EIEF WORKING PAPER SERIES

IEF

Einaudi Institute for Economics and Finance

EIEF Working Paper 30/13
December 2013

Seller - paid Ratings

by

Sergei Kovbasyuk

(EIEF)

Seller-paid Ratings

Sergei Kovbasyuk*

November 11, 2013

Abstract

We study the interaction between the seller of a product, the buyers who are uncertain about the product quality and a rating agency who observes the quality and sends signals about it. Assuming the seller-pays model of rating agency, we analyze the cases in which the payment to the rater is publicly disclosed and fixed, publicly disclosed and rating-contingent, private and rating-contingent. First, we characterize all possible equilibrium partitions of the underlying quality range into ratings in these three cases. We then characterize the seller's optimal ratings in the three cases. Finally, we perform welfare analysis and discuss regulation.

Keywords: cheap talk, rating agency, mechanism design, transparency, limited commitment. *JEL classification*: D82, G24, G28, L15.

*EIEF (skovbasyuk@gmail.com). This paper benefited from helpful comments by Jonathan Berk, Bruno Biais, Jacques Crémer, Hugo Hopenhayn, Larry Epstein, Jana Friedrichsen, Mikhail Golosov, Bruno Jullien, Jean-Paul L'Huillier, Tymofiy Mylovanov, Marco Ottaviani, Nikita Roketskiy, Ilya Strebulaev, Daniele Terlizzese, Jean Tirole, Nicholas Trachter, Laura Veldkamp and seminar participants at Toulouse School of Economics, CREST, Einaudi Institute for Economics and Finance, Stanford GSB, Boston University, New Economic School, HEC Lausanne, University Pompeu Fabra, University of Zurich, Paris School of Economics, University of Mannheim, European University Institute, University of Copenhagen, CSEF University of Naples, EEA-ESEM 2011, EARIE 2012, NBER Credit Rating Agency Meeting 2013. The remaining errors are mine.

Introduction

In assessing the quality of an unfamiliar product the buyers often rely on the rating provided by a rater. Usually the buyers access the ratings for free: the widespread model for raters is the seller-pays paradigm. A well-known example is that of the three largest credit rating agencies Moody's, Standard & Poor's and Fitch. According to Partnoy (2006), around 90 percent of rating agencies' revenue comes from the fees paid by the issuers. The recent economic crisis raised concerns over the conflict of interest created in the seller-pays business model. Credit rating agencies were blamed for catering to issuers and inflating the ratings for mortgage backed collateralized debt obligations.[1] At first blush, this seems to contradict rationality of buyers: if buyers correctly infer the informational content of a rating there should be no reason for rating inflation.

In this paper we assume that the buyers are fully rational and there are no external forces that would make certain ratings highly valuable.[2] The rating agency receives payments from the seller and is partially liable to the buyers. We ask the following questions. Is there any flaw in this business model? If so, what is the nature of the flaw? Can a regulation mandating a fixed rater's fee improve the situation? Shall a regulation require public disclosure of the rater's fees or, is mandated disclosure redundant because all relevant parties already know the fees?

We show that under publicly disclosed rating-contingent payments, the seller-pays busi-

[1] Griffin and Tang (2012) employ a model used by one of the credit rating agencies and document that just before 2007 the AAA tranches of CDOs were larger than what the rating agency model would deliver. Furthermore, according to the 2009 report by International Monetary Fund of all asset-backed security collateralized debt obligations issued in 2005-2007 in the U.S. and rated AAA by Standard & Poor's, only 10% maintained AAA rating in 2009. Pagano and Volpin (2010) documents similar evidence.

[2] Several recent papers argued that the conflict of interest in credit rating agencies may have been exacerbated by regulation based on ratings Opp et al. (2013) or by seller's forum shopping and trusting nature of some buyers Skreta and Veldkamp (2009), Bolton et al. (2012).

ness model performs well. In this case the optimal contract between the seller and the rater leads to perfectly revealing ratings with positive payments above some quality threshold and an imprecise rating with zero payment below the threshold.[3] In this situation, the welfare is high. On the other hand, when the payments between the seller and the rater are private and rating-contingent, the seller-pays model can lead to low welfare and rating inflation: the highest rating being issued for a wide range of qualities. Mandating an upfront fixed payment does not help. It leads to an uninformative rating and low welfare. A desirable regulation of raters allows rating-contingent fees and requires their disclosure.

The 2010 report on credit rating agencies by the U. S. Government Accountability Office recommended that "An effective compensation model should be transparent to market participants to help them understand it and to increase market acceptance." Yet, current regulation under the Dodd-Franck act does not require credit rating agencies to disclose their fees. This is striking, given that the role of credit ratings in marketing financial products is almost as important as the role of underwriters. Underwriters of securities are required to disclose their commissions under the 1933 Securities Act. Buyers of securities can check underwriters' commissions in the prospectus published on the SEC website (www.sec.gov). Our analysis suggests that similar requirements can be imposed on credit rating agencies.

In our model, the seller wants to sell a product of unknown quality θ. The rater can learn the quality and issue a rating r. When the seller applies for a rating she commits to pay a fee $t(r)$, possibly contingent on a rating r. The rating is public and free for the buyers. The buyers use the rating to update their beliefs about the quality of the

[3]The optimal payments above the threshold are approximately proportional to the volume of the issue $t \sim pq$. In reality the fees of credit rating agencies correspond to 3-4 basis points of the volume of the issue for corporate bonds. Similarly, according to Chen and Ritter (2000) investment banks when underwrite an initial public offering of shares charge a fee equal to around 7% of the volume of the issue.

product, the seller sets the price $p(r)$, and the buyers purchase the desired quantity $q(r)$. All agents are Bayesian. The rater is partially liable to the buyers. We characterize feasible rating schemes (all possible equilibria) under three regulatory environments: 1) publicly disclosed fixed payment, 2) publicly disclosed and rating-contingent payment, 3) private rating-contingent payment.[4] Then, we solve for the seller's optimal rating scheme in these regulatory regimes and perform welfare analysis.

First, we consider the fixed fee regulation. In our model the rater' payoff is not affected by ratings directly and the ratings are cheap-talk. We show that a rater partially liable to the buyers has a natural downward bias in the spirit of Crawford and Sobel (1982). The bias arises endogenously. Intuitively, if the buyers were to take the rater's rating at face value the rater would understate the true quality so as to generate price moderation. The product price $p(r)$ is set after the rating r is issued. If ratings were perfectly revealing, for small changes in r the changes in the buyers' payoff would be proportional to changes in $\theta - p(r)$, the per unit surplus.[5] Under a fixed fee the rater's interests are congruent with the interests of buyers, meaning that the rater has an incentive to downplay the quality in an attempt to lower the price. Of course, in a Bayesian equilibrium the ratings must be correct on average and cannot be systematically lower than the true quality. Yet, the informativeness of ratings is affected by the bias. In the base-line setup the bias is extreme and leads to the uninformative rating under a fixed fee.[6]

Second, we consider the regulation mandating public disclosure of payments. In this

[4]When the payments is private to the seller and the rater, neither the fee schedule nor the actual payments are disclosed to the public. Bribe is an example of a private payment. Note that disclosure is irrelevant for a fixed fee as long as it is common knowledge that it is indeed fixed.

[5]The effect of changes in quantities $q(r)$ on the buyers' payoff is of second order because $q(r)$ is optimally chosen by the buyers and the envelope theorem applies.

[6]In principle, in the sender-receiver cheap-talk games inspired by Crawford and Sobel (1982), partially informative equilibria may happen if the sender's bias is small. We discuss this possibility in section 5.4.

environment perfectly revealing ratings are feasible. We solve for the seller's optimal rating scheme. It results in a coarse rating and no payment for qualities below some threshold (which one can interpret as rejections) and perfectly revealing ratings with positive and increasing payments above the threshold. The optimal payment balances the rater's bias and improves information revelation, resembling the profit-sharing fee adopted by credit rating agencies for bond ratings. We show that, contrary to conventional wisdom, an optimal rating scheme under a public contingent payment leads to more precise ratings and to a higher welfare than a rating scheme under a fixed payment.

Third, we consider the environment where rating-contingent payments are private to the seller and the rater. Since the buyers do not observe actual payments their perception of the ratings does not depend on actual payments. If the ratings were perfectly revealing the seller would have an incentive to elicit the highest rating by inflating the payment for this rating. This would bias the rater's reports away from truthtelling. To prevent such manipulations the ratings must be imprecise. Intuitively, when the highest rating is issued for a wide range of qualities the seller's expected cost of inflating the corresponding payment is high because with a substantial probability the quality will be in the range and the seller will have to make the inflated payment. At the same time the expected gain is limited because the buyers' perception of the rating is pinned down by the average quality over the corresponding range. The highest rating being issued for a wide range of qualities is akin to "rating inflation," even though rational buyers correctly infer the underlying quality. We show that any feasible rating scheme under private payments consists of partially informative ratings. Surprisingly, the seller's optimal rating scheme in this case is uninformative and the welfare is low. Intuitively, the expected payment under any feasible partially informative rating scheme is so high that the seller would prefer not to have the rater around. We conclude that contingent private payments are detrimental for welfare and information production.

Our paper also makes a methodological contribution to the literature on cheap talk with multiple audiences and to the literature on contacting for information. The next section discusses the methodological contribution in a greater detail.

Related literature

Our paper contributes to two strands of literature. On the methodological side we extend the literature on cheap talk communication in several dimensions. First, we extend it to games among multiple receivers and show that strategic interaction among multiple receivers can endogenously give rise to the properties of payoffs in the spirit of Crawford and Sobel. The sender's bias, which is often assumed mechanically in cheap talk games, arises naturally in our setup. The rater's announcement affects both the buyers' purchasing decision and the seller's pricing decision. The pricing decision in turn affects the buyers' payoff and makes the buyer-protective rater willing to bias the announcement. The communication with two strategic parties, the buyer and the seller, relates our paper to the literature on cheap talk with multiple audiences. In Farrell and Gibbons (1989), a sender S is informed about the true state θ and sends a message m to two uninformed receivers: P and Q, who take actions p and q respectively. Farrell and Gibbons show that the informativeness of communication between the sender S and receiver Q can be limited if receiver P also reacts to the sender's messages. They do not justify why receivers have different preferences and do not allow receivers to interact: receiver P's action does not affect receiver Q's payoff and vice versa. Differently from Farrell and Gibbons (1989), our receivers, the buyers and the seller, naturally have different preferences and interact strategically. We also allow one of the receivers (the seller) to contract with the sender.

The seller's ability to sign a rating-contingent contract with the rater also distinguishes our paper from standard sender-receiver cheap talk games and relates it to Krishna and Morgan (2008). Krishna and Morgan show that contracts facilitate communication in the single-receiver Crawford and Sobel (1982) set-up. In contrast to Krishna and Morgan, we analyze multiple strategic receivers and recover their payoffs from primitives. Furthermore, the single-receiver setup of Krishna and Morgan (2008) cannot distinguish between public and private contracts, as there is no third party. In our set-up the seller and the rater sign a contact which may not be observed by the buyers. The analysis of private contracts is thus natural in our setup, it constitutes our third contribution to the cheap-talk literature.

Our paper is also related to Inderst and Ottaviani (2012). They analyze recommendations by an information intermediary who is paid by two competing sellers A and B. The sellers differ in production costs and in suitability of their products for the buyers. The intermediary cares about product's suitability but not about the production costs. If payments are publicly disclosed he recommends the product the buyers would like, even when its production is very costly. Inderst and Ottaviani show that hidden kickbacks can tilt the recommendations towards the cost efficient product and improve social welfare compared to public payments. They do not fully study the production of information by the intermediary: the intermediary can only recommend product A or product B but cannot assess the quality of each of the products. In our model, the rater evaluates the quality of a product and communicates it to the buyers. Our model allows for multiple ratings and stresses the adverse effect private payments have on the number and on the informativeness of ratings. To the best of our knowledge, our paper is the first to study the endogenous partition of information in multiple ratings under private contracts.

Lerner and Tirole (2006) analyzes forum shopping by sellers when raters intrinsically care about product buyers and sellers can make costly concessions. In their model, as in Inderst and Ottaviani (2012), the buyers' decision is binary: they either adopt the product or not, which makes pass-fail examination optimal. They consider raters' incentives as given and focus on forum shopping by the sellers, finding that weak applicants go to tougher raters and make more concessions. In our model, the buyers can purchase any non-negative quantity, which makes pass-fail recommendations not optimal in general. We allow the seller to offer contingent payments to the rater, which enables us to analyze endogenous information production by the rater. We evaluate regulations that require the payments to the rater to be fixed, publicly disclosed and rating-contingent, or private and rating-contingent. We also shed some light on the issue of forum shopping for ratings, which was the main objective of Lerner and Tirole (2006), of Skreta and Veldkamp (2009), and of Bolton et al. (2012). A working paper by Goel and Thakor shows that a rater who maximizes a weighted average of the social value of the rating and the seller's profit uses imprecise ratings. Differently from our paper, Goel and Thakor do not solve for optimal contract between the seller and the rater and do not study how different disclosure regimes affect the precision of ratings.

The paper also relates a strand of literature on certification started by Lizzeri (1999), which assumes that raters can commit to a disclosure rule and ignores the issue of credibility. These papers are silent about how information production, credibility and welfare are affected when the rater faces a contingent contract. Lizzeri (1999) finds that a monopolistic rater who charges a fixed fee for ratings discloses minimum information to ensure efficient market exchange. Similarly, Kartasheva and Yilmaz (2013) shows that imprecise ratings allow a monopolistic rater who can't price discriminate between applicants to raise profits by equalizing the applicants' wiliness to pay for a rating. Farhi et al. (2013) argue that raters, regardless of their market power, have no incentive to disclose rejections. They conclude that requiring raters to reveal rejections benefits sellers. We also advocate for transparency, showing that public disclosure of sellers' payments to raters improves welfare. The focus of our study is different, as we look at the rater's ability to communicate information given the incentives he faces.

The paper is organized as follows. In section 1 we introduce a general environment. In section 2 we first prove that ratings partition the information available to the rater into potentially coarse grades (modified revelation principle). Secondly, we characterize feasible and optimal rating scheme when the payment from the seller to the rater is publicly disclosed. In section 3, we proceed with the analysis of private payments and characterize feasible and optimal rating scheme in this case. Section 4 is dedicated to welfare analysis. Section 5 presents several extensions of the basic model: the case of the informed seller, forum shopping by the seller, the case when optimal rating scheme is negotiated by the seller and the rater, and finally an extended analysis of feasible ratings under a fixed payment. Section 6 concludes.

1 Environment

There is a unit mass of potential buyers of a product of unknown quality θ. The quality is drawn from $F(.)$ supported on $[0,1]$. When buyers buy q units of a quality θ product at price p they receive a payoff

Assumption 1. $S(\theta, q, p) = \theta q - q^{\gamma}/\gamma - pq$, $\gamma \geq 2$.

The seller of the product is monopolistic and has zero marginal cost of production. She pays for ratings. Her profit net of the payment t given to the rater is $\pi(\theta, q, p) = pq - t$. For simplicity the seller is also uninformed about θ and always applies to the rater (we relax this assumption in section 5.1).

Assumption 2. *The product's quality θ is private to the rater.*

The rater learns θ at no cost and issues a public rating m from an arbitrary set M. The rating does not affect the rater per se, instead the rater puts weight $\lambda > 0$ on the buyers' surplus $S(\theta, q, p)$ and receives monetary payments $t(m) \geq 0$ for ratings from the seller.[7] The rater's payoff is

Assumption 3. $U(\theta, q, p, m) = \lambda S(\theta, q, p) + t(m)$.

