On the Effect of Aggregate Uncertainty on Certification Intermediaries' Incentives to Distort Ratings*

Petra Loerke[†] Andras Niedermayer[‡]

September 28, 2016

Abstract

We analyze a certification intermediary's incentives to distort ratings in a model with a monopolistic profit maximizing certification intermediary, a continuum of heterogeneous sellers, and a competitive market of risk-neutral buyers. The value of a seller's good is known to the seller and observable by the certifier, but not by buyers. Sellers can choose to get a rating. The certification intermediary can reveal a signal of arbitrary precision about the quality of the good. In contrast to the existing literature, we allow aggregate uncertainty. As in the existing literature, one rating class is optimal. However, the certification intermediary does not choose a socially optimal cutoff: the certifier is more likely to be too lenient if the distribution of aggregate uncertainty has a lower mean, a higher variance, and is more left skewed. It is more likely to be too strict if the opposite holds.

Keywords: Certification intermediaries, aggregate uncertainty
JEL-Classification: C72, D42, D82

*We want to thank Dirk Hoffmann, Raphaël Levy, Leonard Nakumura, Martin Peitz, André Stenzel, Joel Shapiro, Ernst-Ludwig von Thadden, Thomas Tröger, and participants of IIOC 2011 in Boston, EARIE 2011 in Stockholm, the SFB TR 15 Meeting 2011 in Tutzing, the MaCCI Competition and Regulation Day 2014 in Mannheim and seminars at the University of Mannheim and at Concordia University for their helpful comments. The authors acknowledge financial support from the Deutsche Forschungsgemeinschaft through SFB-TR 15. The opinions expressed in this paper are those of the authors and should not be regarded as those of NERA Economic Consulting.

†NERA Economic Consulting, Berlin, Germany.

‡Economics Department, University of Mannheim, L7, 3-5, D-68131 Mannheim, Germany. Email: aniederm@rumms.uni-mannheim.de. Corresponding author.

1 Introduction

Ratings and other quality certifications by third parties play an important role in today's economy. For instance, the volume of rated debt issues was over $8,000 billion in 2006. Ratings are used by investors to guide their investment decisions. They are also crucial for financial regulation: Basel III includes ratings as one criterion for the calculation of the capital adequacy requirements for banks. So does the Solvency II Directive of the European Union, passed on March 11, 2014, which harmonizes insurance regulation in the European Union and came into effect on January 1, 2016.

However, ratings as a basis of regulation have been viewed controversially, especially after the financial crisis. The major concern is that the ratings used for regulation are given by rating agencies, which may have an incentive to distort ratings in order to maximize profit. These concerns have been expressed both in public policy debates and in investigations and lawsuits against rating agencies.[1] As a reaction to this concern, Section 939A of the Dodd–Frank Act (effective since 2010) requires that all federal agencies "must remove any reference to or requirement of reliance on credit ratings."

The present paper addresses the question of incentives to distort ratings by a profit maximizing certification intermediary with particular consideration of aggregate uncertainty. Aggregate uncertainty plays a major role in many markets. As an example, for subprime mortgages, the question was not only how good were the subprime mortgages that one particular financial institution invested in. The question was whether subprime mortgages as a whole were a sufficiently safe investment. While the relevance of the link between aggregate uncertainty and ratings was most prominent during the crisis, it plays a role for other types of intermediaries as well: as an example, the controversy surrounding the health effects of saturated fats af-

[1] As the most recent example, see the settlement between Standard and Poor's, the U.S. Department of Justice, and 19 state attorneys general that includes a payment of $1.5 billion by S&P in the context of the rating of securities backed by subprime mortgages. While "S&P did not admit to any violations of law," the rating agency "did sign a statement of facts acknowledging that its executives in 2005 delayed implementing new models that produced more negative ratings" (Viswanatha and Freifeld, 2015).

fects the perceived quality of all products containing such fats. Certification of food is hence affected by such aggregate uncertainty.[2]

To investigate the effect of aggregate uncertainty on the incentives to distort, we consider a model in which all other possible incentives to distort are shut down. In particular, we consider a monopolistic certification intermediary that can credibly commit to a rating strategy in a one period model. This shuts down effects such as forum shopping, reneging on the ratings strategy, or reputational cycles.

Besides the certification intermediary, there is a continuum of sellers selling goods. There is a continuum of buyers seeking to buy these goods. For the sake of concreteness, one can think of the certification intermediary as a rating agency, the sellers as firms selling bonds, and the buyers as investors. The mass of investors is larger than the mass of sellers, so that competition leads to prices being bid up to the expected value of a bond. The quality of a seller's bond is perfectly known to the seller, but unknown to investors. The rating agency has a technology to perfectly observe the seller's quality. Sellers can decide whether they want to be rated. The aggregate distribution of the sellers' types is initially unknown to all market participants, except for a common prior about the distribution of the aggregate states of the world. The states of the world differ by a different aggregate distribution of sellers' types. After sellers get rated, the aggregate state of the world is revealed to all market participants and investors buy the bonds. The price depends on the expected quality in a rating class for the realized aggregate state of the world.

We show that in accordance with the existing literature, a profit maximizing rating agency will choose a coarse binary rating: either satisfactory or unsatisfactory (which can be thought of as investment grade or junk bonds). However, in sharp contrast to the existing literature, aggregate uncertainty leads to the cutoff not being at the socially optimal level. Whether the rating agency has an incentive to be too lenient (a negative cutoff) or too strict (a positive cutoff) is pinned down by three moments of the aggregate belief distribution. The aggregate belief distribution is

[2]As an example, the organization "Stiftung Warentest" provides ratings for a wide range of products, including food, which takes into account the health effects of a product.

defined as follows: take for every state of the world the mean quality of bonds that would be bought if quality were publicly known. The distribution of the beliefs of the market participants about these means is the aggregate belief distribution. The rating agency has more of an incentive to be too lenient if the distribution has a low mean, a high variance, and a low higher order skewness (defined as the sum of the third and higher moments). A low higher order skewness can be thought of as a left skewed distribution, i.e., bonds have with a high probability a mean quality above average, but the distribution has a fat tail at the bottom, which implies that bonds have a very low mean quality with a small probability. The opposite result holds for a larger mean, lower variance, and a larger higher order skewness. These results can be interpreted as two opposite effects on the rating agency's incentive to distort ratings. One effect is procyclical: there is an incentive to be too lenient before the realization of unfavorable aggregate uncertainty (which can be thought of as before the outbreak of a crisis, interpreting this period as a period with a large variance and left skewness of aggregate uncertainty) and an incentive to be too strict after the realization of unfavorable aggregate uncertainty (after the outbreak of the crisis). The other effect is countercyclical: a higher mean in market beliefs about aggregate uncertainty (likely to occur before a crisis) gives the rating agency an incentive to be too strict and a lower mean (after a crisis) to be too lenient. While anecdotal evidence suggests that the procyclical effect is stronger,[3] it is ultimately an empirical question which effect dominates.

This sheds light on a disturbing aspect of using credit ratings for capital adequacy regulation: they may introduce procyclicality into the system. Capital adequacy requirements based on ratings (such as in Basel III and the Solvency II Directive) may be too lenient before and too strict after the crisis. Our theory can be seen to justify two possible policies to deal with this problem. One policy, as in Section 939A of the Dodd–Frank Act, is to remove from the regulations any reference to or requirement of reliance on credit ratings. This approach has the advantage of

[3]In hindsight, observers of financial markets considered the ratings of agencies to have been too lenient before and too strict after the crisis.

having a clear unambiguous rule. However, this is also viewed controversially, since it may be too costly for smaller banks to replace external credit ratings with internal credit rating systems.[4] An alternative policy would be to use credit ratings, but take into account their cyclicality in regulation. In particular, if one believes that the procyclical element dominates, capital adequacy requirements based on ratings should include countercyclical elements to counterbalance procyclicality.

We provide two extensions of our main result. First, we outline an empirical strategy to determine whether the procyclical or the countercyclical effect dominates. While an empirical analysis is outside the scope of this paper, we show how the moments of the distribution of aggregate uncertainty can be identified from the prices of financial derivatives.

Second, we extend the model to a setup with risk aversion. A model with risk aversion explains why there are multiple rating categories and not just one (i.e., investment grade, and possibly a second, speculative grade). The reason is that with risk aversion, investors value more precise information about the quality of an asset to reduce risk. We provide numerical examples to illustrate that a hybrid model of risk aversion and aggregate uncertainty preserves the key insights about the rating agency being too lenient or too strict, but additionally predicts multiple rating categories.

Our paper relates to a large literature on rating agencies, experts, and reputation. We differ from all papers mentioned below by having market participants' uncertainty about the aggregate distribution of qualities as the driving force that determines the rating strategy.

If one were to remove aggregate uncertainty from our model, it would reduce to the model in Lizzeri (1999)'s seminal contribution on certification intermediaries. Lizzeri (1999) shows the by now well known result that certification intermediaries choose two categories (corresponding to investment grade and junk bonds) and set a cutoff at 0 which is the first-best level. (Note that this result can also be viewed as

[4] See, for example, http://www.americanbanker.com/bankthink/an-easy-fix-to-dodd-franks-credit-ratings-rule-1063396-1.html.

only one rating category being chosen—investment grade—and other assets not being rated.) We show that introducing aggregate uncertainty fundamentally changes the predictions of the theory with respect to the welfare implications of intermediaries: while without aggregate uncertainty, the cutoff is chosen at the first-best level, in the presence of aggregate uncertainty, the cutoff introduces a distortion compared to first-best, leading to either underinvestment or overinvestment.

Lizzeri (1999)'s work has been extended in a number of directions, including Doherty, Kartasheva, and Phillips (2012)'s work on risk-averse buyers. With risk-averse buyers, it can be optimal to have more than two categories.

Certification intermediaries[5] have been considered from various other perspectives besides the disclosure policy, such as reputational effects (Bolton, Freixas, and Shapiro, 2012, Bar-Isaac and Shapiro (2013), Strausz (2005), Ottaviani and Sørensen (2006), Bouvard and Levy (2009),Mariano (2008), Pollrich and Wagner (2013), Mathis, McAndrews, and Rochet (2009)), the effect of competition (Lerner and Tirole (2006), Skreta and Veldkamp (2009)), whether the buyer or the seller should pay for a rating (Stahl and Strausz (2014)), and the effect of the regulatory status of ratings (Opp, Opp, and Harris (2013)). The main difference in this paper is the focus on aggregate uncertainty.[6]

This paper is structured as follows. Section 2 describes the model. Section 3 shows that it is optimal to rate according to a simple cutoff rule and Section 4 derives conditions under which this cutoff is positive or negative. Section 5 describes a

[5]Some of the following papers consider credit rating agencies specifically, other papers consider more generally experts, who provide information about the quality of a good.

[6]Our aggregate uncertainty has some similarities to the business cycle effects considered by Bar-Isaac and Shapiro (2013) (where it is more difficult to hire qualified raters during booms) and Opp et al. (2013) (where there is more rating inflation during booms). In our analysis, the key differences from these papers are that we show that multiple moments of the aggregate distribution matter (the variance and the skewness and not just the mean) and that ratings sometimes may be too strict. Additionally, we differ from Bar-Isaac and Shapiro (2013) by not having naive investors and from Opp et al. (2013) by having rating distortions without ratings used for regulation. The differences are more than technical: if ratings can be too strict under some conditions, then regulation (such as capital adequacy requirements) based on ratings should be more lax under these conditions. Further, if not just the mean of aggregate fluctuations matters, but also the higher order moments, then a regulation that seeks to balance rating distortions should be contingent on the variance and skewness and not just the mean. Whether these higher moments matter in practice is an empirical question that has not been answered yet. However, as we will discuss later, there are reasons that suggest that they should matter.

stylized empirical identification strategy. Section 6 shows that with risk-averse buyers, several rating classes can be optimal but that the effects of aggregate uncertainty on the optimal cutoff remain. Section 7 concludes.

2 The Model

There is one certification intermediary, a continuum of sellers, and a continuum of possible buyers. Each firm sells a good of quality t, where t is a random variable with support $[\underline{t}, \overline{t}]$ with $\underline{t} < 0 < \overline{t}$. The seller has private information about the quality. Buyers are risk neutral and a buyer's expected gross utility from buying the good is equal to the quality t. A seller's utility of retaining the good is normalized to zero.

For the sake of concreteness, we will, throughout this paper, refer to the certification intermediary as the rating agency, to the sellers selling a good as firms selling a bond, and to buyers as investors. The quality t of the bond can be thought of as capturing the probability of default, and the loss given default.

