

NBER WORKING PAPER SERIES

WHO SHOULD PAY FOR CREDIT RATINGS AND HOW?

Anil K Kashyap  
Natalia Kovrijnykh

Working Paper 18923  
<http://www.nber.org/papers/w18923>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
March 2013

We have benefited from suggestions and comments from 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, USC, 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research. For information on Kashyap's outside compensated activities see <http://faculty.chicagobooth.edu/anil.kashyap/>.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2013 by Anil K Kashyap and Natalia Kovrijnykh. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Who Should Pay for Credit Ratings and How?

Anil K Kashyap and Natalia Kovrijnykh

NBER Working Paper No. 18923

March 2013, Revised August 2015

JEL No. D82,D83,D86,G24

**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.

Anil K Kashyap

Booth School of Business

University of Chicago

5807 S. Woodlawn Avenue

Chicago, IL 60637

and NBER

[anil.kashyap@chicagobooth.edu](mailto:anil.kashyap@chicagobooth.edu)

Natalia Kovrijnykh

Department of Economics

W. P. Carey School of Business

Arizona State University

P.O. Box 873806

Tempe, AZ 85287-3806

[Natalia.Kovrijnykh@asu.edu](mailto:Natalia.Kovrijnykh@asu.edu)

# 1 Introduction

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, p. xxv) concludes that “this crisis could not have happened without the rating agencies”. Likewise, the United States Senate Permanent Subcommittee on Investigations (2011, p. 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”. But the details of the indictments differ slightly across the analyses. For instance, the Senate report points to inadequate staffing as a critical factor, the Financial Crisis Inquiry Commission highlights the business model that had firms seeking to issue securities pay for ratings as a major contributor, while the U.S. Department of Justice lawsuit identifies the desire for increased revenue and market share as a critical factor.<sup>1</sup> In this paper we explore the role that these and other factors might play in creating inaccurate ratings.

We study a one-period environment 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 rating agency 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

---

<sup>1</sup>The United States Senate Permanent Subcommittee on Investigations (2011) 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”. Financial Crisis Inquiry Commission (2011) 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”. The United States Department of Justice Complaint (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”.

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 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 which party 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 exist for more established types of securities. Finally, revisions to ratings are slow.

We begin our analysis by characterizing the optimal compensation arrangement for the CRA. The need to provide incentives for effort requires setting the fees that are contingent on outcomes — the issued rating and the project’s performance —, which can be interpreted as rewarding the CRA for establishing a reputation for accuracy.<sup>2</sup> Moreover, as is often the case in this kind of models, the problem of effort under-provision 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, 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 beliefs about the quality of the project. The ambiguity about the total surplus arises because even though investors induce the CRA to exert more effort, they will 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 four ways. The first extension explores the implications of allowing rating agencies to compete for business. An immediate implication of competition is a tendency to reduce fees in order to win business. But with lower fees comes lower effort in evaluating projects. Hence, this framework predicts that competition tends to lead to less accurate ratings.

---

<sup>2</sup>We discuss this interpretation of outcome-contingent fees in more detail in Section 4.2.

Second, we analyze the case when the CRA can misreport its information. We show that although the optimal compensation scheme is different than without the possibility of misreporting, our other main findings extend to this case.

The third extension considers the accuracy of ratings for different types of securities. We suppose that 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 demonstrate that in this case it is inevitable that the ratings will deteriorate.

Finally, we allow for a second period in the model and posit that investment is needed in each of the two periods, so that there is a role for ratings in both periods. The need to elicit effort in both periods creates a dilemma. The most powerful way to provide incentives for the accuracy of the initial rating requires paying the CRA only when it announces identical ratings in both periods and the project’s performance matches these ratings. Paying the CRA if it makes a ‘mistake’ in the initial rating (when a high rating is followed by the project’s failure) would be detrimental for the incentives in the first period’s effort. However, refusing to pay 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 fees in the second period after a ‘mistake’ being too low ex-post, which leads to the CRA being slow to acknowledge mistakes.

While we find that our simple model is very powerful in that it explains the five aforementioned observations 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 rating shopping — a common criticism of the issuer-pays model.<sup>3</sup> 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); 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 credit rating

---

<sup>3</sup>See the literature review below for discussion of papers that do generate rating shopping. Notice, however, that even without rating shopping we are able to identify some problems with the issuer-pays model.

agencies will recognize that there were several problems and that moral hazard was likely one of them.

The remainder of the paper is organized as follows. The next section documents the empirical regularities that motivate our analysis, and compares our model to others in the literature. Section 3 introduces the baseline model. Section 4 presents our main results about the CRA compensation as well as comparison between the issuer-pays and investor-pays models. Section 5 covers the four extensions just described. Section 6 concludes. The Appendix contains proofs and further discussion of some of the model's extensions.

## 2 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 ratings agencies, it is not surprising that there has been an explosion of research on credit rating agencies. White (2010) offers a concise description of the rating industry and recounts its role in the crisis. To understand our contribution, we find it helpful to separate the recent literature into three sub-areas.

### 2.1 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 takes as given. So we will point to specific contributions that document these particular facts.

First, the question of who pays for a rating does seem to matter. The rating industry is currently dominated by Moody's, S&P, and Fitch Ratings which are each compensated by issuers. So comparisons of their recent performance does not speak to this issue. But Cornaggia and Cornaggia (2012) provide some evidence on this question by comparing Moody's ratings to those of Rapid Ratings, a small rating agency which is funded by subscription fees from investors. They find that Moody's ratings are slower to reflect bad news than those of Rapid Ratings.

Jiang et al. (2012) provide complementary evidence by analyzing data from the 1970s when Moody's and S&P were using different compensation models. In particular, from 1971 until June 1974 S&P was charging investors for ratings, 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 from that point forward over the next three years their ratings essentially matched Moody's.<sup>4</sup>

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 — see Griffin and Tang (2012) for a description of this ratings process is conducted. As Pagano and Volpin (2010) note, the volume of these new securities increased tenfold between 2001 and 2010. As Mason and Rosner emphasize, the mistakes that happened for these new products were not found for corporate bonds where CRAs had much more experience. In addition, Morgan (2002) 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.

Third, some of the mistakes in the structured products seem to be due to insufficient monitoring and effort on the part of the analysts. For example, Owusu-Ansah (2012) shows that downgrades by Moody's tracked movements in aggregate Case-Shiller home price indices much more than any private information that CRAs had about specific deals. In the context of our model, this is akin to the CRAs not investigating enough about the underlying securities to make informed judgments about their risk characteristics.

Interestingly, the Dodd-Frank Act in the U.S. 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 early in the last decade 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.

---

<sup>4</sup>See also Griffin and Tang (2011) who show that the ratings department of the agencies, which are responsible for generating business, use more favorable assumptions than the compliance department (which is charged with assessing the ex-post accuracy of the ratings).

Finally, it appears that competition among rating agencies reduces the accuracy of ratings. Very direct evidence on this comes from Becker and Milbourn (2011) who study how the 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 typically industry. Becker and Milbourn 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.<sup>5</sup> 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 under-performing 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 decline.

## 2.2 Theoretical Models of the Rating Business

Our paper is also related to the many theoretical papers on rating agencies that have been proposed to explain these and other facts.<sup>6</sup> However, we believe our paper is the only one that simultaneously accounts for the five observations described above.

The paper by Bongaerts (2013) is closest to ours in that it focuses on an optimal contracting arrangement where a CRA must be compensated for exerting effort to generate a signal about the quality of project to be funded by outside investors. As in our model, a

---

<sup>5</sup>Bongaerts et al. (2012) identify another interesting competitive effect. If two of the firms disagree about whether a security qualifies as an investment grade, then the security does not qualify as 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 to serve as a tie-breaker creates an incentive for an issuer to seek an opinion from Fitch when the other two disagree. Bongaerts et al. 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.

<sup>6</sup>While not applied to rating agencies, there are a number of theoretical papers on delegated information acquisition, see, for example, Chade and Kovrijnykh (2012), Inderst and Ottaviani (2009, 2011) and Gromb and Martimort (2007).



central challenge is to set fees for the CRA to induce sufficient effort to produce informative ratings. He solves a planning problem, where in the base case the issuer pays for the ratings. He then explores the effect of changing the institutional arrangements by having investors pay for ratings, mandating that CRAs co-invest in the securities that they rate, or relying on ratings assessments that are more costly to produce than those created by CRAs.

Bongaerts' analysis differs from ours in three important ways. First, the risky investment projects produce both output and private benefits for the owner of the technology, which create incentives for owners to fund bad projects. This motive is absent in our model. Second, the CRAs effort is ex-post verifiable, while effort cannot be verified in our model. Third, he studies a dynamic problem, where the CRA is infinitely lived, but investment projects last for one period, investors have a one-period horizon, and new investors and new projects arrive each period. This creates interesting and important changes in the structure of the optimal contract. We see his results complementing ours by extending the analysis to a dynamic environment, though the assumptions that render his more complicated problem tractable also make comparisons between his results and ours very difficult.

Opp et al. (2013) explain rating inflation by building a model where ratings not only provide information to investors, but are also used for regulatory purposes. As in our model, expectations are rational and a CRA's effort affects rating precision. But unlike us, they assume that the CRA can commit to exert effort (or, equivalently, that effort 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.

Cornaggia and Cornaggia (2012) find evidence directly supporting the prediction of the Opp et al. (2013) model. Specifically, it seems that Moody's willingness to grant inflated ratings (relative to a subscription-based rating firm) is concentrated on the kinds of marginal investment grade bonds that regulated entities would be prevented from buying if tougher ratings were given by Moody's.

A recent paper by Cole and Cooley (2014) also argues that distorted ratings during the financial crisis were more likely caused by regulatory reliance on ratings rather than by the issuer-pays model. We agree that regulations can influence ratings, but we see our results complementing the analysis in Opp et al. (2013) and Cole and Cooley (2014) and providing additional insights on issues they do not explore.

Bolton et al. (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 (misreport) ratings in booms, when there are more naive investors, and/or when the risks of failure which could damage CRA reputation are lower. In their model, both the rating precision and reputation costs are exogenous. In contrast, in our model the rating precision is chosen by the CRA; also, our optimal contract with performance-contingent fees can be interpreted as the outcome of a system in which reputation is endogenous. 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 more shirking by CRAs.

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 at 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 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 set up the free-rider problem can then potentially eliminate the ratings market completely.

Sangiorgi and Spatt (2012) have a model that generates rating 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 equilibrium. While we conjecture that a similar result might hold in our model, the analysis of the case where multiple ratings are acquired in equilibrium is hard since, unlike in Sangiorgi and Spatt, the rating technology is endogenous in our setup.

Similar to us, Faure-Grimaud et al. (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 endogenously embodying reputational effects. Other papers that model reputational concerns of rating agencies include, for example, Bar-Isaac and Shapiro (2013), Fulghieri et al. (2011), and Mathis et al. (2009).

### 2.3 Policy Analysis of the Rating Business

Finally, the third body of research which relates to our paper includes the many policy-oriented papers that discuss potential reforms of the credit rating agencies. Medvedev and Fennell (2011) provide an excellent summary of these issues. Their survey is also representative of most of the papers on this topic in that it identifies the intuitive conflicts of interest that arise from the issuer-pays model, and compares them to the alternatives problems that arise under other schemes (such as the investor-pays, or having a government agency issue ratings). But all of these analyses are partly limited by the lack of microeconomic foundations underlying the payment models being contrasted. By deriving the optimal compensation schemes, we believe we help clarify these kinds of discussions.

## 3 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 quality of the project can be good or bad,  $q \in \{g, b\}$ , and is unobservable to everyone.<sup>7</sup> A project of 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

---

<sup>7</sup>We discuss what happens if the issuer has private information about its type in the conclusions.

project but not a low-quality one. The prior belief that the project is of high quality is denoted by  $\gamma$ , 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 exerts. Specifically,

$$\Pr\{\theta = h|q = g, e\} = \Pr\{\theta = \ell|q = b, e\} = \frac{1}{2} + e, \quad (1)$$

where  $e$  is restricted to be between 0 and  $1/2$ . Note that if effort is zero, the conditional distribution of the signal is the same regardless of the project's quality, and therefore 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 the agent's effort. So higher effort makes the signal more informative in Blackwell's sense.<sup>8</sup>

Exerting effort is costly for the CRA, where  $\psi(e)$  denotes the cost of effort  $e$ , in units of the consumption good. The function  $\psi$  satisfies  $\psi(0) = 0$ ,  $\psi'(e) > 0$ ,  $\psi''(e) > 0$ ,  $\psi'''(e) > 0$  for all  $e > 0$ , and  $\lim_{e \rightarrow 1/2} \psi(e) = +\infty$  (which is a sufficient but not necessary condition to guarantee that the project's quality is never learned with certainty). The assumptions on the second and third derivatives of  $\psi$  guarantee that the CRA's and planner's problems, respectively, are strictly concave in effort. 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.