In the first part of the analysis we abstract from the question who defines M, t and characterize possible equilibria taking $M, t(.)$ as given. Later on we turn to the optimal design of $M, t(.)$.

The timing is as follows. A set of ratings M and payment schedule $t(.)$ are defined. The seller applies for a rating. The rater observes θ and publishes a rating $m \in M$: his strategy $[0,1] \to \Delta(M)$ assigns to each quality θ a probability distribution over possible ratings M with a density function $\sigma(m|\theta)$, $\int_M \sigma(m|\theta)dm = 1$. Having observed m the seller sets the price p, her strategy is $\hat{p} : [0,1] \times M \to R^+$. The buyers and the seller form beliefs $M \to \Delta([0,1])$ that specify for each rating m a probability distribution over the possible qualities $\theta \in [0,1]$ with a density function $\mu(\theta|m)$, $\int_0^1 \mu(\theta|m)d\theta = 1$. Having observed p and m the buyers decide on quantity q, their decisions generate a market demand $\hat{q} : R^+ \times M \to R^+$. Finally payoffs are realized.

[7] If the rater cares about the seller's gross profit as well $U(b, m) = \lambda S(\theta, q, p) + \eta pq + t(m) = \lambda(\theta q - q^\gamma/\gamma) - (\lambda - \eta)pq + t(m)$, $\eta > 0$ the qualitative results are the same (for details see remark 1).

We assume that under a public payment all aspects of the game except θ are common knowledge. Under a private payment, before applying for a rating, the seller privately proposes payment schedule $t(.)$ to the rater, so both θ and $t(.)$ are not known to the buyers.

Assumption 1 allows us to obtain closed form solutions. It can be viewed as an approximation of the mean-variance preferences. To illustrate this, consider a product with a random unknown value. The rater gets a private signal about the value $\theta \in [0, 1]$. Conditional on θ, the value is distributed with mean θ and variance 1. Suppose the buyers have mean-variance preferences. If they knew θ, then their expected payoff from buying q units of the product at price p would be $S(\theta, q, p) = \theta q - \frac{\rho}{2}q^2 - pq$, with ρ being a measure of variance aversion. This expected payoff is similar to assumption 1 with $\gamma = 2$.

Assumption 2 simplifies the analysis. In practice, often the seller does not know the rater's methodology and cannot predict the rater's assessment of her product. Moreover, the raters get private information from many sellers they rate; for example, credit rating agencies and securities underwriters get confidential information about many issuers of securities. This enables raters to provide good evaluations of the *relative* quality of sellers, which is hard to do for an individual seller. In other words, a seller may have private information about her product, but the rater has private information about many products and, therefore, has an informational advantage over any seller about the *relative* quality of her product. In section 5.1, we study the case of an informed seller. We show that all equilibria obtained when the seller is uninformed can be replicated when the seller is informed, if the set of ratings M and transfers t are decided under the veil of ignorance.

Assumption 3 is crucial, because a rater with $\lambda = 0$ does not care about the true quality

θ and simply picks the highest transfer.[8] Even small positive λ allows our model to fit fees of real raters. For instance, $\lambda \approx 0.0003$ generates payments (as a part of optimal rating scheme) that match a typical credit rating agency fee of 3-4 basis points. In reality, parameter λ measures how liable the rater is for the buyers' losses in case of misleading ratings. It can be due to direct litigation costs, reputation concerns, or buyers' interests being explicitly represented within raters' decision bodies.[9] If the rater's expected litigation costs with the buyers were proportional to the buyers' expected loss, the rater's expected payoff would be equivalent to assumption 3. For example, $\lambda = 0.0003$ would imply that the buyers' expected loss of \$1 million would cost the rater \$300 in expectation.[10]

2 Public payments

First, we characterize a potentially feasible rating scheme and then analyze the optimal rating scheme. Intuitively, a rating scheme that is feasible when $t(.)$ is private should also be feasible when $t(.)$ is public. Therefore, we start with a less demanding case of public payments. Let M and $t(.)$ be publicly known and consider a Perfect Bayesian Equilibrium.

[8] Also, if discovering the true θ required an effort, then $\lambda > 0$ would be necessary to induce the rater to bear the effort cost.

[9] In governmental bodies such as the Food and Drug Administration, it is stated clearly that their aim is to serve the interests of consumers. Committees of Standard Setting Organizations that decide on new technological standards, on the other hand, are often composed of potential technology buyers.

[10] In practice lawsuits against raters are not uncommon. For example, according to The Wall Street Journal in April 2013 Moody's, Standard & Poor's, and Morgan Stanley settled two lawsuits with a group of investors led by Abu Dhabi Commercial Bank and King County in Washington. The lawsuits were related to concealing the risk of mortgage related deals. The plaintiffs sought \$708 million in damages, but the parties settled for \$225 million that were split equally among the defendants. Moody's shares rose 8.3% on the next trading day after the settlement, even though \$75 million payment was equivalent to 2.7% of Moody's total revenue in 2012. For details see wsj.com.

Modified revelation principle

We start by proving a property of a feasible rating scheme that greatly simplifies the analysis. The property states that an uninformed party needs to infer from a rating only the quality range (grade) for which the rating is issued. It builds on a modified revelation principle introduced in Krishna and Morgan (2008) and Bester and Strausz (2001).

Definition 1. A set of ratings G is a *grading* if each rating $r \in G$ corresponds to a convex set of qualities (grade) g_r and $r = E_F[\theta|\theta \in g_r]$. Grades satisfy $g_{r'} \cap g_r = \emptyset$ for any $r' \neq r$ and $\underset{r \in G}{\cup} g_r = [0, 1]$.

For example, a completely uninformative grading has a single rating $r = E_F[\theta]$ and a single grade $g_r = [0, 1]$. A perfectly informative grading, in contrast, has a continuum of perfectly revealing ratings and grades $r = g_r = \theta \in [0, 1]$. In the cheap-talk literature partition equilibria are common. Grading is a generalization of the partition concept, which permits intervals of perfect revelation alongside coarse intervals.

Lemma 1. *For any equilibrium under any $\{M, t(.)\}$, there exists an outcome-equivalent equilibrium in pure strategies under a grading $\{G, t'(.)\}$ in which the rater announces rating r whenever θ is in grade g_r.*

The proof is in the appendix. Intuition is the following. First, each rating gives rise to a market outcome $q(m)$, $p(m)$. Second, the single-crossing condition $U_{\theta q} > 0$ guarantees that in equilibrium, the quantity $q(\theta) = q(m(\theta))$ is weakly increasing in quality. The monotone function $q(\theta)$ has at most a countable number of jumps and naturally defines grades $\{g_r\}_{r \in G}$ over the interval $[0, 1]$, so that $q(\theta)$ is constant within each grade g_r.

The buyers perceive all products within the same grade to be of the same quality. In other words, they do not have more precise information than the grade itself. They also cannot have less precise information than the grade, since they know $q(\theta)$ changes with grades. Under any rating scheme, therefore, in equilibrium, the buyers learn merely the grade, and considering gradings is without loss of generality.[11] Considering gradings makes the analysis of a cheap-talk game simple and intuitive: the beliefs are naturally fixed by a given grading and it is enough to consider pure reporting strategies.

2.1 Feasible rating scheme under public payments

A feasible rating scheme is essentially a grading G and a payment $t(.)$ that are compatible with the parties' incentives: conditions (1), (2), (3), (4) and (5) listed below are satisfied.

Given a rating $r \in G$, the buyers believe the quality to be in the grade g_r. Therefore, for a price p they must buy a quantity q that maximizes their expected surplus:

$$\hat{q}(r,p) = \arg\max_{q \geq 0} E_F[\theta q - q^\gamma/\gamma - pq | \theta \in g_r], \ \forall r \in G. \tag{1}$$

For any $r \in G$ the seller takes into account the buyers' reaction and sets a price p that maximizes her profit:

$$\hat{p}(r) \in \arg\max_{p \geq 0} p\hat{q}(r,p), \ \forall r \in G. \tag{2}$$

Given the actual quality θ and the market reaction $\hat{p}(r)$, $\hat{q}(r, \hat{p}(r))$, the rater issues his preferred rating $r \in G$. His incentive compatibility constraint is

$$\hat{r}(\theta) \in \arg\max_{r \in G} [\lambda S(\theta, \hat{q}(r, \hat{p}(r)), \hat{p}(r)) + t(r)], \ \forall \theta \in [0,1]. \tag{3}$$

[11] One might wonder what happens if the rater instead of reporting from G should report something else $a \notin G$ or would not report at all. For these announcements, one can set $t'(a) = 0$ and buyers' beliefs such that $E[\theta|a] = E_F[\theta|\theta \in g_r]$ for some $r \in G$. Then for any $\theta \in [0,1]$, rating r dominates any $a \notin G$ because $\lambda S(\theta, q(r), p(r)) + t'(r) \geq \lambda S(\theta, q(r), p(r))$. Alternatively, we could dispose with the limited liability assumption and assume that the rater is very risk averse, his utility from money is t for $t \geq 0$ and $-\infty$ for $t < 0$. Then it suffices to specify $t'(a) < 0$ for $a \notin G$.

The reports must be consistent with the grading: the rater must truthfully report r if and only if the true quality is in the appropriate grade g_r

$$\hat{r}(\theta) = r,\ \forall \theta \in g_r,\ \forall r \in G. \tag{4}$$

The rater has limited liability: the payments are non-negative.[12] Non-negative payments also guarantee the participation constraint for the rater.

$$t(r) \geq 0,\ \forall r \in G. \tag{5}$$

Rater's intrinsic bias

First note that in any feasible rating scheme a rating $r \in G$ pins down equilibrium market price $p(r)$ and quantity $q(r)$. The first order condition for (1) gives $E_F[\theta|\theta \in g_r] - p - q^{\gamma-1} \leq 0$. In a grading $E_F[\theta|\theta \in g_r] = r$ and for $p \leq r$ the buyer's equilibrium strategy is $\hat{q}(r,p) = (r-p)^{\frac{1}{\gamma-1}}$. The seller' problem (2) gives $p(r) = \hat{p}(r) = \frac{\gamma-1}{\gamma}r$. After substituting for $\hat{p}(r)$ in $q(r) = \hat{q}(r, \hat{p}(r))$ one gets

$$p(r) = \frac{(\gamma-1)r}{\gamma},\ q(r) = \left(\frac{r}{\gamma}\right)^{\frac{1}{\gamma-1}}. \tag{6}$$

The market reaction to a rating is natural: both equilibrium price and equilibrium quantity increase with a rating.

One may think that the rater's strategy (3) can also be easily described. For instance, one may conjecture that under a fixed payment ($t = const$) the rater reports his information

[12] Negative payments are not plausible as they may provoke bogus applications.

perfectly $\hat{r}(\theta) = \theta \in [0,1]$. This conjecture is false. The rater's reporting incentives are not trivial. To see why perfect reports conflict the rater's incentives, imagine for a second that the buyers were not Bayesian: the buyers were to take the rater's report $r \in [0,1]$ at face value, that is their posterior would be r.

Definition 2. Suppose the buyers were to take a report $r \in [0,1]$ at face value. If for a given θ the rater would report $r < \theta$ he is *downward* biased, if he would report $r > \theta$ he is *upward* biased.

Lemma 2. *Under a fixed payment the rater is downward biased for any $\theta \in (0,1]$.*

Intuitively a public rating r affects both equilibrium quantity $q(r)$ and price $p(r)$, that in turn affect the buyers' surplus. The marginal effect of quantity on buyers' surplus is zero because of the envelope theorem (the buyers choose the quantity); consequently only the effect of price is relevant. A price set by the seller increases with a rating which causes a loss of buyers' welfare. Therefore under a fixed fee the buyer-protective rater would underreport quality: this would lead to a first order gain due to reduced price and only to a second order loss due to a distorted quantity. Formally, given θ the rater's report must satisfy $\frac{\partial[\lambda S(\theta,q(r),p(r))+t(r)]}{\partial r} = 0$. Suppose he reports $r = \theta$, from the buyers' problem (1) we have $\frac{\partial S(\theta,q(r),p(r))}{\partial q}|_{r=\theta} = 0$. For a constant fee $\frac{\partial t}{\partial r} = 0$ and we get $\frac{\partial[\lambda S(.)+t(r)]}{\partial r}|_{r=\theta} = \lambda \frac{\partial S(.)}{\partial p}\frac{\partial p}{\partial r}|_{r=\theta} = -\lambda q(\theta)\frac{\partial p}{\partial r} = -\lambda\frac{\gamma-1}{\gamma}q(\theta) < 0$, that is the rater prefers not to report $r = \theta$. One can show that he reports $r < \theta$ for $\theta \in (0,1]$.

The publicity of the ratings is key for this result. As was pointed out in Farrell and Gibbons (1989) in cheap-talk models with multiple audiences private communication with one of the audiences may be easier than public communication.[13] Certain raters do advise

[13] If the rater privately reports to buyers the seller cannot react to a report $\frac{\partial \hat{p}}{\partial r} = 0$, hence a fully informative private communication may be feasible under a fixed payment.

their clients privately; investment banks or consultants are among the examples. The analysis of private raters is out of the scope of this paper.

Remark 1. If the rater were to internalize profits $U = \lambda S(\theta, q, p) + \eta pq + t$ he would be downward biased for $\theta \in (0, 1]$ if $\eta < \eta^* = \frac{\gamma-1}{\gamma}\lambda$.

If $\eta > \eta^*$ the rater is intrinsically upward biased. This is not natural because balancing the rater's reporting incentives would require the payment $t(r)$ to be decreasing: that is the seller would pay more for low rating than for high ratings. The goal of the paper is to study how a general payment $t(r)$, of which $\hat{t}(r) = \eta p(r)q(r)$ is a special case, affects the rater's incentives. Therefore, we consider $\eta = 0$. We find that the payment $\hat{t}(r) = \eta p(r)q(r)$ is not a part of the optimal rating scheme.[14]

From now on the buyers are Bayesian, therefore in a feasible rating scheme the ratings must be correct on average. It is easy to see that the uninformative rating scheme with a single rating $r = E_F[\theta]$ and zero payment $t(r) = 0$ is always feasible. This natural benchmark corresponds to a market without a rater. In order to solve the model explicitly, we introduce an assumption common to the cheap-talk literature

Assumption 4. *The distribution of quality is uniform: $F(.) = U[0, 1]$.*

Proposition 1. (Monotonicity of payments.) *In a feasible rating scheme with at least two ratings the payments strictly increase with ratings: for any $r > r'$ one has $t(r) > t(r')$.*

The proof is in the appendix. The idea is that the rater's intrinsic bias in favor of a lower rating is so strong, that it is necessary to pay him an extra amount for a higher rating in

[14]Interestingly, if the rater cared about the utilitarian social welfare ($\eta = \lambda$) he would be upward biased. A price $p > 0$ leads to a transfer pq between the buyers and the seller which does not affect the utilitarian social welfare, but $p > 0$ discourages the buyers from consuming the socially optimal quantity. Indeed if the rater reports $r = \theta$ then $q(r) = (r/\gamma)^{\frac{1}{\gamma-1}} < \arg\max\,[\theta q - q^\gamma/\gamma] = \theta^{\frac{1}{\gamma-1}}$. Therefore, the utilitarian rater would report $r = \gamma\theta > \theta$ in an attempt to restore efficiency.

order to encourage him to issue the higher rating. Monotonicity of payments implies that the rater's limited liability constraint may bind only for the lowest rating, which greatly simplifies the analysis.