There are N different states of the world. The probability of the world being in state i is ϵ_i. Having a two dimensional distribution (different states of the world, different distributions of qualities in each state of the world) adds a considerable amount of complexity. To still have a tractable model, we impose a restriction on this two dimensional distribution. We assume that there is a mass κ_i of sellers whose quality $\underline{t}$ is so low that one would never want to rate them (we will formalize this later on in Assumption 2). There is a mass μ_i of sellers whose quality $\overline{t}$ is so high that one would always want to rate them. And then there is a mass λ_i of sellers with intermediate qualities $t \in (\underline{t}, \overline{t})$. We allow for arbitrary distributions of κ_i, λ_i, μ_i (with the only restrictions being that the sum $\kappa_i + \lambda_i + \mu_i$ is constant and Assumption 2), but restrict the distribution conditional on being in $(\underline{t}, \overline{t})$ to be a distribution F which is the same for all states. We assume that F is continuously differentiable with density $f(t) > 0$ for all t in $(\underline{t}, \overline{t})$.

Further, define the expected masses on $(\underline{t}, \overline{t})$ as $\tilde{\mu} := \sum_i \epsilon_i \mu_i$ and on $\overline{t}$ as $\tilde{\lambda} :=$

$\sum_i \epsilon_i \lambda_i$. Normalize $\tilde{\lambda}$ to 1. The probabilities ϵ_i and the distributions of quality are known to all players.

A firm can choose to pay an upfront fee P to the rating agency in order to get rated before the state of the world becomes known to the market participants. The agency rates firms that paid for a rating according to a precommitted rating strategy.[7]

The timing of the moves is as follows:

- The agency sets the rating fee P and commits to a rating strategy s, $s(t) = r$, $s : \mathbb{R} \to \mathbb{R} \cup \{\emptyset\}$.
- Nature draws the state of the world i and quality t of each firm.
- The firms observe their own qualities, but not the state of the world, and decide whether to go to the agency to ask for a rating or not. This decision depends on the own type t, the strategy of the agency s and the price P.
- The agency observes the quality of each firm asking for a rating, and gives the ratings according to its strategy. The ratings are publicly observable. However, investors do not observe whether a firm went to the rating agency if the firm gets no rating ($\emptyset$).
- Observing the state of the world, the buyers decide how much to bid in a second price auction for a good. Since it is a second price auction, buyers bid their own expected valuation which depends on their belief about the expected quality given the information (s, P, r, i). Assuming that there are more investors than firms, investors will pay exactly the expected quality in equilibrium.

We analyze the Perfect Bayesian Equilibrium of this game. We restrict the strategy of the firms to pure strategies and look at symmetric equilibria.

The profits of the agency in one state of the world equal the rating fee P times the mass of firms asking for a rating. This mass depends on P and the rating strategy s.

[7] It does not matter in equilibrium whether the strategy is known at the beginning or not.

The agency is risk neutral and chooses s and P to maximize expected profits before knowing the state of the world.

The rating agency's rating strategy s partitions the set $[\underline{t}, \bar{t}]$ into M subsets, with each subset $m = 1, ..., M$ being the set of types $T_m = \{t|s(t) = r_m\}$ with M distinct r_m.[8] We will call these subsets rating classes in the following. Since in the end only the M distinguishable classes $\{T_m\}_{m=1}^M$ matter, and not the labels $\{r_m\}_{m=1}^M$ attached to them, the following analysis will focus on $\{T_m\}$.

It is useful to define the expected quality in state i conditional on t being above a threshold $x > \underline{t}$ as

$$E_i(x) := \frac{\lambda_i \int_x^{\bar{t}} t dF + \mu_i \bar{t}}{\lambda_i \int_x^{\bar{t}} dF + \mu_i}.$$

A firm in $(\underline{t}, \bar{t})$ attaches probability $\hat{\epsilon}_i := \epsilon_i \lambda_i / \tilde{\lambda}$ to being in state i. Consequently, from a $(\underline{t}, \bar{t})$ firm's perspective, the expected quality above a threshold x over all states is

$$\tilde{E}(x) := \sum_i \hat{\epsilon}_i E_i(x).$$

In the following, we will assume that the virtual valuation function attached to $\tilde{E}(x)$ is monotone in x for $x \in (\underline{t}, \bar{t})$.

Assumption 1. *The virtual valuation function $\tilde{E}(x) - \tilde{E}'(x)\frac{1-F(x)+\tilde{\mu}}{f(x)}$ is monotone in x for $x \in (\underline{t}, \bar{t})$.*

This assumption basically ensures that the second-order condition is fulfilled whenever the first-order condition is fulfilled and it excludes the corner solution that it is optimal to only rate $\bar{t}$. Monotonicity of the virtual valuation function is a standard assumption, see e.g., Myerson (1981).

We further assume that $\underline{t}$ is sufficiently small:

[8] Technically speaking, there are $M + 1$ subsets because there can be types which do not receive any rating, $s(t) = \emptyset$. We will show later in this paper that it cannot be optimal to have more than two rating categories. Therefore, for the sake of notational simplicity, we do not consider an uncountable infinity of rating classes. To take into account the possibility of an uncountable infinity of rating classes, e.g., full disclosure, one could use the correspondence $T(r) = \{t|s(t) = r\}$ with $r \in \mathbb{R} \cup \{\emptyset\}$ instead of the sets $\{T_m\}_{m=1}^M$.

Assumption 2.

$$\underline{t} < -\frac{\lambda_i \int_0^{\bar{t}} t dF(t) + \mu_i \bar{t}}{\kappa_i}, \qquad \forall i = 1, ..., N$$

Assumption 2 makes sure that we do not have to deal with the uninteresting corner solution in which the rating agency wants to rate all firms, including $\underline{t}$ firms (see Lemma 1 below).[9]

Our setup would be intractable even for simple aggregate distributions. To achieve tractability, we have made the following assumption: there are two mass points at $\underline{t}$ and $\bar{t}$, respectively, and the density on $(\underline{t}, \bar{t})$ moves proportionally in different states. This simplification buys us a surprising level of generality in terms of the distribution of aggregate uncertainty: we do not need to make any assumptions about a first-order stochastic dominance ranking of different states of the world and do not need to assume the monotone likelihood ratio property. Our assumption can be viewed as the distribution above a threshold being collapsed to the mass point $\bar{t}$ and the distribution below a threshold being collapsed to $\underline{t}$.

3 Optimality of a Threshold Rating Strategy

In the following, we will show that it is optimal to rate all firms in an interval $[x, \bar{t}]$ in one rating class and to not give a rating to any firm with $t < x$. Formally, $s(t) = 1$ for all $t \geq x$ and $s(t) = \emptyset$ for all $t < x$.[10] We will show this in four steps. First, we show that it cannot be optimal to exclude type $\bar{t}$. Second, we show that the price of a rating is determined by firms with $t < \bar{t}$. Third, given that $\bar{t}$ is included, it is optimal to have only one rating class rather than multiple classes. Fourth, given that there is only one rating class, the set of types belonging to this class has to be convex.

While the following lemmas are intuitive, their proofs are surprisingly long. We

[9]Without Assumption 2, our results for the interior solutions given by the first-order conditions would still hold, but we would have to distinguish between the cases with an interior solution and a corner solution.

[10]This is equivalent to $s(t) = 1$ for all $t \geq x$ and $s(t) = 0$ for all $t < x$ because firms with $t < x$ are not rated in equilibrium.

therefore provide an intuition in the main text and relegate the proofs to the Appendix.

Lemma 1. *(i) It cannot be optimal that* $\underline{t} \in \cup_{m=1}^{M} T_m$. *(ii) It cannot be optimal that* $\bar{t} \notin \cup_{m=1}^{\tilde{M}} \tilde{T}_m$.

Part (i) of the lemma holds by Assumption 1. The intuition for part (ii) of the lemma is that $\bar{t}$ should be included in the rating because it increases the mass of rated firms as well as, due to its high type, other firms' willingness to pay for a rating.

Next, we state a lemma which will be useful throughout our analysis. The lemma states that if firms with either $t \in \left(\underline{t}, \bar{t}\right)$ or with $t = \bar{t}$ are in the same rating class, then firms with $t \in \left(\underline{t}, \bar{t}\right)$ have a lower willingness to pay for a rating than firms with $t = \bar{t}$.

Lemma 2. *Take an arbitrary rating class* T *that includes both firms with* $t \in \left(\underline{t}, \bar{t}\right)$ *and those with* $t = \bar{t}$. *The willingness to pay for a rating is higher for those with* $\bar{t}$ *than for those with* $t \in (\underline{t}, \bar{t})$.

The reason is that firms update $\hat{\epsilon}_i$ differently and we show that firms of type $\bar{t}$ assign a higher probability to states with higher expected quality than do firms with $t \in (\underline{t}, \bar{t})$. Lemma 2 can be used to prove the next lemma, which states that if there are multiple rating classes and the highest type $\bar{t}$ is included, then it is better to merge all rating classes into one single class.

Lemma 3. $M = 1$ *with* $T_1 = \cup_{m=1}^{\tilde{M}} \tilde{T}_m$ *is better than* $\left\{\tilde{T}_m\right\}_{m=1}^{\tilde{M}}$ *with* $\tilde{M} > 1$ *if* $\bar{t} \in \cup_{m=1}^{\tilde{M}} \tilde{T}_m$.

Considering the types that the agency intends to attract, the rating fee is always determined by the type with the lowest willingness to pay for a rating. Merging the rating class with a lowest willingness to pay with classes with a higher willingness to pay, the expected quality and thus, also the minimum willingness to pay, will increase.

The next lemma states that all firms above a threshold are rated, which means that no types in between are excluded.

Lemma 4. *If $M = 1$ and $\bar{t} \in T_1$, then T_1 has to be convex.*

If the set were not convex, there would be at least one unrated hole in the middle and the agency could rate firms in the hole instead of rating some other types below with the same mass. This would increase the expected type in every state and, therefore, also the rating fee the agency could charge from the firms. Therefore, a nonconvex T_1 cannot be optimal.

Lemmas 1, 3, and 4 together lead to the following proposition.

Proposition 1. *It is optimal to choose $M = 1$ with $T_1 = [x, \bar{t}]$ for some x.*

Proposition 1 shows that the best equilibrium for the rating agency is such that the agency offers the following ratings strategy:

$$s(t) = \begin{cases} 1 & \text{if } t \geq x, \\ \emptyset & \text{otherwise,} \end{cases}$$

that is, all firms above some cutoff x get a positive rating. Subsequently, all firms with $t \in [x, \bar{t}]$ get rated and investors pay the expected quality over $[x, \bar{t}]$.

As usual in such models, there is a multiplicity of equilibria in the subgame following the ratings agency's announcement of its price P and rating strategy s. For example, there is the trivial equilibrium in which no firm applies for a rating and investors have the off-equilibrium belief that firms that do get a rating are of the worst possible rated quality x. Since x is less than the price of a rating P, it is a best response for firms to stay unrated.

The usual arguments for selecting the equilibrium we described apply: the rating agency has a first-mover advantage, hence, it is reasonable that the equilibrium most favorable to the rating agency will be selected. Further, by a small perturbation of its strategy, the rating agency can get rid of undesired equilibria. For example, if no firm gets a rating, the agency might incentivize the first firms who apply for a rating in order to jump-start the market.[11]

[11] A simple, albeit extreme example is the following: as long as not all firms with a quality $t \in [x, \bar{t}]$ enter, firms get their rating fees refunded and get an additional small compensation. This makes sure that any equilibrium in which not all firms in $[x, \bar{t}]$ get rated breaks down, so that the refund never has to be paid in equilibrium.

4 Optimal Threshold

By Proposition 1 we can restrict our attention to threshold rules which consist of all types above a cutoff x being pooled in one class and all types below not being rated. This has the advantage that we can write the intermediary's profits $\Pi(x)$ and its total welfare $W(x)$ as functions of the cutoff x.

Profits are

$$\Pi(x) := \left[\sum_{i=1}^{N}(\lambda_i(1-F(x))+\mu_i)\epsilon_i\right]\left[\sum_{j=1}^{N}E_j(x)\hat{\epsilon}_j\right],$$

where the first expression in square brackets is the mass of the rated firms and the second expression in square brackets is the willingness to pay of a rated firm.

The welfare generated by a single firm of type t that is rated is simply t. The total welfare can be obtained by integrating over t and summing over the states of the world:

$$W(x) := \sum_{i=1}^{N}(\lambda_i(1-F(x))+\mu_i)E_j(x)\epsilon_i$$

For the first step, it is useful to rearrange the expressions for profits and welfare for the purpose of a comparison. This should give a first idea of where the tension between profit maximization and welfare maximization is coming from.

Profits can be written as

$$\Pi(x) :=(1-F(x)+\tilde{\mu})\tilde{E}(x)$$

where $\tilde{E}(x)$ is the expected value of a rating from a firm's perspective which assigns the probabilities $\hat{\epsilon}_i$ to different states.

