$. Next, in the problem with constraint (4), define $v^* = V(e^*)$. By construction, e^* can be implemented at $v = v^*$. For $v > v^*$, e^* can be implemented by paying the same outcome-contingent fees as at v^* plus an flat fee equal to $v - v^*$. Since $u(v) \leq \tilde{u}(v)$ for all v and $u(v) = \tilde{u}(v) = S^{FB} - v$ for $v \geq \max\{v^*, \tilde{v}^*\}$, it follows that $v^* \geq \tilde{v}^*$. Thus $u(v^*) \leq \tilde{u}(v^*) \leq \tilde{u}(\tilde{v}^*)$. Therefore, to prove the result, it is enough to show that $\tilde{u}(\tilde{v}^*) < 0$ if $\gamma \geq \hat{\gamma}$ and $\tilde{u}(\tilde{v}^*) < -1 + \pi_1 y$ otherwise.

The first-order condition for effort in the first-best case is $-\psi'(e) - \pi_h'(e) + \pi_{h1}'(e)y = 0$. Hence, for the first-best effort to be implemented in the problem without misreporting, it must be the case that $\pi_i' f_i^* = -\pi_h' + \pi_{h1}' y$. Substituting this into the firm's payoff, obtain $\tilde{u}(\tilde{v}^*) = -\pi_h + \pi_{h1} y - \pi_i f_i^* = -\pi_h + \pi_{h1} y - (-\pi_h' + \pi_{h1}' y)\pi_i/\pi_i'$. If $\gamma \geq \hat{\gamma}$ so that $i = h1$, then the right-hand side becomes $-\pi_h + \pi_h' \pi_{h1}/\pi_{h1}' < 0$, where the inequality can be verified with straightforward algebra. And if $\gamma < \hat{\gamma}$, then it equals $-1 + y(\pi_{h1} - \pi_{h1}' \pi_\ell/\pi_\ell') = -1 + \pi_1 y + y(\pi_{\ell 1}' \pi_\ell/\pi_\ell' - \pi_{\ell 1}) < -1 + \pi_1 y$ as $\pi_{\ell 1}'/\pi_{\ell 1} < \pi_\ell'/\pi_\ell$.

(ii) Consider maximizing the firm's payoff while omitting constraint (2). Reformulating the problem in terms of $\tilde{f}_{h1} \equiv f_{h1} - f_{h0}$, $\tilde{f}_\ell \equiv f_\ell - f_{h0}$, and $f \equiv f_{h0}$, reveals that f drops out of all the constraints except (2) and (5). Since we are omitting constraint (2), it is easy to see that $f \equiv f_{h0} = 0$ is optimal. Then the firm's payoff is $-\pi_h(e) + \pi_{h1}(e)y - \pi_{h1}(e)f_{h1} - \pi_\ell(e)f_\ell = -\pi_h(e) + \pi_{h1}(e)y - C(e)$. The first-order condition with respect to effort is $0 = [-\pi_h'(e) + \pi_{h1}'(e)y] - C'(e)$. The term in the square brackets is strictly positive, while $C'(e)$ equals zero at $e = 0$ by Lemma 3. Thus $e = 0$ cannot maximize the firm's profits, and the effort level e_0 that solves the above equation is strictly positive. To see that constraint (2) does not bind for v low enough, consider the payoff to the CRA, $V(e) = -\psi(e) + C(e)$. By Lemma 3, $V(0) = 0$ and $V'(e) > 0$ for $e > 0$. Evaluating $V(e)$ at $e_0 > 0$, $v_0 \equiv V(e_0) > 0$. Thus, for $v \leq v_0$ constraint (2) does not bind, and the optimal level of effort implemented on this interval equals $e_0 = e(v_0) > 0$.

(iii) For $v \leq v_0$ effort is constant at $e(v_0)$, and for $v \geq v^*$ it is constant at e^*. Let $v \in (v_0, v^*)$. Since $V(e)$ is strictly increasing in e by Lemma 3, the implemented effort is strictly increasing in v on this interval. Since the total surplus $-\psi(e) - \pi_h(e) + \pi_{h1}(e)y$ is strictly increasing in e for $e < e^*$, the total surplus is also strictly increasing in v. □

Proof of Lemma 2. Suppose first that $-1 + \pi_1 y > 0$. We want to show that financing the project without a rating cannot happen in equilibrium. In particular, we will demonstrate that it cannot happen that no investor orders a rating when fees are sufficiently low, and thus the CRA can sell a rating to investors by posting fees low enough.

Suppose that investors do not order a rating regardless of the fees. In such an equilibrium, the CRA and investors earn zero profits, while the firm captures all the surplus, $-1+\pi_1 y$. Investors always finance the project, and charge $\hat{R} = 1/\pi_1$ that solves $-1+\pi_1\hat{R} = 0$. Suppose the CRA were to offer a flat fee f plus outcome-contingent fees f_{h1} and f_ℓ, which implement the effort level e that solves $\psi'(e) = \pi'_{h1}(e)f_{h1} + \pi'_\ell(e)f_\ell$, and ensure truthful disclosure of the signal. Let $C(e)$ be the corresponding expected cost of implementing e (net of f) as given by equation (A.3). Consider a deviation by one investor who orders a rating, only invests if it is high, and offers the same interest rate as uninformed investors. Assume for simplicity that if the firm is indifferent between investors' offers, it obtains an equal amount of funds from each investor. (If each investor can fund the project alone, this is also equivalent to the firm randomizing with equal probabilities over which investor to borrow from.)[35] Net of the flat fee, the profits to the CRA and the investor are $V(e) = -\psi(e) + C(e)$ given by equation (A.2) and $\Pi(e) = [-\pi_h(e) + \pi_{h1}(e)/\pi_1]/n - C(e)$, respectively. Using Lemma 3, $V(0) = \Pi(0) = V'(0) = 0$. Also, $\Pi'(e) = [-\pi'_h(e) + \pi'_{h1}(e)/\pi_1]/n - C'(e)$. The second term is zero at $e = 0$, while straightforward algebra shows that the first term is strictly positive. Thus $\Pi'(0) > 0$, as the marginal cost of implementing an arbitrarily small level of effort is zero, while the marginal benefit is positive. Therefore the deviating investor can generate strictly positive profits by requesting a rating, and will agree to any strictly positive flat fee f that is strictly lower than these profits. This in turn means that the CRA can sell a rating by setting fees low enough.