Corollary 1. *Under a fixed payment only the uninformative rating scheme is feasible.*

Corollary follows from proposition 1. The rater's bias is powerful enough to prevent informative ratings under a fixed payment, which echoes the Crawford and Sobel's babbling equilibrium when the sender's bias is extreme. In principle, under a contingent public payment the rater's bias can be undone altogether with an appropriate choice of payments. For example, perfectly reveling ratings are feasible with the following payment schedule $t(0) \geq 0$, $t(r) = \lambda \frac{(\gamma-1)^2}{\gamma} (r/\gamma)^{\frac{\gamma}{\gamma-1}} + t(0)$, $\forall r \in [0,1]$.

2.2 Optimal rating scheme under public payments

Having described a feasible rating scheme $G, t(.)$ under a contingent public payment we characterize a rating scheme maximizing the seller's expected profit. In section 5.3 we let the seller and the rater negotiate $G, t(.)$ and find similar results. The rating scheme optimal for the seller is a natural benchmark for the analysis.[15]

An optimal rating scheme must be feasible and maximize the seller's expected profit:

$$\max_{\{G,t(.)\}} E[pq - t|G, t(.)],\ s.t.(1), (2), (3), (4), (5). \tag{7}$$

Under a fixed payment only the uninformative rating scheme is feasible and, a fortiori, is

[15] If the rater uses the same grading and payment schedule to rate a representative sample of sellers, then the rating scheme we solve for delivers the highest profit to a representative seller. This benchmark corresponds to competitive rating industry where several identical raters compete for sellers. Also, if the seller markets a set of products of different qualities, then she must push the rater to adopt a rating scheme that maximizes her profit from selling a representative set of products.

optimal. When payments can depend on ratings one obtains

Proposition 2. *If $\lambda < \lambda^*$ an optimal rating scheme under a public payment commands a single imprecise rating with no payment for low qualities $\theta < \theta^*(\lambda)$ and perfect ratings with positive payments for high qualities $\theta \geq \theta^*(\lambda)$, if $\lambda \geq \lambda^*$ the uninformative rating scheme is optimal.*

The proof is in the appendix. Low quality applicants with $\theta < \theta^*(\lambda)$ pay nothing and receive low, imprecise quality assessments that can be interpreted as rejections. Applicants above some minimal threshold $\theta \geq \theta^*(\lambda)$ receive perfectly revealing ratings and pay for ratings. Optimal rating scheme is intuitive. Payments must increase with ratings (proposition 1). In the absence of a limited liability constraint the seller would implement perfectly reveling ratings for all qualities in $[0, 1]$ with a zero expected payment by asking the rater to pay a certain amount upfront. Under limited liability a net payment from the seller to the rater cannot be negative, hence perfect ratings become expensive for the seller. The payments increase additively with ratings, therefore the best way to economize on expected payments is to pay nothing for the lowest rating and increase the likelihood $\theta^*(\lambda)$ of this rating being issued. Contrary to conventional wisdom, a contingent payment, supposedly feeding the rater's conflict of interest, causes no apparent harm to the buyers. Quite the opposite in fact; a public contingent payment can implement perfectly revealing ratings while a fixed payment leads to uninformative rating and information loss.

Somewhat surprisingly, under a public payment optimal rating scheme becomes more informative when the rater cares little about the buyers: the threshold $\theta^*(\lambda)$ increases with λ and the interval with perfect ratings $[\theta^*(\lambda), 1]$ shrinks when $\lambda < \lambda^*$. Hence, if λ goes down the interval with perfect ratings widens. In the limit case $\theta^*(0) = 0$ and the optimal rating scheme is fully informative. Intuitively when λ increases, the rater internalizes the buyers' surplus to a greater extent and a larger compensation is necessary to balance his intrinsic downward bias. Perfect ratings become more expensive and the seller prefers a rating scheme where a large set of qualities is associated with the cheap imprecise rating. Interpreting the lowest rating as rejections we conclude that a tougher rater $\lambda' > \lambda$ rejects a higher fraction of applicants $\theta^*(\lambda') > \theta^*(\lambda)$.

The subsequent analysis shows that low λ is problematic if the seller pays the rater privately. For instance, $\lambda = 0$ would make the rater unable to provide informative ratings. Another reason for $\lambda > 0$ may be the rater's moral hazard: a costly effort may be necessary to learn the true quality θ. Clearly, a rater with $\lambda = 0$ would only care about the payment but not about the true quality and would not bear the effort cost.

A rough assessment of λ for real raters can be made using the information about their fees. The optimal rating scheme requires a fee $t(r) \approx \lambda p(r)q(r)$, where $p(r)q(r)$ is the gross revenue of the seller. According to Partnoy (2006) main credit rating agencies charge around 3 basis points of the volume of the issue (which is an analog of the gross revenue $p(r)q(r)$) as a fee for rating corporate bonds, this implies $\lambda \approx 0.0003$. Investment banks when underwrite an initial public offering of equity charge a fee of 7% of the offering volume, this gives $\lambda \approx 0.07$.

3 Private payments

Certain raters do not disclose the compensation they receive from sellers, for instance main credit rating agencies do not reveal the issuers' payments for ratings. Essentially, the payments are private: they are not observed by the buyers. The privacy of payments may prompt the seller to secretly skew the rater's compensation for high ratings (loosely speaking the seller may bribe the rater). For simplicity we still assume that the set of feasible ratings G is publicly known. This assumption can be easily relaxed. In real world the number of grades (ratings) typically is observable, while the exact borders between grades are not. Any rating scheme feasible under a private contract is also feasible under a public contract, hence the modified revelation principle applies. However, a rating scheme feasible under a public contract may be prone to manipulations under privacy.

The timing is as follows. Given a publicly known grading G the seller privately proposes payments $t(.)$ to the rater. The rater learns actual θ and reports $r \in G$. Given the rating r the buyers and the seller believe the quality to be in the corresponding grade g_r, the seller sets a price p and buyers buy a quantity q.

3.1 Feasible rating scheme under private payments

Here we characterize a feasible rating scheme under a contingent private payment $t(.)$. When the seller proposes $t(.)$ she takes into account the market reaction (1), (2) the rater's incentive compatibility constraint (3) and the limited liability constraint (5).[16] However, she ignores the rater's truth-telling constraint (4). With a private payment the seller may want to deceive the buyers and manipulate the rater's actual reports away from the implied grading G. Her maximization problem is

$$\max_{\{t(.)\}} \int_0^1 E[pq - t|G, t(.)]d\theta, \ s.t.(1), (2), (3), (5). \tag{8}$$

[16] One should not take our model literally, where a single seller contracts with the rater. Instead one can think of a representative sample of sellers, each of whom is uncertain about her product and attempts to tilt the rater in her favor. The sum of these efforts would result in something close to the problem we analyze. This is coherent with the observation in Partnoy (2006): "Because rating agencies rate thousands of bond issues, they do not depend on any particular issuer, so the concern about conflicts is more systemic than individualized."

Bayesian buyers can't be deceived in an equilibrium, a rating scheme $G, t(.)$ is feasible if and only the private payment solving (8) induces the rater's reports consistent with the truth-telling constraint (4). In other words, the grading G must be robust to hidden payment manipulations by the seller.

Proposition 3. *A feasible rating scheme $G, t(.)$ under a private payment has no intervals of perfect revelation: G contains at most a countable number N of coarse ratings. If $N \geq 2$ at any border point between any two ratings both ratings deliver to the seller the same virtual profit $\tilde{\pi}(\theta, r) = p(r)q(r) + \lambda S(\theta, q(r), p(r)) - \lambda q(r)(1-\theta)$:*

$$\tilde{\pi}(\theta_{i-1}, r_{i-1}) = \tilde{\pi}(\theta_{i-1}, r_i), i = 2, ..., N. \tag{9}$$

The detailed proof is in the appendix. To illustrate the idea behind the proof we reformulate the problem (8) using the rater's indirect utility $U(\theta) = \underset{r \in G}{max}[\lambda S(\theta, q(r), p(r)) + t(r)]$. Using envelope theorem $\frac{dU(\theta)}{d\theta} = S_\theta(\theta, q(r), p(r)) = \lambda q(r) \geq 0$ we express the seller's expected profit as $\Pi = \int\limits_0^1 \tilde{\pi}(\theta, r)d\theta - U(0)$. The seller's revenue $p(r)q(r)$ increases with ratings, thus for each border point θ_i, $i = 1, ..., N-1$ per se the seller prefers a rating r_i over a rating r_{i-1}. If she attempts to marginally increase the chance of receiving r_i instead of r_{i-1} and inflates the payment $t(r_i)$ she will have to raise payments for all ratings above r_i to maintain incentive compatibility, which is costly. This additional cost is captured by the term $\lambda S(\theta, q(r), p(r)) - \lambda q(r)(1-\theta)$ in the virtual profit. Note that the lower the θ_i the higher is the associated cost of payment inflation. For the seller not to be wiling to manipulate the payments she must perceive the same virtual profit from the two bordering ratings (9).

In order to see that perfectly revealing ratings are not possible, suppose there was an interval with perfectly revealing ratings (a,b). Then maximization of Π would require a necessary condition $\frac{\partial\tilde{\pi}(\theta,r)}{\partial r} = (\frac{r}{\gamma})^{\frac{1}{\gamma-1}}\left[1-\lambda\frac{\gamma^2+\gamma-1}{\gamma^2-1}+\frac{\lambda}{(\gamma-1)}\frac{2\theta-1}{r}\right] = 0$ for $r = \theta \in (a,b)$. The ratio $\frac{2\theta-1}{\theta}$ is strictly increasing, the necessary condition would be violated for some $\theta \in (a,b)$ and, hence, an interval with perfect ratings is impossible. In other words, if the ratings were perfect the seller would choose the payments that would induce the rater to misreport the quality. Misreporting is inconsistent with rational expectations, hence perfect ratings are not possible in equilibrium.

Rater's bias under private payments

A private payment may be skewed in favor of high ratings. Consequently, the rater may be endogenously upward biased for high ratings. Unfortunately it is impossible to measure the rater's bias as it depends on a particular payment $t(.)$ (for instance if $t = 0$ the rater has no bias caused by payments). Nevertheless, some properties of a feasible rating scheme can be established. We explicitly characterize a feasible rating scheme and illustrate its properties in case $\gamma = 2$, which gives rise to the buyers' utility function $U = q\theta - q^2/2 - pq$. For brevity we consider $\lambda \in (0, 2]$ as degree to which the rater may internalize the buyers' surplus.

Corollary 2. *If $\gamma = 2$ a feasible rating scheme $G, t(.)$ under a private payment depends on λ.*

1) If $\lambda \in (0, \frac{2}{11}]$ then $G, t(.)$ admits $N = 1$ rating with borders $\theta_0 = 0$, $\theta_1 = 1$.

2) If $\lambda \in (\frac{2}{11}, \frac{2}{3}]$ then $G, t(.)$ admits $N = 1, 2$ ratings. If $N = 2$ the ratings' borders are $\theta_0 = 0$, $\theta_1 = \frac{11\lambda-2}{10\lambda+4}$, $\theta_2 = 1$.

3) If $\lambda \in (\frac{2}{3}, 2]$ *then* $G, t(.)$ *admits* $N = 1, ..., \infty$ *ratings. If* $N \geq 2$ *the ratings' borders are* $\theta_0 = 0$*,* $\theta_N = 1$*,* $\theta_i = \sum_{j=1}^{i} \frac{(8-4\lambda+8\lambda D^N)D^{N+1-j}+(8\lambda+(8-4\lambda)D^N)D^j}{(3\lambda-2)(1-D^{2N})(1-D)}$*,* $i = 1, ..., N-1$*,* $D = \frac{5\lambda+2-4\sqrt{\lambda(\lambda+2)}}{3\lambda-2}$.

The proof is a solution of (9) for $\gamma = 2$, it is in the appendix. We illustrate the result with a figure. Given λ we depict the boundaries of ratings for a rating scheme with the highest number of ratings feasible. In the figure, $\theta_1^*(\lambda)$ is the upper boundary of the low rating (the thick solid line) and $\theta_{N-1}^*(\lambda)$ is the lower boundary of the high rating (the thick dashed line).

If the rater is lax $\lambda \leq \frac{2}{11}$ only the uninformative rating scheme is feasible. Indeed, if the rater were to use more than one rating then, given that λ is small, the seller would offer a private payment which would induce the rater to report a high rating even when the quality is low. Such a manipulation would be profitable because a small premium would bias the rater's reports away from what is implied by the grading. Therefore, informative ratings are not possible in equilibrium.

When $\lambda \in (\frac{2}{11}, \frac{2}{3}]$ the rater is sufficiently tough to implement an informative rating scheme with two ratings. The high rating is issued for a wide range of qualities and involves more pooling than the low rating: $\theta_1^*(\lambda) < 1/2$. This is because the rater's reports are affected by a generous payment offered for the high rating. For the same reason a finer grading with more than two ratings in not possible: if the rater were to use more than two ratings the seller would offer a large payment for the highest rating and would induce the rater to misreport.

The most interesting case is $\lambda \in (\frac{2}{3}, 2]$, when infinitely many ratings are feasible. As can be seen on the picture the low and the high ratings are very imprecise, while intermediate ratings are relatively precise. It appears as if the rater is downward "biased" for low qualities and upward "biased" for high qualities. To understand the intuition, notice that the seller may want to manipulate the payment for two reasons. First, the seller does not like to pay for ratings per se and, therefore, she prefers low payments. Second, the seller prefers high ratings over low ratings and she may want to increase payments for high ratings.

The seller has a strong incentive to increase $t(r_N)$, which lowers the threshold θ_{N-1} and increases the likelihood of obtaining the highest rating. An incentive to increase the payment for the second highest rating is not as strong, because an increase in $t(r_{N-1})$ not only lowers the threshold θ_{N-2} but also raises the threshold θ_{N-1}. The second effect is unwelcome since the likelihood of rating r_{N-1} increases at the expense of the likelihood of the highest rating r_N. In order to counterbalance the second effect $t(r_N)$ has to go up. By similar reasoning, incentives to inflate $t(r_{N-2})$ are even lower than incentives to inflate $t(r_{N-1})$, and so on. As a result, under a private payment the seller offers a high compensation for high ratings. This in turn makes the rater endogenously more "upward" biased for high ratings than for low ratings. In fact, for very low ratings the rater's upward bias because of the payments becomes smaller than the rater's intrinsic downward bias (lemma 2). Overall, therefore, the rater is downward biased for low ratings.

One can interpret the imprecise highest rating as a manifestation of so called "rating inflation". In our framework all agents are rational and nobody is deceived, yet "rating inflation" can be harmful as the informational value of the highest rating is limited. The informational value of the rating is low not because buyers fail to interpret it, but because a high payment from the seller induces the rater to give the highest rating even to the products of a quality below the highest.

3.2 Optimal rating scheme under private payments

An ex ante optimal rating scheme under a private payment must be feasible and it must maximize the seller's profit

$$\max_{\{G,t(.)\}} E[pq - t|G, t(.)],\ s.t.(4), (8). \tag{10}$$

Proposition 4. *Under a private payment the uninformative rating scheme with zero payments is optimal.*

Formal proof is in the appendix. The result is striking because the seller is information loving per se. However, informative ratings require the payment to rapidly increase with ratings. As a result, the seller's profit net of the payment becomes a concave function of quality: under a private payment the seller becomes information averse. In other words, if the seller has an opportunity to privately pay the rater, then from an ex ante perspective she prefers not to have an informative rater around. If such a rater was present then the seller would not resist the temptation to offer high payments for high ratings. Consequently, in expectation the seller would end up paying the rater more than the extra profit she would gain from informative ratings.