Welfare can be rearranged to

$$W(x) =(1-F(x)+\tilde{\mu})\hat{E}(x)$$

where

$$\hat{E}(x) = \frac{\sum_i \epsilon_i(\lambda_i(1-F(x))+\mu_i)E_i(x)}{1-F(x)+\tilde{\mu}}.$$

$\hat{E}(x)$ is the expected value of a rating from a welfare perspective which takes into account that the quantity of firms being rated $(\lambda_i(1 - F(x)) + \mu_i)$ is different in every state.

Comparing these expressions for Π and W reveals that the difference between profits and welfare is driven by how valuable a rating is for the intermediary's profits $\tilde{E}$ and how valuable it is for welfare $\hat{E}$.

For the special case when there is no aggregate uncertainty, $\tilde{E}$ and $\hat{E}$ coincide, and the certifier's profit maximization problem coincides with the welfare maximization problem. In this case, the agency chooses the cutoff $x = 0$. To see this, take a model with only one state of the world, e.g., by setting $\mu_i = \tilde{\mu}$ and $\lambda_i = \tilde{\lambda} = 1$ for all i. Then the agency's profit is

$$\Pi = (1 - F(x) + \tilde{\mu})\frac{\int_x^{\bar{t}} tdF(t) + \tilde{\mu}\bar{t}}{1 - F(x) + \tilde{\mu}} = \int_x^{\bar{t}} tdF(t) + \tilde{\mu}\bar{t}.$$

The first derivative is $\Pi'(x) = -xf(x)$, which is equal to 0 if $x = 0$. Therefore, the optimal threshold for the agency is $x = 0$.[12] This special case of our model corresponds to Lizzeri (1999)'s results.

Let us now turn to the case where there is aggregate uncertainty, so that profit maximization and welfare maximization do not coincide.

For the following analysis, it will be useful to denote the expected type in $[x, \bar{t})$ as

$$E_0(x) := \frac{\int_x^{\bar{t}} tdF(t)}{1 - F(x)}.$$

In the following, we will drop the argument x in $E_i(x)$, $E_0(x)$, $\tilde{E}(x)$, $\hat{E}(x)$ when it is unambiguous, in order to simplify notation. $\hat{E}$ and $\tilde{E}$ can be compared in the following way.

Lemma 5. *The value of a rating is larger from a welfare perspective than from a firm's perspective:* $\hat{E} \geq \tilde{E}$ *for all* x.

This implies that $W(x) \geq \Pi(x)$. For nondegenerate distributions of the state of the world, the inequality is strict and in contrast to a one-state-of-the-world setup,

[12] It is easy to check that the second-order condition is also satisfied at $x = 0$.

the rating agency cannot extract the whole surplus: $W(x) > \Pi(x)$.[13]

Next, let us compare the first-order conditions for profit maximization and welfare maximization. The derivative of the profit function with respect to the cutoff is

$$\Pi'(x) = -f(x)\left[\underbrace{\tilde{E}(x)}_{marginal\ effect} - \underbrace{\frac{1-F(x)+\tilde{\mu}}{f(x)}\tilde{E}'(x)}_{inframarginal\ effect}.\right] \quad (1)$$

and we will show later that the first order condition is sufficient for profit maximization. Thus, the profit maximizing cutoff is given by $\Pi'(x) = 0$. Changing the cutoff has two opposite effects on the agency's profit: increasing the cutoff decreases the mass of firms asking to be rated (marginal effect), but it also increases the expected quality of firms being rated and in this way it increases a firm's willingness to pay for being rated (inframarginal effect).

We call the expression in the square brackets in (1) the virtual valuation function for $\tilde{E}$.[14] By Assumption 1 it is monotone and, thus, the first order condition is sufficient to find an optimum.[15] This also implies that there is a unique solution of the first order condition.

We are interested in comparing the profit maximizing cutoff with the welfare maximizing cutoff. Thus, we also have to determine the socially optimal cutoff. The derivative of welfare with respect to the threshold is

$$W'(x) = -f(x)\left(\underbrace{\hat{E}(x)}_{marginal\ effect} - \underbrace{\frac{1-F(x)+\tilde{\mu}}{f(x)}\hat{E}'(x)}_{inframarginal\ effect}\right). \quad (2)$$

A comparison of (1) with (2) provides an intuition for the different incentives at

[13] Even if $\hat{\epsilon}_i = \epsilon_i$ for all i, the inequality is strict for nondegenerate distributions. Besides by the updating of $\hat{\epsilon}_i$, the difference between $\hat{E}$ and $\tilde{E}$ is caused by the different mass of firms being rated in different states of the world.

[14] We can rewrite the virtual valuation in terms of E_i as $\sum_i \epsilon_i \lambda_i \left(E_i - E_i' \frac{1-F+\tilde{\mu}}{f}\right)$.

[15] The second order condition follows directly from Assumption 1. That we do not have a corner solution at $x = \bar{t}$ can be seen by observing that $\Pi(x)$ is continuous at $x = \bar{t}$ and $\lim_{x \to \bar{t}} \Pi'(x) < 0$. Assumption 2 implies that there is no corner solution at $x = \underline{t}$ (see the proof of Lemma 1).

work for a profit maximizing and a welfare maximizing rating agency. Both types of agencies care about the marginal effect (a change in the number of rated firms) and the inframarginal effect (the willingness to pay for a rating changes for all firms that get a rating). However, the two types of rating agencies differ in that a profit maximizing rating agency considers $\tilde{E}$, the expected value of a rating from a profit maximization perspective, whereas a welfare maximizing rating agency considers $\hat{E}$, the expected value of a rating from a welfare maximization perspective.

Observe that (2) simplifies to

$$W'(x) = -xf(x)$$

which is the same as for one state of the world. Again, the derivative is 0 at $x = 0$ and thus, the welfare maximizing cutoff is at 0.

The following analysis is simplified by observing that, with some algebra, $\Pi(x)$ can be rewritten[16] as

$$\Pi(x) = W(x) + L(x)$$

where $L(x) := -\sum_j E_j \hat{\epsilon}_j (\mu_j - E[\mu])$ is the non-extractable part of the surplus ("loss" compared to extracting total surplus). Recall that $W'(0) = 0$, which implies that $L'(0) = \Pi'(0)$. Thus, the incentive for the agency to distort the rating compared to the welfare maximizing rating is given by the sign of $L'(0)$. This simplifies the following analysis: instead of dealing with $\Pi'(x)$, we can deal with $L'(x)$. The

[16]The intermediate steps leading to the expression below are

$$\begin{aligned}
\Pi(x) =& \left(\sum_i (\lambda_i (1 - F(x)) + \mu_i)\epsilon_i\right) \left(\sum_j E_j(x)\hat{\epsilon}_j\right) \\
=& \sum_j \left(E_j(x)\hat{\epsilon}_j \left(\sum_i (\lambda_i (1 - F(x)) + \mu_i)\epsilon_i\right)\right) \\
=& \sum_j E_j(x)(1 - F(x) + \mu_j)\hat{\epsilon}_j - \sum_j E_j(x)\hat{\epsilon}_j \left(\mu_j - \sum_i \mu_i \epsilon_i\right) \\
=& W(x) + L(x)
\end{aligned}$$

optimal cutoff is positive if $L'(0) > 0$ and it is negative if $L'(0) < 0$.

Proposition 2. *The derivative $L'(0)$ is given by*

$$L'(0) = \frac{f(0)\tilde{E}}{\bar{t} - \hat{E}} \left(\hat{E} - \tilde{E} - \frac{(\sum_i \hat{\epsilon}_i E_i)^2 - \sum_i \hat{\epsilon}_i E_i^2}{\tilde{E}} \right). \tag{3}$$

The proof of Proposition 2 is provided in the Appendix.

Since the expression before the parenthesis is always positive, the sign of $L'(0)$ and, therefore, the sign of the profit maximizing cutoff, depends on the sign of $\left(\hat{E} - \tilde{E} - \frac{(\sum_i \hat{\epsilon}_i E_i)^2 - \sum_i \hat{\epsilon}_i E_i^2}{\tilde{E}} \right)$. $\hat{E} - \tilde{E}$ is positive and can be interpreted as the difference between the expected value of a rating from the social planner's point of view and a firm's perspective. An intuition for $\frac{(\sum_i \hat{\epsilon}_i E_i)^2 - \sum_i \hat{\epsilon}_i E_i^2}{\tilde{E}}$ is that it is the variance divided by the mean of the posterior distribution of E_i. It reflects the uncertainty about the state of the world: if this uncertainty is sufficiently large, the cutoff is negative. The reason for this is that firms care less about the effect of the cutoff x on the expected quality of a rated firm if the expected quality is to a large extent driven by uncertainty about the state of the world. Thus, the sign of $L'(0)$ is determined by the difference between the expected value of a rating and the ratio of the variance to the mean of the posterior distribution of E_i.

While the above expression for $L'(0)$ provides some insight into the determinants of the optimal cutoff, it is difficult to use it for comparative statics, since a change of the mean and variance of E_i will also change $\hat{E}$. Therefore, in the following, we will express $L'(0)$ in terms of the moments of the posterior distribution of E_i.

For this purpose, first define the scaled value of a rating $e_i := E_i/\bar{t}$. Next, define the kth moment of the posterior distribution of e_i as $m_k := \sum_i \hat{\epsilon}_i e_i^k$. Further, define the higher-order skewness as $m_{3+} := \sum_{k=3}^{\infty} m_k$, that is, the sum of the third and all higher order moments. With these definitions at hand, we can state our main result in the following proposition.

Proposition 3. *Define*

$$S := 1 - \frac{1}{1 + m_1 + m_2 + m_{3+}} - \frac{m_2}{m_1}. \tag{4}$$

The optimal cutoff is positive if $S > 0$, negative if $S < 0$, and zero if $S = 0$.

Proof. Because $\mu_i = \lambda_i(1 - F(x))\frac{E_i - E_0}{\bar{t} - E_i}$, we can write

$$\begin{aligned}
\tilde{\mu} &= \sum_i \epsilon_i \lambda_i (1 - F(x)) \frac{E_i - E_0}{\bar{t} - E_i} \\
&= (1 - F(x)) \sum_i \hat{\epsilon}_i \frac{E_i - \bar{t} + \bar{t} - E_0}{\bar{t} - E_i} \\
&= (1 - F(x)) \left(-1 + (\bar{t} - E_0) \sum_i \hat{\epsilon}_i \frac{1}{\bar{t} - E_i} \right).
\end{aligned} \tag{5}$$

Plugging (5) into the definition of $\hat{E}$ we get

$$\begin{aligned}
\hat{E} =& \frac{(1 - F(x))E_0 + \tilde{\mu}\bar{t}}{1 - F(x) + \tilde{\mu}} \\
=& \frac{E_0 + \left(-1 + (\bar{t} - E_0) \sum_i \hat{\epsilon}_i \frac{1}{\bar{t} - E_i} \right) \bar{t}}{1 + \left(-1 + (\bar{t} - E_0) \sum_i \hat{\epsilon}_i \frac{1}{\bar{t} - E_i} \right)} \\
=& \bar{t} - \frac{1}{\sum_i \hat{\epsilon}_i \frac{1}{\bar{t} - E_i}}.
\end{aligned}$$

Observe that $\sum_i \hat{\epsilon}_i \frac{1}{\bar{t} - E_i} = \sum_i \hat{\epsilon}_i \frac{1}{\bar{t} - e_i \bar{t}}$. The kth derivative of $\frac{1}{\bar{t}(1 - e_i)}$ with respect to e_i is

$$\frac{\partial^k}{\partial e_i^k} \left[\frac{1}{\bar{t}(1 - e_i)} \right] = k! \frac{1}{\bar{t}(1 - e_i)^{k+1}}.$$

Using these derivatives one can construct the Taylor series of $\frac{1}{\bar{t}(1 - e_i)}$ with respect to e_i around $e_i = 0$. This yields

$$\begin{aligned}
\frac{1}{\bar{t}(1 - e_i)} =& \sum_{k=0}^{\infty} \frac{e_i^k}{k!} \frac{\partial^k}{\partial e_i^k} \left[\frac{1}{\bar{t}(1 - e_i)} \right] \Bigg|_{e_i = 0} \\
=& \sum_{k=0}^{\infty} \frac{e_i^k}{\bar{t}}.
\end{aligned}$$

Taking expectations over the state of the world yields

$$\sum_i \hat{\epsilon}_i \left[\frac{1}{\bar{t}(1 - e_i)} \right] = \frac{1}{\bar{t}} \left(1 + m_1 + m_2 + \sum_{k=3}^{\infty} m_k \right),$$

This implies that we can write

$$\hat{E} = \bar{t} - \frac{\bar{t}}{1 + m_1 + m_2 + \sum_{k=3}^{\infty} m_k}. \tag{6}$$

Observe that (3) simplifies to

$$L'(0) = \frac{f\tilde{E}}{\bar{t} - \hat{E}} \left(\hat{E} - \frac{\sum_i \hat{\epsilon}_i E_i^2}{\tilde{E}} \right). \tag{7}$$

Plugging (6), $\tilde{E} = m_1 \bar{t}$ and $\sum_i \hat{\epsilon}_i E_i^2 = \bar{t}^2 m_2$ into (7) yields

$$\begin{aligned} L'(0) =& \frac{f m_1 \bar{t}}{\frac{\bar{t}}{1+m_1+m_2+m_{3+}}} \left(\bar{t} - \frac{\bar{t}}{1 + m_1 + m_2 + m_{3+}} - \frac{\bar{t}^2 m_2}{m_1 \bar{t}} \right) \\ =& f m_1 \bar{t} \left(\sum_{k=0}^{\infty} m_k \right) \left[1 - \frac{1}{1 + m_1 + m_2 + m_{3+}} - \frac{m_2}{m_1} \right]. \end{aligned}$$

Observing that everything outside of the square brackets is positive and that the optimal cutoff has the same sign as $L'(0)$ yields the result in the proposition. □

The intuition for Proposition 3 is the same as the one provided after Proposition 2: the expression in parentheses in (3) can be rewritten as S in (4), where $\hat{E} - \tilde{E} - (\sum_i \hat{\epsilon}_i E_i)^2/\tilde{E}$ becomes $1/(m_1 + m_2 + m_{3+})$ and $\sum_i \hat{\epsilon}_i E_i^2/\tilde{E}$ becomes m_2/m_1.