We have shown that all investors not asking for a rating and always financing the project cannot be part of equilibrium. Thus, at least some investors must be ordering a rating, and it must be informative, that is, $e^{inv} > 0$. As we will demonstrate in the proof of Proposition 5, it in fact must be the case that in equilibrium all investors ask for a rating. $\square$

Proof of Proposition 5. (i) Suppose that $-1+\pi_1 y \leq 0$. We want to show that in this case $S^{inv} = S^{SB}$. Suppose first that if X is the planner, then asking for a rating and financing only after the high rating results in a negative total surplus. In this case, it is optimal not to ask for a rating and never finance, so that $S^{SB} = 0$. If X is each investor, then by definition $S^{inv} \leq S^{SB}$. Ordering a rating cannot be part of an equilibrium strategy, since it would result in a negative payoff to at least one player. Hence in this case investors do not order a rating and never finance, so that $S^{inv} = S^{SB}$. Now suppose that $S^{SB} = \bar{v} + u(\bar{v})$. In the second-best case the CRA captures all the surplus, and the firm and investors earn zero. Clearly, this is also an equilibrium when investors order a rating, and the one that

[35] This assumption is not crucial here, and with minor modifications the proof goes through by having the deviating investor offer $\hat{R} - \varepsilon$, where $\varepsilon > 0$ is arbitrarily small.

maximizes the total surplus. Thus in this case $S^{inv} = S^{SB}$.

(ii) Suppose that $-1+\pi_1 y > 0$. We first prove that in equilibrium investors who ask for a rating earn zero profits if the rating is low (not taking into account possible fee payments), and charge $\hat{R} = 1/\pi_1$ (and earn positive profits) conditional on the high rating.

To show that investors who ask for a rating must earn zero profits after the low rating, suppose not. If the profit after the low rating is negative, then there is a profitable deviation of offering no financing conditional on the low rating. And if the profit after the low rating is positive and investors who ask for a rating charge R_ℓ after the low rating, then there is a profitable deviation to not ask for a rating and always offer R_ℓ.

Suppose that investors who ask for a rating charge $R_h > \hat{R}$ conditional on the high rating. Then there is a profitable deviation by one investor, namely, do not order a rating and offer $R' \in (\hat{R}, R_h)$. The firm prefers R' to R_h, and this investor makes positive profits: $-1 + \pi_1 R_h > -1 + \pi_1 \hat{R} = 0$. In the investor-pays model, interest rates are used to finance rating fees. So in the equilibrium where the CRA charges the highest fees, the interest rate conditional on the high rating must also be the highest possible, that is, exactly equal to $1/\pi_1$.

Next, we will show that in equilibrium all investors must ask for a rating. To the contrary, suppose that there is an equilibrium where $k < n$ investors ask for a rating and $n-k$ investors do not and always finance. Uninformed investors must earn zero profit, and hence must charge $\hat{R} = 1/\pi_1$. Informed investors also charge $\hat{R}$ (conditional on the high rating). Again, assume for simplicity that if the firm is indifferent between investors' offers, it obtains an equal amount of funds from each investor.[36] Hence the firm borrows equally from all investors (informed and uninformed) when the rating is high, and borrows equally from all uninformed investors when the rating is low. The expected profit of an uninformed investor is $[-\pi_h + \pi_{h1}/\pi_1]/n + [-\pi_\ell + \pi_{\ell 1}/\pi_1]/(n-k) < [-\pi_h + \pi_{h1}/\pi_1]/n + [-\pi_\ell + \pi_{\ell 1}/\pi_1]/n = 0$, where the inequality follows from the fact that $-\pi_\ell(e) + \pi_{\ell 1}(e)/\pi_1 < 0 < -\pi_h(e) + \pi_{h1}(e)/\pi_1$ for any $e > 0$. This is a contradiction.

Let $\bar{f}_i$ denote the fee conditional on outcome i charged by the CRA, so that the total fee collected from n investors after outcome i is $f_i = n\bar{f}_i$. Each investor earns $[-\pi_h(e) + \pi_{h1}(e)/\pi_1 - \sum_i \pi_i(e) f_i]/n$. Define $I = \{h1, h0, \ell 1, \ell 0\}$. The CRA's problem can be written as $\max_{e \geq 0, \{f_i \geq 0\}_{i \in I}} -\psi(e) + \sum_{i \in I} \pi_i(e) f_i$ subject to $-\pi_h(e) + \pi_{h1}(e)/\pi_1 - \sum_{i \in I} \pi_i(e) f_i = 0$, $\psi'(e) = \sum_{i \in I} \pi_i'(e) f_i$, $-\psi(e) + \sum_{i \in I} \pi_i(e) f_i \geq \max\{\pi_1 f_{h1} + \pi_0 f_{h0}, \pi_1 f_{\ell 1} + \pi_0 f_{\ell 0}\}$, and $f_{\ell 0} = f_{\ell 1}$

[36] Without this assumption, the proof applies with the modification that informed investors must charge $1/\pi_1 - \varepsilon$, where $\varepsilon > 0$ is arbitrarily small.

if $-\pi_\ell(e) + \pi_{\ell 1}(e)y < 0$. Taking the first-order condition with respect to f_i and following the same arguments as in the proof of Proposition 1, one can show that the optimal fee structure (assuming $e < e^*$) satisfies $f_{h1} > 0$, $f_{\ell 0} > 0$ and $f_{h0} = f_{\ell 1} = 0$ if financing takes place after both ratings (and it is as described in Proposition 1 if financing takes place only after the high rating).

To find the equilibrium level of effort in the investor-pays model, we need to find e that solves the zero-profit condition for the investors, $-\pi_h(e) + \pi_{h1}(e)/\pi_1 = \hat{C}(e)$. Here $\hat{C}(e)$ equals $C(e)$ given by equation (A.3) if the project is financed only after the high rating, and if the project is financed after both ratings, then the expression for the cost is a straightforward modification of equation (A.3):

$$\hat{C}(e) \equiv \psi(e) + \begin{cases} \pi_1 \left[\dfrac{\psi'(e) - \psi(e)\pi'_{\ell 0}(e)/\pi_{\ell 0}(e)}{\pi_{\ell 1}(e)\pi'_{\ell 0}(e)/\pi_{\ell 0}(e) - \pi'_{\ell 1}(e)} \right], & \text{if } \gamma \geq \hat{\gamma}, \\ \pi_0 \left[\dfrac{\psi'(e) - \psi(e)\pi'_{h1}(e)/\pi_{h1}(e)}{\pi_{h0}(e)\pi'_{h1}(e)/\pi_{h1}(e) - \pi'_{h0}(e)} \right], & \text{if } \gamma < \hat{\gamma}. \end{cases}$$

The function $\hat{C}$ has the same properties as the function C as stated in Lemma 3, that is, $\hat{C}(0) = \hat{C}'(0) = 0$, and $\hat{C}'(e) > 0$ for all $e > 0$ (the proof is analogous).