4 Welfare analysis

For any feasible rating scheme $G, t(.)$ denote the ex ante expected buyers' surplus as $S(G)$, the seller's gross expected profit as $P(G)$ and the expected payment to the rater as $T(G, t(.)) = E[t|G, t(.)]$. The seller's net expected profit is $\Pi = P(G) - T(G, t(.))$ and the rater's expected payoff is $U = \lambda S(G) + T(G, t(.))$. We define the social welfare as a sum of the buyers' surplus, the seller's profit and the rater's payoff $W(G, t(.)) = (1+\lambda)S(G) + P(G)$.

Remark 2. A social planner with the objective functions W is information loving.

Intuitively information is beneficial for the buyers because high quality products have higher value and should be purchased in larger quantities than low quality products. At the same time, the seller also benefits from information because higher quality products have higher equilibrium prices and are sold in larger quantities.

Remark 3. If the social planner with the utilitarian objective function W were to choose rating scheme, she would induce perfectly revealing ratings for all $\theta \in [0, 1]$.

In reality, a regulator can hardly impose a particular rating scheme. Nevertheless, the regulator may be able to impose restrictions on contracts between the parties. For instance, require a fixed payment or mandate full disclosure of payments from the seller to the rater. We proceed assuming that the social planner can choose between three contracting regimes: private contingent payments, public contingent payments and fixed payments; we also assume that under any regime the rating scheme optimal for the seller takes place. Corollary 1 and proposition 4 state that the uninformative rating scheme prevails under a fixed-fee and under a private contingent payment. According to proposition 2, under a public contingent payment the optimal rating scheme is informative if and only it $\lambda < \lambda^*$, thus we have

Proposition 5. *If $\lambda < \lambda^*$ then a a public contingent payment strictly dominates a fixed-fee or a private contingent payment from the social welfare perspective. If $\lambda \geq \lambda^*$ all three regimes lead to the uninformative rating scheme and deliver the same social welfare.*

Generally the rating scheme preferred by the seller produces less information than socially desirable. The utilitarian social planer does not care about the cost of implementing the fully informative rating scheme because the payment to the rater is a redistribution which does not affect the total welfare. The seller, on the other hand, bears all the cost of ratings but does not internalize a surplus which accrues to the buyers, hence she prefers not the fully informative rating scheme. One way to get closer to the social optimum is to ask buyers to pay the rater for his services. The free-rider problem and lack of commitment on the buyers' side make such a scheme dubious. An alternative regulation, which might appear controversial at first glance, calls for subsidizing the seller's expenses on ratings. Indeed, under a public payment if the seller were to receive a subsidy for each dollar spent on ratings, she would induce a more informative rating scheme and the social welfare would improve.

5 Extensions

In this section we present several extensions of the basic model. First, we generalize our analysis to the case of informed seller. Second, we let the seller choose among multiple raters. Then, we let the seller and the rater negotiate optimal rating scheme. Finally, we extend our analysis to markets with a moderate price reaction to ratings and find that informative ratings can be feasible and optimal under a fixed payment.

5.1 Informed seller

Here we generalize our results to the case when the seller is informed. We rule out other ways to signal quality than ratings.[17] Suppose the rating scheme maximizing the ex ante

[17] Other signalling ways like warranties or advertising may not be credible in the context of financial products. For instance an issuer of bonds can't issue a meaningful warranty on her own bonds.

expected profit of an uninformed seller $G, t(.)$ is established as a common business practice. The seller learns her quality before applying for a rating. It is natural to think that the buyers hold pessimistic beliefs: if the seller does not apply they perceive her product to be of quality $\underline{r} = \min G$. In reality the lowest rating is often a rejection which is not disclosed. Hence, not certified sellers are pooled in the same category with low quality sellers and are perceived in the same way.[18] The rest of the game proceeds as before.

A rating scheme optimal under a fixed payment and under a private contingent payment requires zero payments and is uninformative. When $\lambda \geq \lambda^*$ the optimal rating scheme under a public contingent payment is also uninformative. It is easy to see that the uninformative rating scheme can be sustained as an equilibrium when the seller is informed. Indeed the seller is indifferent about applying for a rating and it is an equilibrium strategy to apply.

Proposition 6. *If $\lambda < \lambda^*$ a rating scheme $G, t(.)$ optimal for an uninformed seller under a public contingent payment can be sustained as an equilibrium when the seller is informed.*

The proof is in the appendix. Intuitively, the seller's profit increases with ratings. A seller who has not applied for a rating is believed to have a low quality product. Given that the payment for the lowest rating is zero it is a weakly dominant strategy for the seller to apply: in the worst case she will get the lowest rating and will not pay anything, if she will get a higher rating her profit will be higher. Hence, it is an equilibrium strategy to apply.[19]

[18] For instance credit rating agencies do not disclose their rejections and most of the issuers without a rating are perceived to be of low quality. Also, a hotel with no stars is perceived as a low quality hotel.

[19] The signalling equilibrium presented here is not unique. For instance, if a seller who has no rating is believed to have a quality $E[\theta] = \frac{1}{2}$, while a seller who has applied is believed to have $\underline{r} < E[\theta] = \frac{1}{2}$, then no seller would apply. The equilibrium described in proposition 6 delivers the highest profit to an average seller, and can be thought of as a benchmark.

5.2 Forum shopping

In the original set-up with the uninformed seller we allow the seller to choose a rater out of a continuum of raters with different degrees of buyer-protectiveness $\lambda \in [0, \infty)$ known to all parties. This environment can be seen as an approximation of a competitive rating industry. If the payments between the seller and the rater are either private or fixed, then the optimal rating scheme is uninformative independently of λ. In these cases the analysis is trivial. Consider public contingent payments. Assume each of the raters, depending on his λ, proposes a rating scheme preferred by the seller.

Remark 4. Under a public contingent payment the seller hires the rater with $\lambda = 0$.

It follows from proposition 2, low λ leads to low $\theta^*(\lambda)$ and high information production. The seller is information loving, therefore she chooses a rater with the smallest λ. Intuitively, a rater who cares little about the buyers can be cheaply compensated for his intrinsic bias and induced to perfectly rate the product. As a result, in the absence of other concerns, such as private payments or rater's moral hazard in information acquisition, the seller picks the rater which cares about the buyers the least.

5.3 Optimal negotiated rating scheme

In sections 2.2 and 3.2 we have analyzed rating scheme optimal for the seller. In real world the rater often decides on the grading and on the associated payments. To accommodate this possibility we let the rating scheme $G, t(.)$ be jointly decided by the rater and the seller before any party learns the actual quality. Given a feasible rating scheme $G, t(.)$ the seller's ex ante expected profit is equal to the gross expected profit net of the expected payment $\Pi = P(G) - T(G, t(.))$. In a similar way, the rater's expected payoff is a sum of a weighted buyers' expected surplus and the expected payment $U = \lambda S(G) + T(G, t(.))$. If the parties fail to agree on $G, t(.)$ the uninformative rating $E[\theta] = 1/2$ with $t = 0$ obtains; and the parties get their reservation utilities $\Pi_0 = P(\frac{1}{2})$ and $U_0 = \lambda S(\frac{1}{2})$. To model negotiations we use a generalized Nash bargaining solution where the seller's bargaining power is $\nu \in [0, 1]$:

$$\max_{\{G,t(.)\ feasible\}} (\Pi - \Pi_0)^{\nu}(U - U_0)^{1-\nu},\ s.t.\ \Pi \geq \Pi_0,\ U \geq U_0.$$

It is easy to see that an optimal negotiated rating scheme under a fixed payment or under a private payment is uninformative and requires no payments. Under a fixed payment only the uninformative rating scheme is feasible hence it is optimal. Under a private payment in any feasible informative rating scheme the seller gets a profit $\Pi < \Pi_0$ (for details see the proof of proposition 4), therefore the negotiation leads to the uninformative rating scheme with no payments. The same is true in case of a public contingent payment if $\lambda \geq \lambda^*$.

Proposition 7. *If $\lambda < \lambda^*$ the negotiated rating scheme under a public contingent payment is as follows.*

1) If the rater is not very buyer-protective $\lambda \leq \lambda_\Delta$ and the seller has a low bargaining power $\nu \leq \nu^*$ the optimal rating scheme is fully informative.

2) If $\lambda > \lambda_\Delta$ or $\nu > \nu^*$ the optimal rating scheme leads to a coarse rating and zero payment for $\theta \in [0, \theta^*(\nu, \lambda)]$, and perfect ratings with positive payments for $\theta \in (\theta^*(\nu, \lambda), 1]$. The threshold $\theta^*(\nu, \lambda)$ is increasing in ν.

The detailed proof can be found in the appendix. The idea is the following. For a given total expected transfer $T(G, t(.))$ the seller and the rater have congruent preferences about the grading. As a result, an optimal negotiated rating scheme is qualitatively similar to the rating scheme maximizing the seller's expected profit (which obtains for $\nu = 1$): a threshold $\theta^*(\nu, \lambda)$ separates a low imprecise rating from high perfect ratings. The difference is that the threshold $\theta^*(\nu, \lambda)$ depends on the bargaining power of the seller ν. Only ratings above the threshold $\theta^*(\nu, \lambda)$ command positive payments to the rater. If the seller has a low bargaining power ν, the rater gets a high expected payment $T(G, t(.))$. This in turn implies a wide interval of perfect ratings with positive payments and, hence, a low threshold $\theta^*(\nu, \lambda)$. If the rater's bargaining power is high enough $\nu \leq \nu^*$ fully informative ratings are implemented for $\theta \in [0, 1]$.

5.4 Partially informative ratings under fixed payments

Throughout the analysis we have assumed $\gamma \geq 2$ which guaranteed a significant price reaction to a rating $\frac{\partial p(r)}{\partial r} = \frac{\gamma - 1}{\gamma} \geq \frac{1}{2}$. We showed that in this case informative ratings were not feasible under a fixed payment. The assumption $\gamma \geq 2$ is reasonable in a context of financial markets, where competition among investors (potential buyers) drives the product's price close to it's expected value $p(r) \approx r$. Nevertheless, in consumer goods markets one may expect a less competitive demand and a lower reaction of price to quality changes. Here we let $\gamma \in (1, 2)$ and analyze a rating scheme under a fixed public payment. We do not analyze contingent payments because of technical difficulties. A rating scheme is feasible whenever (1), (2), (3), (4) and $t(r) = t \geq 0$, $\forall r \in G$ hold.

Proposition 8. *A feasible rating scheme G, t under a fixed payment is a partition $\{\theta_i\}_{i=0,\ldots,N}$ of interval $[0, 1]$ into N ratings, such that $\theta_0 = 0$, $\theta_N = 1$. If $N \geq 2$ then*

$$S(\theta_{i-1}, q(r_i)p(r_i)) = S(\theta_{i-1}, q(r_{i-1}), p(r_i)),\ i = 2, ..., N, \tag{11}$$

i) the number of ratings is not greater than $N(\gamma) < \infty$*,* $N(\gamma) = 1$ *if* $\gamma \geq \gamma^*$*,* $N(\gamma) \geq 2$ *if* $\gamma < \gamma^*$, $\gamma^* \approx 1.74$.

ii) if $N \geq 2$ *high ratings are more precise than low ratings :* $\theta_i - \theta_{i-1} < \theta_{i-1} - \theta_{i-2}$*,* $i = 1, ..., N$.

The detailed proof is in the appendix. Observe that γ determines the mark-up $p(r) = \frac{\gamma-1}{\gamma}r$. When $\gamma \geq \gamma^* \approx 1.74$ the mark-up is high and, hence, the rater's intrinsic bias is high. As a result the rater fails to communicate product quality, in this case only the uninformative rating is feasible.

An optimal rating scheme under a fixed payment maximizes the seller's expected profit. The seller's problem is (7) with an additional constraint $t(r) = t$, for any $r \in G$.

Proposition 9. *An optimal rating scheme under a fixed payment is a partition* $\{\theta_i\}_{i=0,...,N(\gamma)}$ *with the maximum feasible number of ratings* $N(\gamma)$ *and zero payments.*

Formal proof is in the appendix. The result is intuitive. A fixed fee does not influence the rater's reports and the seller sets a zero fee. Equilibrium prices and quantities are higher when a high quality is certified, hence the seller's profit is a convex function of quality. The seller is information loving and prefers the most informative feasible rating scheme.

Our findings accord with practices of some real raters. Often the raters that charge fixed fees issue few coarse ratings. For instance the Food and Drug Administration sets standard application fees and publishes pass-fail quality assessments. Similarly raters of consumer goods such as the European New Car Assessment Programme or the Michelin Guide typically charge fixed fees and evaluate products on five- and three-star scales. The analysis stresses that a moderate price reaction to ratings, which is common to consumer goods, is the prerequisite of an informative rating scheme under a fixed payment. In case of a financial product, on the other hand, the price is very close to the perceived product's value and is very sensitive to a rating change. Our theory predicts that in the latter case informative ratings under a fixed payment are problematic.

6 Conclusion

This paper has studied the issuer-pays business model of a rating agency. Our main conclusion is that the issuer-pays model performs well as long as payments between the seller and the rater are publicly disclosed. If a rational buyer observes that a seller has paid an unusually high fee for a rating, then his perception of the product's quality goes down. This in turn prevents the seller from inflating the payments and improves the precision of the ratings. By contrast, a phenomenon akin to "rating inflation" can occur when payments are private for the seller and the rater. In this case the seller may manipulate the payments in an attempt to bias the rater's reports. In order to prevent such manipulations the rater's ratings must be imprecise.

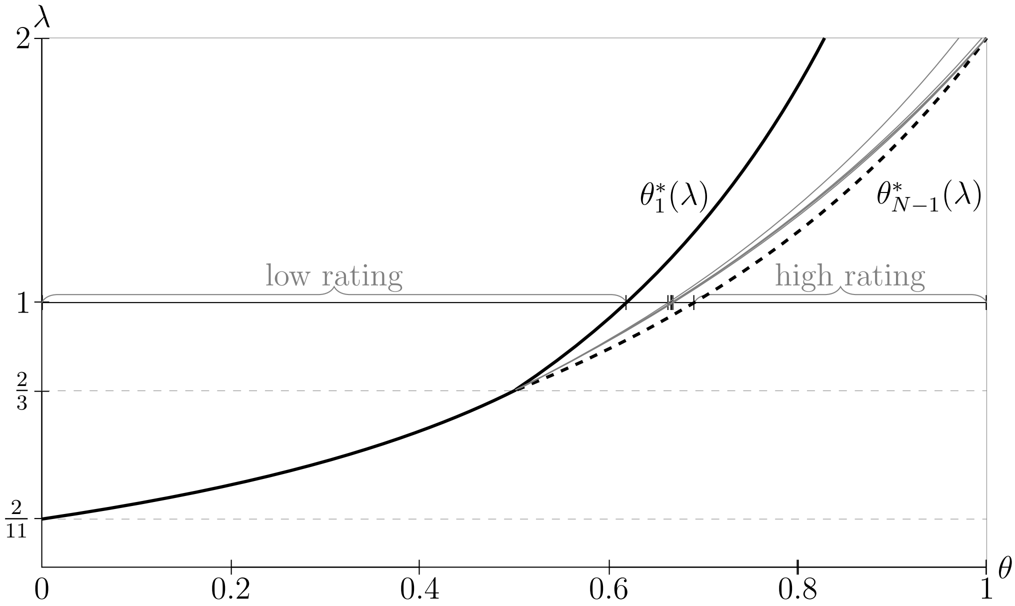
Our analysis points to possible improvements in regulation. We show that regulation requiring fixed fees is not desirable. A desirable regulation requires the rater to disclose the payments he receives from the seller.

In reality disclosure regimes differ. Food and Drug administration discloses the fees it charges to applicants. Underwriters of shares in initial public offerings are required to disclose their compensations in the prospectus. On the other hand, main Credit Rating Agencies do not disclose their fees.