To keep the analysis simple, we will assume that the CRA cannot lie about a signal realization so the rating it announces will be the same as the signal. We describe what happens if we dispose of this assumption in Section 5.2. While allowing for misreporting changes the form of the optimal compensation to the CRA, it does not affect any other key results, as we illustrate in the Appendix. 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 therefore needs investors to finance the project.<sup>9</sup> Investors are deep-pocketed so that there is never a shortage of funds.<sup>10</sup> They behave competitively and will make zero profits in equilibrium.

---

<sup>8</sup>See Blackwell and Girshick (1954), chapter 12.

<sup>9</sup>We make this assumption for expositional convenience. Our results would not change if the firm had initial wealth which is strictly smaller than one — the amount of funds needed to finance the project.

<sup>10</sup>It is not necessary for our results to assume that each investor has enough funds to finance the project alone. As long as each investor has more funds than what the firm borrows from him in equilibrium, our results still apply.

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 rating fee schedule — the fees to be paid at the end of the period, conditional on the history. When  $X$  is the firm, it might not be able to pay for a rating if the fee structure requires payments when no output is generated, as it has no internal funds. Thus we assume that in this case each investor offers rating financing terms that specify the return paid by the issuer when it has output in exchange for the investor paying the fee on the issuer’s behalf. Then  $X$  decides whether to ask for a rating, and chooses whether to reveal to the public that a rating has been ordered. If a rating is ordered, the CRA exerts effort and announces the rating to  $X$ , who then decides whether it should be published (and hence made known to other agents). Given the rating or the absence of one, each investor announces project financing terms (interest rates). The firm decides whether to borrow from investors in order to finance the project.<sup>11</sup> If the project is financed, its success or failure is observed. The firm repays investors, and the CRA collects its fees.<sup>12</sup>

We are interested in analyzing Pareto efficient perfect Bayesian equilibria in this environment. We will compare effort and total surplus depending on who orders a rating. 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. In our static environment, we will not always be able to Pareto rank equilibria depending on who orders the rating. However, it can be shown that constraints that lead to a lower total surplus in the static model, lead to Pareto dominance in a repeated infinite horizon version of the model.

## 4 Analysis and Results

Before deriving any results, it will be convenient to introduce some notation. First, let  $\pi_1$  denote the ex-ante probability of success (before observing a rating), so  $\pi_1 = p_g\gamma + p_b(1-\gamma)$ . Next, let  $\pi_h(e)$  denote the probability of observing a high rating given effort  $e$ , that is,

---

<sup>11</sup>We assume that if the firm is indifferent between investors’ financing terms, 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.

<sup>12</sup>We assume that  $X$  can commit to paying the fees due to the CRA, and that the firm can commit to paying investors.

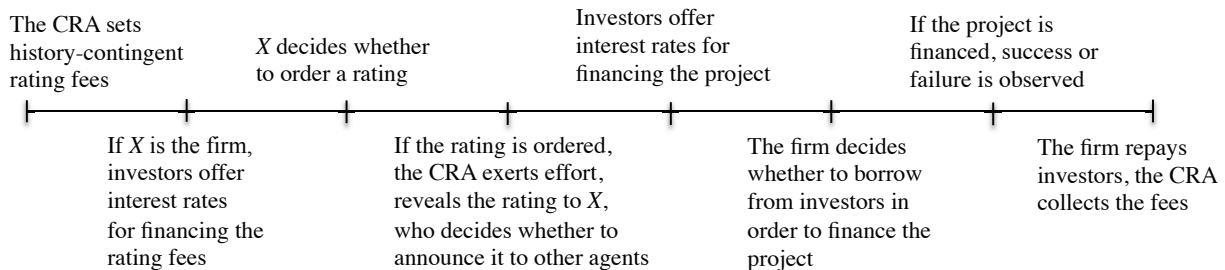


Figure 1: Timing.

$\pi_h(e) = (1/2 + e)\gamma + (1/2 - e)(1 - \gamma)$ . 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  $\pi_{h0}$  denote the probabilities of observing a high rating followed by the project's success/failure given effort  $e$ :  $\pi_{h1}(e) = p_g(1/2 + e)\gamma + p_b(1/2 - e)(1 - \gamma)$  and  $\pi_{h0}(e) = (1 - p_g)(1/2 + e)\gamma + (1 - p_b)(1/2 - e)(1 - \gamma)$ . Similarly, the probabilities of observing a low rating followed by success/failure given  $e$  are  $\pi_{\ell1}(e) = p_g(1/2 - e)\gamma + p_b(1/2 + e)(1 - \gamma)$  and  $\pi_{\ell0}(e) = (1 - p_g)(1/2 - e)\gamma + (1 - p_b)(1/2 + e)(1 - \gamma)$ .

The probability of observing a high rating bears directly on the earlier discussion of the possibility that rating agencies issued inflated ratings for securities that eventually failed. In our model, when the CRA puts insufficient effort, its ratings will be unreliable. Thus, for bad projects, the under-provision of effort will lead to a more likely (incorrect) assignment of high ratings. The assumed connection between the CRA's effort and the signal distribution given by (1) implies that the probability of giving a high rating to a bad-quality project is the same as the probability of giving a low rating to a good-quality project. Thus *unconditionally* the high rating is produced more often if less effort is put in whenever  $\gamma < 1/2$ . (Formally,  $\pi'_h(e) < 0$  for  $\gamma < 1/2$ .) The cutoff value of  $1/2$  arises because of the symmetric structure of (1). If instead we had adopted a more general signal structure such as  $\Pr\{\theta = h|q = g, e\} = \alpha + \beta_h e$  and  $\Pr\{\theta = h|q = b, e\} = \alpha - \beta_\ell e$ , then the cutoff value for  $\gamma$  that governs when erroneous ratings will be too high would differ. In particular, the lower the ratio  $\beta_h/\beta_\ell$  (i.e., the more important is the CRA's effort in detecting bad projects relative to recognizing good ones), the higher will be the cutoff.<sup>13</sup>

<sup>13</sup>Formally,  $\pi_h(e) = (\alpha + \beta_h e)\gamma + (\alpha - \beta_\ell e)(1 - \gamma)$ , which is decreasing in  $e$  if and only if  $\gamma < \beta_\ell/(\beta_\ell + \beta_h)$ .

## 4.1 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.<sup>14</sup> Given a rating (or the absence of one), the project is financed if and only if it has a positive NPV. Thus, the total surplus in the first-best case is

$$S^{FB} = \max_e -\psi(e) + \pi_h(e) \max \left\{ 0, -1 + \frac{\pi_{h1}(e)}{\pi_h(e)} y \right\} + \pi_\ell(e) \max \left\{ 0, -1 + \frac{\pi_{\ell1}(e)}{\pi_\ell(e)} y \right\},$$

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  $\pi_{h1}(e)/\pi_h(e) \geq \pi_{\ell1}(e)/\pi_\ell(e)$ , 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 from the structure of the problem that in cases (i) and (ii) the optimal choice of effort is zero. This result is very intuitive — it cannot be efficient to expend effort if the information it produces is not used. In case (iii), the optimal effort is strictly positive — denote it by  $e^*$  — and (given our assumptions)  $e^*$  uniquely solves  $\max_e -\psi(e) - \pi_h(e) + \pi_{h1}(e)y$ . Thus, the problem of finding the first-best surplus can be simplified to

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

The following lemma shows which of the three alternatives (i)–(iii) the planner chooses depending on the prior  $\gamma$  (where we denote the first-best effort by  $e^{FB}$ ).

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

- (i)  $e^{FB} = 0$  for  $\gamma \in [0, \underline{\gamma}]$ , and the project is never financed;
- (ii)  $e^{FB} = 0$  for  $\gamma \in [\bar{\gamma}, 1]$ , and the project is always financed;
- (iii)  $e^{FB} > 0$  for  $\gamma \in (\underline{\gamma}, \bar{\gamma})$ , and the project is only financed after the high rating.

The intuition behind this result is quite simple. If the prior belief about the project 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 additional information about the quality of the project.

---

<sup>14</sup>In fact, it is easy to check that when effort is observable, the total surplus is the same regardless of who orders a rating.



We now turn to the analysis of the more interesting case when the CRA's effort is unobservable, and payments are subject to limited liability. The CRA will now choose its effort privately, given the fees it expects to receive at the end of the period.

## 4.2 Second Best – the Social Planner Orders a Rating

To understand the logic of our model it is simplest to start by analyzing the case where the planner gets to decide whether to order a rating and in doing so sets the fee structure. This construct allows us to write a standard optimal contracting problem. In this setup, we will characterize the constrained Pareto frontier (optimal contract), and consider an equilibrium on the frontier where the total surplus is maximized. Then we will demonstrate that the resulting equilibrium is the same as when the CRA chooses the fees (which is the actual assumption in our analysis).

To find the optimal contract (or the optimal fee structure) that provides the CRA with incentives to exert effort, we want to allow for as rich a contract space as possible. This is accomplished by supposing that fees can be made contingent on possible outcomes. Just as in the first-best case, it is straightforward to show that 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 non-trivial problem of finding the optimal fee structure. In this case, 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  $f_\ell$  denote the fees that the CRA receives in each scenario.

On the Pareto frontier, the payoff to one party is maximized subject to delivering at least certain payoffs to other parties. Investors behave competitively and thus always earn zero profits. Therefore, we can maximize the value to the firm subject to delivering at least a certain value  $v$  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 only financed after the high rating. Since investors earn zero profits, the firm extracts all the surplus generated in production, net of the expected fees paid to the CRA. Then the Pareto frontier can be written as



$\max\{0 - v, -1 + \pi_1 y - v, u(v)\}$ , where

$$u(v) = \max_{e, f_{h1}, f_{h0}, f_\ell} -\pi_h(e) + \pi_{h1}(e)y - \pi_{h1}(e)f_{h1} - \pi_{h0}(e)f_{h0} - \pi_\ell(e)f_\ell \quad (2)$$

$$\text{s.t. } -\psi(e) + \pi_{h1}(e)f_{h1} + \pi_{h0}(e)f_{h0} + \pi_\ell(e)f_\ell \geq v, \quad (3)$$

$$\psi'(e) = \pi'_{h1}(e)f_{h1} + \pi'_{h0}(e)f_{h0} + \pi'_\ell(e)f_\ell, \quad (4)$$

$$e \geq 0, f_{h1} \geq 0, f_{h0} \geq 0, f_\ell \geq 0. \quad (5)$$

Constraint (3) ensures that the CRA's profits are at least  $v$ . Constraint (4) is the CRA's incentive constraint, which reflects the fact that CRA chooses its effort privately. Accordingly, this constraint is obtained by maximizing the left-hand side of (3) with respect to  $e$ . 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, i.e.,  $u(v) \geq 0$  (which restricts the values of  $v$  that can be promised).

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

**Proposition 1 (Optimal Compensation Scheme)** *Suppose the project is financed only after the high rating. Define the cutoff value  $\hat{\gamma} = 1/(1 + \sqrt{p_g/p_b})$ .*

(i) *If  $\gamma \geq \hat{\gamma}$ , then it is optimal to set  $f_{h1} > 0$  and  $f_\ell = f_{h0} = 0$ .*

(ii) *If  $\gamma \leq \hat{\gamma}$ , then it is optimal to set  $f_\ell > 0$  and  $f_{h1} = f_{h0} = 0$ .*

The proposition states that there is a threshold level for the prior belief, above which the CRA should be rewarded only if it announces the high rating and it is followed by success, and below which the CRA should be rewarded only if it announces the low rating. Notice that, quite intuitively, the CRA is never paid for announcing the high rating if it is followed by the project's failure. 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 agents' prior.

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

Many papers in this literature model such reputation mechanisms as exogenous; for example, Bolton et al. (2012) (see also references therein) introduce exogenous reputation

costs — the discounted sum of future CRA profits, which, in their case, are available when the CRA is not caught lying. In our paper, the payoff to the CRA that varies depending on the outcome can be interpreted in precisely this way, except that the reputation costs are endogenous because the compensation structure is endogenous.

To see that outcome-contingent payments can be interpreted as the CRA’s future profits in a more elaborate dynamic model, consider instead the following repeated setting. In each period the CRA charges an upfront (flat) fee, but the fee can vary over time depending on how the firm’s performance compared with the announced ratings. Technically, the optimal compensation structure written in a recursive form will require the CRA’s ‘continuation values’ (future present discounted profits) to depend on histories.<sup>15</sup> Thus even if the fees are restricted to be paid upfront in each period, the CRA will be motivated to exert effort by the prospect of higher future profits — via higher future fees — that follow from developing a ‘reputation’ by correctly predicting the firm’s performance. To put it differently, equilibrium strategies and expectations of market participants in a Pareto optimal equilibrium depend on histories in such a way that the CRA expects to be able to charge higher fees and earn higher profits (because market participants are willing to pay those higher fees given their beliefs about the CRA’s diligence) if the market observes outcomes that are most consistent with the CRA exerting effort.