Note that S depends only on the moments of e_i. More precisely, it depends only on the mean m_1, the second moment m_2 and the sum of all higher moments m_{3+}. For example, assume that $L'(0) < 0$. Keeping the mean and the second moment constant and increasing the sum of higher moments, S increases and $L'(0)$ can switch signs from negative to positive.

We can calculate the threshold $\overline{m}_{3+}$ for which $L'(0)$ is 0. Setting S to 0 and rearranging (4) yields

$$\overline{m}_{3+} = \frac{m_2^2 + m_2 - m_1^2}{m_1 - m_2}.$$

Observe that $\overline{m}_{3+}$ is always positive because $m_1 > m_2$ and $m_2 - m_1^2$ is the variance of e_i. This implies that for $m_{3+} < \overline{m}_{3+}$ the expression S is negative and thus $L'(0) = \Pi'(0) < 0$.

Proposition 4. *The optimal cutoff for the rating agency is negative if* $m_{3+} < \overline{m}_{3+}$

and positive if $m_{3+} > \overline{m}_{3+}$.

We also derive thresholds for m_1 and m_2. First, observe that S is increasing in m_1 and decreasing in m_2 given that $m_1, m_2, m_{3+} > 0$. Second, by setting the expression in square brackets to zero and solving for m_1 and m_2, respectively, one gets thresholds for m_1 and m_2 that determine whether the cutoff of the rating agency is positive or negative. The thresholds are stated in the following two propositions.

Proposition 5. *The optimal cutoff for the rating agency is negative if* $m_1 < \overline{m}_1$ *and positive if* $m_1 > \overline{m}_1$*, where*

$$\overline{m}_1 := \frac{1}{2}\left(-m_{3+} + \sqrt{4m_2 + (2m_2 + m_{3+})^2}\right)$$

Proposition 6. *The optimal cutoff for the rating agency is negative if* $m_2 > \underline{m}_2$ *and positive if* $m_2 < \underline{m}_2$*, where*

$$\underline{m}_2 := \frac{1}{2}\left(-1 - m_{3+} + \sqrt{2m_{3+} + 1 + (2m_1 + m_{3+})^2 + 2m_{3+} + 1}\right)$$

Both thresholds, $\overline{m}_1$ and $\underline{m}_2$, are positive given that $m_1, m_2, m_{3+} > 0$.

An intuition for these results is that with a high variance and a small higher order skewness (i.e., a left-skewed distribution), the value of a rating is highly volatile, not because of the rating strategy, but because of aggregate uncertainty. Hence the rating agency worries less about keeping the value of the rating high by being restrictive on which firms get a rating than if the variance were low or the higher order skewness were large.

Propositions 4, 5, and 6 have a striking implication: the rating agency has more of an incentive to be too lenient if the distribution of aggregate uncertainty is more left-skewed (in the sense of a smaller higher order skewness, i.e., a lower m_{3+}), the mean is smaller, or the variance is larger. Left skewness and a high variance can be reasonably considered as being associated with a period preceding the beginning of a crisis. For moments that can be reasonably associated with a period shortly after a crisis (right skewness, low variance), the incentive of the rating agency moves in the opposite direction: the rating agency has an increasing incentive to be too strict. This gives the rating agency an incentive to rate procyclically: excessively lenient ratings expand investments during booms, excessively restrictive ratings restrict investments during recessions. Observe that the mean of the aggregate uncertainty has a counter cyclical effect: a small expected average, which can be associated with a period shortly after a crisis, gives the rating agency an incentive to be too lenient. The opposite holds for a high expected average.

How to best view cyclicality in terms of the moments of the aggregate distribution is of course ultimately an empirical question. We are not aware of any clearcut evidence one way or the other. Having a theory that highlights the potential importance of such moments can be useful for future empirical work. We will also discuss empirically testable predictions concerning the moments of the aggregate distribution.

However, even without having at hand formal empirical estimates of the moments, there is reason to suspect that higher order moments should matter. For one thing, asset managers' bonuses are often based on returns corrected by the variance of the investment.[17] This gives asset managers incentives to increase revenues at the cost of more risk in terms of higher order losses. Further, the originators of the mortgages in mortgage backed security funds were typically required to hold on to the junior tranches of the fund. This meant that they were exposed to small risks. However, once the risk was so high that it wiped out the entire junior tranche, the originator did not have any skin in the game any more. Additionally, the most senior tranches were not affected by small risks, but were severely affected by large risks that were sufficiently large to wipe out all the junior and mezzanine tranches. More generally, in the case of limited liability, firms care less about the lower tail of the

[17]The most common measures used to evaluate asset managers' performance are based on the returns and on the volatility (i.e., the square root of the variance): the Sharpe ratio, the Treynor ratio, and Jensen's alpha. See Jensen (1968) and the subsequent literature on the performance evaluation of funds.

distribution, since small losses are borne by shareholders, but large losses by debtors. Another point worth mentioning is that a left skewness was also introduced by chain reactions in capital markets and feedback loops, e.g., by AIG insured lenders of subprime mortgage against credit default to a certain degree. However, once the crisis hit, AIG went bankrupt.

4.1 Example of Beta Distributions

It is illustrative to parametrize the posterior distribution of E_i as a Beta distribution with support $[E_0, \bar{t}]$, i.e., E_i has a density $h(y) \propto y^{\alpha-1}(1-y)^{\beta-1}$, where $y = (E_i - E_0)/(\bar{t} - E_0)$. The distribution of $E_i/\bar{t}$ is determined by the three parameters α, β, and $e_0 := E_0/\bar{t}$. (The upper bound of the support of $E_i/\bar{t}$ is 1.) These three parameters pin down m_1, m_2, and m_{3+}:

$$m_1 = \frac{\alpha + \beta e_0}{\alpha + \beta}, \qquad m_2 = \frac{(1-e_0)^2 \alpha\beta}{(\alpha+\beta)^2(1+\alpha+\beta)} + m_1^2$$

$$m_{3+} = \frac{\alpha+\beta-1}{(1-e_0)(\beta-1)} - 1 - m_1 - m_2$$

It can be shown that this is a one-to-one mapping from (α, β, e_0) to (m_1, m_2, m_{3+}).[18] One can use this one-to-one mapping for comparative statics with respect to, say, m_{3+} while keeping m_1 and m_2 constant. Figure 1 shows a Beta distribution with $\alpha = 3$, $\beta = 5$ and $e_0 = 0.1$ (dashed line). For this distribution, $L'(0) = 0$, i.e., the rating agency sets the cutoff at exactly the socially optimal level $x = 0$. For the

[18] The mapping in the opposite direction can be derived in closed form, but the resulting expressions are rather long and uninformative and therefore omitted. m_1 and m_2 are the well-known first two moments of the Beta distribution. m_{3+} can be derived by observing that $E[(1-y)^{-1}(1-e_0)^{-1}] = E[\sum_{k=0}^{\infty}(e_0 + (1-e_0)y)^k] = E[\sum_{k=0}^{\infty} e^k] = 1 + m_1 + m_2 + m_{3+}$, where $e = E_i/\bar{t} = e_0 + (1-e_0)y$. For a Beta distribution with density $h(y) = y^{\alpha-1}(1-y)^{\beta-1}/B(\alpha,\beta)$, the expected value is

$$E\left[\frac{1}{1-y}\right] = \int_0^1 \frac{y^{\alpha-1}(1-y)^{\beta-2}}{B(\alpha,\beta)} dy = \frac{B(\alpha,\beta-1)}{B(\alpha,\beta)} = \frac{\alpha+\beta-1}{\beta-1},$$

where the last equality follows from the relation of the Beta to the Gamma function $B(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$, and the property $\Gamma(x+1) = x\Gamma(x)$ of the Gamma function, which imply

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta-1)(\beta-1)}{\Gamma(\alpha+\beta-1)(\alpha+\beta-1)} = \frac{\beta-1}{\alpha+\beta-1} B(\alpha,\beta-1).$$

Putting this together yields the expression for m_{3+}.

dotted line, m_1 and m_2 are kept constant and m_{3+} is reduced by 0.01. The dotted line has a fatter lower tail, which means that it has more mass at the bottom of the distribution. The mean and variance remain the same, but if a crisis hits, it is more likely to be severe. For the dotted distribution, $L'(0) < 0$, and hence the cutoff is negative, $x < 0$, which means that the rating criterion is too loose compared to the socially optimal ones. For the solid line, m_{3+} is increased by 0.01 while keeping m_1 and m_2 constant. For this distribution, $L'(0) > 0$ and hence $x > 0$, that is, the ratings are too strict compared to the socially optimal ones.

Figures 2, 3, and 4 illustrate the change of $L'(0)$ as m_{3+}, m_1, and m_2 are changed, respectively, while keeping the other parameters constant. The optimal cutoff for example can switch from a negative to a positive cutoff if the mean or the higher order skewness increase or if the variance decreases. For all values of m_1, m_2, and m_{3+}, the parameters α, β, and e_0 are within the permissible ranges.[19]

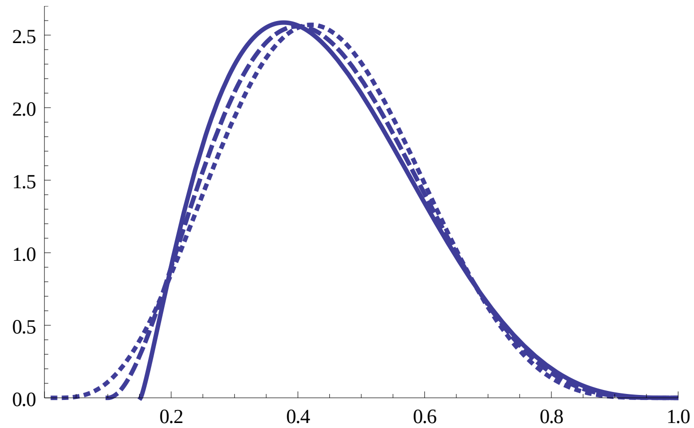

Figure 1: Density of $E_i/\bar{t}$ for $\alpha = 3$, $\beta = 5$, $e_0 = 0.1$ (dashed line). For the dotted line, m_{3+} is reduced by 0.01, for the solid line, m_{3+} is increased by 0.01, while m_1 and m_2 are kept constant. (The corresponding parameters are $\alpha = 4.4322$, $\beta = 5.8781$, $e_0 = 0.013363$ for the dotted and $\alpha = 2.23$, $\beta = 4.38985$, $e_0 = 0.151755$ for the solid distribution.)

5 Empirical Implications

Our model shows how the rating agency's incentive to be too lenient or too strict depends on the moments of the distribution of beliefs about aggregate uncertainty.

[19]The permissible ranges are $\alpha > 0$, $\beta > 0$, and $e_0 \in (0, 1)$.

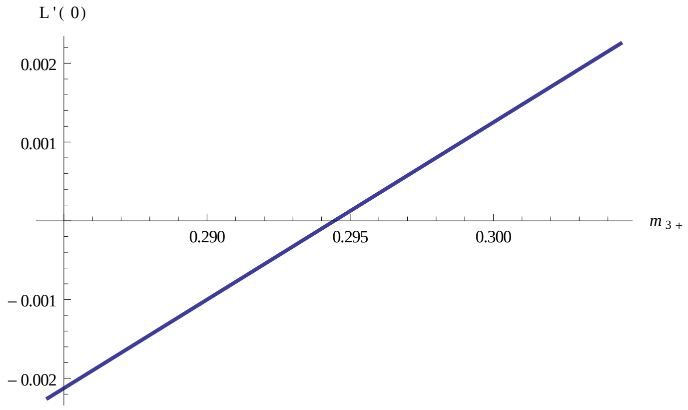

Figure 2: Values of $L'(0)$ as m_{3+} is changed and m_1 and m_2 are kept constant. Starting point is $\alpha = 3$, $\beta = 5$, $e_0 = 0.1$ (which corresponds to $m_1 = 0.4375$, $m_2 = 0.2125$, and $m_{3+} = 0.294444$) for which $L'(0) = 0$. Further parameters are normalized to $\bar{t} = 1$ and $f(0) = 1$.