Even when the payments are not publicly disclosed some parties know them and, probably, this information may reach some of the potential buyers. Intuitively, the less transparent are the payments between the seller and the rater the lower are the chances of buyers to learn the payments and, hence, the higher are the seller's incentives to manipulate the payments. Naturally, one can expected ratings with limited information value in markets where the buyers have few opportunities to learn about actual payments.

The paper highlighted the benefits associated with regulation mandating disclosure of payments. Clearly, disclosure may also have undesirable consequences. For example, it may limit competition and foster collusion among raters. Also, disclosure may make it hard for the raters to price discriminate among different sellers. Disclosure may also lower the rater's profit and reduce the rater's incentives to invest in information acquisition. These aspects of disclosure deserve a separate investigation.

Figure 1: The rating scheme with the highest feasible number of ratings given λ.



7 Appendix

Proof of lemma 1. Under $(M, t(.))$ in a Perfect Bayesian equilibrium for a given price p and realized beliefs $\mu(\theta|m)$ the buyers buy $\hat{q}(m,p) = \arg\max_{q\geq 0} \int_0^1 (\theta - p)q - q^\gamma/\gamma)\mu(\theta|m)d\theta$. Given the buyers' demand $\hat{q}(m,p)$ and realized beliefs $\mu(\theta|m)$ the seller sets $\hat{p}(m) = \arg\max_{p\geq 0} \int_0^1 \hat{q}(m,p)p\mu(\theta|m)d\theta$. Denote $E[\theta|m] = \int_0^1 \theta\mu(\theta|m)d\theta$. The buyers' problem has a unique solution $\hat{q}(m,p) = (E[\theta|m] - p)^{\frac{1}{\gamma-1}}$ if $p \leq E[\theta|m]$ and $\hat{q}(m,p) = 0$ otherwise. If $E[\theta|m] > 0$, the solution to the seller's problem is unique $\hat{p}(m) = \frac{\gamma-1}{\gamma}E[\theta|m]$. The equilibrium quantity is $q(m) = \hat{q}(m, \hat{p}(m)) = (E[\theta|m]/\gamma)^{\frac{1}{\gamma-1}}$. If $E[\theta|m] = 0$ any $p \in [0, \infty)$ can happen in equilibrium. In this case $q = 0$ and all equilibria deliver the same payoffs to the parties; we set $\hat{p}(m) = 0$ with no loss of generality. It follows that for each $m \in M$ the belief function $\mu(\theta|.)$ induces a unique market reaction $p(m)$, $q(m)$ and considering pure strategies is without loss of generality.

The rater maximizes his payoff given θ, $p(m)$, $q(m)$ and the buyers' beliefs function $\mu(\theta|.)$. Given θ the rater issues a rating m with positive probability $\sigma(m|\theta) \geq 0$ if $m(\theta) \in \arg\max_{m\in M} [\lambda S(\theta, q(m), p(m)) + t(m)]$, otherwise $\sigma(m|\theta) = 0$. Bayes' rule, when applicable, requires $\mu(\theta|m) \int_0^1 \sigma(m|\theta')dF(\theta') = \sigma(m|\theta)f(\theta)$.

The rest of the proof builds on Krishna and Morgan (2008). In equilibrium the quantity is weakly increasing in quality. For $\theta_2 > \theta_1$ take any $m_2 : \sigma(m_2|\theta_2) > 0$ and any $m_1 : \sigma(m_1|\theta_1) > 0$. The rater's revealed preference writes $\lambda S(\theta_2, q(m_2), p(m_2)) + t(m_2) \geq \lambda S(\theta_2, q(m_1), p(m_1)) + t(m_1)$, $\lambda S(\theta_1, q(m_1), p(m_1)) + t(m_1) \geq \lambda S(\theta_1, q(m_2), p(m_2)) + t(m_2)$. Combining inequalities and substituting for S we get $\lambda(q(m_2) - q(m_1))(\theta_2 - \theta_1) \geq 0$ and $q(m_2) \geq q(m_1)$.

For every $\theta \in [0, 1]$ the set of quantities implemented with positive probability $q(\theta) =$

$\{q(m) : \sigma(m,\theta) > 0\}$ contains at most three distinct elements. Suppose for some $\tilde{\theta}$ the set $q(\tilde{\theta})$ contains four elements $\underline{q} < q_1 < q_2 < \overline{q}$. For any $m_1 : q(m_1) = q_1$ one must have $\sigma(m_1,\theta') = 0$ if $\theta \neq \tilde{\theta}$. Indeed if $\sigma(m_1,\theta) > 0$ for some $\theta < \tilde{\theta}$ one would have $q(\theta) > \underline{q} \in q(\tilde{\theta})$ which contradicts the weakly increasing quantity. If $\sigma(m_1,\theta) > 0$ for some $\theta > \tilde{\theta}$ one would have $q(\theta) < \overline{q} \in q(\tilde{\theta})$ also a contradiction. Therefore $\sigma(m_1,\theta) > 0$ iff $\theta = \tilde{\theta}$, the Bayes' rule implies $E(\theta|m_1) = \tilde{\theta}$ and $q_1 = q(m_1) = (\tilde{\theta}/\gamma)^{\frac{1}{\gamma-1}}$. By analogous reasoning for any $m_2 : q(m_2) = q_2$ one obtains $q_2 = q(m_2) = (\tilde{\theta}/\gamma)^{\frac{1}{\gamma-1}} = q_1$, which contradicts $q_2 > q_1$. Hence $q(\theta)$ contains at most three elements.

For almost every $\theta \in [0,1]$ the set of $q(\theta)$ is single valued, that is randomization could happen at points that have measure zero. In equilibrium quantity weakly increases with θ and $\overline{q}(\theta) = \sup\{q(\theta)\}$ is monotone. Since $q(\theta)$ contains at most three elements $\overline{q} \in q(\theta)$. At any point θ where $q(\theta)$ is not single valued function $\overline{q}(\theta)$ is discontinuous (for any $\theta' < \theta$ and $\theta'' > \theta$ one has $\overline{q}(\theta') \leq \inf\{q(\theta)\} < \overline{q}(\theta) \leq \inf\{q(\theta'')\}$). A monotone $\overline{q}(\theta)$ can be discontinuous only at a countable number of points: for almost every point $q(\theta)$ is single valued.

In an equilibrium denote by Θ_d the set of θs where $q(\theta)$ and $\overline{q}(\theta)$ differ. For each $\theta \in \Theta_d$ find m such that $\sigma(m,\theta) > 0$ and $q(m) = \overline{q}(\theta)$, set $\sigma'(m,\theta) = 1$. Set $\sigma'(m,\theta) = \sigma(m,\theta)$ for $\theta \notin \Theta_d$ so that the modified reporting strategy would implement $\overline{q}(\theta)$ for any $\theta \in [0,1]$. $\sigma'(m,\theta)$ is consistent with the rater's incentives if the meanings of messages do not change, that is for any m such that $\sigma(\theta|m) > 0$ for some $\theta \in \Theta_d$ we have $E(\theta|m) = \int\limits_0^1 \theta\mu'(\theta|m)d\theta = \int\limits_0^1 \theta\mu(\theta|m)d\theta$.

If m is issued for a single point $\theta \in \Theta_d$, that is $\sigma(\theta,m) > 0$ and $\sigma(x,m) = 0$ for $x \neq \theta$, then $\int\limits_0^1 x\mu(x|m)dx = \theta$. For all such m we set beliefs so that $\int\limits_0^1 x\mu'(x|m)dx = \theta$, these beliefs are consistent with σ'. If m is issued for $\theta \in \Theta_d$ and some $\theta' \neq \theta$, then $q = q(m)$ is implemented for an interval of θs because $q(\theta)$ is weakly monotone. Using the notation $\Theta_q = \{\theta : \overline{q}(\theta) = q\}$ one can show $\int\limits_0^1 x\mu(x|m)dx = E_F[x|x \in \Theta_q]$. It implies that if we change the beliefs for a single point $\theta \in \Theta_d$ the expectation would not change: $\int\limits_0^1 x\mu'(x|m)dx = \int\limits_0^1 x\mu(x|m)dx$. It follows that for any equilibrium which implements $q(\theta)$ we can construct an equilibrium which implements $\overline{q}(\theta)$.

Let Q be the set of qs such that $\overline{q}(\theta) = q$ for some $\theta \in [0, 1]$. For every $q \in Q$ define a set of qualities that induce the same $\overline{q}(\theta)$, that is $\Theta_q = \{\theta : \overline{q}(\theta) = q\}$. Since $\overline{q}(\theta)$ is monotone each $\Theta_q \subseteq [0, 1]$ is either an interval or a point, moreover $\underset{q\in Q}{\cup}\Theta_q = [0, 1]$ and $\Theta_q \cap \Theta_{q'} = \emptyset$ if $q' \neq q$. For each $q \in Q$ define $r_q = E_F[\theta|\theta \in \Theta_q]$, then $G = \{r_q\}_{q\in Q}$ is a grading by construction. For each $q \in Q$ pick a message m_q which induces q in equilibrium: $q(m_q) = q$ and let $t'(r_q) = t(m_q)$.

It is immediate to see that for an equilibrium under $\{M, t(.)\}$ which implements $\overline{q}(\theta)$ there exists an equilibrium in pure strategies under $\{G, t'(.)\}$ which also implements $\overline{q}(\theta)$.

The market reaction conforms to relation $p(\theta) = (\gamma - 1)q(\theta)^{\gamma-1}$. Therefore, given θ functions $q(\theta)$ and $t'(r(\theta))$ pin down the parties' payoffs $S(\theta, q, p) = \theta q(\theta) - q(\theta)^\gamma/\gamma - p(\theta)q(\theta) + t'(r(\theta))$ and $\pi(q, p) = p(\theta)q(\theta) - t'(r(\theta))$. In any equilibrium $q(\theta)$ and $\overline{q}(\theta)$, coincide for almost every point. Payments $t'(r(\theta))$ and $t(m(\theta))$ also coincide for almost every point. Thus for any equilibrium under $M, t(.)$ there exists an outcome equivalent equilibrium in pure strategies under $G, t'(.)$ QED.

Proof of proposition 1. Take two consecutive ratings $r_{i-1} = E_F[\theta|\theta \in [\theta_{i-2}, \theta_{i-1})]$ and $r_i = E_F[\theta|\theta \in [\theta_{i-1}, \theta_i)]$ (considering intervals is without loss of generality). The incentive compatibility and truthful reports (3), (4) require $\lambda S(\theta_{i-1}, r_{i-1}) + t(r_{i-1}) = \lambda S(\theta_{i-1}, r_i) + t(r_i)$. Using (6) we express $S(\theta, r) = (\theta - \frac{\gamma^2-\gamma+1}{\gamma^2}r))r^{\frac{1}{\gamma-1}}\gamma^{\frac{-1}{\gamma-1}}$ and we get $t(r_i) - t(r_{i-1}) =$

$\lambda\frac{\gamma^2-\gamma+1}{\gamma^2}(r_i^{\frac{\gamma}{\gamma-1}} - r_{i-1}^{\frac{\gamma}{\gamma-1}})\gamma^{\frac{-1}{\gamma-1}} - \lambda\theta_{i-1}(r_i^{\frac{1}{\gamma-1}} - r_{i-1}^{\frac{1}{\gamma-1}})\gamma^{\frac{-1}{\gamma-1}}$, $i = 2, ..., N$. Observe that $r_i{}^{\frac{\gamma}{\gamma-1}} - r_{i-1}{}^{\frac{\gamma}{\gamma-1}} = (r_i + r_{i-1})(r_i{}^{\frac{1}{\gamma-1}} - r_{i-1}{}^{\frac{1}{\gamma-1}}) + r_i r_{i-1}{}^{\frac{1}{\gamma-1}} - r_{i-1} r_i{}^{\frac{1}{\gamma-1}} \geq (r_i + r_{i-1})(r_i{}^{\frac{1}{\gamma-1}} - r_{i-1}{}^{\frac{1}{\gamma-1}})$ for $\gamma \geq 2$ since $r_{i-1} < r_i$. Thus we get $t(r_i) - t(r_{i-1}) \geq \lambda(r_i{}^{\frac{1}{\gamma-1}} - r_{i-1}{}^{\frac{1}{\gamma-1}})(\frac{\gamma^2-\gamma+1}{\gamma^2}(r_i + r_{i-1}) - \theta_{i-1})\gamma^{\frac{-1}{\gamma-1}} \geq \lambda(r_i{}^{\frac{1}{\gamma-1}} - r_{i-1}{}^{\frac{1}{\gamma-1}})(\frac{\gamma^2-\gamma+1}{\gamma^2} r_{i-1} - \frac{\gamma-1}{\gamma^2}\theta_{i-1})\gamma^{\frac{-1}{\gamma-1}}$ since $r_i \geq \theta_{i-1}$. Moreover for $F(.) = U[0,1]$ (4) implies $r_{i-1} = \frac{\theta_{i-1}+\theta_{i-2}}{2}$ therefore $(\gamma^2 - \gamma + 1)r_{i-1} - (\gamma - 1)\theta_{i-1} \geq \frac{\gamma^2-3\gamma+3}{2}\theta_{i-1} > 0$ since $\gamma \geq 2$. Thus $t(r_i) > t(r_{i-1})$ QED.

Proof of proposition 2. Using the rater's indirect utility $U(\theta) = \underset{r\in G}{max}[\lambda S(\theta, q(r), p(r)) + t(r)]$ and the envelope theorem

$\frac{dU(\theta)}{d\theta} = S_\theta(\theta, q(r), p(r)) = \lambda q(r) \geq 0$ the seller's expected profit can be expressed as $\Pi = \int_0^1 \tilde{\pi}(\theta, r))dF(\theta) - U(0)$, $\tilde{\pi}(\theta, r) = p(r)q(r) + \lambda S(\theta, q(r), p(r)) - \lambda q(r)\frac{1-F(\theta)}{f(\theta)}$. Since $\frac{dU(\theta)}{d\theta} \geq 0$ the incentive constraint (3) holds. Due to proposition 1 the constraint (5) can be replaced with $t(r_1) \geq 0 \Leftrightarrow U(0) \geq \lambda S(0, q(r_1), p(r_1))$. The seller's problem is

$$\max_{\{r(.),U(0)\}} \Pi,\ s.t.\ (1),(2),(4),\ U(0) \geq \lambda S(0, q(r_1), p(r_1)). \tag{12}$$

First, solution has $U(0) = \lambda S(0, q(r_1), p(r_1)) \Leftrightarrow t(r_1) = 0$. Second, $F(.) = U[0,1]$, (1), (2) for any r imply $E_F[\tilde{\pi}(\theta, r), \theta \in g_r)] = \tilde{\pi}(r, r) = p(r)q(r) + \lambda S(r, q(r), p(r)) - \lambda q(r)(1-r)$ is convex for $r \in (0, 1]$. Indeed denote $\beta = \frac{\gamma-1}{\gamma} + \lambda\frac{\gamma^2+\gamma-1}{\gamma^2}$, then $\tilde{\pi}(r, r) = (\beta r - \lambda)r^{\frac{1}{\gamma-1}}\gamma^{\frac{-1}{\gamma-1}}$

$\frac{d^2\tilde{\pi}(r,r)}{dr^2} = (\gamma\beta r - (2-\gamma)\lambda)r^{\frac{3-2\gamma}{\gamma-1}}\gamma^{\frac{-1}{\gamma-1}}\frac{1}{(\gamma-1)^2} > 0$ for $r \in (0, 1]$ since $\gamma \geq 2$. Maximize $\Pi = \sum_{r\in G}\int_{\theta\in g_r} \tilde{\pi}(r, r)dF(\theta) - \lambda S(0, q(r_1), p(r_1))$ subject to (4). Given that $\tilde{\pi}(r, r)$ is convex the solution to the last problem has at most one coarse rating r_1 corresponding to $[0, \theta_1]$. Indeed otherwise replacing an extra coarse rating $r' \neq r_1$ with an interval of perfect revelation $r = \theta \in G_{r'}$ would raise Π. (4) requires $r_1 = \frac{\theta_1}{2}$, $\theta \in [0, \theta_1]$ and $r = \theta$, $\theta \in (\theta_1, 1]$. Substitute $S(0, q(r_1), p(r_1)) = -\frac{1}{\gamma}q(r_1)^\gamma - p(r_1)q(r_1)$. Optimizing $\Pi(\theta_1) =$

$\tilde{\pi}(\frac{\theta_1}{2}, \frac{\theta_1}{2}))\theta_1 + \int\limits_{\theta_1}^{1} \tilde{\pi}(\theta, \theta) dF(\theta) + \lambda \frac{1}{\gamma} q(\frac{\theta_1}{2})^{\gamma} + \lambda p(\frac{\theta_1}{2}) q(\frac{\theta_1}{2})$ over $\theta_1 \in [0,1]$ we obtain $\theta^* = \min\left[\frac{\lambda}{\beta} \frac{\gamma(\gamma-1)2^{\gamma/(\gamma-1)} - (\gamma^2+\gamma-1)}{\gamma(\gamma-1)2^{\gamma/(\gamma-1)} - \gamma(2\gamma-1)}, 1\right] \geq 0$ since $\gamma \geq 2$.