The fee structure in our static model can then be viewed as a shortcut for such a reputation mechanism. We set up the repeated model with up-front fees in Section A.3 of the Appendix, and briefly discuss which predictions of our model still apply in that model. The dynamic model is much more complicated to analyze, and at the same time does not offer any new important insights, which is why we choose to focus on the static model instead. So outcome-contingent fees should not be interpreted literally, but instead should be recognized as a simplification to bring reputational considerations into the analysis in a tractable way. It is worth pointing out that even though we model reputation in such a ‘reduced form’ way, we are able to match several facts about the rating business, which suggests that the mechanisms operating in our model are certainly not unreasonable.

With this interpretation in mind, let us return to the analysis of the static model. The next proposition derives several properties of the Pareto frontier which will be important for our subsequent analysis.

---

<sup>15</sup>Claim 4 in the Appendix, which is the analog of Proposition 1 in the dynamic case, shows how the continuation values optimally vary depending on histories.

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

- (i) *There exists  $v^*$  such that for all  $v \geq v^*$   $e(v) = e^*$ , but  $u(v) < 0$ .*
- (ii) *There exists  $v_0 > 0$  such that (3) is slack for  $v < v_0$  and binds 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) of the proposition says that there exists a threshold promised value,  $v^*$ , above which the first-best effort is implemented. However, the resulting profit to the firm is strictly negative, violating individual rationality, and so this arrangement cannot be sustained in equilibrium. It will be handy to denote the highest value that can be delivered to the CRA without leaving the firm with negative profits by  $\bar{v} \equiv \max\{v|u(v) = 0\} < v^*$ .

There is an interesting economic reason why implementing the first-best effort requires the firm's profits to be negative. Suppose for concreteness  $\gamma \geq \hat{\gamma}$  (the other case is similar), so that the CRA gets paid after history  $h1$ . Then the intuition is as follows. When effort is observable, the problem can be recast as saying that the firm chooses to acquire information itself rather than delegating this task to the CRA. But when the firm is making the effort choice, it accounts for two potential effects of increasing effort. One benefit is the increased *probability* that a surplus is generated. The other is that investors will lower the interest rate to reflect a more accurate rating, leading to an increase in the *size* of the surplus. When the CRA is doing the investigation and its effort is unobservable, the CRA internalizes the fact that more effort generates a higher probability of the fee being paid. But it cannot get a higher fee based on higher effort. So the only way to induce the CRA to exert the first-best level effort is to set an extraordinarily generous fee that leaves the firm with negative profits.<sup>16</sup>

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. It immediately follows from (ii) that for  $v \leq v_0$   $u(v)$  does not depend on  $v$  and hence is constant; while if  $v > v_0$ , constraint (3) binds, which means that  $u(v)$  must be strictly decreasing in  $v$ .

---

<sup>16</sup>Formally, the firm's problem in the first-best case is  $\max_e -\psi(e) + \pi_{h1}(e)(y - R(e))$ , where the interest rate  $R(e)$  solves the investors' break even condition  $-\pi_h(e) + \pi_{h1}(e)R(e) = 0$ . This implies that  $1/R(e)$  equals the conditional probability of success given the high rating,  $\pi_{1|h}(e) = \pi_{h1}(e)/\pi_h(e)$ , which is strictly increasing in effort. The CRA's problem is  $\max_e -\psi(e) + \pi_{h1}(e)f_{h1}$ , where  $f_{h1}$  does not depend on  $e$ . Thus, in order to induce  $e^{FB}$ ,  $f_{h1}$  must exceed  $y - R(e)$ , leaving the firm with negative profits:  $\pi_{h1}(e)(y - R(e) - f_{h1}) < 0$ .

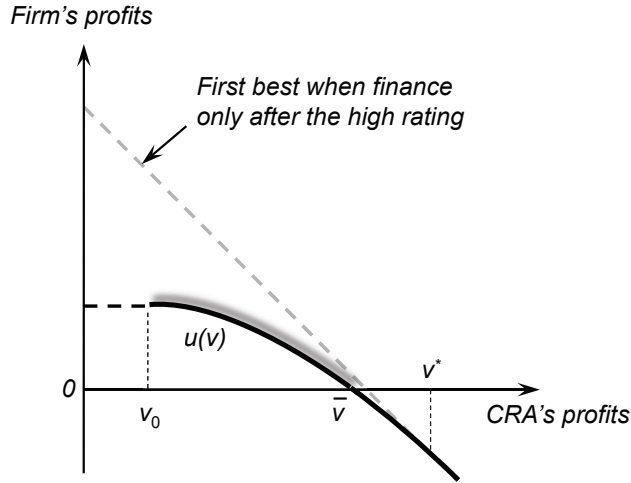


Figure 2: The Pareto frontier (the shaded area of the  $u(v)$  curve) when the project is financed only after the high rating.

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 it follows because unobservability of effort leads to its under-provision. To implement the highest possible effort, one needs to set the fees as high as possible, extracting all surplus from the firm and giving it to the CRA. However, as part (i) implies, implementing the first-best level of effort would result 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.

Notice also that the firm's profits are maximized at  $v_0$ . This follows immediately from part (ii) of Proposition 2. Thus the firm prefers a less informative rating than is socially optimal (as effort at  $v_0$  is lower than that at  $\bar{v}$  or  $v^*$ ), but the firm still prefers to have an informative rating (because effort is positive at  $v_0$ ).

The function  $u(v)$  is graphed in Figure 2. Recall that  $u(v)$  only describes the part of the Pareto frontier which corresponds to the situation when the project is financed after the high rating and not financed after the low rating. The whole Pareto frontier is given by  $\max\{-v, -1 + \pi_1 y - v, u(v)\}$ , and the corresponding total surplus is  $\max\{0, -1 + \pi_1 y, v + u(v)\}$ .

To summarize, the fact that the CRA chooses its effort privately (and is protected by limited liability) delivers two important results. First, the optimal compensation must involve outcome-contingent fees, which can be interpreted as rewards for establishing a

good reputation. Second, the CRA exerts less effort, and hence there are more rating errors compared to the case when the CRA's effort is observable. These results are general — they do not depend on who orders a rating, and they will also hold in the extensions of the basic model that we will consider in Section 5.

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.<sup>17</sup>

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. The implemented effort, which we denote by  $e^{SB}$  (where  $SB$  stands for the second best), is strictly smaller than  $e^{FB}$ .

To close the loop, let us return to the issue of what happens if instead of the planner setting the fees, the CRA does. As we showed, when the planner sets the fees (and the project is financed only after the high rating), the CRA captures all the surplus. Therefore the fees set by the CRA will choose the same ones as selected by the planner.

We summarize our results in the following proposition.

**Proposition 3 ( $X = \text{Planner}$ )** *If the social planner is the one who decides whether a rating should be ordered, then*

- (i) *The maximum total surplus in equilibrium is  $S^{SB} = \max\{0, -1 + \pi_1 y, \bar{v} + u(\bar{v})\}$ ;*
- (ii)  *$e^{SB} \leq e^{FB}$ ,  $S^{SB} \leq S^{FB}$ , with strict inequalities if  $e^{FB} > 0$ .*

Figure 3 uses a numerical example to compare the total surplus and effort in the first- and second-best cases as functions of  $\gamma$ , depicted with solid blue and dashed gray lines, respectively.<sup>18</sup> The thin dotted line in the left panel is  $-1 + \pi_1 y$ , the total surplus if the project is financed without a rating. The total surplus if the project is not financed without a rating is zero. Therefore, the total surplus in the first-best case,  $S^{FB}$ , is the

---

<sup>17</sup>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.

<sup>18</sup>Notice slight kinks in the second-best surplus and effort that occur at  $\hat{\gamma}$  — which equals .366 for the given parameter values — due to the different fee structures above and below  $\hat{\gamma}$ .

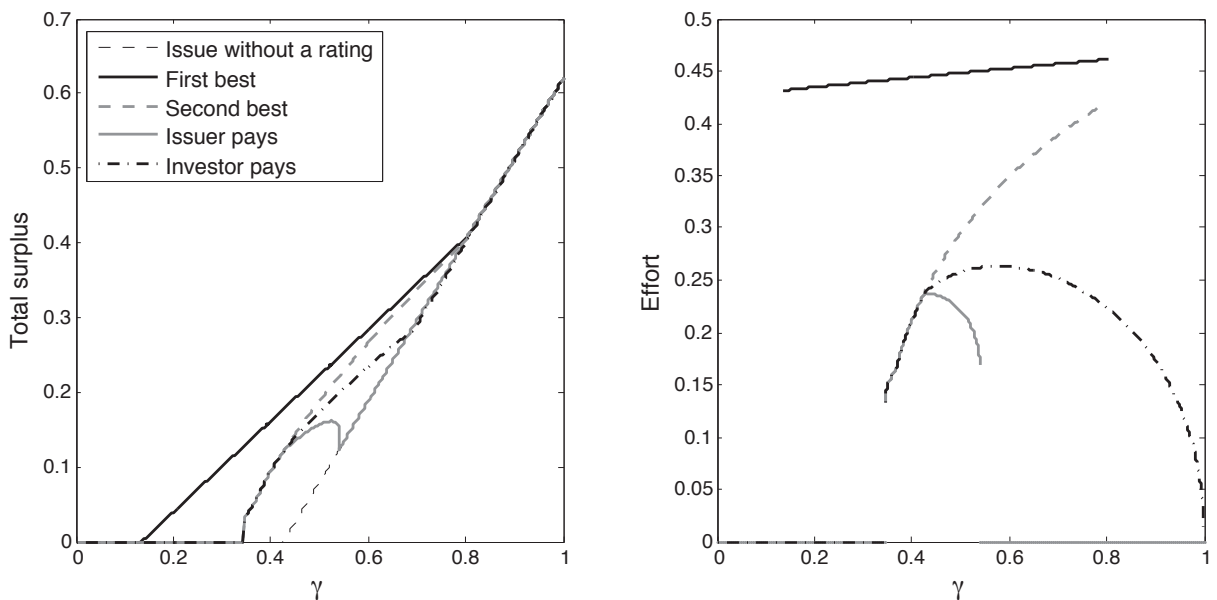


Figure 3: The total surplus (left) and effort (right) as functions of the prior belief  $\gamma$ . Parameter values:  $y = 1.8$ ,  $p_g = .9$ ,  $p_b = .3$ ,  $\psi(e) = 1.5e^4$ .

upper envelope of three lines,  $0$ ,  $-1 + \pi_1 y$ , and  $v^* + u(v^*)$ . Similarly, the total surplus in the second-best case,  $S^{SB}$ , is the upper envelope of  $0$ ,  $-1 + \pi_1 y$ , and  $\bar{v} + u(\bar{v})$ .

From Figure 3 it is apparent that the planner decides not acquire a rating for some values of  $\gamma$  when one would be acquired if effort were observable. The reduced propensity to get the rating occurs because the total surplus from acquiring the rating is lower. Thus, graphically, the interval on which the upper envelope of the three lines equals  $\bar{v} + u(\bar{v})$  is smaller than that in the first-best case.

Next, we will analyze how maximum total surplus and the corresponding effort in cases where the issuer or investors order ratings, compare to the second best case. We will ask ourselves: does it matter who orders ratings? We will show that the answer is no in “bad times”, when the average project has negative NPV (formally,  $-1 + \pi_1 y \leq 0$ ), and the answer is yes in “good times”, when the expected NPV is positive ( $-1 + \pi_1 y > 0$ ).

### 4.3 The Issuer Orders a Rating

Consider the case where the firm decides whether to order a rating(which, as we will see, will be very similar to the case when the planner chooses 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 equals 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  $\underline{u} \equiv \max\{0, -1 + \pi_1 y\}$ .<sup>19</sup> Thus, if a rating is acquired 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  $-1 + \pi_1 y > 0$  since  $u(v)$  is strictly decreasing in  $v$  for  $v > v_0$ . Denote the total surplus and effort in the issuer-pays case by  $S^{iss}$  and  $e^{iss}$ , respectively. Recall from Proposition 2 that the total surplus and effort are increasing in  $v$ . This leads us to the following result:

**Proposition 4 ( $X = \text{Issuer}$ )** *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  $-1 + \pi_1 \leq 0$ , 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 if  $e^{SB} > 0$ .*

As usual, the firm will decide not to ask for a rating if the prior belief  $\gamma$  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.

The total surplus and implemented effort in the case when the issuer orders a rating are depicted with solid gray lines on Figure 3. As described in Proposition 4, when  $-1 + \pi_1 y > 0$ , the total surplus and effort are lower than when the planner orders a rating. Notice that  $e^{iss}$  decreases with  $\gamma$  when  $-1 + \pi_1 y > 0$  because the firm's outside option is  $-1 + \pi_1 y$ , and  $\pi_1$  increases with  $\gamma$ .