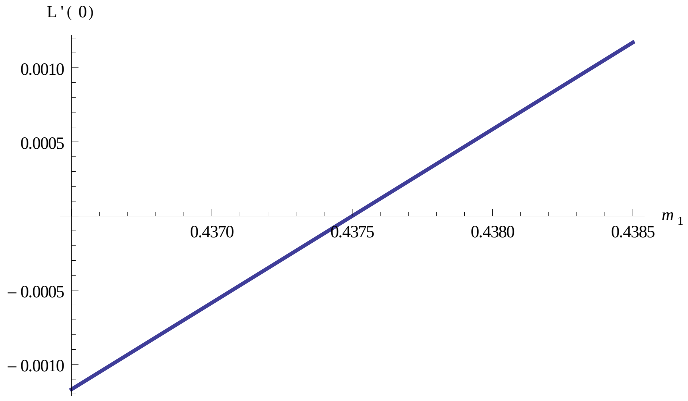

Figure 3: Values of $L'(0)$ as m_1 is changed and m_2 and m_{3+} are kept constant. Starting point is $\alpha = 3$, $\beta = 5$, $e_0 = 0.1$ (which corresponds to $m_1 = 0.4375$, $m_2 = 0.2125$, and $m_{3+} = 0.294444$) for which $L'(0) = 0$. Further parameters are normalized to $\bar{t} = 1$ and $f(0) = 1$.

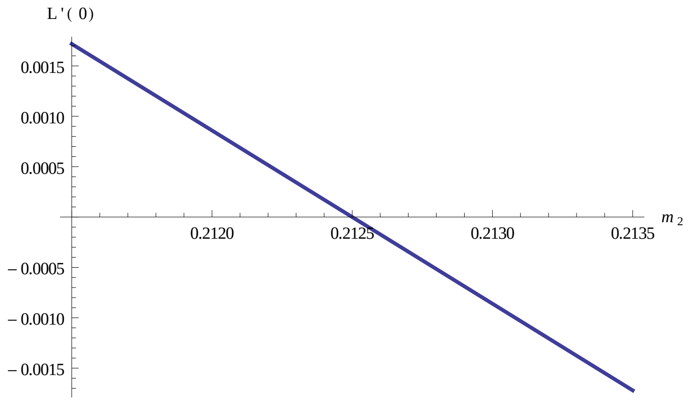

Figure 4: Values of $L'(0)$ as m_2 is changed and m_1 and m_{3+} are kept constant. Starting point is $\alpha = 3$, $\beta = 5$, $e_0 = 0.1$ (which corresponds to $m_1 = 0.4375$, $m_2 = 0.2125$, and $m_{3+} = 0.294444$) for which $L'(0) = 0$. Further parameters are normalized to $\bar{t} = 1$ and $f(0) = 1$.

Since these moments cannot be observed directly, one may wonder about the empirical content of our model.

First, it should be noted that an empirical estimate of the distribution of aggregate uncertainty is nontrivial, especially if the main concern is about the distribution of aggregate uncertainty shortly before a crisis. The reason is that few crises occur, so it is difficult to have larger amounts of data.

However, an empirical estimate of market participants' beliefs about the distribution of aggregate uncertainty can be obtained. We illustrate the basic idea of how to estimate these moments in a strongly stylized setup containing the core idea of the empirical strategy.

Consider the following stylized setup. There is an index for the bonds being sold by the firms in the market. Further, there is a market for financial derivatives based on this index. As an example, one can think of a subprime mortgage-backed securities index, such as ABX.HE.[20] An example of financial derivatives would be credit default options. Call options on the index can be bought in the first period of the model, before aggregate uncertainty is realized.[21] The options expire in the second period after the realization of aggregate uncertainty. Time is discrete and the options are European options.[22] Further, aggregate uncertainty is such that the mid-quality firms' beliefs are the same as the general market beliefs, formally, $\hat{\epsilon}_i = \epsilon_i$ for all i.[23] Suppose that the cut-off of the agency is close to 0 ($x \approx 0$), so that the value of the index $E_i(x)$ is well approximated by $E_i(0)$.

Further, assume that there exists a call option with strike price $y_i = E_i$ for each state of the world i. Without loss of generality, order the states of the world increasingly, i.e., $E_j > E_i$ if $j > i$. The second-period value of a call option with strike price y_j in state i is $E_i - y_j$ if $E_i > y_j$ and 0 if $E_i \leq y_j$. Denote the first-period

[20] ABX.HE was launched by CDS Indexco and Markit in 2006.

[21] For the sake of simplicity, we focus on call options on an index. One could also think of put options on an index or, e.g., mortgage credit default swap ABX indexes.

[22] In a discrete two-period model, it does not matter whether the option is European or American. In a continuous time model, calculations for American options are somewhat more complex, but standard and well known in the literature.

[23] A sufficient condition is that $\lambda_i = \tilde{\lambda}$ for all i, that is, aggregate uncertainty enters through changes of κ_i and μ_i for different states of the world i.

price of option j with strike price y_j by O_j. O_j is given by the market's expected value of the second period value (ignoring discounting):

$$O_j = \sum_{i=1}^{N} \epsilon_i \max\{E_i - y_j, 0\}. \tag{8}$$

The next proposition shows that given a set of call options, the information on their strike prices y_j and first-period prices O_j identifies the market's beliefs about the distribution of aggregate uncertainty: it identifies the probability ϵ_i of the expected quality E_i. The proof is provided in the Appendix.

Proposition 7. *Given strike prices and first period prices $\{(y_j, O_j)\}_{j=1}^{N}$, the probability mass function of the distribution of aggregate uncertainty is given by*

$$\epsilon_j = \frac{O_j - O_{j+1}}{y_{j+1} - y_j} - \frac{O_{j-1} - O_j}{y_j - y_{j-1}} \tag{9}$$

for interior states ($1 < j < N$) and

$$\epsilon_N = \frac{O_{N-1}}{y_N - y_{N-1}}, \qquad \epsilon_1 = 1 - \sum_{i=2}^{N} \epsilon_i.$$

at the boundaries.

In the following, we will derive an analogous, but more tractable, result for continuous distributions of E_i, since this makes the subsequent analysis clearer.

For the continuous distribution version, drop the index in E_i and denote the distribution of E by G. Assume that prices $O(y)$ for call options with a continuum of strike prices $y \in [\underline{t}, \overline{t}]$ are observed. Then $O(y)$ is given by

$$O(y) = \int_{\underline{t}}^{\overline{t}} \max\{E - y, 0\} dG(E) = \int_{y}^{\overline{t}} (E - y) dG(E).$$

It is easy to see that the distribution of aggregate uncertainty G is nonparametrically identifiable from the call option prices $O(y)$ by observing that the derivative can be rearranged to

$$O'(y) = -(1 - G(y)).$$

This result can also be related to the discrete version by considering the second derivative $O''(y) = g(y)$, which is the continuous analog of (9), since (9) is the discrete version of a second derivative.

In practice, one expects to observe fewer options than there are states of the world, so parametric assumptions are required to be able to estimate the distribution of E.

In the following, we make the parametric assumption that the distribution G is a polynomial with lower bound of support E_0 and upper bound $\bar{t}$. As an example, consider a cubic function

$$G(E) = a_1 + 2a_2E + 3a_3E^2 + 4a_4E^3.$$

The price of a call option will also be a polynomial function of the strike price y, since

$$O(y) = \int_y^{\bar{t}} (1 - G(E))dE = a_0 + a_1y + a_2y^2 + a_3y^3 + a_4y^4,$$

where

$$a_0 = -\sum_{i=1}^{4} a_i\bar{t}^i.$$

Suppose we observe data for five call options with strike prices $\{y_j\}_{j=1}^5$ and option prices $\{O(y_j)\}_{j=1}^5$. In this case, the parameters $\{a_i\}_{i=0}^4$ are given by the system of linear equations

$$O(y_j) = a_0 + a_1y_j + a_2y_j^2 + a_3y_j^3 + a_4y_j^4, \qquad j = 1, ..., 5. \tag{10}$$

As long as the matrix $\left[y_j^i\right]_{j=1,...,5;i=0,...,4}$ is nonsingular, the system of equations (10) yields a unique solution for the variables $\{a_i\}_{i=0}^4$. Note that E_0 and $\bar{t}$ are uniquely pinned down by the parameters $\{a_i\}_{i=0}^4$ and by the equations $G(E_0) = 0$ and $G(\bar{t}) = 0$.[24]

Given the distribution G of E, we can obtain the distribution of $e = E/\bar{t}$ and

[24]While $G(E)$ has several roots due to G's being a polynomial, the solution of $G(E_0) = 0$ is unique nonetheless. This is because of the constraints $G'(E) > 0$ for $E \in [E_0, \bar{t}]$ and $y_j \in [E_0, \bar{t}]$ for all j. By the same reasoning, there is a unique solution of $G(\bar{t}) = 1$.

the moments m_1, m_2, m_{3+} of e. This in turn yields

$$S = 1 - \frac{1}{1 + m_1 + m_2 + m_{3+}} - \frac{m_2}{m_1}$$

from equation (4) in Proposition 3 and determines the sign of the marginal profit $\Pi'(0)$ at $x = 0$. Table 1 provides examples of hypothetical observed prices of call options and the corresponding estimated parameters, moments, and S. For the first set of observations (first line), the rating agency has an incentive to choose the cutoff at the first best level $x = 0$. For the second line the agency has an incentive to choose a negative cutoff, and for the third line, a positive cutoff.

observed prices					estimated parameters				moments			S
O_1	O_2	O_3	O_4	O_5	a_1	a_2	a_3	a_4	m_1	m_2	m_{3+}	
11	8.0	5.7	3.9	2.6	-0.37	0.0096	-0.000030	3.4×10^{-8}	0.35	0.15	0.30	0.0
5.9	4.1	2.7	1.8	1.1	0.0	0.0077	-0.000026	3.4×10^{-8}	0.24	0.095	0.12	-0.073
14	10	7.2	4.9	3.2	-0.89	0.014	-0.000043	5.1×10^{-8}	0.41	0.20	0.37	0.019

Table 1: Example of parameter estimates for data on call option prices $O_j = O(y_j)$ for strike prices $(y_1, y_2, y_3, y_4, y_5) = (80, 90, 100, 110, 120)$.

Two remarks are in order. First, by a symmetric argument, the above results also hold for put options and not just call options. Second, for mortgage backed securities, one can construct a mortgage backed security index using the prices of mortgage backed securities. One can further construct synthetic derivatives for this index using credit default swaps and credit default options.

We have illustrated the basic idea behind an empirical strategy to estimate the moments of the distribution of beliefs about aggregate uncertainty. To practically apply this strategy, several additional steps are required, which are outside the scope of this paper. First, one needs to construct synthetic call options for the index of the bonds being rated. Second, the pricing of options in a multi-period environment is more complicated than the simple two-period setup we used for illustrative purposes. These problems are far from trivial, but well studied in Finance, see e.g., Hull (2009). Additionally, one could use a different parametrization for G instead of the polynomial parametrization or, if sufficiently many observations are available, one could possibly even use a nonparametric estimate of the function $O(y)$ given the observations $\{y_j, O(y_j)\}_j$. Further, one would also want to estimate the confidence interval for S.

6 Risk Aversion

In the main part of this paper, we have assumed that investors are risk neutral and we have shown that it is optimal for the agency to pool all types above a cutoff into one rating class. Doherty, Kartasheva, and Phillips (2012) extend the model of Lizzeri (1999) by allowing investors to be risk averse, and they show that if the level of risk aversion is sufficiently high, the rating agency rates types above a cutoff in several rating classes.

First, following the paper of Doherty et al. (2012), we provide a simplified hybrid model incorporating risk aversion and aggregate uncertainty. We show that introducing risk aversion in a model with several states of the world can also yield several rating classes. In this case, our previous analysis can be interpreted as determining the optimal cutoff of the lowest investment grade rating class (e.g., BBB). Second, we provide a numerical example to show that the effects of the moments of the expected quality distribution on the optimal cutoff have the same sign as before even with risk aversion and several rating classes.

We provide the simplest possible setup which is rich enough to illustrate the idea. Assume that buyers are risk averse. Their utility of an asset is equal to t but their expected utility depends on both the mean and the variance of the quality of the asset they buy. We include a second mass point at $\bar{t}_2$, $\bar{t}_2 \geq \bar{t}$, with mass γ_i in state i. To avoid confusion, write $\bar{t}$ as $\bar{t}_1$. See Figure 6.