Reporting $r(\theta) = \theta^*$ for $\theta \in [0, \theta^*]$ and $r(\theta) = \theta$ for $\theta \in (\theta^*, 1]$ is a unique solution. Of course there exists an equivalent solution with a pooling interval $[0, \theta^*)$; the expected parties' payoffs are the same in both cases and we refer to them as a single solution. The corresponding grading is $G = \{\frac{\theta^*}{2} \cup (\theta^*, 1]\}$; condition (3) pins down the payment $t(r(\theta)) = \int\limits_{0}^{\theta} \lambda q(r(x)) dx - \lambda S(\theta, q(r(\theta), p(r(\theta))) + \lambda S(0, q(\frac{\theta^*}{2}), p(\frac{\theta^*}{2}))$, $\theta \in [0,1]$ QED.

Proof of proposition 3.

Perfect ratings are not possible (see the argument in the text). Without loss of generality each rating $r_i \in G$, $i = 1, ..., N-1$ corresponds to a grade $[\theta_{i-1}, \theta_i)$, $\theta_{i-1} < \theta_i$, r_N corresponds to $[\theta_{N-1}, 1]$. Denote r_1 the lowest rating in G. Due to proposition 1 constraint (5) in problem (8) can be replaced with $t(r_1) \geq 0 \Leftrightarrow U(0) \geq \lambda S(0, q(r_1), p(r_1))$. Problem (8) is equivalent to

$$\max_{\{r(.), U(0)\}} \int\limits_{0}^{1} \tilde{\pi}(\theta, r) d\theta - U(0), \ s.t. \ (1), (2), (3), \ U(0) \geq \lambda S(0, q(r_1), p(r_1)). \tag{13}$$

Here $\tilde{\pi}(\theta, r) = p(r)q(r) + \lambda S(\theta, q(r), p(r)) - \lambda q(r)(1-\theta)$. A unique solution to this problem is $U(0) = \lambda S(0, q(r_1), p(r_1)) \Leftrightarrow t(r_1) = 0$, and an increasing function $r^*(\theta) = \arg\max_{\{r \in G\}} [\tilde{\pi}(\theta, r)]$, $\forall \theta \in [0,1]$ since $\frac{\partial^2 \tilde{\pi}}{\partial \theta \partial r} = 2\lambda \frac{\partial q}{\partial r} > 0$. Border points between any two ratings issued with positive probability r_{i-1}, r_i are given by $\tilde{\pi}(\theta^*_{i-1}, r_{i-1}) = \tilde{\pi}(\theta^*_{i-1}, r_i)$, $i = 2, ..., N$.

Condition (4) requires $\theta^*_{i-1} = \theta_{i-1}$, $i = 2, ..., N$. Hence we obtain (9): $\tilde{\pi}(\theta_{i-1}, r_{i-1}) = \tilde{\pi}(\theta_{i-1}, r_i)$, $i = 2, ..., N$.

Given that $U_{\theta r} = \frac{\partial q}{\partial r} > 0$ and $r^*(\theta)$ is increasing from $U(\theta) = \underset{r \in G}{max}[\lambda S(\theta, q(r), p(r)) + t(r)]$ and $U(0) = \lambda S(0, q(r_1), p(r_1)) \Leftrightarrow t(r_1) = 0$ we can find $t(.)$ satisfying (3) QED.

Proof of corollary 2. For $\gamma = 2$ and $F(.) = U[0,1]$ condition (9) becomes $(3\lambda - 2)\theta_i - 2(5\lambda+2)\theta_{i-1} + (3\lambda-2)\theta_{i-2} = -8\lambda$, $i = 2, ..., N$. Since $\theta_0 = 0$, $\theta_N = 1$ a solution with $N = 2$ exists iff $\lambda \in (\frac{2}{11}, 6)$ and is characterized by $\theta_1 = \frac{11\lambda-2}{10\lambda+4}$. Suppose $N \geq 3$ and take the first difference $(3\lambda-2)d\theta_i - 2(5\lambda+2)d\theta_{i-1} + (3\lambda-2)d\theta_{i-2} = 0$. One must have $d\theta_i \geq 0$ for $i = 1, .., N$, which is not possible if $\lambda \leq \frac{2}{3}$, therefore $N < 3$ for $\lambda \leq \frac{2}{3}$. Suppose $\lambda > \frac{2}{3}$, denote $x = \frac{5\lambda+2}{3\lambda-2}$, then $1 + x = \frac{8\lambda}{3\lambda-2}$ and we obtain a difference equation $d\theta_i - 2xd\theta_{i-1} + d\theta_{i-2} = 0$, $i = 3, ..., N$. Characteristic polynomial $D^2 - 2xD + 1 = 0$ delivers $D = x - \sqrt{x^2-1} < 1$ and $D' = 1/D$. Solution to the difference equation is $d\theta_i = AD^i + A'D^{-i}$, $i = 1, ..., N$. Conditions $\theta_0 = 0$, $\theta_N = 1$ require $(d\theta_2 + d\theta_1) - 2xd\theta_1 = -(1+x)$ and $1 - 2x(1 - d\theta_N) + (1 - d\theta_N - d\theta_{N-1}) = -(1+x)$ correspondingly. Substituting for $d\theta_1$, $d\theta_2$, $d\theta_{N-1}$ and $d\theta_N$ we obtain $A' = D(A - \frac{1+x}{1-D})$, $A' = D(AD^{2N} + \frac{x-3}{1-D}D^N)$ which implies $A = \frac{(1+x)+(x-3)D^N}{(1-D)(1-D^{2N})}$, $A' = D\frac{(x-3)D^N+(1+x)D^{2N}}{(1-D)(1-D^{2N})}$ after substitutions we get $d\theta_i = \frac{(8-4\lambda+8\lambda D^N)D^{N+1-i}+(8\lambda+(8-4\lambda)D^N)D^i}{(3\lambda-2)(1-D)(1-D^{2N})}$ and $\theta_i = \sum_{j=1}^{i} d\theta_j$, $i = 1, ..., N$. Partition with N ratings is admissible iff $d\theta_i \geq 0$, $i = 1, ..., N$ that is if $(8-4\lambda)D^N + 8\lambda\frac{D^{2N+1}+D^{2i}}{D+D^{2i}} \geq 0$ for any $i = 1, ..., N$, which is equivalent to $8 - 4\lambda + 8\lambda\frac{D^N(1+D)}{D+D^{2N}} \geq 0$. If $\lambda \in (\frac{2}{3}, 2]$ any $N = 2, ..., \infty$ is admissible, if $\lambda > 2$ then the highest admissible $N^* < \infty$ satisfies $8 - 4\lambda + 8\lambda\frac{D^{N^*}(1+D)}{D+D^{2N^*}} \geq 0$ and $8 - 4\lambda + 8\lambda\frac{D^{1+N^*}(1+D)}{D+D^{2(1+N^*)}} < 0$. As was shown before $N = 2$ is admissible only for $\lambda \in (\frac{2}{11}, 6)$, that is for $\lambda \geq 6$ and $\lambda \leq \frac{2}{11}$ the uninformative rating scheme prevails QED.

Proof of proposition 4. Due to proposition 3 a grading G has at most countable number of coarse ratings and, hence, the seller's virtual profit $\tilde{\pi}(\theta) \equiv \tilde{\pi}(\theta, r(\theta))$ is piecewise differentiable on $[0,1]$. Using $F(.) = U[0,1]$ we get $\tilde{\pi}'(\theta) = 2\lambda q(r(\theta))$. Due to (9) $\tilde{\pi}(\theta)$ is continuous at border points θ_i, $i = 1, ..., N-1$. Therefore we express $\Pi = p(r_1)q(r_1) -$

$\lambda q(r_1) + \int\limits_0^1 2\lambda q(r(\theta))(1-\theta)d\theta$. Denote by $\hat{\pi}(r) = 2\lambda(1-r)(\frac{r}{\gamma})^{\frac{1}{\gamma-1}}$ the modified virtual profit. Since $q(r) = (\frac{r}{\gamma})^{\frac{1}{\gamma-1}}$ and any $r_i = E[\theta|\theta \in [\theta_{i-1}, \theta_i)]$ we obtain $\Pi = p(r_1)q(r_1) - \lambda q(r_1) + \sum\limits_{i=1}^{N} \int\limits_{\theta_{i-1}}^{\theta_i} \hat{\pi}(r(\theta))d\theta$. Function $\hat{\pi}(r)$ is concave, hence $\sum\limits_{i=1}^{N} \int\limits_{\theta_{i-1}}^{\theta_i} \hat{\pi}(r(\theta))d\theta \leq 2\lambda \int\limits_0^{\theta_1}(1-\theta)(\frac{\theta_1}{2\gamma})^{\frac{1}{\gamma-1}}d\theta + 2\lambda\int\limits_{\theta_1}^{1}(1-\theta)(\frac{1+\theta_1}{2\gamma})^{\frac{1}{\gamma-1}}d\theta$.

Substitute $\alpha = \frac{1}{\gamma-1} \in (0,1]$ and get $\Pi \leq \overline{\Pi}(\theta_1) = \frac{1}{1+\alpha}(\frac{\theta_1}{2})^{1+\alpha}\gamma^{-\alpha} + \lambda(1-\theta_1)^2[(\frac{1+\theta_1}{2})^{\alpha} - (\frac{\theta_1}{2})^{\alpha}]\gamma^{-\alpha}$. Function $\overline{\Pi}(y)$ is convex for $y \in (0,1)$ because $\frac{\partial^2(1-y)^2((1+y)^{\alpha}-y^{\alpha})}{\partial y^2} = (1+y)^{\alpha} - y^{\alpha} + 2(1-y)\alpha(y^{\alpha-1} - (1+y)^{\alpha-1}) + (1-y)^2\alpha(1-\alpha)(y^{\alpha-2} - (1+y)^{\alpha-2}) > 0$.

Let's show that for any admissible $\theta_1 < 1$ we have $\overline{\Pi}(\theta_1) < \overline{\Pi}(1)$. Let x be the smallest admissible $\theta_1 > 0$. Function $\overline{\Pi}(\theta_1)$ is convex therefore it suffices to show that $\overline{\Pi}(x) < \overline{\Pi}(1)$ whenever $x < 1$.

$\overline{\Pi}(1) - \overline{\Pi}(x) = \frac{\gamma^{-\alpha}}{(1+\alpha)2^{1+\alpha}}[(1-x^{1+\alpha})) - 2(1+\alpha)\lambda(1-x)^2((1+x)^{\alpha} - x^{\alpha})]$.

From $(1+x)^{\alpha} \leq 1+\alpha x$ for $x \in [0,1]$ follows $(1+x)^{\alpha} - x^{\alpha} \geq 1 + \alpha x - x^{\alpha}$ and $(1-x^2)((1+x)^{\alpha} - x^{\alpha}) \leq (1-x^2)(1+\alpha x - x^{\alpha}) \leq 1 - x^{1+\alpha}$. Indeed the last inequality is equivalent to $(x^{\alpha} - \alpha x)(1-x) + (1-\alpha-x^{\alpha}+\alpha x)x^2 \geq 0$ which holds because $(x^{\alpha} - \alpha x) \in [0, 1-\alpha]$ for $x \in [0,1]$, $\alpha \leq 1$. In turn $1 - x^{1+\alpha} \geq (1-x^2)(1+\alpha x - x^{\alpha})$ implies $\overline{\Pi}(1) - \overline{\Pi}(x) \geq \frac{\gamma^{-\alpha}}{(1+\alpha)2^{1+\alpha}}(1-x^2)(1+\alpha x - x^{\alpha})(1 - 2\lambda(1+\alpha)\frac{1-x}{1+x}) \geq 0$ for $x \geq \underline{x} = \frac{2\lambda(1+\alpha)-1}{2\lambda(1+\alpha)+1}$.

Denote $z = \frac{1}{(1+\alpha)\lambda} - \frac{1+\alpha+\alpha^2}{(1+\alpha)^2}$, condition (9) writes

$$z(r_i^{1+\alpha} - r_{i-1}^{1+\alpha}) = (1-2\theta_{i-1})(r_i^{\alpha} - r_{i-1}^{\alpha}),\ i = 2, ..., N. \tag{14}$$

Consider $\lambda > \frac{1+\alpha}{1+\alpha+\alpha^2} \Leftrightarrow z < 0$. Given that $\frac{(\theta_2+\theta_1)^{1+\alpha} - \theta_1^{1+\alpha}}{(\theta_2+\theta_1)^{\alpha} - {\theta_1}^{\alpha}} \geq (\theta_2+\theta_1)+\theta_1$. From (14) we get $1 - 2\theta_1 \leq z(\theta_1 + \frac{1}{2}\theta_2)$. Since $\theta_2 \geq \theta_1$ we get $\theta_1 \geq x \geq \overline{x} = \frac{2}{4+3z}$. To prove $\overline{\Pi}(1) \geq \overline{\Pi}(x)$ it suffices to show $\overline{x} \geq \underline{x} \Leftrightarrow 2(1+\alpha^3)\lambda^2 - 3(1+\alpha+\alpha^2)\lambda + 3 + 3\alpha \geq 0$ which holds for any

λ because the discriminant $-15(1-\alpha^2+\alpha^4)-6\alpha(1-\alpha)^2<0$. Therefore, $\overline{\Pi}(x)\leq\overline{\Pi}(1)$ and $\overline{\Pi}(x)<\overline{\Pi}(1)$ if $x<1$.