To summarize, in “good times” the issuer-pays model leads to lower rating precision and total surplus than the planner would attain, because the option of receiving financing without a rating reduces the firm's willingness to pay for a rating. That is, our model predicts that the issuer-pays model is indeed associated with more rating errors than is socially optimal. As we will see in the next section, in some cases the rating errors are also larger than when the investors order ratings.

---

<sup>19</sup>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 Appendix.

## 4.4 Investors Order a Rating

Consider finally the case when each investor decides whether to order a rating. We will show that this case results in a lower total surplus relative to the planner's case because investors are competing over the interest rates that they offer to the firm conditional on the rating. As we will see, the comparison of the total surplus and effort relative to the issuer-pays case will depend on the prior  $\gamma$ .

The assumption that investors who do not pay for a rating can be excluded from learning it is critical. If this is not the case and 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 became increasingly impractical, which arguably resulted in the switch to the issuer-pays model.

The following lemma shows an important inefficiency specific to the investor-pays model. Recall that for  $\gamma$  sufficiently close to one, it is socially optimal 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.<sup>20</sup>

**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, while the CRA earns nothing. The CRA can try to sell a rating; it would not succeed if the planner controls whether it should be ordered, unless the generated surplus is at least  $-1 + \pi_1 y$  (or unless the firm's profit is at least that amount, in case the firm orders a rating). However, when investors order a rating, they are not concerned with the total or the firm's surplus. They make zero profits, and they can always pass along the costs of getting a rating to the firm, while the CRA generates profits.

But why do investors necessarily choose to order a rating if they earn zero profits either way? To show that this must be the case, suppose instead that no one asks for a rating regardless of what the fees are. Then we prove that if fees are low enough, one investor

---

<sup>20</sup>Thus in this case equilibrium payoffs actually lie inside the (constrained) Pareto frontier.



could generate profits by ordering a rating, hiding it from other investors, only investing if it is high, but charging the same or a slightly lower rate of return as other investors. Knowing this, the CRA can set fees low enough to entice someone to ask for a rating and hence break any equilibrium where no one is ordering a rating.

Building upon Lemma 2, the following proposition describes the total surplus and effort in the investor-pays case.

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

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

(ii) *Suppose that  $-1 + \pi_1 y > 0$ . If  $-\pi_\ell(e^{inv}) + \pi_{\ell 1}(e^{inv})y \leq 0$ , then  $S^{inv} = v^{inv} + u(v^{inv})$ , where  $v^{iss} < v^{inv} < \bar{v}$ . If  $-\pi_\ell(e^{inv}) + \pi_{\ell 1}(e^{inv})y > 0$ , then  $S^{inv} = -\psi(e^{inv}) - 1 + \pi_1 y < 1 + \pi_1 y$ .<sup>21</sup> In both cases,  $S^{inv} < S^{SB}$ . Furthermore,  $e^{iss} < e^{inv}$ , and  $e^{inv} < e^{SB}$  as long as  $e^{SB} > 0$ .*

When 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 decides whether to order a rating.

Important differences arise only when the project is ex-ante profitable. As we showed in Lemma 2, investors necessarily ask for a rating in this case. But, investors, may choose to fund the project even with a low rating if the precision of the rating is sufficiently low (so that the project's NPV after the low rating is positive).<sup>22</sup> So when  $\gamma$  is high enough, the outcome is worse than in the second-best or issuer-pays cases where the project is also always financed, but no effort is wasted.

Put differently, in the second-best and issuer-pays cases, not ordering a rating and always financing the project is an equilibrium, and it dominates ordering a rating and always financing the project. The problem in the investor-pays model, illustrated by Lemma 2, is that financing without a rating is *not* an equilibrium.

Another important result contained in Proposition 5 is that when the project is only financed after the high rating, the value to the CRA in the investor-pays case, denoted by

<sup>21</sup>In this case,  $S^{inv}$  is no longer equal to  $v + u(v)$  for some  $v$ , as  $u(v)$  is defined as the Pareto frontier, while it is not Pareto optimal to implement positive effort and then finance the project regardless of the rating.

<sup>22</sup>Since financing takes place after both ratings, the set of possible histories now is  $\{h1, h0, \ell1, \ell0\}$ . We show in the proof of Proposition 5 that the only positive fee in equilibrium is either  $f_{h1}$  or  $f_{\ell0}$  depending on whether  $\gamma$  is above or below a certain threshold. For simplicity, all of our subsequent results are stated and proven for the case when the set of possible outcomes after the CRA exerts positive effort is  $\{h1, h0, \ell\}$ , but appropriate modifications for the case when this set is  $\{h1, h0, \ell1, \ell0\}$  would be trivial.

$v^{inv}$ , is strictly between  $v^{iss}$  and  $\bar{v}$ . Therefore by Proposition 2, the implemented effort (and hence the rating precision) if investors ask for a rating is lower than if the planner asks for a rating, but higher than if the issuer does. The reason for  $v^{inv} < \bar{v}$  is that the option to finance without a rating caps interest rates, and therefore caps fees that investors are willing to pay to the CRA. (This interest rate cap is  $\hat{R} = 1/\pi_1$ , which solves  $-1 + \pi_1 \hat{R} = 0$ .) And  $v^{inv} > v^{iss}$  because the firm pays the same rate of return to investors as if there was no rating ( $\hat{R}$ , defined above), but receives financing less often — only when the rating is high (without a rating, it would be financed with probability one). Hence the firm’s profits are lower when the investor pays than when the issuer does,  $u(v^{inv}) < u(v^{iss})$ , which in turn implies that  $v^{inv} > v^{iss}$ .

The total surplus and effort in the case when investors order a rating are plotted with dashed-dotted blue lines on Figure 3. As one can see, the comparison between the total surplus in the issuer-pays and investor-pays cases depend on the prior belief about the project’s quality. In “bad times”, when the project is not profitable to finance ex-ante, i.e., when  $-1 + \pi_1 y \leq 0$ , the total surplus and effort in both models are equal, and coincide with what the planner achieves. However, in “good times”, when  $-1 + \pi_1 y > 0$ , the issuer-pays model leads to a lower total surplus than the investor-pays model for intermediate values of  $\gamma$ , but performs better if  $\gamma$  is sufficiently high.

To summarize, the investor-pays model yields higher rating accuracy than the issuer-pays model, but lower than under the planner. The reason is that investors do not care about the firm’s outside option, but the option to finance without a rating caps interest rates, and hence fees. On the other hand, investors ask for a rating too often, even when it is socially inefficient to do so.

## 5 Extensions

We now consider four variants of the baseline model. Our first extension explores the effect of allowing more than one rating agency. Next, we consider the implications of allowing the CRA to misreport its information. Third, we look at differences in ratings for securities which differ in their ease of monitoring. The last modification introduces a second period in the model so that the propensity to downgrade a security can be studied.

## 5.1 Multiple CRAs

If multiple ratings are acquired in equilibrium, the problem becomes quite complicated. In particular, contracts will depend on CRAs' relative performance (i.e., a CRA's compensation would in part depend on other CRAs' ratings).<sup>23</sup> In fact, it may be advantageous to order an extra rating only to fine-tune the contract structure, while planning to ignore that rating for the purpose of the financing decision. Furthermore, 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 modelling complications, but also implies the benefits of having more information will be smaller if the signals are more correlated. Finally, if ratings are acquired sequentially and are only published 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 sufficiently complicated 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 rating agencies, only one rating is acquired in equilibrium.<sup>24</sup> (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 by 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. CRAs compete in fees, which leads to maximizing the issuer's profits. Recall from Proposition 2 that the firm's profits are maximized at  $v_0$ .<sup>25</sup> Hence, the total surplus in this case, denoted by  $S_{many}^{iss}$ , equals  $\max\{0, -1 + \pi_1 y, v_0 + u(v_0)\}$ . Let  $e_{many}^{iss}$  denote the corresponding level of effort. Since  $v_0 < v^{iss}$ , it immediately follows from part (iii) of Proposition 2 that  $e_{many}^{iss} \leq e^{iss}$  and  $S_{many}^{iss} \leq S^{iss}$ , with strict inequalities if  $e^{iss} > 0$ .

We find this extension interesting because it suggests that a battle for market share and

---

<sup>23</sup>An example of a paper that considers relative performance incentives is Che and Yoo (2001).

<sup>24</sup>As we argue in the Appendix, even with this assumption analyzing competition in a dynamic model is very complicated. So while the other results we emphasize in the body of the paper more or less carry over to a dynamic setup, these would not necessarily apply.

<sup>25</sup>Of course, now there are more players in the game. If there are  $N$  CRAs and the firm randomizes between whom to ask for a rating if it is indifferent, then each CRA receives  $v_0/N$  in expectation. The frontier on Figure 2 shows the surplus division between the CRA whose rating is ordered and the firm after the outcome of the randomization is observed, with other CRAs (as well as investors) receiving zero profits.

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 the possibility of Bertrand competition, the CRAs still make positive profits, because  $v_0 > 0$ .

It is instructive to compare the outcomes of the issuer-pays model and the planner's problem with multiple CRAs. The planner will always want to order the most precise rating 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.<sup>26</sup>

## 5.2 Misreporting a Rating

We next return to our original model with one CRA. So far we assumed that the CRA cannot misreport its signal; now we relax this assumption and suppose that the CRA can lie. In addition to moral hazard, this creates an adverse selection problem. Solving for the optimal contract requires imposing additional constraints to our optimal contracting problem (2)–(5).

It is easy to see that if the CRA intends to lie, the most profitable way to do so is a double deviation: exert no effort, and always report whatever rating yields the highest expected fee. This should not be surprising because if the CRA intends to misreport, exerting effort is wasteful. Thus, the additional constraint that needs to be imposed in order to deliver a truthful report is

$$-\psi(e) + \pi_{h1}(e)f_{h1} + \pi_{h0}(e)f_{h0} + \pi_\ell(e)f_\ell \geq \max\{\pi_1f_{h1} + \pi_0f_{h0}, f_\ell\}, \quad (6)$$

which is equivalent to imposing the following two constraints:

$$-\psi(e) + \pi_{h1}(e)f_{h1} + \pi_{h0}(e)f_{h0} + \pi_\ell(e)f_\ell \geq \pi_1f_{h1} + \pi_0f_{h0}, \quad (7)$$

$$-\psi(e) + \pi_{h1}(e)f_{h1} + \pi_{h0}(e)f_{h0} + \pi_\ell(e)f_\ell \geq f_\ell. \quad (8)$$

The left-hand side of (6) shows the CRA's payoff if it exerts effort and truthfully reports

---

<sup>26</sup>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 (i.e., analyzing out of equilibrium behavior where different investors acquire ratings from different CRAs).

the acquired signal. The right-hand side is the value from exerting no effort and always reporting the rating that delivers the highest expected fee (or randomizing between the two, if the fees are the same).

The next proposition shows how the optimal compensation must be structured if the possibility of misreporting is present.

**Proposition 6 (Optimal Compensation under Misreporting)** *Suppose the project is financed only after the high rating. Then for each  $\gamma$  it must be the case that  $f_{h1} > 0$ ,  $f_\ell > 0$ , and  $f_{h0} = 0$ . Furthermore, (7) will bind for  $\gamma \geq \hat{\gamma}$  and (8) will bind for  $\gamma < \hat{\gamma}$ , so long as the implemented effort is below the first-best level  $e^*$ .*

Recall from Proposition 1 that when the CRA cannot misreport its signal, only one of the two fees,  $f_{h1}$  or  $f_\ell$ , is strictly positive. The situation is different with the possibility of misreporting: both  $f_{h1}$  and  $f_\ell$  must be strictly positive. The reason for paying in both cases is intuitive. In particular, without misreporting the CRA would only be paid for issuing a high rating followed by success if the prior about the project's quality is high enough. But if the CRA can misreport its signal, it would always issue a high rating given this compensation scheme. To prevent the CRA from lying, it must be also paid for issuing a low rating.

Since constraint (6) binds, the total surplus generated if the rating is ordered when the CRA can lie is lower than in the case when it cannot lie. Also, the range of priors for which the rating will be ordered (by any agent) is smaller than when the CRA cannot lie. This is not surprising since essentially the option to lie gives the CRA leverage that allows it to extract fees in order to tell the truth. These fees were previously unnecessary and mean that the agents now become more cautious about using the CRA.

While the optimal compensation scheme is affected by the possibility of misreporting, our other results still apply — the proofs that require modification are provided in the Appendix.

### 5.3 New Securities

We will now use our results from Section 5.2 to analyze the case where the CRA must rate 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, and there is no adequate rating model that has been developed yet.

One way to analyze this in our framework is to parametrize the cost of effort as  $\psi(e) = A\varphi(e)$ , with  $A > 0$ , and think of a new type of security as the one with a higher value of  $A$ .<sup>27</sup> 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, paying the same cost would produce a less accurate rating.