The investors will finance the project whenever the project's value is positive given the rating. At $\gamma = \gamma_0 \equiv (1/y - p_b)/(p_g - p_b)$ such that $-1 + \pi_1 y = 0$, for any $e > 0$ we have that $-1 + [\pi_{\ell 1}(e)/\pi_\ell(e)]y < 0$, that is, the project's NPV after the low rating is negative. By continuity, since $e^{inv} > 0$, this is also true for $\gamma > \gamma_0$ close enough to γ_0, and thus investors only finance after the high rating. In this case the total surplus is given by $u(v^{inv}) + v^{inv}$, where v^{inv} is such that the payoff to the firm is $u(v^{inv}) = \pi_{h1}(e^{inv})(y - 1/\pi_1)$, as it receives financing after the high rating only and pays the interest rate of $1/\pi_1$. Notice that $u(\bar{v}) = 0 < \pi_{h1}(e^{inv})(y - 1/\pi_1) < \pi_1(y - 1/\pi_1) = -1 + \pi_1 y = u(v^{iss})$, where the first inequality follows from $-1 + \pi_1 y > 0$, and the second inequality follows from $\pi_{h1}(e) < \pi_1$ for any $e > 0$. Thus $u(\bar{v}) < u(v^{inv}) < u(v^{iss})$, which, as long as $v^{inv} < v^*$, by part (iii) of Proposition 2 implies $v^{iss} < v^{inv} < \bar{v}$ and $e^{iss} < e^{inv} < e^{SB}$. For high enough γ, the project's value after the low rating is positive, and hence the investors provide financing after both ratings. The total surplus in this case is $-\psi(e^{inv}) - 1 + \pi_1 y < -1 + \pi_1 y$.

Notice that whenever $-1 + \pi_1 y > 0$, $S^{iss} \geq -1 + \pi_1 y$. Thus whenever $S^{inv} < -1 + \pi_1 y$, we have $S^{inv} < S^{iss}$. But when the investors acquire a rating and finance only after the high rating (in which case $S^{inv} > -1 + \pi_1 y$), the investor-pays model delivers effort closer to the second-best level than the issuer-pays model, and thus $S^{inv} > S^{iss}$. $\square$

References

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

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

Blackwell, D., and M. Girshick. 1954. *Theory of games and statistical decisions.* New York: Dover.

Bolton, P., X. Freixas, and J. Shapiro. 2012. The credit ratings game. *Journal of Finance* 67:85–112.

Bongaerts, D. 2014. Alternatives for issuer-paid credit rating agencies. European Central Bank Working Paper 1703.

Bongaerts, D., M. Cremers, and W. Goetzmann. 2012. Tiebreaker: Certification and multiple credit ratings. *Journal of Finance* 67:113–152.

Chade, H., and N. Kovrijnykh. Forthcoming. Delegated information acquisition with moral hazard. *Journal of Economic Theory.*

Che, Y.-K., and S.-W. Yoo. 2001. Optimal incentives for teams. *American Economic Review* 91:525–541.

Cole, H., and T. Cooley. 2014. Rating agencies. NBER Working Paper No. 19972.

Cornaggia, J., and K. Cornaggia. 2013. Estimating the costs of issuer-paid credit ratings. *Review of Financial Studies* 26:2229–2269.

Coval, J., J. Jacob, and E. Stafford. 2009. The economics of structured finance. *Journal of Economic Perspectives* 23:3–25.

Covitz, D., and P. Harrison. 2003. Testing conflicts of interest at bond rating agencies with market anticipation: Evidence that reputation incentives dominate. Finance and Economics Discussion Series 2003-68. Board of Governors of the Federal Reserve System (U.S.).

Faure-Grimaud, A., E. Peyrache, and L. Quesada. 2009. The ownership of ratings. *RAND Journal of Economics* 40:234–257.

Financial Crisis Inquiry Commission. 2011. *Financial Crisis Inquiry report: Final report of the National Commission on the causes of the financial and economic crisis in the United States*. Washington, DC: Government Printing Office.

Fulghieri, P., G. Strobl, and H. Xia. 2014. The economics of solicited and unsolicited credit ratings. *Review of Financial Studies* 27:484–518.

Goel, A., and A. Thakor. 2011. Credit ratings and litigation risk. Mimeo.

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

Griffin, J., and D. Tang. 2012. Did subjectivity play a role in CDO credit ratings? *Journal of Finance* 67:1293–1328.

Gromb, D., and D. Martimort. 2007. Collusion and the organization of delegated expertise. *Journal of Economic Theory* 137:271–299.

Jewitt, I., O. Kadan, and J. Swinkels. 2008. Moral hazard with bounded payments. *Journal of Economic Theory* 143:59–82.

Jiang, J. X., M. H. Stanford, and Y. Xie. 2012. Does it matter who pays for bond ratings? Historical evidence. *Journal of Financial Economics* 105:607–621.

Mason, J., and J. Rosner. 2007. Where did the risk go? How misapplied bond ratings cause mortgage backed securities and collateralized debt obligation market disruptions. Mimeo.

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.

Morgan, D. P. 2002. Rating banks: Risk and uncertainty in an opaque industry. *American Economic Review* 92:874–888.

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

Pagano, M., and P. Volpin. 2010. Credit ratings failures and policy options. In *Macroeconomic Stability and Financial Regulation: Key Issues for the G20* edited by Mathias Dewatripont, Xavier Freixas, and Richard Portes, 129–148.

Sangiorgi, F., and C. Spatt. 2015. Opacity, credit rating shopping and bias. Mimeo.

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

Strausz, R. 2005. Honest certification and the threat of capture. *International Journal of Industrial Organization* 23:45–62.

United States Department of Justice. 2013. Department of Justice sues Standard & Poor's for fraud in rating mortgage-backed securities in the years leading op to the financial crisis. February 5, 2013. http://www.justice.gov/opa/pr/2013/February/13-ag-156.html.

United States Senate Permanent Subcommittee on Investigations. 2011. Wall Street and the financial crisis: Anatomy of a financial collapse. Majority and Minority Staff Report.

White, L. 2010. Markets: The credit rating agencies. *Journal of Economic Perspectives* 24:211–226.

B Online Appendix

B.1 Omitted proofs

Proof of Lemma 1. The total surplus in the first-best case is $S^{FB} = \max\{0, -1 + \pi_1 y, \max_e -\psi(e) - \pi_h(e) + \pi_{h1}(e)y\}$, where the third term can be rewritten as $\max_e -\psi(e) + (\alpha + \beta_h e)(-1 + p_g y)\gamma + (\alpha - \beta_\ell e)(-1 + p_b y)(1 - \gamma)$. At $\gamma = 0$, the first term in the expression for S^{FB} exceeds the other two terms: $0 > -1 + \pi_1 y = -1 + p_b y$ and $0 > \max_e -\psi(e) + (\alpha - \beta_\ell e)(-1 + p_b y)$. At $\gamma = 1$, the second term exceeds the other two terms: $-1 + \pi_1 y = -1 + p_g y > 0$ and $-1 + p_g y > \max_e -\psi(e) + (\alpha + \beta_h e)(-1 + p_g y)$. Hence at $\gamma = 0$ ($\gamma = 1$) it is optimal not to acquire a rating and never (always) finance the project.