If buyers are risk averse, a welfare maximizing rating strategy needs to perfectly disclose the type of each asset with a positive value, because any kind of pooling and being vague about a firm's quality leads to a welfare loss. However, such a strategy cannot be optimal for the rating agency.[25] To analyze a general model with risk

[25] To ensure that all firms with $t \geq 0$ are willing to pay the rating fee under full disclosure, the rating fee has to be 0.

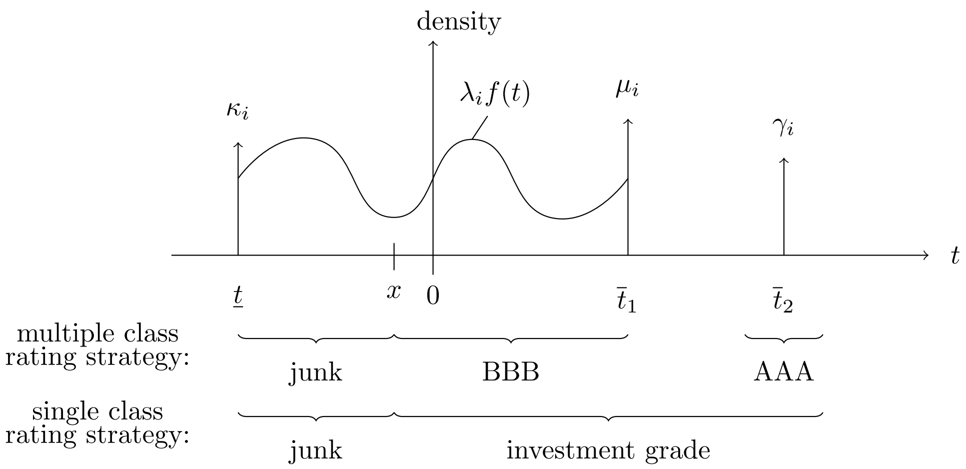

Figure 5: κ_i, μ_i and γ_i are the mass points at $\underline{t}$, $\bar{t}_1$ and $\bar{t}_2$ in state i. λ_i is the mass in state i that is allotted to the types $t \in (\underline{t}, \bar{t})$ with the distribution F.

aversion is outside the scope of this paper. In the following, we compare two rating strategies: (i) pooling all types above a cutoff into one rating class, which is the optimal strategy without risk aversion, and (ii) a strategy in which the agency only pools the low types but rates the high types separately. Doherty et al. (2012) show that strategy (ii) is optimal in a model with one state of the world if the level of risk aversion is sufficiently high.

Analogously to the case without risk aversion, we derive the profit of the agency if it pools all types above a cutoff x into one class. The expected type above a cutoff x is

$$Q_i(x) := E[t|t \geq x]$$

and the variance is

$$\sigma_i(x) := \mathrm{Var}[t|t \geq x].$$

The buyer's valuation of the asset of a seller in this rating class is

$$Q_i(x) - a\sigma_i(x)$$

where a is the measure of risk aversion. If $a = 0$, the buyers are risk neutral and the model is equivalent to the one before.

The profit of the agency if it pools all types is

$$\hat{\Pi}(x) := \left(\sum_i (\lambda_i(1-F(x)) + \mu_i + \gamma_i)\epsilon_i\right)\left(\sum_i \hat{\epsilon}_i(Q_i(x) - a\sigma_i(x))\right)$$
$$=(1-F(x)+\tilde{\mu}+\tilde{\gamma})\sum_i \hat{\epsilon}_i(Q_i(x) - a\sigma_i(x))$$

where $\tilde{\gamma}$ is the expected value of γ: $\tilde{\gamma} = \sum_i \epsilon_i \gamma_i$.

Alternatively, the rating agency can pool $t \in [x, \bar{t}_1]$ and rate $\bar{t}_2$ separately, as shown in Figure 6. If the agency rates types $\bar{t}_2$ in a separate class, these sellers are willing to pay a high rating fee (up to $\bar{t}_2$), and therefore the rating fee is determined by the sellers in the class $t \in [x, \bar{t}_1]$. Keeping the cutoff x constant, the mass of rated firms is the same for both strategies and the rating fee determines which rating strategy yields higher profits. If the agency pools types $t \in [x, \bar{t}_1]$, the expected type in this rating class is smaller than $Q_i(x)$ but the variance is also smaller than $\sigma_i(x)$. Thus, it is not straightforward to see under which strategy the rating fee can be higher.

Now, we derive sufficient conditions such that the agency prefers to rate $\bar{t}_2$ separately instead of pooling all types above x in one rating class. Define $z_i := \gamma_i \bar{t}_2$ and $\tilde{z}$ as the expected value of z_i, $\tilde{z} := \sum_i \epsilon_i z_i$. Rewrite $\tilde{z}$ as $\tilde{z} = \bar{t}_2 \tilde{\gamma}$, which can be interpreted as the agency's profit if it charges a rating fee of $\bar{t}_2$ and rates only firms with type $\bar{t}_2$. Recall that $\Pi(x)$ is defined as the profit if the agency rates only $t \in [x, \bar{t}_1]$ and pools them all in one class. The sufficient condition is given in the following proposition, the proof being provided in the Appendix.

Proposition 8. *Take an arbitrary cutoff x. For any $\tilde{z}$ with $\tilde{z} \leq \Pi(x)$, there exists a $\overline{T}_2$ such that for all $\bar{t}_2 \geq \overline{T}_2$ the rating agency is better off pooling $t \in [x, \bar{t}_1]$ and rating $\bar{t}_2$ in a separate class than pooling all types above x in one rating class.*

Since investors are risk averse, their expected utility buying an asset in a given rating class decreases if the variance inside this rating class becomes larger. If the variance is sufficiently large, investors are not willing to pay any positive price for an asset even if the expected quality is positive. Thus, if the variance is large, the agency is better off splitting the types into several rating classes in order to reduce the variance inside one class and to increase investors' willingness to pay for an asset. The condition that $\tilde{z} \leq \Pi(x)$ ensures that the agency does not prefer to charge a rating fee of $\bar{t}_2$ and to exclude firms with $t < \bar{t}_2$ from the rating.

Risk aversion not only has the effect of multiple rating classes becoming optimal, but it also has an additional effect on the optimal cutoff. Increasing the cutoff reduces the variance in a rating class and this can give additional incentives to increase the cutoff.[26] In the following we provide numerical examples in which we show that the effects of the first, second, and higher moments are similar to our analysis without risk aversion.[27] In the numerical example we have four states of the world. We take the Generalized Pareto distribution $F(t) = 1 - ((1-t)/2)^3$ for $t \in (-1, 1)$ and fix $\bar{t}_1 = 1$. This gives us $E_0 = 1/4$. We fix $\lambda = 5$, $\bar{t}_2 = 110$, and $\nu_i = 0.0001$ for all i. The states only differ in the weights μ_i at the mass point at $\bar{t}_1$, with $\mu_1 = 0.03$, $\mu_2 = 0.2$, $\mu_3 = 0.4$, and $\mu_4 = 0.7$. Changing the moments of the aggregate distribution, we keep the distribution inside a state constant (and therefore also the expected type) and only vary the probabilities for the different states. There is a one-to-one mapping from $(\epsilon_1, \epsilon_2, \epsilon_3)$ to (m_1, m_2, m_{3+}) and the fourth probability is pinned down by $\epsilon_4 = 1 - \epsilon_1 - \epsilon_2 - \epsilon_3$. For all values of the example, the probabilities are in $[0, 1]$.

Figures 6, 7, and 8 illustrate the change of the optimal cutoff as m_{3+}, m_1 and m_2 are changed while keeping the other moments constant. The solid line is the optimal cutoff for $a = 0$, the dashed line for $a = 0.01$ and the dotted-dashed line for $a = 0.02$. If investors are risk neutral, $a = 0$, the agency pools all types above the cutoff in one class. For $a = 0.01$ and $a = 0.02$, investors are risk averse and the agency prefers to pool all types $t \in [x, \bar{t}_1]$ in one class and to rate $\bar{t}_2$ separately. Note that increasing the level of risk aversion leads to an increase in the optimal cutoff

[26] Doherty et al. (2012) show that the optimal cutoff can be positive even with only one state of the world if the level of risk aversion is sufficiently high.

[27] In the main part of the paper the moments were defined for the distribution of the expected type in $[0, \bar{t}]$ (scaled by $\bar{t}$). For the sake of comparison, in the numerical examples the moments are defined for the distribution of the expected type in $[x, \bar{t}_1]$ and thus, the expected type is not influenced by changes in the mass on $\bar{t}_2$. We deviate from our previous analysis by taking the threshold x as the lower bound of the interval. In this way we can determine the optimal cutoff explicitly and not only its sign.

x^*. The figures show that our results from the main part of the paper carry over to a setup including risk aversion: keeping the other moments constant, a higher mean, a lower variance, or an increase in the higher order skewness lead to an increase in the optimal cutoff. For changes with the opposite sign, the optimal cutoff decreases.

Hence, results from the risk-neutral setup carry over qualitatively to a setup with risk aversion. A setup with risk aversion comes at the cost of being analytically intractable, but has the desirable feature that it makes the more realistic prediction of multiple rating classes.

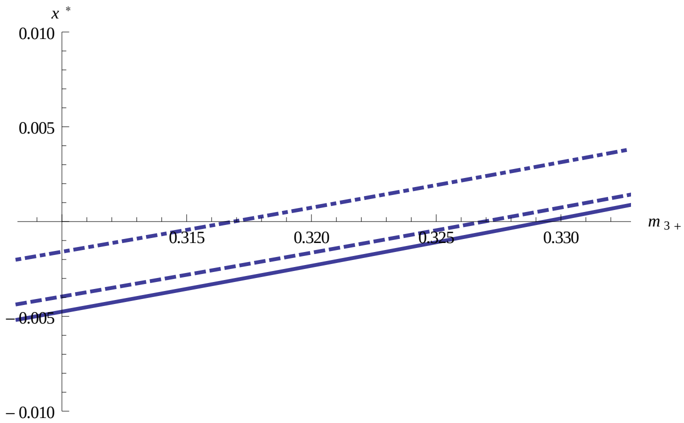

Figure 6: Values of the optimal threshold x^* as m_{3+} is changed and m_1 and m_2 are kept constant. For the solid line $a = 0$, for the dashed line $a = 0.01$ and for the dotted-dashed line $a = 0.02$. The rating strategy for the solid line is to pool all types above x. For the dashed and dotted-dashed line all types in $[x, \bar{t}_1]$ are pooled and $\bar{t}_2$ is rated separately. (The starting values are $\epsilon_i = 1/4$ for all i. This implies $m_1 = 0.47627$, $m_2 = 0.244859$ and as a starting value $m_{3+} = 0.321538$.)

7 Conclusions

We have considered the profit maximizing rating strategy of a rating agency in the face of aggregate uncertainty. We have shown that with risk-neutral investors, it is still optimal for the rating agency—as in a setup without aggregate uncertainty—to choose only one rating class for rated firms and to not rate the remaining firms.

The model's predictions about the cutoff for the rating class differ strikingly from the predictions of a model without aggregate uncertainty: the rating agency has more of an incentive to be too lenient if the expected average quality is small, the variance

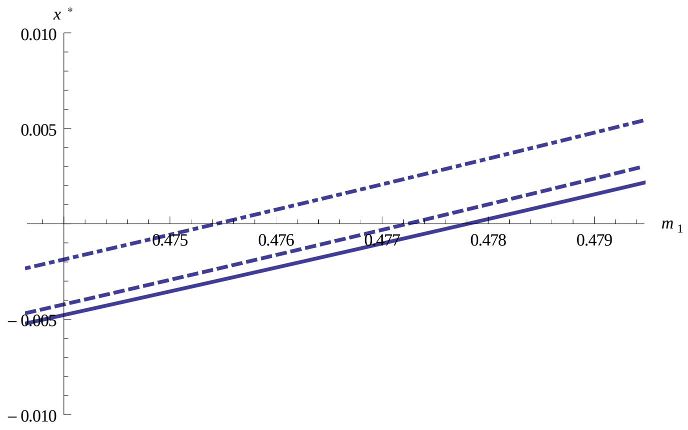

Figure 7: Values of the optimal threshold x^* as m_1 is changed and m_2 and m_{3+} are kept constant. For the solid line $a = 0$, for the dashed line $a = 0.01$ and for the dotted-dashed line $a = 0.02$. The rating strategy for the solid line is to pool all types above x. For the dashed and dotted-dashed line all types in $[x, \bar{t}_1]$ are pooled and $\bar{t}_2$ is rated separately. (The starting values are $\epsilon_i = 1/4$ for all i. This implies $m_2 = 0.244859$, $m_{3+} = 0.321538$ and as a starting value $m_1 = 0.47627$.)