Suppose that  $A$  increases to  $A'$ . We consider two scenarios. In the first scenario, the increase in  $A$  is unanticipated. In this case, fees remain unchanged. If the CRA cannot misreport a rating, the CRA's incentive constraint immediately implies that it will exert less effort. Now consider a more interesting case when the CRA can misreport its rating. Claim 2 in the Appendix shows that in this case constraint (6) with  $A'$  instead of  $A$  will be violated (recall from Proposition 6 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  $\ell$ , depending on the prior. In particular, when the prior is above  $\hat{\gamma}$ , the CRA always reports the high rating.

Now consider the second scenario where the shift in  $A$  is anticipated, and thus rating fees change appropriately. Claim 3 in the Appendix shows that it is optimal to implement lower effort with  $A'$  than with  $A$ , which results in more rating inaccuracies. (This result holds regardless of whether the CRA can or cannot misreport its rating.) Intuitively, since the marginal cost of information acquisition is higher, it is optimal to implement a lower level of effort.<sup>28</sup>

Thus, our model predicts that under both scenarios the quality of ratings deteriorates for new securities.

## 5.4 Delays in Downgrading

Finally, we demonstrate that a straightforward extension of our model can explain delays in downgrading. Suppose that there are two periods. The project requires investment in both periods, and the project quality is the same 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

---

<sup>27</sup>We do assume that everything else, in particular, parameters  $p_b$ ,  $p_g$ ,  $y$ , and  $\gamma$  remain the same.

<sup>28</sup>The result that information acquisition is decreasing in the cost parameter is also obtained in Opp et al. (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. This result is similar to our result in the case of an unanticipated change without the possibility of misreporting. With an anticipated change, our result is less straightforward since fees are optimally chosen, but nonetheless the new optimal fee structure results in lower effort.

made at the end of the second period, conditional on the history. Denote these payments by  $f_{i,j}$ , where  $i, j \in \{h1, h0, \ell\}$ , whenever positive effort is exerted in both periods.

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 Proposition 1, to provide incentives for the second period effort at this point, the CRA should be paid either  $f_{h0,h1} > 0$  or  $f_{h0,\ell} > 0$ , depending on whether the posterior after  $h0$  is above or below  $\hat{\gamma}$ . We are interested in the scenario when after a mistake the market is sufficiently pessimistic about the security that the CRA is expected to *downgrade* it, i.e., announce the low rating in period 2. That is, after  $h0$  the best way to provide incentives for effort is to reward 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 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.<sup>29</sup> (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 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.

---

<sup>29</sup>See the Appendix for the formal analysis.



## 6 Conclusions

We develop a parsimonious optimal contracting model that addresses multiple issues regarding ratings performance. We show that when the CRA's effort is unobservable, a rating is less precise, and is acquired less often (on a smaller set of priors) than in the first-best case. 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 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.

While we view the mileage that is possible with our very parsimonious framework as impressive, there are many ways in which the modelling can be extended. Perhaps most natural would be to allow the firm to have superior information about its investment opportunities relative to other agents. While a general analysis of moral hazard combined with adverse selection is typically quite complicated, there are a few things we can see in some interesting 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 net present value. 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 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 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 fees and thus produce ratings with different degrees of precision. In particular, suppose that 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 and the associated issues for future work.

## A Appendix

**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  $\max_e -\psi(e) + (1/2 + e)(-1 + p_g y)\gamma + (1/2 - e)(-1 + p_b y)(1 - \gamma)$ . 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  $0 > \max_e -\psi(e) + (1/2 - e)(-1 + p_b y)$ . At  $\gamma = 1$ , the second term exceeds the other two terms:  $-1 + \pi_1 y = -1 + p_g y > 0$  and  $-1 + p_g y > \max_e -\psi(e) + (1/2 + e)(-1 + p_g y)$ . Hence at  $\gamma = 0$  ( $\gamma = 1$ ) it is optimal not to acquire a rating and never (always) finance the project.

Define  $\gamma^*$  such that  $-1 + \pi_1 y = (-1 + p_b y)\gamma + (-1 + p_b y)(1 - \gamma) = 0$  at  $\gamma = \gamma^*$ . We claim that at  $\gamma = \gamma^*$ , 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,

$$\psi'(e) = (-1 + p_g y)\gamma - (-1 + p_b y)(1 - \gamma). \quad (9)$$

The right-hand side of this equation is strictly positive at  $\gamma = \gamma^*$ . Hence (9) has a unique solution  $e > 0$  at  $\gamma^*$ . 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  $\gamma$ . It will then follow that it must single-cross 0 at  $\underline{\gamma} \in (0, \gamma^*)$  and  $-1 + \pi_1 y$  at  $\bar{\gamma} \in (\gamma^*, 1)$ , proving the interval structure stated in the lemma. Indeed, by the Envelope theorem,  $\partial[-\psi(e) - \pi_h(e) + \pi_{h1}(e)y]/\partial\gamma = (1/2 + e)(-1 + p_g y) - (1/2 - e)(-1 + p_b y) > 0$ . Differentiating again yields  $\partial^2[-\psi(e) - \pi_h(e) + \pi_{h1}(e)y]/\partial\gamma^2 = (-1 + p_g y - 1 + p_b y)\partial e/\partial\gamma = (-1 + p_g y - 1 + p_b y)^2/\psi''(e) \geq 0$ , where the last equality follows from differentiating (9) with respect to  $\gamma$ , which completes the proof.  $\square$

**Proof of Proposition 1.** Let  $\lambda$  and  $\mu$  denote the Lagrange multipliers on constraints (3) and (4), respectively. The first-order condition of problem (2)–(5) with respect to  $f_i$ ,  $i \in \{h1, h0, \ell\}$  is

$$(-1 + \lambda)\pi_i(e) + \mu\pi'_i(e) \leq 0, \quad f_i \geq 0,$$

with complementary slackness. Dividing by  $\pi_i(e)$ , one can see that the first-order condition that will hold with equality (resulting in the strictly positive corresponding fee) is the one that corresponds to the highest likelihood ratio,  $\pi'_i(e)/\pi_i(e)$ . Straightforward algebra shows

that

$$\begin{aligned} \frac{\pi'_{h1}}{\pi_{h1}} \geq \frac{\pi'_\ell}{\pi_\ell} &\Leftrightarrow \gamma \geq \frac{1}{1 + \sqrt{p_g/p_b}}, \\ \frac{\pi'_{h1}}{\pi_{h1}} &> \frac{\pi'_{h0}}{\pi_{h0}} \quad \text{for all } \gamma, \end{aligned}$$

which completes the proof.  $\square$

**Proof of Proposition 2.** In this proof, we only consider the case when  $\gamma \geq \hat{\gamma}$ , as the other case is analogous.

(i) Define  $f_{h1}^* = \psi'(e^*)/\pi'_{h1}(e^*)$  — the fee that implements  $e^*$  — and let  $v^* = -\psi(e^*) + \pi_{h1}(e^*)f_{h1}^*$ . Thus by construction  $e^*$  can be implemented at  $v = v^*$ . For  $v > v^*$ , it can be implemented by paying the fee  $f_{h1}^*$  plus an upfront fee equal to  $v - v^*$ .

We now show that  $u(v^*) < 0$ . Since  $u(v) = u(v^*) - (v - v^*)$  for  $v \geq v^*$ , it will follow that  $u(v) < 0$  for  $v \geq v^*$ . Using (9), for the first-best effort to be implemented it must be the case that  $\pi'_{h1}f_{h1}^* = (-1 + p_g y)\gamma - (-1 + p_b y)(1 - \gamma)$ . Substituting this into the firm's payoff, obtain

$$\begin{aligned} u(v^*) &= -\pi_h + \pi_{h1}y - \pi_{h1}f_{h1}^* = (1/2 + e)(-1 + p_g y)\gamma + (1/2 - e)(-1 + p_b y)(1 - \gamma) \\ &\quad - [(1/2 + e)p_g \gamma + (1/2 - e)p_b(1 - \gamma)] \frac{(-1 + p_g y)\gamma - (-1 + p_b y)(1 - \gamma)}{p_g \gamma - p_b(1 - \gamma)} \\ &= \gamma(1 - \gamma)[p_b - p_g]y < 0, \end{aligned}$$

where the last equality follows from straightforward algebra.

(ii) Consider maximizing the firm's payoff while omitting constraint (3). Using the incentive constraint  $\psi'(e) = \pi'_{h1}(e)f_{h1}$ , the firm's payoff can be written as  $-\pi_h(e) + \pi_{h1}(e)y - \pi_{h1}(e)f_{h1} = (-1 + p_g y)(1/2 + e)\gamma + (-1 + p_b y)(1/2 - e)(1 - \gamma) - \pi_{h1}(e)\psi'(e)/\pi'_{h1}(e)$ . The first-order condition with respect to effort is  $0 = [(-1 + p_g y)\gamma - (-1 + p_b y)(1 - \gamma)] - \psi'(e) - \psi''(e)\pi_{h1}(e)/(p_b \gamma - p_g(1 - \gamma))$ . The term in the square brackets is strictly positive, while the last two terms are zero at  $e = 0$  by our assumptions  $\psi'(0) = \psi''(0) = 0$ . Thus  $e = 0$  cannot maximize the firm's profits, and hence the optimal level of effort when (3) does not bind is strictly positive.

To see that (3) indeed does not bind for  $v$  low enough, consider the CRA's payoff  $\Pi(e) = -\psi(e) + \pi_{h1}(e)\psi'(e)/\pi'_{h1}(e)$ , where  $\Pi(0) = 0$ . Differentiating yields  $\Pi'(e) = \psi''(e)\pi_{h1}(e)/(p_b \gamma - p_g(1 - \gamma)) \geq 0$ , with strict inequality for  $e > 0$ . Therefore, at  $v = 0$  (3) cannot bind, and by continuity it will not bind for  $v$  below some threshold value, which we

denote by  $v_0$ . Thus for  $v \in [0, v_0]$  the optimal contract is the same, and the optimal level of effort is strictly positive.

(iii) For  $v \leq v_0$ , constraint (3) does not bind, and hence the total surplus and effort are constant. For  $v \geq v^*$ ,  $e = e^*$  and the total surplus equals  $S^{FB}$ . Suppose that  $v \in (v_0, v^*)$ . Then (3) holds with equality:  $-\psi(e) + \pi_{h1}(e)\psi'(e)/\pi'_{h1}(e) = v$ , where we used the incentive constraint to substitute for  $f_{h1}$ . Differentiating the left-hand side with respect to  $e$  and using the fact that  $\pi'_{h1}(e) = p_g\gamma - p_b(1 - \gamma)$  is independent of  $e$ , yields  $\pi_{h1}(e)\psi''(e)/\pi'_{h1}(e)$ , which is strictly positive for  $e > 0$  as  $\psi''(e) > 0$  and  $\pi'_{h1} > 0$  for  $\gamma \geq \hat{\gamma}$ . Thus the optimal choice of  $e$  must be strictly increasing in  $v$ . Since the total surplus  $-\psi(e) - \pi_h(e) + \pi_{h1}(e)y$  is strictly increasing in  $e$  for  $e < e^*$ , it follows that the total surplus is also strictly increasing in  $v$ .  $\square$

**Proof of Proposition 4.** Part (i) is shown in the main text. Part (ii) immediately follows from part (iii) of Proposition 2 and the fact that  $v^{iss} \leq \bar{v}$ , with strict inequality if  $-1 + \pi_1 y > 0$ . Given our assumption that the firm can credibly announce that it did not get rated, it immediately follows that if  $-1 + \pi_1 y > 0$ , then  $\underline{u} = -1 + \pi_1 y$ , and thus  $v^{iss} < \bar{v}$ , as described in the main text. Claim 1 in this Appendix shows how the results change if the firm could not credibly reveal to investors that it did not order a rating.  $\square$

**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 the gross rate of return  $\hat{R} = 1/\pi_1$  that solves  $-1 + \pi_1 \hat{R} = 0$ .