Define γ_0 such that $-1 + \pi_1 y = (-1 + p_b y)\gamma + (-1 + p_b y)(1 - \gamma) = 0$ at $\gamma = \gamma_0$. We claim that at $\gamma = \gamma_0$, the third term in the expression for S^{FB} exceeds the other two terms, and hence it is optimal to acquire a rating and only finance the project after the high rating. To see this, consider the first-order condition of the maximization problem in the third term,

$$\psi'(e) = \beta_h(-1 + p_g y)\gamma - \beta_\ell(-1 + p_b y)(1 - \gamma). \tag{B.1}$$

The right-hand side of this equation is strictly positive at $\gamma = \gamma_0$. Hence (B.1) has a unique solution $e > 0$ at γ_0. Moreover, it is always possible to obtain zero surplus by choosing $e = 0$. Since the problem is strictly concave in effort, $-\psi(e) - \pi_h(e) + \pi_{h1}(e)y$ must be strictly positive at the optimal e.

Next, we show that the term $\max_e -\psi(e) - \pi_h(e) + \pi_{h1}(e)y$ is strictly increasing and convex in γ. It will then follow that it must single-cross 0 at $\underline{\gamma}^{FB} \in (0, \gamma_0)$ and $-1 + \pi_1 y$ at $\bar{\gamma}^{FB} \in (\gamma_0, 1)$, proving the interval structure stated in the lemma. Indeed, by the Envelope theorem, $\partial[-\psi(e) - \pi_h(e) + \pi_{h1}(e)y]/\partial\gamma = (\alpha + \beta_h e)(-1 + p_g y) - (\alpha - \beta_\ell e)(-1 + p_b y) > 0$. Differentiating again yields $\partial^2[-\psi(e) - \pi_h(e) + \pi_{h1}(e)y]/\partial\gamma^2 = [\beta_h(-1 + p_g y) + \beta_\ell(-1 + p_b y)]\partial e/\partial\gamma = [\beta_h(-1 + p_g y) + \beta_\ell(-1 + p_b y)]^2/\psi''(e) \geq 0$, where the last equality follows from differentiating (B.1) with respect to γ, which completes the proof. $\square$

Proof of Lemma 3. Differentiating and collecting terms, $V'(e) = \psi''(e)/D(e)$, where $D(e)$ equals $[\pi_{\ell 1}(e)\pi'_\ell(e)/\pi_\ell(e) - \pi'_{\ell 1}(e)]/\pi_1$ if $\gamma \geq \hat{\gamma}$ and $\pi_h(e)\pi'_{h1}(e)/\pi_{h1}(e) - \pi'_h(e)$ otherwise. The numerator is strictly positive by our assumption of strict convexity of ψ, and straightforward algebra shows that the denominator D is also strictly positive for any e. Thus $V'(e) > 0$ and $C'(e) = \psi'(e) + V'(e) > 0$ for any $e > 0$. Given our assumptions $\psi(0) = \psi'(0) = \psi''(0) = 0$, evaluating (A.2), (A.3), and their first derivatives at $e = 0$ implies the result. $\square$

The argument behind the proof of Proposition 4 relies on the assumption that the firm can credibly announce that it did not get rated. Claim 1 below demonstrates how the results change if we dispose of this assumption. First, when the project is financed only after the high rating, the issuer-pays model delivers the same rating accuracy as the investor-pays model. But unlike the investors, the issuer does not ask for a rating when γ is high enough, in which case the project always receives financing.

Claim 1. Suppose that the firm cannot credibly reveal to investors that it did get rated. Then the maximum total surplus in the issuer-pays case is $\max\{0, -1+\pi_1 y, v^{inv}+u(v^{inv})\}$.

Proof. When $-1+\pi_1 y \leq 0$, the analysis is the same as before. Suppose that $-1+\pi_1 y > 0$. First we show that it is an equilibrium for the firm to not ask for a rating and for the investors to always finance the project. Notice that in such an equilibrium, it must be the case that regardless of the fees charged by the CRA, investors finance without a rating and the firm does not ask for a rating. (Indeed, if this was not true, then the CRA would charge fees for which a rating is ordered, and earn profits.) Investors charge the interest rate that breaks them even, that is, $\hat{R} = 1/\pi_1$. We also need to specify the investors' off-equilibrium-path beliefs if they do see a rating (while the firm's equilibrium strategy is not to order one). Since this happens off the equilibrium path, Bayes' rule does not apply. We assume that in this equilibrium, no matter what the posted fees are, whenever investors see a rating, they assume it is uninformative, so they still finance the project with the interest rate equal to $\hat{R}$.[1] It is easy to see that it is then the best response for the firm to not ask for a rating regardless of the fees.

Suppose γ is such that the issuer asks for a rating. Given that the implemented effort is $e(>0)$, investors finance after the high rating at the gross interest rate $\pi_{h1}(e)/\pi_h(e)(< 1/\pi_1)$, and do not finance if the rating is low or if there is no rating. (If investors financed the project without a rating but not with a low rating, then the firm with the low rating would choose not to announce it. The investors' beliefs if they see no rating is that the firm received a low rating and did not reveal it.) What are the highest fees that the

[1] One might argue that when the issuer's strategy is not to ask for a rating yet a high rating is reported, investors should form their expectations about the rating precision based on the fees they observe. So if the posted fees implement strictly positive effort, investors should finance the project at a lower interest rate than if they see no rating. With this assumption on the off-equilibrium-path beliefs, the equilibrium is such that the firm asks for a rating even for high γ, and investors finance the project whenever it has a positive NPV given the rating, so that $S^{iss} = S^{inv}$ and $e^{iss} = e^{inv}$. However, such an equilibrium would not survive in a dynamic model, where investors form their expectations about rating precision based on their beliefs about the CRA's future profits given the equilibrium played.