Consider $\lambda\leq\frac{1+\alpha}{1+\alpha+\alpha^2}$, then $z\geq 0$ and only $N\leq 2$ is admissible. Suppose $N\geq 3$, then for any three consecutive ratings condition (14) requires $z\left[\frac{r_i^{1+\alpha}-r_{i-1}^{1+\alpha}}{r_i^{\alpha}-r_{i-1}^{\alpha}}-\frac{r_{i-1}^{1+\alpha}-r_{i-2}^{1+\alpha}}{r_{i-1}^{\alpha}-r_{i-2}^{\alpha}}\right]=-2\lambda(\theta_{i-1}-\theta_{i-2})<0$ which is impossible because the left hand side is not negative. The left hand side is proportional to $h=r_i^{1+\alpha}r_{i-1}^{\alpha}-r_i^{1+\alpha}r_{i-2}^{\alpha}+r_{i-1}^{1+\alpha}r_{i-2}^{\alpha}-r_i^{\alpha}r_{i-1}^{1+\alpha}+r_i^{\alpha}r_{i-2}^{1+\alpha}-r_{i-1}^{\alpha}r_{i-2}^{1+\alpha}\geq 0$. To see this compute $\frac{\partial^2 h}{\partial r_{i-1}^2}=-\alpha(1-\alpha)r_{i-1}^{\alpha-2}(r_i^{1+\alpha}-r_{i-2}^{1+\alpha})-\alpha(1+\alpha)r_{i-1}^{\alpha-1}(r_{i-1}^{\alpha}-r_{i-2}^{\alpha})<0$ since $\alpha\leq 1$ and $r_{i-2}<r_{i-1}<r_i$. Given that h is concave in r_{i-1} and $h(r_{i-1}=r_{i-2})=h(r_{i-1}=r_i)=0$, we have $h(r_{i-1}\in(r_{i-2},r_i))\geq 0$.

For $N=2$ (14) implies $\theta_1\leq\frac{1}{2}$ and $\lambda(\theta_1)=\frac{1}{2(1+\alpha)}\frac{(1+\theta_1)^{1+\alpha}-\theta_1^{1+\alpha}}{(1-2\theta_1)((1+\theta_1)^{\alpha}-\theta_1^{1+\alpha})+\frac{\kappa}{2}((1+\theta_1)^{1+\alpha}-\theta_1^{1+\alpha})}$, here $\kappa=\frac{1+\alpha+\alpha^2}{(1+\alpha)^2}$. Substituting for $\lambda(\theta_1)$ in $\overline{\Pi}(\theta_1)$ we find that $\overline{\Pi}(\theta_1)<\overline{\Pi}(1)$ whenever

$$((1+\theta_1)^{1+\alpha}-\theta_1^{1+\alpha})(1-\theta_1)^2<(1-\theta_1^{1+\alpha})\left[(1-2\theta_1)+\frac{\kappa}{2}\frac{(1+\theta_1)^{1+\alpha}-\theta_1^{1+\alpha}}{(1+\theta_1)^{\alpha}-\theta_1^{\alpha}}\right].$$

Provided that $\frac{(1+\theta_1)^{1+\alpha}-\theta_1^{1+\alpha}}{(1+\theta_1)^{\alpha}-\theta_1^{\alpha}}\geq 1+2\theta_1$ and $\kappa=\frac{1+\alpha+\alpha^2}{(1+\alpha)^2}\geq\frac{3}{4}$ a sufficient condition for $\overline{\Pi}(\theta_1)<\overline{\Pi}(1)$ is $((1+\theta_1)^{1+\alpha}-\theta_1^{1+\alpha})(1-\theta_1)^2<(1-\theta_1^{1+\alpha})(\frac{11}{8}-\frac{5}{4}\theta_1)$. The condition holds for $\theta_1=0$. For $\theta_1>0$ denote $y=\frac{1}{\theta_1}\in[2,\infty)$, an equivalent sufficient condition is $A(y,\alpha)=(y^{1+\alpha}-1)(11y-10)y-8((1+y)^{1+\alpha}-1)(y-1)^2>0$.
Take $\frac{\partial A}{\partial\alpha}=-\ln(1+y)\left[8(y-1)^2(1+y)^{1+\alpha}-\ln(\frac{y}{1+y})(11y-10)y^{2+\alpha}\right]<0$ for $y\geq 2$. Given that $\alpha\in(0,1]$, a stronger sufficient condition for $\overline{\Pi}(\theta_1)<\overline{\Pi}(1)$ is $A(y,1)>0$. $A(y,1)=y^2(3y^2-10y+10)+3y(y-2)>0$ for $y\in[2,\infty)$ therefore $\overline{\Pi}(\theta_1)<\overline{\Pi}(1)$ for $\lambda\leq\frac{1+\alpha}{1+\alpha+\alpha^2}$ if $\theta_1<1$.

We have shown that under a private payment any feasible rating scheme with $N\geq 2$, $\theta_1<1$ delivers a lower expected profit than the uninformative rating scheme with $N=1$,

$\theta_1 = 1$: $\Pi(\theta_1) \leq \overline{\Pi}(\theta_1) < \overline{\Pi}(1) = \Pi(1)$. Consequently, $N = 1$, $\theta_1 = 1$, $G = E[\theta] = \frac{1}{2}$, $t = 0$ is the optimal rating scheme QED.

Proof of proposition 6. An optimal rating scheme has a coarse rating $r_1 = \frac{\theta^*}{2}$ with $t = 0$ for $\theta \in [0, \theta^*]$ and perfect ratings $r = \theta$ with $t(r) = \int_{\theta^*}^{r} \lambda q(x)dx - \lambda S(r, q(r), p(r)) + \lambda S(\theta^*, q(r_1), p(r_1))$ for $\theta \in (\theta^*, 1]$. A seller of quality $\theta \in [0, \theta^*]$ is indifferent about applying for a rating because in any case her product is perceived to be of quality r_1. We assume she applies. A seller of quality $\theta \in (\theta^*, 1]$ will apply only if $p(\theta)q(\theta) - t(\theta) \geq p(r_1)q(r_1)$. Consider a seller $\theta \to \theta^*$, she applies iff $p(\theta^*)q(\theta^*) - p(\theta^*/2)q(\theta^*/2) \geq \lambda S(\theta^*, q(\theta^*/2), p(\theta^*/2)) - \lambda S(\theta^*, q(\theta^*), p(\theta^*))$ which is equivalent to

$\lambda \leq \lambda' = \frac{\gamma(\gamma-1)2^{\gamma/(\gamma-1)} - \gamma^2 + \gamma}{\gamma^2 + \gamma - 1 - (\gamma-1)2^{\gamma/(\gamma-1)}}$. The latter condition holds because

$\lambda < \lambda^* = \frac{\gamma(\gamma-1)2^{\gamma/(\gamma-1)} - 2\gamma^2 + \gamma}{\gamma^2 + \gamma - 1 - (\gamma-1)2^{\gamma/(\gamma-1)}} < \lambda'$. It remains to verify that sellers with $\theta > \theta^*$ also prefer to apply, which is guaranteed by $\frac{\partial(p(\theta)q(\theta) - t(\theta))}{\partial\theta} = (1 - \lambda\frac{\gamma-1}{\gamma})\theta^{\frac{1}{\gamma-1}}\gamma^{\frac{-1}{\gamma-1}} \geq 0$ for $\theta \in [\theta^*, 1]$. The latter condition holds because for $\gamma \geq 2$ we have $2^{\gamma/(\gamma-1)} \leq 2(1 + \frac{1}{\gamma-1})$ which implies $\lambda^* \leq \frac{\gamma}{\gamma^2 - \gamma - 1} \leq \frac{\gamma}{\gamma-1}$. It follows that it is an equilibrium strategy for any seller $\theta \in [0, 1]$ to apply for a rating. After a seller has applied the game proceeds as before and the players play the same equilibrium strategies. Thus, we have constructed an equilibrium which implements the optimal rating scheme $G, t(.)$ QED.

Proof of proposition 7. Denote $\Delta(G) = P(G) - P(\frac{1}{2}) \geq 0$. Given that $P(G) = \gamma S(G)$ we have $\Pi - \Pi_0 = \Delta(G) - T(G, t)$ and $U - U_0 = \lambda\Delta(G) + \gamma T(G, t)$. Optimal negotiated rating scheme must solve

$$\max_{\{G, t\ feasible\}} (\Delta(G) - T(G, t))^{\nu}(\lambda\Delta(G) + \gamma T(G, t))^{1-\nu}, \ s.t. \ \Delta(G) \geq T(G, t). \quad (15)$$

According to proposition 2 if $\lambda \geq \lambda^*$ the maximum for $\Pi = \Delta(G) - T(G, t)$ under a public contingent payment obtains in case of the uninformative rating with $t = 0$, which results in $\Pi = 0$. Consequently, the uninformative rating with $t = 0$ is the solution to (15) for $\lambda \geq \lambda^*$.

Consider $\lambda < \lambda^*$. Denote by T^* the expected payment under the optimal rating scheme preferred by the seller $\nu = 1$. Naturally in an optimal negotiated rating scheme $T(G, t) \geq T^*$.

The highest $\Delta(G)$ corresponds to perfectly informative ratings $G = [0, 1]$, denote it $\overline{\Delta} = \frac{\gamma-1}{\gamma}(\frac{\gamma-1}{2\gamma-1} - 2^{\frac{-\gamma}{\gamma-1}})\gamma^{\frac{-1}{\gamma-1}}$. Compute a minimal expected transfer compatible with a fully informative rating scheme $\overline{T} = \lambda\frac{(\gamma-1)^3}{\gamma^2(2\gamma-1)}\gamma^{\frac{-1}{\gamma-1}}$. A solution to (15) results in $T(G, t) \in [T^*, \overline{\Delta}]$. Note that $\overline{\Delta} \geq \overline{T} \Leftrightarrow \lambda \leq \lambda_\Delta = \frac{\gamma}{\gamma-1}(1 - \frac{2\gamma-1}{\gamma-1}2^{\frac{-\gamma}{\gamma-1}})$. One can check that $\lambda_\Delta < \lambda^*$.

Define $\Delta(T) = \max\limits_{\{G,t\ feasible\}} \Delta(G)$, s.t.: $T(G, t) = T$ for $T \geq T^*$. The Lagrangian is $L = \Delta(G) - \eta(T(G, t) - T)$. Any G feasible for T is also feasible for $T' > T$ hence $\eta \geq 0$. A problem $\max\limits_{\{G,t\ feasible\}} \Delta(G) - \eta T(G, t)$ is equivalent to problem (7) with $\lambda' = \lambda\eta$. This is obvious for $\eta = 1$. In general λ only enters (3) hence the above problem with η, λ is equivalent to a problem with $\eta' = 1$, $\lambda' = \eta\lambda$ and, therefore, is equivalent to (7) with λ'.

A solution with $\eta = 0$ leads to a fully informative rating scheme $\Delta(T) = \overline{\Delta}$, it happens for $T \geq \overline{T}$. Consider $T < \overline{T}$. A solution with $\eta > 0$ is given by proposition 2. Note that $\eta\lambda \geq \lambda^*$ is not possible, because in this case G is uninformative and $t = T = 0$, which contradicts $T \geq T^* > 0$. Consider $\eta\lambda \in (0, \lambda^*)$. G and t are fully characterized by a threshold $\theta^*(\eta\lambda) = \frac{\gamma^2\eta\lambda}{\gamma(\gamma-1)+\eta\lambda(\gamma^2+\gamma-1)} \frac{\gamma(\gamma-1)2^{\gamma/(\gamma-1)}-(\gamma^2+\gamma-1)}{\gamma(\gamma-1)2^{\gamma/(\gamma-1)}-\gamma(2\gamma-1)}$: a coarse rating and zero payment $t(\theta; \eta) = 0$ for $\theta \in [0, \theta^*(\eta\lambda)]$, and perfect ratings with positive payments $t(\theta; \eta) = t_0 + \lambda(\gamma - 1)\gamma^{-\frac{\gamma}{\gamma-1}} \int\limits_{\theta^*(\eta\lambda)}^{\theta} x^{\frac{1}{\gamma-1}} dx$ for $\theta \in (\theta^*(\eta\lambda), 1]$, here $t_0 = \frac{\lambda}{\gamma}\left(\frac{\theta^*(\eta\lambda)}{2\gamma}\right)^{\frac{\gamma}{\gamma-1}}(\gamma^2 + \gamma - 1 - 2^{\frac{\gamma}{\gamma-1}}(\gamma - 1)) > 0$. There is continuous correspondence between η and the expected payment T. Threshold $\theta^*(\eta\lambda)$ increases with η, hence, $T(\theta^*(\eta\lambda)) = T(\eta) = \int\limits_{\theta^*(\eta\lambda)}^{1} t(\theta;\eta)d\theta$ decreases with η. Indeed $\frac{\partial T(\theta^*)}{\partial\theta^*} = -t_0 + (1-\theta^*)[\frac{\partial t_0}{\partial\theta^*} - \lambda(\gamma-1)\gamma^{-\frac{\gamma}{\gamma-1}}\theta^{*\frac{1}{\gamma-1}}]$ and after substitutions $\frac{\partial T(\theta^*)}{\partial\theta^*} = -t_0 + (1-\theta^*)\lambda(2\gamma)^{-\frac{\gamma}{\gamma-1}}\theta^{*\frac{1}{\gamma-1}}[\frac{\gamma^2+\gamma-1}{\gamma-1} - \gamma 2^{\frac{\gamma}{\gamma-1}}] < 0$ because $t_0 > 0$, $2^{\frac{\gamma}{\gamma-1}}(\gamma-1) \geq 2\gamma-1$ and $2\gamma^2-\gamma \geq \gamma^2+\gamma-1$. This implies $\frac{\partial\eta(T)}{\partial T} < 0$ for $T \in [T^*, \overline{T}]$. Moreover $\eta(T^*) = 1$ and $\eta(T) = 0$ for $T \geq \overline{T}$. We obtain that function $\Delta(T) = \int\limits_0^T \eta(x)dx$ is concave for $T \geq T^*$ and $\Delta(T) = \overline{\Delta}$ for $T \geq \overline{T}$. Consider

$$\max_{\{T\geq T^*\}} (\Delta(T)-T)^\nu(\lambda\Delta(T)+\gamma T)^{1-\nu},\ s.t.\ \Delta(T) \geq T. \tag{16}$$

If $\nu = 1$ the solution is $T = T^*$. For $\nu < 1$ the first order condition is necessary and sufficient. The unique solution is given by

$$\frac{(1-\eta(T))T}{\Delta(T)-T} = \frac{\gamma+\lambda\eta(T)}{\nu(\lambda+\gamma)} - 1 \tag{17}$$

The left hand side is increasing in T because $\eta(T) \leq 1$ and $\Delta(T) - T \geq 0$ both decrease with T. The right hand side is decreasing in T and ν. It follows that $T(\nu)$ is decreasing. $\Delta(T)$ and $\eta(T)$ are differentiable, hence $T(\nu)' < 0$ and $\eta(T)' > 0$.

1) If $\overline{\Delta} < \overline{T} \Leftrightarrow \lambda > \lambda_\Delta$ then for any $\nu \in [0,1]$ we must have $T < \overline{T}$. In this case the solution is pinned down by $\eta(\nu) > 0$. The rating scheme leads to a coarse rating and zero payment $t(\theta;\eta(\nu)) = 0$ for $\theta \in [0, \theta^*(\eta(\nu)\lambda)]$, and perfect ratings with positive payments $t(\theta;\eta(\nu)) = t_0(\nu) + \lambda(\gamma-1)\gamma^{-\frac{\gamma}{\gamma-1}} \int\limits_{\theta^*(\eta(\nu)\lambda)}^{\theta} x^{\frac{1}{\gamma-1}}dx$ for $\theta \in (\theta^*(\eta(\nu)\lambda), 1]$, here $t_0(\nu) = \frac{\lambda}{\gamma}\left(\frac{\theta^*(\eta(\nu)\lambda)}{2\gamma}\right)^{\frac{\gamma}{\gamma-1}}(\gamma^2+\gamma-1-2^{\frac{\gamma}{\gamma-1}}(\gamma-1)) > 0$. $T(\nu)$ and $\eta(\nu)$ satisfy (17), $T(\nu) = \int\limits_{\theta^*(\eta(\nu)\lambda)}^{1} t(\theta;\eta(\nu))d\theta$, $\Delta(T(\nu)) = \frac{\gamma-1}{\gamma}\gamma^{\frac{-1}{\gamma-1}}(\frac{\gamma-1}{2\gamma-1} - 2^{\frac{-\gamma}{\gamma-1}})(1-\theta^*(\eta(\nu)\lambda)^{\frac{2\gamma-1}{\gamma-1}})$ and $\theta^*(\eta(\nu),\lambda) = \frac{\gamma^2\eta(\nu)\lambda}{\gamma(\gamma-1)+\eta(\nu)\lambda(\gamma^2+\gamma-1)}\frac{\gamma(\gamma-1)2^{\gamma/(\gamma-1)}-(\gamma^2+\gamma-1)}{\gamma(\gamma-1)2^{\gamma/(\gamma-1)}-\gamma(2\gamma-1)}$.