Suppose the CRA were to offer a flat fee  $f$  plus a history-contingent fee  $f_i$ , where  $\pi'_i(e)/\pi_i(e) (> 0)$  is the highest likelihood ratio for all possible equilibrium histories  $i$  (see the proof of part (ii) of Proposition 5 for details). The level of effort that these fees induce solves  $\psi'(e)/\pi'_i(e) = f_i$ . Consider a deviation by one investor who orders a rating, only invests if the rating is high, and asks the issuer for the same rate of return as everyone else. Net of the flat fee  $f$ , the profits to the CRA and the investor are  $\Pi_{CRA}(e) = -\psi(e) + \pi_i(e)\psi'(e)/\pi'_i(e)$  and  $\Pi_{inv}(e) = [-\pi_h(e) + \pi_{h1}(e)/\pi_1]/n - \pi_i(e)\psi'(e)/\pi'_i(e)$ , respectively, where

$\pi'_i(e)$  is independent of  $e$  since  $\pi(e)$  is linear in  $e$ .<sup>30</sup> Given our assumptions,  $\Pi_{CRA}(0) = \Pi_{inv}(0) = 0$ . Furthermore,  $\Pi'_{CRA}(e) = \psi''(e)\pi_i(e)/\pi'_i(e)$ , which equals zero at  $e = 0$  since  $\psi''(0) = 0$ . Also,  $\Pi'_{inv}(e) = [-\pi'_h(e) + \pi'_{h1}(e)/\pi_1]/n - [\psi'(e) + \psi''(e)\pi_i(e)/\pi'_i(e)]$ . Since  $\psi'(0) = \psi''(0) = 0$ , the second term is zero at  $e = 0$ , while straightforward algebra shows that the first term is strictly positive. Thus  $\Pi'_{inv}(0) > 0$ , 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, in equilibrium at least some investors must be ordering a rating, and the rating must be informative, i.e.,  $e^{inv} > 0$ . 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 the planner is the one who orders a rating, then asking for the 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 investors are the ones who order a rating, 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  $S^{SB} = \bar{v} + u(\bar{v})$ . All surplus in the second-best case is captured by the CRA, and the firm and investors earn zero. Clearly, this is also an equilibrium when investors order a rating, and the one that 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 after the low rating (not taking into account possible fee payments), and charge the gross rate of return equal to  $\hat{R} = 1/\pi_1$  after the high rating.

To see 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

---

<sup>30</sup>The expression for  $\Pi_{inv}(e)$  uses the assumption that if the firm is indifferent between investors' offers, it obtains an equal amount of funds from each investor (or randomizes between whom to borrow from with equal probabilities). 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.

of offering the same interest rate as after the high rating, and offering no financing after the high rating. And if the profit after the low rating is positive and the investors who ask for a rating charge  $R_\ell$  after the low rating, then there is a profitable deviation to not ask for a rating and always offer  $R_\ell$ .

Suppose that investors who ask for a rating charge  $R_h > \hat{R}$  after 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 (which yields to the the highest payoff to the CRA and the highest surplus), the interest rate after the high rating must also be the highest possible, i.e., 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. Investor who do not ask for a rating (uninformed investors) must earn zero profit, and hence must charge  $\hat{R} = 1/\pi_1$ . Investor who ask for a rating and only invest if the rating is high (informed investors) also charge  $\hat{R}$ . Recall our assumption that the firm borrows equal amounts (or one unit with equal probabilities) from investors between whose offers it is indifferent.<sup>31</sup> 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. Denote  $\pi_{\ell 1}(e) \equiv \pi_1 - \pi_{h1}(e)$ , the probability that the low rating is followed by the project's success. Then the expected profit of an uninformed investor is  $[-\pi_h + \pi_{h1}/\pi_1]/n + [-\pi_\ell + \pi_{\ell 1}/\pi_1]/(n - k) < [-\pi_h + \pi_{h1}/\pi_1]/n + [-\pi_\ell + \pi_{\ell 1}/\pi_1]/n = 0$ , as  $-\pi_\ell + \pi_{\ell 1}/\pi_1 < 0 < -\pi_h + \pi_{h1}/\pi_1$ . A contradiction.

Let  $\tilde{f}_i$  denote the fee conditional on outcome  $i$  charged by the CRA, so that the total fee collected from  $n$  investors after outcome  $i$  is  $f_i = n\tilde{f}_i$ . Each investor earns  $[-\pi_h(e) + \pi_{h1}(e)/\pi_1 - \sum_i \pi_i(e)f_i]/n$ . 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  $-\pi_h(e) + \pi_{h1}(e)/\pi_1 - \sum_{i \in I} \pi_i(e)f_i = 0$ , and  $\psi'(e) = \sum_{i \in I} \pi'_i(e)f_i$ , where  $I = \{h1, h0, \ell\}$  if  $-\pi_\ell(e) + \pi_{\ell 1}(e) \geq 0$  and  $I = \{h1, h0, \ell1, \ell0\}$  otherwise. Taking first-order conditions with respect to  $f_i$  and following the same arguments as in the proof of Proposition 1, one can show that the only positive fee is the one for which the likelihood ratio  $\pi'_i(e)/\pi_i(e)$  is the largest. When  $I = \{h1, h0, \ell\}$ , the optimal fee structure is as described in Proposition 1. If  $I = \{h1, h0, \ell1, \ell0\}$ , then straightforward algebra shows that

<sup>31</sup>Without this assumption, the proof applies with the modification that informed investors must charge  $1/\pi_1 - \varepsilon$ , where  $\varepsilon > 0$  is arbitrarily small.

$f_{h1} > 0$  for  $\gamma > 1/(1 + \sqrt{p_g(1 - p_g)/[p_b(1 - p_b)]}) (> \hat{\gamma})$  and  $f_{\ell 0} > 0$  otherwise.

Let  $i$  denote the equilibrium history for which the likelihood ratio  $\pi'_i(e)/\pi_i(e)$  is the highest, so that  $f_i$  is the only positive fee. From the CRA's incentive constraint  $f_i = \psi'(e)/\pi'_i(e)$ . Hence 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,  $-\pi_h(e) + \pi_{h1}(e)/\pi_1 = \pi_i(e)\psi'(e)/\pi'_i(e)$ . Denote the left- and right-hand sides of this equation by  $l(e)$  and  $r(e)$ , respectively. Given our assumptions,  $l(0) = r(0) = 0$ ,  $l'(e) > 0$  for all  $e$ ,  $l''(e) = 0$  for all  $e$ ,  $r'(0) = 0$ ,  $r'(e) > 0$  and  $r''(e) > 0$  for all  $e > 0$ . Thus there are two solutions to the above equation, one equal to zero and the other  $e > 0$ . Since zero effort (i.e., not asking for a rating) cannot be an equilibrium, there is a unique equilibrium effort level  $e^{inv}$ , and it is strictly positive.

Recall that the investors will finance the project whenever the project's value is positive given the rating. At  $\gamma = \gamma^* \equiv (1/y - p_b)/(p_g - p_b)$  such that  $-1 + \pi_1 y = 0$ , for any  $e > 0$  we have that  $-1 + [\pi_{\ell 1}(e)/\pi_{\ell}(e)]y < 0$ , i.e., the project's NPV after the low rating is negative. By continuity, since  $e^{inv} > 0$ , this is also true for  $\gamma > \gamma^*$  close enough to  $\gamma^*$ , 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 by part (iii) of Proposition 2 implies  $v^{iss} < v^{inv} < \bar{v}$  and  $e^{iss} < e^{inv} < e^{SB}$ . For high enough  $\gamma$ , 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  $-\psi(e^{inv}) - 1 + \pi_1 y < -1 + \pi_1 y$ . The comparison of effort levels and surpluses stated in this proposition then follows immediately.  $\square$

The argument behind the proof of Proposition 4 relied on the assumption that the firm can credibly announce that it did not get rated. The claim below demonstrates how the results change if we dispose of this assumption.

**Claim 1** *Suppose that the firm cannot credibly reveal to investors that it did not order a rating. Then the maximum total surplus and effort in the issuer-pays case are the same as in the investor-pays case,  $S^{inv}$  and  $e^{inv}$ .*

**Proof.** When  $-1 + \pi_1 y \leq 0$ , the analysis is the same as before. Suppose that  $-1 + \pi_1 y > 0$ . If the issuer cannot credibly announce that it did not order a rating, investors' contracts



cannot distinguish between events when a rating has not been ordered, and when it has been ordered, but the firm chose not to reveal it. Furthermore, 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. Therefore the argument provided in the main text for showing that  $\underline{u} = -1 + \pi_1 y$  when  $-1 + \pi_1 y > 0$  does not apply.

We first show that financing without a rating cannot happen in equilibrium. Suppose the opposite is true. Then 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, i.e.,  $\hat{R} = 1/\pi_1$ . Now, suppose the CRA posts fees that would implement a positive level of effort if a rating was ordered. (Specifically, let  $i$  be the outcome corresponding to the highest likelihood ratio, so that  $f_i$  is the only positive fee.) Given the fees, if investors were to observe the high rating, they would charge a different, lower, interest rate than if they see no rating:  $R_h(e) = \pi_h(e)/\pi_{h1}(e)$ , where  $e$  is the level of effort implemented given the fees. But then suppose the firm orders a rating and only discloses it if it is high. The firm receives financing at the rate  $1/\pi_1$  if the rating is low (in which case the firm will not show it to investors) and at  $R_h(e)$  if the rating is high. The firm therefore earns  $\Pi_{iss}(e) = \pi_{h1}(e)(y - R_h(e)) + \pi_{\ell 1}(e)(y - 1/\pi_1) - \pi_i(e)\psi'(e)/\pi'_i(e) = -1 + \pi_1 y - \pi_h(e) + \pi_{h1}(e)/\pi_1 - \pi_i(e)\psi'(e)/\pi'_i(e)$ . Notice that  $\Pi_{iss}(0) = -1 + \pi_1 y$  and straightforward algebra reveals that  $\Pi'_{iss}(0) > 0$ . Thus if the fee is small enough, the firm earns profits strictly above  $-1 + \pi_1 y$  by ordering a rating. Thus it cannot be the case that the rating is not ordered regardless of the fees when  $-1 + \pi_1 y > 0$ .

Then what is the equilibrium? Suppose the implemented effort is  $e(> 0)$ . Consider the following equilibrium candidate. If  $-\pi_\ell(e) + \pi_{\ell 1}(e)y \leq 0$ , then 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. And if  $-\pi_\ell(e) + \pi_{\ell 1}(e)y > 0$ , then investors finance after the high rating at  $\pi_{h1}(e)/\pi_h(e)$ , and finance at  $\pi_{\ell 1}(e)/\pi_\ell(e) (> 1/\pi_1)$  if the rating is low or if there is no rating. What are the highest fees that the 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  $\Pi_1 \equiv \pi_{h1}(e)(y - \pi_h(e)/\pi_{h1}(e)) - \sum_i \pi_i(e)f_i$ . If the firm



borrow 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.<sup>32</sup> This restriction imposes an upper bound  $\tilde{e}$  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  $\Pi_1$  (so long as the payoff to the CRA exceeds  $v_0$ ). The highest fee/effort for which  $\Pi_1$  just equals  $\Pi_2$  can be found by setting  $\Pi_2$  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, denoting the outcome with the highest likelihood ratio by  $i$  and using the CRA's incentive constraint  $\psi'(\tilde{e}) = \pi'_i(\tilde{e})f_i$ , the equation that  $\tilde{e}$  solves becomes  $\pi_{h1}(\tilde{e})[y - \pi_h(\tilde{e})/\pi_{h1}(\tilde{e})] - \pi_i(\tilde{e})\psi'(\tilde{e})/\pi'_i(\tilde{e}) = \pi_{h1}(\tilde{e})(y - 1/\pi_1)$  or  $-\pi_h(\tilde{e}) + \pi_{h1}(\tilde{e})/\pi_1 = \pi_i(\tilde{e})\psi'(\tilde{e})/\pi'_i(\tilde{e})$ . Notice that this condition, which pins down the equilibrium level of effort, is exactly the same as in the investor-pays case — see the proof of Proposition 5. Thus  $\tilde{e} = e^{inv}$ .

Finally, recall that when  $-\pi_\ell(\tilde{e}) + \pi_{\ell 1}(\tilde{e})y > 0$ , investors finance after the low rating or no rating at  $\pi_{\ell 1}(\tilde{e})/\pi_\ell(\tilde{e})$ . We need to verify that in this case the firm has no incentives not to order a rating and receive financing at this rate. By ordering a rating, the firm's gain from a lower interest rate after the high rating is  $\pi_{h1}(\pi_h/\pi_{h1} - \pi_\ell/\pi_{\ell 1}) = -\pi_h + \pi_{h1}\pi_\ell/\pi_{\ell 1} = -\pi_h + (\pi_{h1}/\pi_1)(\pi_1\pi_\ell/\pi_{\ell 1}) > -\pi_h + \pi_{h1}/\pi_1$ , where the right-hand side equals  $\pi_i f_i = \pi_i \psi' / \pi'_i$  at the equilibrium level of effort. Thus the gain from ordering a rating is indeed larger than the cost, implying that it is not profitable for the firm to deviate.