CRA can charge? It must be the case that no investor finds it profitable to deviate by offering financing regardless of the rating. Such a deviation yields negative profits if the firm orders a rating, borrows from other investors if the rating is high, and only borrows from the deviating investor if the rating is low. Consider the profits that the firm earns after the high rating depending on whom it chooses to borrow from. If the firm borrows from investors who finance after the high rating at $\pi_{h1}(e)/\pi_h(e)$, it earns $\Pi_1 \equiv \pi_{h1}(e)[y - \pi_h(e)/\pi_{h1}(e)] - C(e)$. If the firm borrows from the deviating investor after the high rating, it earns $\Pi_2 = \pi_{h1}(e)(y - R_d)$, where R_d is the gross interest rate charged by the deviating investor. The issuer will order a rating and choose the first option after the high rating—and thus the deviating investor will earn negative profits—if the first payoff exceeds the second one.[2] This restriction imposes an upper bound $\check{e}$ on the effort level that can be implemented in equilibrium. The higher the fees and the effort, the higher the payoff to the CRA and the lower the payoff Π_1. The highest fee/effort for which Π_1 just equals Π_2 can be found by setting Π_2 as low as possible. The lowest interest rate that such a deviating investor can charge is $\hat{R} = 1/\pi_1$ (charging anything lower would earn him negative profits). Thus, the equation that $\check{e}$ solves becomes $\pi_{h1}(\check{e})[y - \pi_h(\check{e})/\pi_{h1}(\check{e})] - C(\check{e}) = \pi_{h1}(\check{e})(y - 1/\pi_1)$ or $-\pi_h(\check{e}) + \pi_{h1}(\check{e})/\pi_1 = C(\check{e})$. Notice that this condition is exactly the same as in the investor-pays case—see the proof of Proposition 5. Thus $\check{e} = e^{inv}$ and $S^{iss} = v^{inv} + u(v^{inv}) = S^{inv}$ in this case. Since financing without a rating is equilibrium when $-1 + \pi_1 y > 0$, we have $S^{iss} = \max\{0, -1 + \pi_1 y, v^{inv} + u(v^{inv})\}$. □

Claim 2. Suppose $\{f_i\}_{i \in \{h1, h0, \ell\}}$ is the optimal compensation scheme in problem (1)–(5), where $\psi(e) = A\varphi(e)$. If the CRA chooses effort facing such compensation and $A' < A$, then (4) is violated, and hence the optimal response of the CRA is to exert zero effort and always report h if $\gamma \geq \hat{\gamma}$ and ℓ otherwise.

Proof. The CRA's profits if it chooses to exert effort are $\pi(A) \equiv \max_e -A\varphi(e) + \pi_{h1}(e) f_{h1} + \pi_{h0}(e) f_{h0} + \pi_\ell(e) f_\ell$. By the Envelope theorem, $\pi'(A) = -\varphi(e) < 0$. Therefore the left-hand side of (4) with A' is strictly lower than that with A. Since the right-hand side of (4) does not change, and the constraint was binding with A, it now becomes violated. Which report the CRA makes then follows from Proposition 1. □

Claim 3. Suppose that $\psi(e) = A\varphi(e)$. Then the optimal level of effort in problem (1)–(5) strictly decreases with A.

[2] An implicit assumption here is that the firm cannot commit not to borrow at a lower interest rate if one is available, and thus cannot commit to borrow from an uninformed investor in all states.

Proof. We use strict monotone comparative statics results from Edlin and Shannon (1998) to show that e is strictly decreasing in A. Let $a = 1/A$, and define $V_\varphi(e)$ as $V(e)$ given in (A.2) where ψ is replaced by φ. Then the maximization problem (1)−(5) can be written as $\max_e -\pi_h(e) + \pi_{h1}(e)y - [V_\varphi(e) + \varphi(e)]/a$ subject to $V_\varphi(e) \geq va$. Denote the objective function by $F(e, a)$. Differentiating with respect to a, $F_a = [V_\varphi(e) + \varphi(e)]/a^2$. Since $V_\varphi(e)$ and $\varphi(e)$ are both strictly increasing in e, $F_{ea} > 0$. In addition, the constraint can be written as $e \in \Gamma(a)$, where Γ is nondecreasing in a in the strong set order. Therefore the optimal choice of effort is strictly increasing in a, or strictly decreasing in A. □

B.2 Delays in downgrading

Consider the extension to two periods. To simplify the analysis, assume that if the project is not financed in the first period, it is not productive in the second period. This implies that if in the first period the project is not financed after the low rating, the CRA will not rate the security again in the second period. Finally, again for simplification purposes, we will assume that the CRA cannot misreport its signals.

Let e and e_i denote effort levels in the first period and in the second period after the outcome $i \in \{h1, h0\}$, respectively. Also, introduce the following notation for the probabilities of outcomes $\{i, j\}$ occurring, where $i \in \{h1, h0\}$ and $j \in \{h1, h0, \ell\}$:

$$
\begin{aligned}
\pi_{h1,h1} &= p_g^2 \left(\alpha + \beta_h e_{h1}\right)\left(\alpha + \beta_h e\right)\gamma + p_b^2 \left(\alpha - \beta_\ell e_{h1}\right)\left(\alpha - \beta_\ell e\right)(1-\gamma),\\
\pi_{h1,h0} &= (1-p_g)p_g \left(\alpha + \beta_h e_{h1}\right)\left(\alpha + \beta_h e\right)\gamma + (1-p_b)p_b \left(\alpha - \beta_\ell e_{h1}\right)\left(\alpha - \beta_\ell e\right)(1-\gamma),\\
\pi_{h1,\ell} &= p_g \left(1 - \alpha - \beta_h e_{h1}\right)\left(\alpha + \beta_h e\right)\gamma + p_b \left(1 - \alpha + \beta_\ell e_{h1}\right)\left(\alpha - \beta_\ell e\right)(1-\gamma),\\
\pi_{h0,h1} &= p_g(1-p_g) \left(\alpha + \beta_h e_{h0}\right)\left(\alpha + \beta_h e\right)\gamma + p_b(1-p_b) \left(\alpha - \beta_\ell e_{h0}\right)\left(\alpha - \beta_\ell e\right)(1-\gamma),\\
\pi_{h0,h0} &= (1-p_g)^2 \left(\alpha + \beta_h e_{h0}\right)\left(\alpha + \beta_h e\right)\gamma + (1-p_b)^2 \left(\alpha - \beta_\ell e_{h0}\right)\left(\alpha - \beta_\ell e\right)(1-\gamma),\\
\pi_{h0,\ell} &= (1-p_g)\left(1 - \alpha - \beta_h e_{h0}\right)\left(\alpha + \beta_h e\right)\gamma + (1-p_b)\left(1 - \alpha + \beta_\ell e_{h0}\right)\left(\alpha - \beta_\ell e\right)(1-\gamma).
\end{aligned}
$$