2) Consider $\overline{\Delta} \geq \overline{T} \Leftrightarrow \lambda \leq \lambda_\Delta$.

a) Solution with $\eta(\nu)) = 0$ requires $T \geq \overline{T}$, that is $\nu \leq \nu^*$. It results in perfectly revealing ratings and transfers $t(\theta;\nu) = t_0(\nu) + \lambda(\gamma-1)\gamma^{-\frac{\gamma}{\gamma-1}} \int\limits_0^\theta x^{\frac{1}{\gamma-1}} dx$ for $\theta \in [0,1]$. Transfers satisfy $t_0(\nu) = T(\nu) - \lambda(\gamma-1)\gamma^{-\frac{\gamma}{\gamma-1}} \int\limits_0^1 \int\limits_0^\theta x^{\frac{1}{\gamma-1}} dx d\theta$ and $T(\nu) = \overline{\Delta}(1-\nu^{\frac{\lambda+\gamma}{\lambda}})$, $\overline{\Delta} = \frac{\gamma-1}{\gamma}(\frac{\gamma-1}{2\gamma-1} - 2^{\frac{-\gamma}{\gamma-1}})\gamma^{\frac{-1}{\gamma-1}}$. From $T(\nu) \geq \overline{T}$ we get $\nu^* = (1-\frac{\overline{T}}{\overline{\Delta}})\frac{\lambda}{\lambda+\gamma}$. b) For $\nu > \nu^*$ one has $T < \overline{T}$ and $\eta(\nu) > 0$; the same formulas as in 1) deliver the solution.

We have characterized solution $T(\nu)$ to (16) for any $\nu \in [0,1]$ when $\lambda < \lambda^*$. The corresponding $G(\nu), t(\nu)$ is pinned down by $\eta(T(\nu))$. Denote the maximal value of the objective function in (16) by $\overline{\Psi}$. Denote the objective function in the initial problem (15) by $\Psi(G,t)$. On the set of feasible G, t such that $\Delta(G) \geq T(G,t)$ the value of this function is in the interval $[0, \overline{\Psi}]$, therefore its maximum is $\overline{\Psi}$. The rating scheme $G(\nu), t(\nu)$ solving (16) delivers this maximum. We say that $G(\nu), t(\nu)$ is the solution to (15) because any other rating scheme which delivers $\overline{\Psi}$ is outcome equivalent to $G(\nu), t(\nu)$ QED.

Proof of proposition 8. A rating scheme is feasible iff (1), (2), (3), (4), (5) hold. First, take any $t \geq 0$ to satisfy (5). Second, lemma 2 implies that under a fixed fee a feasible grading G cannot have intervals of perfect revelation $r = \theta$. However, stand alone perfect ratings $r = \theta$ are possible. These ratings $r = \theta$ correspond to a set of qualities of a zero measure, therefore any feasible grading G is equivalent to a partition $\{\theta_i\}_{i=0,\ldots,N}$ of interval $[0,1]$.

Consider an informative rating scheme $N \geq 2$, conditions (1), (2) deliver $q(r_i) = (\frac{r_i}{\gamma})^{\frac{1}{\gamma-1}}$, $p(r_i) = \frac{\gamma-1}{\gamma} r_i$, $r_i = \int_{\theta_{i-1}}^{\theta_i} \frac{xdF(x)}{F(\theta_i)-F(\theta_{i-1})}$, $i = 1, \ldots, N$. Condition (3) requires that at each point separating two ratings the rater is indifferent between the ratings:

$$\lambda S(\theta_{i-1}, q(r_i), p(r_i)) + t = \lambda S(\theta_{i-1}, q(r_{i-1}), p(r_{i-1})) + t, \ i = 2, .., N,$$

and $\theta_0 = 0$, $\theta_N = 1$. $S_{q\theta} > 0$, $q(r(\theta))$ is not decreasing, thus the above condition is also sufficient for (3). For $t = const$ it is equivalent to (11).

Substitute for $q(r)$ and $p(r)$ to obtain $S(\theta, r) = (\theta - \frac{\gamma^2-\gamma+1}{\gamma^2} r)) r^{\frac{1}{\gamma-1}} \gamma^{\frac{-1}{\gamma-1}}$. From $F(.) = U[0,1]$ we get $r_i = \frac{\theta_i+\theta_{i-1}}{2}$, $i = 2, ..., N$. Note that if all $\{\theta_i\}_{i=0,...,N}$ are multiplied by a positive scalar the above condition is not affected.

Remark 5. If a grading with $N \geq 2$ ratings is feasible, that is a sequence $\{\theta_i\}_{i=0,...,N}$ solves (11), then a sequence $\{\theta_i/\theta_{N-1}\}_{i=0,...,N-1}$ also solves (11) and, therefore defines a feasible grading with $N-1$ ratings.

For a feasible grading with N ratings take the highest rating r_N. Consider an equation for $\theta_N = 1$: $S(1, r_N) = S(1, y)$. If this equation has a solution $y_N > 1$, which implies $\tilde{\theta}_{N+1} = 2y_N - 1 > 1$, then one can construct a feasible grading with $N+1$ ratings with borders points $\{\theta_i/\tilde{\theta}_{N+1}\}_{i=0,...,N}$ and $\theta_{N+1} = 1$. Since S is homogeneous the new grading satisfies (11) by construction. Note $S_y(1, y) = (1 - y(\gamma - 1 + \frac{1}{\gamma})) \frac{1}{\gamma-1} y^{\frac{2-\gamma}{\gamma-1}} \gamma^{\frac{-1}{\gamma-1}} \gtreqless 0$ whenever $\frac{1}{\gamma-1+\frac{1}{\gamma}} \gtreqless y$, therefore $y^* = \frac{1}{\gamma-1+\frac{1}{\gamma}} < 1$ is the maximum of S. It follows that the solution y to the equation, if it exists, is unique. In other words a grading with N ratings can be extended to a grading with $N+1$ in a unique way. It follows that for each N there is at most one feasible grading.

Let $\bar{r} < y^*$: $S(1, \bar{r}) = S(1, 1)$ then equation $S(1, r_N) = S(1, y)$ has a solution $y_N > 1$ iff $r_N < \bar{r}$. The lowest value of r_N corresponds to uninformative rating scheme $r_N = \frac{1}{2}$, it follows that an informative rating scheme is feasible iff $\bar{r} > \frac{1}{2}$. Since $S(1, y)$ is increasing for $y < y^*$ this condition is equivalent to $S(1, \frac{1}{2}) < S(1, 1)$: $2^{\frac{1}{\gamma-1}} > \frac{1}{2}(2 + \gamma + \frac{1}{\gamma-1}) \Leftrightarrow \gamma < \gamma^* \approx 1.742$.

Suppose $r_N < \bar{r}$ (which implies $\gamma < \gamma^*$), let's prove $y_N - 1 < 1 - r_N$. It is enough to show

$S(1,r_N) > S(1,2-r_N)$ for $r_N < \overline{r}$ since $S(1,y)$ decreases for $y > 1$. $\frac{\partial S(1,2-r_N)}{\partial r_N} = -(1-(2-r_N)(\gamma-1+\frac{1}{\gamma}))\frac{1}{\gamma-1}(2-r_N)^{\frac{2-\gamma}{\gamma-1}}\gamma^{\frac{-1}{\gamma-1}} > -(1-(2-r_N)(\gamma-1+\frac{1}{\gamma}))\frac{1}{\gamma-1}r_N^{\frac{2-\gamma}{\gamma-1}}\gamma^{\frac{-1}{\gamma-1}}$ since $r_N < 1$, $\gamma < 2$ and $\gamma-1+\frac{1}{\gamma} > 1$. It follows $\frac{\partial(S(1,r_N)-S(1,2-r_N))}{\partial r_N} < 2(1-(\gamma-1+\frac{1}{\gamma}))\frac{1}{\gamma-1}r_N^{\frac{2-\gamma}{\gamma-1}}\gamma^{\frac{-1}{\gamma-1}} < 0$ and, since $S(1,r_N) - S(1,2-r_N) = 0$ for $r_N = 1$, we get $S(1,r_N) - S(1,2-r_N) > 0$ for $r_N < 1$.

Given that $y_N - 1 < 1 - r_N$ for any N such that solution $y_N > 1$ exists we must have $\theta_{i+1} - \theta_i < \theta_i - \theta_{i-1}$, $i = 1, ..., N-1$ in any feasible rating scheme. It follows that the maximum number of ratings $N \leq N(\gamma) < \infty$. Suppose $N \to \infty$ then $r_N \to 1$ and (11) implies $S(1, r_{N-1}/\theta_{N-1}) = S(1, r_N/\theta_{N-1})$. The latter equation has a solution $y = r_N/\theta_{N-1} > 1$ iff $r_{N-1}/\theta_{N-1} < \overline{r} < y^* = \frac{1}{\gamma-1+\frac{1}{\gamma}} < 1$, therefore $\theta_N - \theta_{N-1} = \theta_{N-1}2(1-\overline{r}) > 0$. This in turn implies $\theta_i - \theta_{i-1} > \theta_{N-1}2(1-\overline{r})$, $i = 1, ..., N-1$ which is impossible when $N \to \infty$ since $\sum\limits_{i=1}^{N}(\theta_i - \theta_{i-1}) = 1$, thus there exists $N(\gamma) < \infty$ such that only gradings with $N \leq N(\gamma)$ ratings are feasible QED.

Proof of proposition 9. Clearly $t = 0$ is optimal. For a feasible partition $\{\theta_i\}_{i=0,...,N}$ the seller's gross expected profit is $\Pi = \frac{\gamma-1}{\gamma}\sum\limits_{i=1}^{N}\left(\frac{\theta_{i-1}+\theta_i}{2}\right)^{\frac{\gamma}{\gamma-1}}(\theta_i - \theta_{i-1})\gamma^{\frac{-1}{\gamma-1}} = \frac{\gamma-1}{\gamma}\Sigma\gamma^{\frac{-1}{\gamma-1}}$. Since $\theta_0 = 0$ we rewrite $\Sigma = \sum\limits_{i=1}^{N-1}((\theta_{i-1}+\theta_i)^{\frac{\gamma}{\gamma-1}} - (\theta_{i+1}+\theta_i)^{\frac{\gamma}{\gamma-1}})\theta_i + \theta_N(\theta_N+\theta_{N-1})^{\frac{\gamma}{\gamma-1}} = \sum\limits_{i=1}^{N-1}((\theta_{i-1}+\theta_i)^{\frac{1}{\gamma-1}} - (\theta_{i+1}+\theta_i)^{\frac{1}{\gamma-1}})\theta_i{}^2 + \theta_N{}^2(\theta_N+\theta_{N-1})^{\frac{1}{\gamma-1}}$. Using $S(\theta,r) = (\theta-\kappa r))r^{\frac{1}{\gamma-1}}\gamma^{\frac{-1}{\gamma-1}}$, $\kappa = \frac{\gamma^2-\gamma+1}{\gamma^2}$ and (11) get $(\theta_{i-1}+\theta_i)^{\frac{\gamma}{\gamma-1}} - (\theta_{i+1}+\theta_i)^{\frac{\gamma}{\gamma-1}} = \frac{2}{\kappa}\left((\theta_{i-1}+\theta_i)^{\frac{1}{\gamma-1}} - (\theta_i+\theta_{i+1})^{\frac{1}{\gamma-1}}\right)\theta_i$ for $i = 1, ..., N-1$. Multiply by θ_i and derive

$\sum\limits_{i=1}^{N-1}\left((\theta_{i-1}+\theta_i)^{\frac{1}{\gamma-1}} - (\theta_i+\theta_{i+1})^{\frac{1}{\gamma-1}}\right)\theta_i^2 = -\frac{\kappa}{2-\kappa}\theta_N(\theta_N+\theta_{N-1})^{\frac{1}{\gamma-1}}\theta_{N-1}$ to obtain

$\Sigma = \theta_N(\theta_N+\theta_{N-1})^{\frac{1}{\gamma-1}}\left(\theta_N - \frac{\kappa}{2}(\theta_N+\theta_{N-1})\right)\frac{2}{2-\kappa}$. Substitute $\theta_N = 1$ and note $\Pi \sim \Sigma \sim S(1, \frac{1+\theta_{N-1}}{2})$. It has been shown in the proof of proposition 8 that $S(1, \frac{1+\theta_{N-1}}{2})$ increases with θ_{N-1}, and θ_{N-1} itself increases with N. A fortiori the seller's profit is the highest under the rating scheme with the maximal feasible number of ratings $N(\gamma) < \infty$ QED.

References

Bester, H. and R. Strausz (2001): "Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case," *Econometrica*, 69, 1077–1098.

Bolton, P., X. Freixas, and J. Shapiro (2012): "The Credit Ratings Game," *The Journal of Finance*, 67, 45–83.

Chen, H.-C. and J. R. Ritter (2000): "The Seven Percent Solution," *The Journal of Finance*, 55, 1105–1131.

Crawford, V. P. and J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50, 1431–1451.

Farhi, E., J. Lerner, and J. Tirole (2013): "Fear of Rejection? Tiered Certification and Transparency," *The RAND Journal of Economics*, forthcoming.

Farrell, J. and R. Gibbons (1989): "Cheap Talk with Two Audiences," *The American Economic Review*, 79, 1214–1223.

Goel, A. M. and A. V. Thakor (2013): "Why Are Credit Ratings Coarse?" *Working paper*.

Griffin, J. M. and D. Y. Tang (2012): "Did Subjectivity Play a Role in CDO Credit Ratings?" *The Journal of Finance*, 67, 1293–1328.

Inderst, R. and M. Ottaviani (2012): "Competition through Commissions and Kickbacks," *American Economic Review*, 102, 780–809.

International Monetary Fund (2009): "Global Financial Stability Report: Navigating the Financial Challenges Ahead," Tech. rep., World Economic and Financial Surveys (Washington: International Monetary Fund).

Kartasheva, A. V. and B. Yilmaz (2013): "Precision of Ratings," *Working paper.*

Krishna, V. and J. Morgan (2008): "Contracting for Information under Imperfect Commitment," *The RAND Journal of Economics*, 39, 905–925.

Lerner, J. and J. Tirole (2006): "A Model of Forum Shopping," *American Economic Review*, 96, 1091–1113.

Lizzeri, A. (1999): "Information Revelation and Certification Intermediaries," *The RAND Journal of Economics*, 30, 214–231.

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

Pagano, M. and P. Volpin (2010): "Credit Ratings Failures and Policy Options," *Economic Policy*, 25, 401–431.

Partnoy, F. (2006): "How and why credit rating agencies are not like other gatekeepers," in *Financial Gatekeepers: Can They Protect Investors?*, ed. by Y. Fuchita and R. E. Litan, Washington: Brookings Institution Press and the Nomura Institute of Capital Markets Research, chap. 3.

Skreta, V. and L. Veldkamp (2009): "Rating Shopping and Asset Complexity: A Theory of Ratings Inflation," *Journal of Monetary Economics*, 56, 678–695.