Note that the payoff to the firm is positive. But since the firm's payoff of not ordering a rating is zero when it receives no financing without a rating, why does not the CRA charge higher fees than those implementing  $\tilde{e}$  in that case? The answer is that if it did, then the equilibrium strategy of the firm and investors would be different — investors would finance at  $1/\pi_1$  if they see no rating (and at a lower interest rate if they see a high rating), and the firm would not order a rating. The firm will have no incentives to deviate by ordering a rating and only revealing it if it is high since for  $e > \tilde{e}$ ,  $\pi_{h1}(e)(y - \pi_{h1}(e)/\pi_h(e)) - \pi_i(e)f_i < \pi_{h1}(e)(y - 1/\pi_1)$ , or  $\pi_{h1}(e)(y - \pi_{h1}(e)/\pi_h(e)) - \pi_i(e)f_i + \pi_{\ell 1}(y - 1/\pi_1) < y - 1/\pi_1$ .  $\square$

**Proof of Proposition 6.** The first-order conditions of problem (2)–(5) subject to additional constraints (7)–(8) with respect to  $f_i$ ,  $i \in \{h1, h0, \ell\}$  — with  $\xi_h$  and  $\xi_\ell$  denoting

---

<sup>32</sup>An implicit assumption 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.

the Lagrange multipliers on these constraints — can be written as

$$-1 + \lambda + \xi_h + \xi_\ell - \xi_h \frac{\pi_1}{\pi_{h1}(e)} + \mu \frac{\pi'_{h1}(e)}{\pi_{h1}(e)} \leq 0, \quad f_{h1} \geq 0, \quad (10)$$

$$-1 + \lambda + \xi_h + \xi_\ell - \xi_h \frac{\pi_0}{\pi_{h0}(e)} + \mu \frac{\pi'_{h0}(e)}{\pi_{h0}(e)} \leq 0, \quad f_{h0} \geq 0, \quad (11)$$

$$-1 + \lambda + \xi_h + \xi_\ell - \xi_\ell \frac{1}{\pi_\ell(e)} + \mu \frac{\pi'_\ell(e)}{\pi_\ell(e)} \leq 0, \quad f_\ell \geq 0, \quad (12)$$

all with complementary slackness. Straightforward algebra shows that  $\pi'_{h0}(e)/\pi_{h0}(e) < \pi'_{h1}(e)/\pi_{h1}(e)$  and  $\pi_0/\pi_{h0}(e) < \pi_1/\pi_{h1}(e)$  for all  $e$  and all  $\gamma$ . Thus the left-hand side of (11) is always strictly smaller than the left-hand side of (10), and thus  $f_{h0} = 0$ .

To show that both  $f_{h1}$  and  $f_\ell$  must be strictly positive, suppose, for example, that  $f_\ell = 0$ . Then from (7), using  $f_{h0} = f_\ell = 0$ , we have  $-\psi(e) + \pi_{h1}(e)f_{h1} \geq \pi_1 f_{h1}$ , or  $-\psi(e) - \pi_{\ell1}(e)f_{h1} \geq 0$ , where  $\pi_{\ell1}(e) \equiv \pi_1 - \pi_{h1}(e)$ . But the left-hand side is strictly negative since  $e > 0$  (which is the case when the project is only financed after the high rating). A contradiction. A similar argument supposing  $f_{h1} = 0$  and using (8) also arrives to a contradiction.

Since both  $f_{h1} > 0$  and  $f_\ell > 0$ , constraints (10) and (12) must both hold with equality. Subtracting one from the other, obtain:

$$\mu \left[ \frac{\pi'_{h1}(e)}{\pi_{h1}(e)} - \frac{\pi'_\ell(e)}{\pi_\ell(e)} \right] = \frac{\xi_h \pi_1}{\pi_{h1}(e)} - \frac{\xi_\ell}{\pi_\ell(e)}.$$

Suppose that  $\gamma > \hat{\gamma}$ , so that  $\pi'_{h1}/\pi_{h1} - \pi'_\ell/\pi_\ell > 0$  (see the proof of Proposition 1), and (7) does not bind. Then  $\xi_h = 0$  and the right-hand side of the above equation is non-positive. On the other hand, as long as the incentive constraint binds so that  $\mu > 0$ , the left-hand side of (7) is strictly positive, a contradiction. An analogous argument shows that  $\xi_\ell$  must be strictly positive when  $\gamma < \hat{\gamma}$ . When  $\gamma = \hat{\gamma}$ ,  $\pi'_{h1}(e)/\pi_{h1}(e) = \pi'_\ell(e)/\pi_\ell(e)$ , so that incentives for effort can be provided equally well with  $f_{h1}$  and  $f_\ell$ , and thus (6) can be satisfied without any cost. Without loss of generality, we can assume that (7) is satisfied with equality at  $\gamma = \hat{\gamma}$ .  $\square$

**Claim 2** *Suppose  $f_{h1}$  and  $f_\ell$  are the optimal choices of fees in problem (2)–(6), where  $\psi(e) = A\varphi(e)$ . If the CRA chooses effort facing such fees and  $A' < A$ , then (6) 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  $\ell$  otherwise.*

**Proof.** The CRA's profits if it chooses to exert effort are  $\pi(A) \equiv \max_e -A\varphi(e) + \pi_{h1}(e)f_{h1} + \pi_\ell(e)f_\ell$ . By the Envelope theorem,  $\pi'(A) = -\varphi(e) < 0$ . Therefore the left-hand side of (6) with  $A'$  is strictly lower than that with  $A$ . Since the right-hand side of (6) does not change, and the constraint was binding with  $A$ , it now becomes violated. Which report the CRA makes then follows from Proposition 6.  $\square$

**Claim 3** *Suppose that  $\psi(e) = A\varphi(e)$ . Then the optimal level of effort in problem (2)–(5) strictly decreases with  $A$ .*

**Proof.** We use strict monotone comparative statics results from (Edlin and Shannon, 1998) to show that  $e$  is strictly decreasing in  $A$ . Define  $a = 1/A$ . Consider the case  $\gamma \geq \hat{\gamma}$  (the other case is analogous). Using Proposition 2 and substituting from (4), problem (2)–(5) can be written as  $\max_e -\pi_h(e) + \pi_{h1}(e)y - \pi_{h1}(e)\varphi'(e)/(a\pi'_{h1}(e))$  subject to  $-\varphi(e) + \pi_{h1}(e)\varphi'(e)/\pi'_{h1}(e) \geq va$ , where  $\pi'_{h1}(e) = p_g\gamma - p_b(1 - \gamma)$ . Denote the objective function by  $F(e, a)$ . Differentiating with respect to  $e$ ,  $F_e = -\pi'_h(e) + \pi'_{h1}(e)y - [\varphi'(e) + \pi_{h1}(e)\varphi''(e)/\pi'_{h1}(e)]/a$ . Since  $\varphi'(e) > 0$  and  $\varphi''(e) > 0$  for  $e > 0$ , and  $\pi'_{h1} > 0$  for  $\gamma \geq \hat{\gamma}$ , it follows that  $F_{ea} > 0$ . Next, differentiating the left-hand side of the constraint with respect to  $e$ , obtain  $\partial[-\varphi(e) + \pi_{h1}(e)\varphi'(e)/\pi'_{h1}(e)]/\partial e = \pi_{h1}(e)\varphi''(e)/\pi'_{h1}(e) > 0$  for  $e > 0$ . Thus, the constraint can be written as  $g(e) \geq va$ , where  $g$  is a strictly increasing function, or, equivalently,  $e \in \Gamma(a)$ , where  $\Gamma$  is nondecreasing in  $a$  in the strong set order. Therefore the optimal choice of effort is strictly increasing in  $a$ , or strictly decreasing in  $A$ .  $\square$

## A.1 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. 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  $j \in \{h1, h0, \ell\}$ :

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

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:

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

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

$$-1 + \lambda + \frac{1}{\pi_{i,j}} \left[ \mu \frac{\partial \pi_{i,j}(e, e_i)}{\partial e} + \mu_i \frac{\partial \pi_{i,j}(e, e_i)}{\partial e_i} \right] \leq 0, \quad f_{i,j} \geq 0, \quad (13)$$

with complementary slackness.

It is straightforward to check that

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

i.e., providing incentives for effort in period 1 by paying  $f_{h1,h1}$  is more effective than by paying  $f_{h1,h0}$  or  $f_{h1,\ell}$ . 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

$$\gamma < \left[ 1 + \frac{1 - p_g}{1 - p_b} \frac{1/2 - e_{h0}}{1/2 + e_{h0}} \right]^{-1},$$

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

Finally,

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

that is, providing incentives for effort in period 1 by paying  $f_{h0,h1}$  is more effective than by paying  $f_{h0,h0}$  or  $f_{h0,\ell}$ .

For concreteness, suppose that  $\gamma$  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  $\ell$  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  $\gamma$  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

$$\frac{\partial \pi_{h1,h1}(e, e_{h1}) / \partial e}{\pi_{h1,h1}} > \frac{\partial \pi_{h0,h1}(e, e_{h0}) / \partial e}{\pi_{h0,h1}} \quad (14)$$

if and only if

$$\frac{p_g(1/2 + e_{h1})}{p_g(1/2 + e_{h1}) + p_b(1/2 - e_{h1})} > \frac{(1 - p_g)(1/2 + e_{h0})}{(1 - p_g)(1/2 + e_{h0}) + (1 - p_b)(1/2 - e_{h0})}. \quad (15)$$

That is,  $f_{h1,h1}$  dominates  $f_{h0,h1}$  for incentive provision in period 1 if and only if (15) 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  $e_{h1} \geq e_{h0}$ , (15) holds automatically, but it might be violated if  $e_{h0}$  is sufficiently higher than  $e_{h1}$ . In order for (15) 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,  $\bar{e}$ , so that  $\psi(\bar{e})$  is large enough, for instance,  $\psi_{e \rightarrow \bar{e}}(e) = +\infty$ . To derive this upper bound, set  $e_{h1} = 0$  and  $e_{h0} = \bar{e}$  in (15), and replace the inequality with equality:

$$\frac{p_g}{p_g + p_b} = \frac{(1 - p_g)(1/2 + \bar{e})}{(1 - p_g)(1/2 + \bar{e}) + (1 - p_b)(1/2 - \bar{e})}.$$

Rearranging terms yields

$$\bar{e} = \frac{1 \frac{p_g}{p_g} \frac{1-p_b}{1-p_g} - 1}{2 \frac{p_g}{p_g} \frac{1-p_b}{1-p_g} + 1}.$$

## A.2 Proofs in the Case of Misreporting

The proofs of Propositions 3 and 4 (as well as the proof of Claim 1) extend to the case of misreporting without changes. The proofs of Propositions 2 and 5 and Claim 3 for this case are provided below. When it is important to distinguish functions and variables with and without misreporting, we mark those in the latter case with tilde.

First consider the payoff to the CRA,  $-\psi(e) + \pi_{h1}(e)f_{h1} + \pi_\ell(e)f_\ell$ , which from Proposition 6 equals  $\pi_1 f_{h1}$  if  $\gamma \geq \hat{\gamma}$  and  $f_\ell$  if  $\gamma < \hat{\gamma}$ . In the case of  $\gamma \geq \hat{\gamma}$ , (7) holding with equality implies  $\pi_\ell f_\ell = \psi + \pi_{\ell 1} f_{h1}$ . Substituting this into the incentive constraint  $\psi' = \pi'_{h1} f_{h1} + \pi'_\ell f_\ell$ , obtain  $\psi' = \pi'_{h1} f_{h1} + (\psi + \pi_{\ell 1} f_{h1})\pi'_\ell/\pi_\ell$  or  $\psi' - \psi\pi'_\ell/\pi_\ell = [\pi'_{h1} + \pi_{\ell 1}\pi'_\ell/\pi_\ell]f_{h1}$ . We can then express  $f_{h1}$  and substitute it into the payoff to the CRA to express the latter as a function of effort only. Similarly, for  $\gamma < \hat{\gamma}$ , (8) holding with equality implies  $\pi_{h1} f_{h1} = \psi + \pi_h f_\ell$ . Substituting into the incentive constraint, obtain  $\psi' - \psi\pi'_{h1}/\pi_{h1} = f_\ell[\pi'_\ell + \pi_h\pi'_{h1}/\pi_{h1}]$ . This leads to the following expression for the payoff to the CRA as a function of effort only, which we denote by  $V(e)$ :

$$V(e) \equiv \begin{cases} \pi_1 \left[ \frac{\psi'(e) - \psi(e)\pi'_\ell(e)/\pi_\ell(e)}{\pi'_{h1}(e) + \pi_{\ell 1}(e)\pi'_\ell(e)/\pi_\ell(e)} \right], & \text{if } \gamma \geq \hat{\gamma}, \\ \frac{\psi'(e) - \psi(e)\pi'_{h1}(e)/\pi_{h1}(e)}{\pi'_\ell(e) + \pi_h(e)\pi'_{h1}(e)/\pi_{h1}(e)}, & \text{if } \gamma < \hat{\gamma}. \end{cases} \quad (16)$$

Also denote  $C(e) \equiv \psi(e) + V(e)$ , the expected fees that implement effort  $e$ .

**Proof of Proposition 2 under Misreporting.** (i) Define  $\tilde{v}^* = V(e^*)$ . By construction,  $e^*$  can be implemented at  $v = v^*$ . For  $v > v^*$ ,  $e^*$  can be implemented by paying the same history-contingent fees as at  $v^*$  plus an upfront fee equal to  $v - v^*$ .