For simplicity, we will focus on the case where positive effort is implemented in the second period after $h1$ and $h0$. Denote $I = \{h1, h0\}$, and $J = \{h1, h0, \ell\}$. The problem of finding the optimal fees given that in the first period the project is financed only after the high rating can be written as follows:

$$
\begin{aligned}
\max_{e,e_{h1},e_{h0},f_\ell,\{f_{i,j}\}_{i\in I,j\in J}} & -\pi_h(e)+\pi_{h1}(e)y-\pi_{h1,h}(e,e_{h1})+\pi_{h1,h1}(e,e_{h1})y \\
& -\pi_{h0,h}(e,e_{h0})+\pi_{h0,h1}(e,e_{h0})y-\pi_\ell(e)f_\ell-\sum_{i\in I,j\in J}\pi_{i,j}(e,e_i)f_{i,j} \\
\text{s.t. } & -\psi(e)+\sum_{i\in I}\left[-\pi_i(e)\psi(e_i)+\sum_{j\in J}\pi_{i,j}(e,e_i)f_{i,j}\right]+\pi_\ell(e)f_\ell\geq v, \\
& \psi'(e)=\sum_{i\in I,j\in J}\frac{\partial\pi_{i,j}(e,e_i)}{\partial e}f_{i,j}+\frac{\partial\pi_\ell(e)}{\partial e}f_\ell, \\
& \pi_i(e)\psi'(e_i)=\sum_{j\in J}\frac{\partial\pi_{i,j}(e,e_i)}{\partial e_i}f_{i,j}\text{ for } i\in I, \\
& e\geq 0,\ e_i\geq 0,\ f_\ell\geq 0,\ f_{i,j}\geq 0\text{ for } i\in I, j\in J.
\end{aligned}
$$

Let λ, μ, and μ_i denote the Lagrange multipliers on the first, second, and third constraints, respectively. Then the first-order condition with respect to $f_{i,j}$ is

$$
-1+\lambda+\frac{1}{\pi_{i,j}}\left[\mu\frac{\partial\pi_{i,j}(e,e_i)}{\partial e}+\mu_i\frac{\partial\pi_{i,j}(e,e_i)}{\partial e_i}\right]\leq 0,\qquad f_{i,j}\geq 0,
$$

with complementary slackness.

It is straightforward to check that

$$
\frac{\partial\pi_{i,h1}(e,e_i)/\partial e}{\pi_{i,h1}}>\frac{\partial\pi_{i,j}(e,e_i)/\partial e}{\pi_{i,j}}\text{ for } i\in\{h1,h0\},\ j\in\{h0,\ell\},
$$

that is, providing incentives for effort in period 1 by paying $f_{h1,h1}$ is more effective than by paying $f_{h1,h0}$ or $f_{h1,\ell}$, and similarly for $f_{h0,h1}$ vs. $f_{h0,h0}$ and $f_{h0,\ell}$. It is also easy to verify that

$$
\frac{\partial\pi_{h1,h1}(e,e_{h1})/\partial e}{\pi_{h1,h1}}>\frac{\partial\pi_{h0,\ell}(e,e_{h0})/\partial e}{\pi_{h0,\ell}},
$$

which means that providing incentives for effort in period 1 by paying $f_{h1,h1}$ is always more effective than by paying $f_{h0,\ell}$. In other words, if there was no need to provide incentives for effort after a mistake, the fee $f_{h0,\ell}$ would never be positive. Moreover, $\partial\pi_{h0,\ell}(e,e_{h0})/\partial e<0$ if and only if

$$
\gamma<\left[1+\frac{(1-p_g)(1-\alpha-\beta_h e_{h0})\beta_h}{(1-p_b)(1-\alpha+\beta_\ell e_{h0})\beta_\ell}\right]^{-1},
$$

the right-hand side of which is close to one if $(1-p_g)/(1-p_b)$ and/or β_h/β_ℓ are low enough. This means that unless γ is very high, paying $f_{h0,\ell}$ actually reduces effort in period 1.

For concreteness, suppose that γ is high enough so that

$$\frac{\partial\pi_{h1,h1}(e,e_{h0})/\partial e}{\pi_{h1,h1}} > \frac{\partial\pi_\ell(e)/\partial e}{\pi_\ell},$$

and thus $f_\ell = 0$. (The case when the comparison of the likelihood ratios for $h1h1$ and ℓ is reverse can be treated similarly.) Suppose further that the best way to provide incentives for effort in period 2 after $h1$ and $h0$ is to pay $f_{h1,h1}$ and $f_{h0,\ell}$, respectively. Consider the optimal contract that provides incentives for effort in both periods after all histories. Two scenarios are possible: either $f_{h1,h1} > 0$ and $f_{h0,\ell} > 0$, or $f_{h1,h1} > 0$ and $f_{h0,h1} > 0$. (The latter can only happen if $\partial\pi_{h0,h1}(e,e_{h0})/\partial e_{h0} > 0$.) Consider the first scenario first, and look at the point in time when effort in period 1 has already been exerted. How would the fees change if the contract could be optimally modified at this point given that the CRA's expected profits must be $\pi_{h1}(e)v_{h1} + \pi_{h0}(e)v_{h0} = v + \psi(e)$, where $v_i = -\psi(e_i) + \sum_{j\in J}\pi_{i,j}(e,e_i)f_{i,j}/\pi_i(e)$ for $i \in I$? Notice that both $f_{h1,h1}$ and $f_{h0,\ell}$ cannot decrease, because then v_{h1} and v_{h0} both decrease, and the CRA's profits will be less than $v + \psi(e)$. Similarly, they cannot both increase assuming that the promised value to the CRA in the first period is exactly v (which is true as long as v is above the analog of v_0 in the two-period case). Thus $f_{h1,h1}$ and $f_{h0,\ell}$ will either remain unchanged, or change in the opposite directions.

Given the previous comparisons of likelihood ratios, paying $f_{h1,h1}$ always dominates paying $f_{h0,\ell}$ from the point of view of incentive provision in period 1. Moreover, unless γ is very high, paying $f_{h0,\ell}$ reduces effort in period 1. Thus the optimal contract that takes into account incentive provision in both periods would have a lower $f_{h0,\ell}$ relative to what is ex-post optimal. In other words, once effort in period 1 is sunk, it is optimal to increase $f_{h0,\ell}$—which would lead to higher effort after $h0$—and decrease $f_{h1,h1}$.