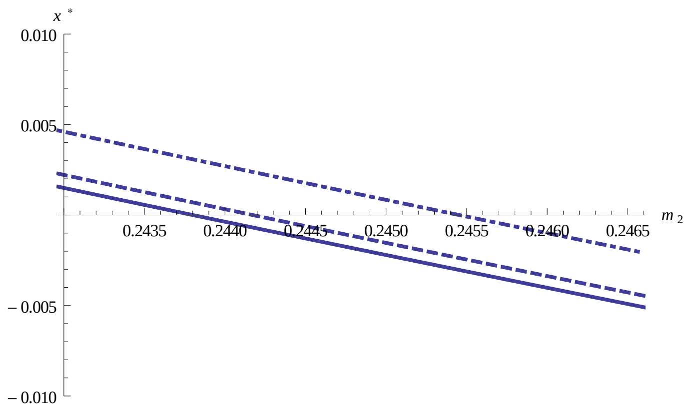

Figure 8: Values of the optimal threshold x^* as m_2 is changed and m_1 and m_{3+} are kept constant. For the solid line $a = 0$, for the dashed line $a = 0.01$ and for the dotted-dashed line $a = 0.02$. The rating strategy for the solid line is to pool all types above x. For the dashed and dotted-dashed line all types in $[x, \bar{t}_1]$ are pooled and $\bar{t}_2$ is rated separately. (The starting values are $\epsilon_i = 1/4$ for all i. This implies $m_1 = 0.47627$, $m_{3+} = 0.321538$ and as a starting value $m_2 = 0.244859$.)

large, and the higher order skewness small. For larger averages, smaller variances, and larger higher order skewness, the opposite holds: the rating agency has more of an incentive to be too strict. These results can be interpreted as ratings' having either a procyclical or an countercyclical effect. We outlined an empirical strategy to estimate the moments of aggregate uncertainty which can be used to determine which effect dominates. Our analysis identifies one up to now unconsidered factor that affects the rating strategy of an agency—aggregate uncertainty—and thereby sheds further light on the behavior of rating agencies. In line with our model, one disturbing effect of using ratings as the basis for financial regulation is that a possible procyclicality of ratings leads to a procyclicality of the capital adequacy requirements for banks, and hence to a procyclicality of lending. One solution is to avoid using ratings as the basis for financial regulation. Another is to counterbalance the procyclicality of ratings by adding a countercyclicality to the capital adequacy requirements that are based on ratings.

The usual disclaimer for the policy implications holds. This paper is about a thorough analysis of the effects of aggregate uncertainty, shutting down other effects such as reputation cycles, imperfect rating technology, and competition between agencies. Further, the implications of the theory depend on the empirical moments of the distribution of aggregate uncertainty. Hence, an empirical analysis is needed to estimate these moments and the relative magnitude of the different effects. Our paper provides a starting point for such an empirical analysis. This paper also serves as a word of caution: using a distribution which is pinned down by its mean and variance (e.g., the often used normal distribution) for an empirical analysis will neglect the impact of the higher order skewness. However, the skewness is crucial for the incentive of the rating agency to distort ratings.

References

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

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

Bouvard, M. and R. Levy (2009): "Humouring both parties: a model of two-sided reputation," Tech. rep., working paper.

Doherty, N. A., A. V. Kartasheva, and R. D. Phillips (2012): "Information effect of entry into credit ratings market: The case of insurers' ratings," *Journal of Financial Economics*, 106, 308–330.

Hull, J. (2009): *Options, Futures and Other Derivatives*, Pearson.

Jensen, M. C. (1968): "The performance of mutual funds in the period 1945–1964," *Journal of Finance*, 23, 389–416.

Lerner, J. and J. Tirole (2006): "A model of forum shopping," *American Economic Review*, 1091–1113.

Lizzeri, A. (1999): "Information revelation and certification intermediaries," *RAND Journal of Economics*, 30, 214–231.

Mariano, B. (2008): "Do reputational concerns lead to reliable ratings?" Tech. rep., Financial Markets Group, London School of Economics and Political Science.

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

Myerson, R. B. (1981): "Optimal auction design," *Mathematics of Operations Research*, 6, 58–73.

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

Ottaviani, M. and P. Sørensen (2006): "Reputational cheap talk," *RAND Journal of Economics*, 37, 155–175.

Pollrich, M. and L. Wagner (2013): "Informational opacity and honest certification," *SFB/TR 15 Discussion Paper No. 481.*

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

Stahl, K. and R. Strausz (2014): "Certification and Market Transparency," Tech. rep., Humboldt-Universität zu Berlin, Wirtschaftswissenschaftliche Fakultät.

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

Viswanatha, A. and K. Freifeld (2015): "S&P reaches $1.5 billion deal with U.S., states over crisis-era ratings," *Reuters*, February 3, 2015.

Appendix

A Omitted Proofs

Proof of Lemma 1. (i) Denote the class containing $\underline{t}$ by T_n. Define an alternative rating class $T_n^* = \{\underline{t}\} \cup \left[0, \bar{t}\right]$. Observe that $\kappa_i \underline{t} + \lambda_i \int_0^{\bar{t}} t dF(t) + \mu_i \bar{t}$ is the expected quality in T_n^* in state i, $E\left[T_{n,i}^*\right]$, and by Assumption 2 this is smaller than zero. The expected quality in class T_n in state i, $E[T_{n,i}]$, can be larger or smaller than $E\left[T_{n,i}^*\right]$. If $E[T_{n,i}]$ is smaller than $E\left[T_{n,i}^*\right]$, it follows directly that $E[T_{n,i}] < 0$.

If $E\left[T_{n,i}\right]$ is larger than $E\left[T_{n,i}^*\right]$, T_n must include types $t \in [E\left[T_{n,i}^*\right], 0]$ to raise the expected quality. Including negative qualities in T_n can increase the expected type in comparison to T_n^* but the expected type $E[T_{n,i}]$ stays negative. Therefore, the willingness to pay for a rating in category T_n is negative and the rating agency prefers not to have category T_n.

(ii) Take a rating strategy $\{\tilde{T}_m\}_{m=1}^{\tilde{M}}$. Assume that $\bar{t}$ is not in any $\tilde{T}_m$. Define for all rating classes $\tilde{T}_m$ the expected value

$$E_m^* = \frac{\int_{t \in \tilde{T}_m} t dF(t)}{\int_{t \in \tilde{T}_m} dF(t)} \tag{11}$$

which is constant over all states of the world. The price is determined by the lowest willingness to pay, $\min_m E_m^*$, and the expected mass of rated firms $\sum_i \epsilon_i \sum_m \int_{t \in \tilde{T}_m} \lambda_i dF(t)$.

Using $\sum_i \epsilon_i \lambda_i = 1$, we get for profits

$$\tilde{\Pi} = \left[\min_m E_m^*\right] \left\{ \sum_{m=1}^{\tilde{M}} \int_{t \in \tilde{T}_m} dF(t) \right\} \tag{12}$$

Now take a rating strategy with $M = \tilde{M} + 1$, $T_m = \tilde{T}_m$ for $m \leq \tilde{M}$ and $T_M = \{\bar{t}\}$. Including the $\bar{t}$ types adds expected mass $\tilde{\mu}$ to the mass of rated firms. Hence, expected profits are

$$\Pi = \left[\min\left(\{\bar{t}\} \cup \{E_m^*\}_{m=1}^{\tilde{M}}\right)\right] \left\{ \tilde{\mu} + \sum_{m=1}^{\tilde{M}} \int_{t \in T_m} dF(t) \right\} \tag{13}$$

Since in (13) the expression in square brackets is weakly greater than in (12) and the expression in curly braces is strictly greater in (13) than in (12), we have $\Pi > \tilde{\Pi}$. □

Proof of Lemma 2. One has that

$$\sum_i \left(\frac{\epsilon_i \mu_i}{\tilde{\mu}} - \frac{\epsilon_i \lambda_i}{\tilde{\lambda}} \right) = 0$$

because $\sum_i \epsilon_i \mu_i / \tilde{\mu} = \sum_i \epsilon_i \lambda_i / \tilde{\lambda} = 1$. Define two sets of states of the world: $i \in A$ if $\frac{\mu_i \epsilon_i}{\lambda_i} > \tilde{\mu}$ and $i \in B$ if $\frac{\mu_i \epsilon_i}{\lambda_i} \leq \tilde{\mu}$. Thus

$$\sum_{i \in A} \left(\frac{\epsilon_i \mu_i}{\tilde{\mu}} - \frac{\epsilon_i \lambda_i}{\tilde{\lambda}} \right) + \sum_{i \in B} \left(\frac{\epsilon_i \mu_i}{\tilde{\mu}} - \frac{\epsilon_i \lambda_i}{\tilde{\lambda}} \right) = 0$$

, and, multiplying by a constant c,

$$\sum_{i \in A} \left(\frac{\epsilon_i \mu_i}{\tilde{\mu}} - \frac{\epsilon_i \lambda_i}{\tilde{\lambda}} \right) c + \sum_{i \in B} \left(\frac{\epsilon_i \mu_i}{\tilde{\mu}} - \frac{\epsilon_i \lambda_i}{\tilde{\lambda}} \right) c = 0. \tag{14}$$

The expected quality in state i is

$$E_i = \frac{\lambda_i \int_{t \in T} t dF + \mu_i \bar{t}}{\lambda_i \int_{t \in T} dF + \mu_i} = \frac{\int_{t \in T} t dF + \mu_i / \lambda_i \bar{t}}{\int_{t \in T} dF + \mu_i / \lambda_i}$$

and is increasing in μ_i / λ_i. Define c as $\frac{\int_{t \in T} t dF + \tilde{\mu} \bar{t}}{\int_{t \in T} dF + \tilde{\mu}}$. The expected quality E_i for $i \in A$

is larger than c and $E_i < c$ for $i \in B$. It follows that

$$\sum_{i\in A}\left(\frac{\epsilon_i\mu_i}{\tilde{\mu}} - \frac{\epsilon_i\lambda_i}{\tilde{\lambda}}\right)c < \sum_{i\in A}\left(\frac{\epsilon_i\mu_i}{\tilde{\mu}} - \frac{\epsilon_i\lambda_i}{\tilde{\lambda}}\right)E_i \tag{15}$$

and

$$\sum_{i\in B}\left(\frac{\epsilon_i\mu_i}{\tilde{\mu}} - \frac{\epsilon_i\lambda_i}{\tilde{\lambda}}\right)c < \sum_{i\in B}\left(\frac{\epsilon_i\mu_i}{\tilde{\mu}} - \frac{\epsilon_i\lambda_i}{\tilde{\lambda}}\right)E_i. \tag{16}$$

Using the inequalities (15) and (16) gives us

$$\begin{aligned}&\sum_{i\in A}\left(\frac{\epsilon_i\mu_i}{\tilde{\mu}} - \frac{\epsilon_i\lambda_i}{\tilde{\lambda}}\right)E_i + \sum_{i\in B}\left(\frac{\epsilon_i\mu_i}{\tilde{\mu}} - \frac{\epsilon_i\lambda_i}{\tilde{\lambda}}\right)E_i\\ >&\sum_{i\in A}\left(\frac{\epsilon_i\mu_i}{\tilde{\mu}} - \frac{\epsilon_i\lambda_i}{\tilde{\lambda}}\right)c + \sum_{i\in B}\left(\frac{\epsilon_i\mu_i}{\tilde{\mu}} - \frac{\epsilon_i\lambda_i}{\tilde{\lambda}}\right)c\end{aligned}$$

which is equal to 0 by (14). Therefore,

$$\sum_i\left(\frac{\epsilon_i\mu_i}{\tilde{\mu}}E_i\right) - \sum_i\left(\frac{\epsilon_i\lambda_i}{\tilde{\lambda}}E_i\right) > 0.$$

Since $\sum_i\left(\frac{\epsilon_i\mu_i}{\tilde{\mu}}E_i\right)$ is the willingness to pay for a rating for type $\bar{t}$ and $\sum_i\left(\frac{\epsilon_i\lambda_i}{\tilde{\lambda}}E_i\right)$ for type $t \in (\underline{t}, \bar{t})$, the lemma follows. □

Proof of Lemma 3. Label the rating class that includes $\bar{t}$ by $\tilde{T}_1$ and the remaining rating classes by $\tilde{T}_{-1} = \cup_{m\neq 1}\tilde{T}_m$. Denote the expected type of $\tilde{T}_1$ conditional on being in state i by $\tilde{E}_i = [\int_{t\in\tilde{T}_1} tdF(t) + \mu_i\bar{t}]/[\int_{t\in\tilde{T}_1} dF(t) + \mu_i]$. Denote the mass of all other classes by $\mu^* = \int_{t\in\tilde{T}_{-1}} dF(t)$ and the expected type by $t^* = [\int_{t\in\tilde{T}_{-1}} tdF(t)]/[\int_{t\in\tilde{T}_{-1}} dF(t)]$.