Next we show that  $\tilde{u}(\tilde{v}^*) < 0$ . Since  $\tilde{u}(v) \leq u(v)$  for all  $v$  and  $\tilde{u}(v) = u(v) = S^{FB} - v$

for  $v \geq \max\{v^*, \tilde{v}^*\}$ , it follows that  $\tilde{v}^* \geq v^*$ . Thus  $\tilde{u}(\tilde{v}^*) \leq \tilde{u}(v^*) \leq u(v^*) < 0$ , where the last inequality follows from part (ii) of Proposition 2.

(ii) Consider maximizing the firm's payoff while omitting constraint (3). The firm's payoff can be written as  $-\pi_h(e) + \pi_{h1}(e)y - \pi_{h1}(e)f_{h1} - \pi_\ell(e)f_\ell = (-1 + p_g y)(1/2 + e)\gamma + (-1 + p_b y)(1/2 - e)(1 - \gamma) - C(e)$ . The first-order condition with respect to effort is  $0 = [(-1 + p_g y)\gamma - (-1 + p_b y)(1 - \gamma)] - C'(e)$ . The term in the square brackets is strictly positive, while straightforward algebra shows that  $C'(e)$  equals zero at  $e = 0$  by our assumptions  $\psi(0) = \psi'(0) = \psi''(0) = 0$ . Thus  $e = 0$  cannot maximize the firm's profits. The CRA's payoff at  $e = 0$  is  $V(0) = -\psi(0) + C(0) = 0$ . Moreover, as we will show in the proof of part (iii) below,  $V(e)$  must be strictly increasing in  $e$  for  $e > 0$ . Therefore, for  $v$  below some threshold value, denoted by  $\tilde{v}_0$ , (3) does not bind. Moreover, the optimal level of effort for  $v \leq \tilde{v}_0$  is strictly positive.

(iii) For  $v \leq \tilde{v}_0$  effort is constant at  $e(\tilde{v}_0)$ , and for  $v \geq \tilde{v}^*$  it is constant at  $e^*$ . Suppose that  $v \in (\tilde{v}_0, \tilde{v}^*)$ . To show that the implemented effort is strictly increasing in  $v$  on this interval, it is enough to show that  $V(e)$  is strictly increasing in  $e$ . Since the total surplus  $-\psi(e) - \pi_h(e) + \pi_{h1}(e)y$  is strictly increasing in  $e$  for  $e < e^*$ , it will then follow that the total surplus is also strictly increasing in  $v$ .

We will only consider the case of  $\gamma \geq \hat{\gamma}$ , as the other case is analogous. The derivative of the numerator in the top expression in (16) with respect to  $e$  is  $\psi'' + (\pi'_\ell/\pi_\ell)^2 - \psi'\pi'_\ell/\pi_\ell$ , which is strictly positive if  $\pi'_\ell \leq 0$ . As for the denominator,  $\pi'_{h1} = p_b\gamma - p_g(1 - \gamma)$  and  $\pi'_\ell = 1 - 2\gamma$  are independent of  $e$ . In addition,  $\pi_{\ell1}(e)/\pi_\ell(e) = \pi_{1|\ell}(e)$ , the probability of success conditional on the low rating, is strictly decreasing in  $e$ . Thus the denominator is strictly decreasing in  $e$ , while the numerator is strictly increasing in  $e$  if  $\pi'_\ell \leq 0$ . Since for  $v \in (\tilde{v}_0, \tilde{v}^*)$  the left-hand side of (16) is equal to  $v$ , the implemented effort is strictly increasing in  $v$  if  $\pi'_\ell \leq 0$ .

Now suppose that  $\pi'_\ell > 0$ , and suppose that as  $v$  increases, the optimal level of effort remains unchanged or falls. The latter is not possible, since it is feasible to increase all  $f_i$ ,  $i \in \{h1, h0, \ell\}$ , by the same amount, thereby keeping  $e$  unchanged; this dominates a lower effort since the total surplus is strictly increasing in effort for  $e < e^*$ . If effort does not change, then an increase in  $v$  can be delivered by increasing all  $f_i$  by the same amount. But as the proof of Proposition 6 shows, it is never optimal to increase  $f_{h0}$  unless  $e < e^*$ . By keeping  $f_{h0}$  unchanged and increasing only  $f_{h1}$  and  $f_\ell$ , effort inevitably increases from (4) since  $\pi'_{h1} > 0$  (for  $\gamma \geq \hat{\gamma}$ ) and  $\pi'_\ell > 0$  (by supposition). A contradiction. The above argument also implies that  $V(e)$  must be strictly increasing in  $e$  (even if  $\pi'_\ell > 0$ ).  $\square$



**Proof of Proposition 5 under Misreporting.** The proof is a straightforward modification of the proof of Proposition 5 without misreporting. The fact that the marginal cost of implementing an arbitrarily uninformative rating with the possibility of misreporting is zero was shown in the proof of part (ii) of Proposition 2 under misreporting:  $C'(0) = 0$ .  $\square$

**Proof of Claim 3 under Misreporting.** The proof is a straightforward extension of the proof of Claim 3 without misreporting. Let  $a = 1/A$ , and define  $V_\varphi(e)$  as  $V(e)$  given in (16) where  $\psi$  is replaced by  $\varphi$ . Then the maximization problem (2)–(6) 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  $\tilde{F}(e, a)$ . 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  $e$  (the former is shown in the proof of Proposition 2 under misreporting), it follows that  $\tilde{F}_{ea} > 0$ . In addition, the constraint can be written as  $e \in \Gamma(a)$ , where  $\Gamma$  is nondecreasing in  $a$  in the strong set order. Therefore the optimal choice of effort is strictly increasing in  $a$ , or strictly decreasing in  $A$ .  $\square$

### A.3 Dynamic Model

As we discussed in the body of the paper, the outcome-contingent fee structure in our static model is essential for providing a CRA with incentives to exert effort. However, in a dynamic model, positive effort can be sustained even if the payment structure is restricted to flat up-front fees.

One quick way to illustrate this is by considering the following two period model. Suppose for simplicity that the project quality is i.i.d. in the two periods, and the CRA's effort is observable in the second period but not in the first period. The CRA charges flat up-front fees in both periods, where the fee in the second period potentially depends on what happened in the first period. Since effort in the second period is observable, the CRA will always exert the first-best level of effort in the second period, and will get paid its fee as long as it is optimal (for  $X$ ) to acquire a rating in the second period. Optimally choosing the fees that the CRA will charge in the second period allows us to give incentives to the CRA to exert effort in the first period, much in the same way as outcome-contingent fees in our original model do.<sup>33</sup>

Perhaps a more realistic — although also much more complicated — model where up-front fees can implement effort is a repeated infinite-horizon model, where effort is

---

<sup>33</sup>In this two period model we would have to assume that the CRA can commit in the first period to the fees that it will charge in the second period.

unobservable in every period. To show that our static model shares important similarities to this set up, let us describe an optimal contracting problem in such a model.

To be specific, consider an environment with one infinitely-lived rating agency and a sequence of short-run players — firms and investors — 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. Project quality is i.i.d. over time. Future profits are discounted at rate  $\beta \equiv (1 + r^f)^{-1} \in (0, 1)$ , where  $r^f$  is the risk-free interest rate.

What can we say about the optimal equilibrium assuming that only up-front fees can be used? 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.<sup>34</sup> Define  $I = \{h1, h0, \ell1, \ell0\}$  — the set of all potentially possible one-period histories. Then the recursive problem for  $\hat{U}_z$  can be written as follows:<sup>35,36</sup>

$$\hat{U}_z(v) = \max_{e, f, \{v_i\}_{i \in I}} -f + E[\max\{-\pi_h(e) + \pi_{h1}(e)y, 0\} + \max\{-\pi_\ell(e) + \pi_{\ell1}(e)y, 0\}] + \beta \sum_{i \in I} \pi_i(e) \hat{U}_z(v_i) \quad (17)$$

$$\text{s.t. } f + E[-\psi(e) + \beta \sum_{i \in I} \pi_i(e)v_i] = v, \quad (18)$$

$$\psi'(e) = \beta \sum_{i \in I} \pi'_i(e)v_i, \quad (19)$$

---

<sup>34</sup>The classic reference is Spear and Srivastava (1987).

<sup>35</sup>The expectation sign in the objective function and the promise-keeping constraint (18) 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.

<sup>36</sup>Note that a solution to this problem may not exist if  $v$  is large enough. For example, suppose that  $z = 0$  and  $\gamma$  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  $\hat{U}_0(0) = 0$ .

$$-f + E[\max\{-\pi_h(e) + \pi_{h1}(e)y, 0\} + \max\{-\pi_\ell(e) + \pi_{\ell1}(e)y, 0\}] \geq z, \quad (20)$$

$$\hat{U}_z(v_i) \geq \frac{z}{1-\beta} \text{ for all } i \in I, \quad (21)$$

$$v_{j1} = v_{j0} \text{ if } -\pi_j(e) + \pi_{j1}(e)y < 0, \quad j \in \{h, \ell\}, \quad (22)$$

$$e \geq 0, f \geq 0, v_i \geq 0 \text{ for all } i \in I. \quad (23)$$

Notice that the incentive constraint (19) is essentially the same as the incentive constraint (4) in the static problem except now the discounted continuation values to the CRA,  $\beta v_i$ , appear instead of the outcome-contingent fees,  $f_i$ . Making the CRA's future profits depend on the history will create incentive to exert effort even though only up-front fees are allowed.<sup>37</sup> The intuition is that the CRA expects to be able to charge different fees in the future depending on today's performance. We can think of choosing  $v_i$ 's optimally as designing the optimal 'reputation system' for the CRA. In particular, the analog of Proposition 1 (that follows straightforwardly from the first-order and envelope conditions of the above problem) is as follows:

**Claim 4** *Suppose the project is financed only after the high rating. Define  $\gamma_1 = 1/2$ ,  $\gamma_2 = 1/(1+p_g/p_b)$ , and  $\gamma_3 = 1/(1+[1-p_g]/[1-p_b])$ . (Recall that  $\hat{\gamma} = 1/(1+\sqrt{p_g/p_b})$  and notice that  $\gamma_1 < \hat{\gamma} < \gamma_2 < \gamma_3$ .)*

(i) *If  $\gamma \leq \gamma_1$ , then it is optimal to set  $v_\ell \geq v$ . If  $\gamma \geq \gamma_2$ , then it is optimal to set  $v_{h1} \geq v$ . If  $\gamma \geq \gamma_3$ , then it is optimal to set  $v_{h0} \geq v$ .*

(ii) *If  $\gamma \geq \hat{\gamma}$ , then it is optimal to set  $v_{h1} \geq v_\ell$ . Setting  $v_{h1} > v_{h0}$  is optimal regardless of  $\gamma$ .*

Next, 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  $U^{SB} = \hat{U}_0$ , 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 (21) 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

---

<sup>37</sup>Constraint (22) ensures that the continuation values do not depend on success/failure event if the project is not financed.

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, 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 (17)–(23), 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 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 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 5, 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 if the CRA can misreport the rating and 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 body of the text.

## 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 Publications Inc.
- Bolton, P., X. Freixas, and J. Shapiro (2012). The credit ratings game. *Journal of Finance* 67, 85–112.
- Bongaerts, D. (2013). Can alternative business models discipline credit rating agencies? mimeo.
- Bongaerts, D., M. Cremers, and W. Goftzmann (2012). Tiebreaker: Certification and multiple credit ratings. *Journal of Finance* 67, 113–152.

- Chade, H. and N. Kovrijnykh (2012). Information acquisition, moral hazard, and rewarding for bad news. mimeo.
- 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 (2012). Estimating the costs of issuer-paid credit ratings. mimeo.
- 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.).
- Edlin, A. and C. Shannon (1998). Strict monotonicity in comparative statics. *Journal of Economic Theory* 81, 201–219.
- 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 (2011). The economics of solicited and unsolicited credit ratings. mimeo.
- Goel, A. and A. Thakor (2011). Credit ratings and litigation risk. mimeo.
- Griffin, J. and D. Tang (2011). Did credit rating agencies make unbiased assumptions on cdos? *American Economic Review: Papers & Proceedings* 101, 125–130.
- 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.

- Inderst, R. and M. Ottaviani (2009). Misseling through agents. *American Economic Review*, 883–908.
- Inderst, R. and M. Ottaviani (2011). Competition through commissions and kickbacks. *American Economic Review* forthcoming.
- 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.
- Medvedev, A. and D. Fennell (2011). An economic analysis of credit rating agency business models and ratings accuracy. Financial Services Authority Occasional Paper 41.
- 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.
- Owusu-Ansah, Y. (2012). What went wrong? examining moodys rated cdo data. mimeo.
- 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 (2012). 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.
- Spear, S. and S. Srivastava (1987). On repeated moral hazard with discounting. *Review of Economic Studies* 54, 599–617.

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

United States Department of Justice Complaint (2013). Department of justice sues standard & poor's for fraud in rating mortgage-backed securities in the years leading up 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.