Now consider the second scenario where $f_{h0,h1} > 0$ instead of $f_{h0,\ell} > 0$ in the optimal contract. Since by assumption $f_{h0,\ell}$ dominates $f_{h0,h1}$ for incentive provision after $h0$, in the renegotiated contract $f_{h0,h1}$ would be replaced with $f_{h0,\ell}$. But would this increase effort after $h0$? The answer is yes if v_{h0} decreases, which happens if $f_{h1,h1}$ is more effective than $f_{h0,h1}$ in providing incentives for effort in period 1. How can we guarantee that this is the case? It is straightforward to check that

$$\frac{\partial\pi_{h1,h1}(e,e_{h1})/\partial e}{\pi_{h1,h1}} > \frac{\partial\pi_{h0,h1}(e,e_{h0})/\partial e}{\pi_{h0,h1}}$$

if and only if

$$\frac{p_g(\alpha + \beta_h e_{h1})}{p_b(\alpha - \beta_\ell e_{h1})} > \frac{(1-p_g)(\alpha + \beta_h e_{h0})}{(1-p_b)(\alpha - \beta_\ell e_{h0})}. \tag{B.2}$$

That is, $f_{h1,h1}$ dominates $f_{h0,h1}$ for incentive provision in period 1 if and only if (B.2) holds. What this condition means is that observing success in the first period followed by the high rating (with effort e_{h1}) necessarily results in a higher posterior belief about the project's quality than observing failure followed by the high rating (with effort e_{h0}). Notice that if $e_{h1} \geq e_{h0}$, (B.2) holds automatically, but it might be violated if e_{h0} is sufficiently higher than e_{h1}. In order for (B.2) to hold, it must be the case that even if the rating in the second period after $h0$ is more precise than after $h1$, the market will still believe that the project is not as good after $h0h$ as it is after $h1h$. In other words, the project's success/failure is always a more informative signal than a high rating.

In order to insure that this is true, it is enough to impose an upper bound on effort, $\tilde{e}$, so that $\psi(\tilde{e})$ is large enough. To derive this upper bound, set $e_{h1} = 0$ and $e_{h0} = \tilde{e}$ in (B.2), and replace the inequality with equality:

$$\frac{p_g}{p_b} = \frac{(1-p_g)(\alpha + \beta_h \tilde{e})}{(1-p_b)(\alpha - \beta_\ell \bar{e})}.$$

Rearranging terms yields

$$\tilde{e} = \alpha \frac{p_g(1-p_b) - p_b(1-p_g)}{p_g(1-p_b)\beta_\ell + p_b(1-p_g)\beta_h}.$$

We have performed our analysis assuming that the CRA cannot misreport ratings. With misreporting, the argument is similar except, for example, after a mistake (outcome $h0$) both $f_{h0,\ell}$ and $f_{h0,h1}$ have to be positive—the fee after a history with the highest likelihood ratio is used to provide incentives, and the other fee is used to prevent misreporting. But the idea remains the same: the payment structure that is best to provide incentives for effort after a mistake in period 2 is suboptimal for incentive provision in period 1.

B.3 Dynamic model

As we discussed in the paper, the outcome-contingent compensation in our static model is necessary to provide incentives for effort. However, in a dynamic model, effort can be sustained even if the payment structure is restricted to flat upfront fees.

To demonstrate this, consider an environment with one infinitely-lived rating agency, infinitely-lived investors, and a sequence of short-run players—firms—each living for one period only, but who are informed of all previous play and correctly form expectations about all future play when choosing their actions. The project's quality is i.i.d. over time. Future profits are discounted by $\beta \equiv (1 + r^f)^{-1} \in (0, 1)$, where r^f is the risk-free interest rate. Suppose that in each period the CRA must set its current-period fee f, to be independent of the rating that is granted or the project's performance in that period. In *equilibrium* the current fee will endogenously depend on the *past* history, in particular, on the extent to which past ratings matched past projects' performance. This past history is summarized by the promised value to the CRA, which we denote by v.

Let $U^{SB}(v)$, $U^{iss}(v)$, and $U^{inv}(v)$ denote the expected present discounted value to all of the firms' profits when the value to the CRA is v, and X is the planner, the issuer, and each investor, respectively. The first two functions will be directly related to the following function. Let $\hat{U}_z(v)$ denote the highest possible expected present discounted value of firms' profits that can be achieved when the value to the CRA is v, when the planner sets the fees and chooses whether to ask for a rating, and z is a parameter that denotes the minimum payoff that each firm must receive. We can formulate the recursive problem for $\hat{U}_z(v)$ using the promised value v as a state variable, and continuation values as control variables.[3] Define $I = \{h1, h0, \ell 1, \ell 0\}$—the set of all potentially possible one-period outcomes. Then the recursive problem for $\hat{U}_z$ can be written as follows:[4,5]

$$\hat{U}_z(v) = \max_{e\geq 0,\ f\geq 0,\ \{v_i\geq 0\}_{i\in I}} -f + E[\max\{-\pi_h(e) + \pi_{h1}(e)y, 0\} \\ + \max\{-\pi_\ell(e) + \pi_{\ell 1}(e)y, 0\} + \beta\sum_{i\in I}\pi_i(e)\hat{U}_z(v_i)] \tag{B.3}$$

$$\text{s.t. } f + E[-\psi(e) + \beta\sum_{i\in I}\pi_i(e)v_i] = v, \tag{B.4}$$

$$\psi'(e) = \beta\sum_{i\in I}\pi_i'(e)v_i, \tag{B.5}$$

[3]The classic reference is Spear and Srivastava (1987).

[4]The expectation sign in the objective function and the promise-keeping constraint (B.4) is added because with deterministic contracts the problem is in general not concave, and thus the use of lotteries over continuation values and the implemented effort can improve welfare. The use of lotteries ensures that the value function is concave, which in turn guarantees that the fixed point of the Bellman operator exists, and also justifies using first-order and Envelope conditions.

[5]Note that a solution to this problem may not exist if v is large enough. For example, suppose that $z = 0$ and γ is very close to zero. Then the only equilibrium is not to order a rating and not to finance the project in every period, and the only v for which the above problem has a solution is $v = 0$, with the corresponding value of the firms' profits equal to $\hat{U}_0(0) = 0$.

$$f + E[-\psi(e) + \beta \sum_{i \in I} \pi_i(e) v_i] \geq f + \beta \max\{\pi_1 v_{h1} + \pi_0 v_{h0}, \pi_1 v_{\ell 1} + \pi_0 v_{\ell 0}\}, \tag{B.6}$$

$$-f + E[\max\{-\pi_h(e) + \pi_{h1}(e) y, 0\} + \max\{-\pi_\ell(e) + \pi_{\ell 1}(e) y, 0\}] \geq z, \tag{B.7}$$

$$\hat{U}_z(v_i) \geq \frac{z}{1-\beta} \text{ for all } i \in I, \tag{B.8}$$

$$v_{j1} = v_{j0} \text{ if } -\pi_j(e) + \pi_{j1}(e) y < 0, \ j \in \{h, \ell\}, \tag{B.9}$$

$$e \geq 0, f \geq 0, v_i \geq 0 \text{ for all } i \in I. \tag{B.10}$$

Notice that the incentive constraint (B.5) is essentially the same as the incentive constraint (3) in the static problem except βv_i's appear instead of f_i's. Indeed, the continuation values in this dynamic model play the same role as the outcome-contingent payments in our static model, and making the continuation values outcome-dependent will create incentive to exert effort (even though only upfront fees are allowed).[6] These continuation values reflect the CRA's expectation of how its future earnings will depend on whether its ratings will match the firms' performance. Embedded in those values are not only future fees that the CRA will be able to charge, but also all players' (rational) expectations about other players' strategies regarding whether to order a rating, whether to provide financing, etc. for all future periods and all possible future histories. So, unlike in our static model, the outcome-contingent compensation structure is not simply the CRA's choice, but is tied to future strategies of all market participants.

It is straightforward to establish properties of the optimal compensation scheme (in particular, the choices of continuation values) by looking at the first-order conditions and the envelope condition. As in Proposition 1, the CRA will be rewarded (with $v_i > v$)/punished (with $v_i < v$) for outcomes whose likelihood ratio is the highest/lowest (while respecting the no-misreporting constraint).

It is worth commenting on the analog of the static limited liability assumption, namely, $v_i \geq 0$. Recall that in the static model $f_{h0} = 0$, that is, limited liability always binds after a 'mistake'. It can be shown that in the dynamic model even though it is optimal to set $v_{h0} < v$ (punish for a 'mistake'), v_{h0} is strictly positive unless the current promised value v is very close to zero. In other words, limited liability seldom binds for continuation values.

Now, consider how U^{SB} and U^{iss} are related to $\hat{U}_z$. When X is the planner, the lowest payoff he can deliver to each firm in equilibrium is zero. Thus $z = 0$ in this case so that

[6] Constraint (B.9) ensures that the continuation values do not depend on success/failure event if the project is not financed.

$U^{SB} = \hat{U}_0$, and the Pareto frontier in this case is $\{(v, U^{SB}(v))|v \geq 0, U^{SB}(v) \geq 0\}$. When X is the issuer, the lowest payoff that each firm can guarantee itself by simply not asking for a rating is $\underline{u} = \max\{0, -1 + \pi_1 y\}$. Thus $z = \underline{u}$ in this case so that $U^{iss} = \hat{U}_{\underline{u}}$, and the Pareto frontier is $\{(v, U^{iss}(v))|v \geq 0, U^{iss}(v) \geq \underline{u}\}$. It is straightforward to show (B.8) binds when v_i is high enough. Using this, one can show that $U^{iss}(v) \leq U^{SB}(v)$ for all v, with strict inequality if $-1 + \pi_1 y > 0$. Notice that the result is stronger than in the static model: the second-best arrangement Pareto dominates the issuer-pays model rather than just yielding a higher total surplus.

As for the investor-pays case, just as in the static model one can show that when $-1 + \pi_1 y > 0$, (i) it is not an equilibrium for investors not to ask for a rating,[7] and (ii) when investors finance after the high rating only, the lowest equilibrium value to the firm is lower than in the issuer-pays case but higher than in the second best. However, carefully writing down the recursive problem in the investor-pays case when $-1 + \pi_1 y > 0$ is rather cumbersome. The reason is that two cases are possible: 1) investors may finance a project only after the high rating, and 2) they may finance it after both ratings. However, in the second case the problem does not have the same structure as the problem (B.3)$-$(B.10), because it is not Pareto optimal to implement positive effort and then finance the project after both ratings. Whether case 1) or case 2) occurs depends on the implemented level of effort, which in turn depends on the promised value v, so which of these cases occurs in the next period, and thus which function the firm's value is given by, depends on v_i. So the equilibrium value function $U^{inv}(v)$ has to combine both of the cases.

Thus, for simplicity we are not going to write down the full problem in the investor-pays case. However, we can still derive many of the results in this case. For instance, (i) and (ii) imply that when $-1 + \pi_1 y > 0$, $U^{inv}(v) < U^{SB}(v)$ for all v. On the other hand, the comparison between $U^{iss}(v)$ and $U^{inv}(v)$ when $-1 + \pi_1 y > 0$ is in general ambiguous. As in the static model, when $-1 + \pi_1 y \leq 0$, equilibrium welfare and effort levels are the same regardless of who orders the rating.

When $-1 + \pi_1 y > 0$, comparing the optimal effort levels analytically for different X is rather complicated. We conjecture that in this case, if for a given v positive effort is implemented in the second-best case, then $e^{iss}(v) < e^{inv}(v) < e^{SB}(v)$, so as in the static case, ratings are less precise when issuers orders them than when investors do, and both

[7] A deviation that breaks an equilibrium where no one asks for a rating involves an arrangement between the CRA and one investor, in which the investor orders ratings in the current and future periods, and the CRA charges future fees in a way that provides incentives for effort in the current and future periods.

models have more rating errors compared to what the planner could achieve. We leave verification of this conjecture for future work.

As for the model extensions that we analyzed in Section 4, results for some of those still go through, while others are more difficult to analyze in the infinite-horizon model. For example, it would be quite challenging to analyze effects of competition in the repeated framework. One of the difficulties is that even if we assume that each firm orders only one rating, in the recursive problem with, say, two CRAs, one has to keep track of promised values to both of them. Moreover, evolution of promised values must be consistent with firms' decisions regarding which CRA to order a rating from. Thus the problem structure becomes rather complicated, and also quite different from the one with a single CRA, which makes it hard to directly compare solutions to the two problems.

Regarding our analysis of new securities, the result that effort drops to zero and the CRA always reports the same rating if an increase in the cost parameter A is unanticipated (shown in Claim 2 in the static case) would go through. This result is quite general and only relies on the binding truth-telling constraint. Showing the analog of Claim 3 is much more challenging in particular because the value function that enters into the problem is endogenous and itself depends on A. However, it still seems rather intuitive that the optimal level effort should be reduced when the rating technology is less productive.

Finally, our result of delays in downgrading is also quite general because it only uses the comparison of likelihood ratios, and it would still apply in an environment where our two-period model is repeated infinitely many times.

Overall, even though less can be said analytically in the repeated infinite-horizon model, most of the main effects that we saw in the static model are still present, and many of our results still apply. At the same time, the dynamic model does not really add any new important insights, while its analysis is considerably more complicated. This is the reason why we analyze the static model in the paper.

References

Edlin, A., and C. Shannon. 1998. Strict monotonicity in comparative statics. *Journal of Economic Theory* 81:201–219.

Spear, S., and S. Srivastava. 1987. On repeated moral hazard with discounting. *Review of Economic Studies* 54:599–617.