The profits for only one rating class $T_1 = \cup_m\tilde{T}_m$ are

$$\Pi = \left[\sum_i \hat{\epsilon}_i E_i\right](\mu^* + \tilde{\mu})$$

where

$$E_i = \frac{\lambda_i(\int_{t\in\tilde{T}_{-1}} tdF(t)) + \int_{t\in\tilde{T}_1} tdF(t) + \mu_i\bar{t}}{\lambda_i(\int_{t\in\tilde{T}_{-1}} dF(t) + \int_{t\in\tilde{T}_1} dF(t)) + \mu_i}$$

is the expected type in state i if there is only one rating class. The profits for separate rating classes $\{\tilde{T}_m\}$ are

$$\tilde{\Pi} = \left[\min\left(\{E_m^*\}_{m=1}^{\tilde{M}} \cup \left\{\sum_i \hat{\epsilon}_i \tilde{E}_i\right\}\right)\right](\mu^* + \tilde{\mu}),$$

where E_m^* is defined as in (11). Further, define the profit in case all rating classes $m \neq 1$ were merged, such that one had two rating classes $\tilde{T}_1$ and $\cup_{m=2}^{\tilde{M}}\tilde{T}_m$, as

$$\hat{\Pi} = \left[\min\left\{t^*, \sum_i \hat{\epsilon}_i \tilde{E}_i\right\}\right](\mu^* + \tilde{\mu}).$$

Since t^* is a weighted average of $\{\tilde{E}_m\}_{m=1}^{\tilde{M}}$, we have $t^* \geq \min\{\tilde{E}_m\}_{m=1}^{\tilde{M}}$ and therefore $\hat{\Pi} \geq \tilde{\Pi}$. (Note that Π, $\hat{\Pi}$, and $\tilde{\Pi}$ only differ in the expressions in square brackets.)

We will prove the lemma by contradiction. Assume to the contrary that separate classes are desirable, i.e., $\tilde{\Pi} > \Pi$. This implies $\hat{\Pi} > \Pi$, which is equivalent to

$$\min\left\{t^*, \sum_i \hat{\epsilon}_i \tilde{E}_i\right\} > \sum_i \hat{\epsilon}_i E_i,$$

by comparison of the expressions in square brackets. This condition is equivalent to both

$$t^* > \sum_i \hat{\epsilon}_i E_i \tag{17}$$

and

$$\sum_i \hat{\epsilon}_i \tilde{E}_i > \sum_i \hat{\epsilon}_i E_i \tag{18}$$

being satisfied at the same time.

The expected value E_i can be written as the weighted average of t^* and $\tilde{E}_i$ for every state i

$$\begin{aligned} E_i =& \frac{\lambda_i(\int_{t\in\tilde{T}_{-1}} t dF(t) + \int_{t\in\tilde{T}_1} t dF(t)) + \mu_i \bar{t}}{\lambda_i(\int_{t\in\tilde{T}_{-1}} dF(t) + \int_{t\in\tilde{T}_1} dF(t)) + \mu_i} \\ =& \frac{\lambda_i t^* \int_{t\in\tilde{T}_{-1}} dF(t) + \tilde{E}_i(\int_{t\in\tilde{T}_1} \lambda_i dF(t) + \mu_i)}{\lambda_i(\int_{t\in\tilde{T}_{-1}} dF(t) + \int_{t\in\tilde{T}_1} dF(t)) + \mu_i}. \end{aligned}$$

Solving for $\tilde{E}_i$, we get

$$\tilde{E}_i = E_i + \frac{\lambda_i \int_{t\in\tilde{T}_{-1}} dF(t)}{\lambda_i \int_{t\in\tilde{T}_1} dF(t) + \mu_i}(E_i - t^*)$$

Plugging $\tilde{E}_i$ into (18), we get

$$\sum_i \hat{\epsilon}_i \left(E_i + \frac{\lambda_i \int_{t\in\tilde{T}_{-1}} dF(t)}{\lambda_i \int_{t\in\tilde{T}_1} dF(t) + \mu_i}(E_i - t^*) \right) > \sum_i \hat{\epsilon}_i E_i$$

or equivalently

$$\sum_i \hat{\epsilon}_i \left(\frac{\int_{t\in\tilde{T}_{-1}} dF(t)}{\int_{t\in\tilde{T}_1} dF(t) + \mu_i/\lambda_i}(E_i - t^*) \right) > 0. \tag{19}$$

Define two sets of states of the world: $i \in A$ if $E_i \geq t^*$ and $i \in B$ if $E_i < t^*$. Then $\mu_i/\lambda_i > \mu_j/\lambda_j$ for all $i \in A$ and $j \in B$, as can be seen by checking that

$$E_i = \frac{(\int_{t\in\tilde{T}_{-1}} tdF(t) + \int_{t\in\tilde{T}_1} tdF(t)) + \mu_i/\lambda_i \bar{t}}{(\int_{t\in\tilde{T}_{-1}} dF(t) + \int_{t\in\tilde{T}_1} dF(t)) + \mu_i/\lambda_i}$$

is increasing in μ_i/λ_i. Put $c_A = \min\{\mu_i/\lambda_i | i \in A\}$ and $c_B = \max\{\mu_i/\lambda_i | i \in B\}$. Note that $c_A > c_B$.

Then (17) can be rewritten as

$$\sum_i \hat{\epsilon}_i (E_i - t^*) < 0$$

which is equivalent to

$$\sum_{i\in A} \hat{\epsilon}_i (E_i - t^*) + \sum_{i\in B} \hat{\epsilon}_i (E_i - t^*) < 0.$$

This implies

$$\left[\frac{\int_{t\in\tilde{T}_{-1}} dF(t)}{\int_{t\in\tilde{T}_1} dF(t) + c_A} \sum_{i\in A} \hat{\epsilon}_i (E_i - t^*) \right] + \left[\frac{\int_{t\in\tilde{T}_{-1}} dF(t)}{\int_{t\in\tilde{T}_1} dF(t) + c_B} \sum_{i\in B} \hat{\epsilon}_i (E_i - t^*) \right] < 0 \tag{20}$$

since

$$\frac{\int_{t\in\tilde{T}_{-1}} dF(t)}{\int_{t\in\tilde{T}_1} dF(t) + c_A} < \frac{\int_{t\in\tilde{T}_{-1}} dF(t)}{\int_{t\in\tilde{T}_1} dF(t) + c_B}$$

and the sum over $i \in A$ is positive and the sum over $i \in B$ is negative. Since

$\mu_i/\lambda_i \geq c_A$ for all $i \in A$ and $\sum_{i\in A} \hat{\epsilon}_i(E_i - t^*)$ positive, the first expression in square brackets in (20) is bounded from below by

$$\frac{\int_{t\in\tilde{T}_{-1}} dF(t)}{\int_{t\in\tilde{T}_1} dF(t) + c_A} \sum_{i\in A} \hat{\epsilon}_i(E_i - t^*) \geq \sum_{i\in A} \hat{\epsilon}_i \left(\frac{\int_{t\in\tilde{T}_{-1}} dF(t)}{\int_{t\in\tilde{T}_1} dF(t) + \mu_i/\lambda_i}(E_i - t^*) \right). \qquad (21)$$

The second expression in square brackets is bounded from below by

$$\sum_{i\in B} \hat{\epsilon}_i(E_i - t^*) \frac{\int_{t\in\tilde{T}_{-1}} dF(t)}{\int_{t\in\tilde{T}_1} dF(t) + c_B} \geq \sum_{i\in B} \hat{\epsilon}_i \frac{\int_{t\in\tilde{T}_{-1}} dF(t)}{\int_{t\in\tilde{T}_1} dF(t) + \mu_i/\lambda_i}(E_i - t^*). \qquad (22)$$

because $\mu_i/\lambda_i \leq c_B$ for all $i \in B$ and the negativity of $\sum_{i\in B} \hat{\epsilon}_i(E_i - t^*)$.

(20),(21) and (22) imply

$$\sum_{i\in A} \hat{\epsilon}_i \left(\frac{\int_{t\in\tilde{T}_{-1}} dF(t)}{\int_{t\in\tilde{T}_1} dF(t) + \mu_i/\lambda_i}(E_i - t^*) \right) + \sum_{i\in B} \hat{\epsilon}_i \frac{\int_{t\in\tilde{T}_{-1}} dF(t)}{\int_{t\in\tilde{T}_1} dF(t) + \mu_i/\lambda_i}(E_i - t^*) < 0$$

which contradicts (19). □

Proof of Lemma 4. Assume that $\hat{T}$ is not convex. Take a convex set T' such that it has the same expected mass of rated firms ($\int_{t\in\hat{T}} dF(t) = \int_{t\in T'} dF(t)$) and both sets include $\bar{t}$. Recall that the profit is

$$\Pi(T) = \left[\sum_i \hat{\epsilon}_i \frac{\lambda_i \int_{t\in T} t dF(t) + \mu_i \bar{t}}{\lambda_i \int_{t\in T} dF(t) + \mu_i} \right] \left(\int_{t\in T} dF(t) + \tilde{\mu} \right).$$

Since $\int_{t\in\hat{T}} dF(t) = \int_{t\in T'} dF(t)$, comparing the profits $\Pi(\hat{T})$ and $\Pi(T')$ boils down to comparing the willingness to pay for $\hat{T}$ and T', which is given in square brackets. Since $\hat{T}$ is not convex, there is at least one unrated hole in the middle and it is possible to rate the mass in the holes instead of rating some types below with the same mass. This increases $\int_{t\in T} t dF(t)$, while the mass of rated types stays the same. It follows that $\frac{\lambda_i \int_{t\in T} t dF(t) + \mu_i \bar{t}}{\lambda_i \int_{t\in T} dF(t) + \mu_i}$ is greater for T' than for $\hat{T}$ and hence $\Pi(T') > \Pi(\hat{T})$. Therefore, it is optimal to rate a set T which is convex and includes $\bar{t}$. □

Proof of Lemma 5. Analogously to the proof of Lemma 2, define two sets of states of the world: $i \in A$ if $\frac{\mu_i}{\lambda_i} > \frac{\tilde{\mu}}{\tilde{\lambda}}$ and $i \in B$ if $\frac{\mu_i}{\lambda_i} \leq \frac{\tilde{\mu}}{\tilde{\lambda}}$. Then $\sum_i \epsilon_i (\mu_i - \lambda_i \tilde{\mu}) = 0$, which we can write as $\sum_A \epsilon_i (\mu_i - \lambda_i \tilde{\mu}) + \sum_B \epsilon_i (\mu_i - \lambda_i \tilde{\mu}) = 0$. Multiplied by a constant

$c = \frac{1}{(1-F)+\tilde{\mu}}$, the expression is still equal to 0. For $i \in A$, $\frac{1}{(1-F)+\mu_i/\lambda_i}$ is smaller than $\frac{1}{(1-F)+\tilde{\mu}}$ and for $i \in B$ it is the other way round. It follows that

$$\sum_A \epsilon_i \left(\mu_i - \lambda_i \tilde{\mu}\right) \frac{1}{(1-F)+\mu_i/\lambda_i} + \sum_B \epsilon_i \left(\mu_i - \lambda_i \tilde{\mu}\right) \frac{1}{(1-F)+\mu_i/\lambda_i} < 0$$

because $\mu_i - \lambda_i \tilde{\mu}$ is positive for $i \in A$ and negative for $i \in B$. This is equivalent to $\frac{(\bar{t}-E_0)(1-F)}{\tilde{\lambda}(1-F)+\tilde{\mu}} \sum_i \hat{\epsilon}_i \frac{\tilde{\lambda}\mu_i - \tilde{\mu}\lambda_i}{\lambda_i(1-F)+\mu_i} < 0$, and thus $\sum_i \hat{\epsilon}_i \frac{\lambda_i(1-F)E_0+\mu_i\bar{t}}{\lambda_i(1-F)+\mu_i} < \frac{\tilde{\lambda}(1-F)E_0+\tilde{\mu}\bar{t}}{\tilde{\lambda}(1-F)+\tilde{\mu}}$. □

Proof of Proposition 2.

$$\begin{aligned} L'(x) =& \Pi'(x) - W'(x) \\ =& f(x)\left((\hat{E} - \frac{1-F(x)+\tilde{\mu}}{f}\hat{E}') - (\tilde{E} - \frac{1-F(x)+\tilde{\mu}}{f}\tilde{E}') \right) \\ =& f(x)\left(\hat{E} - \tilde{E} + \frac{1-F(x)+\tilde{\mu}}{f}(\tilde{E}' - \hat{E}') \right). \end{aligned} \tag{23}$$

We know that $\hat{E} \geq \tilde{E}$ but the sign of $\tilde{E}' - \hat{E}'$ can go in both directions.

Next, we rewrite (23) such that we can express $L'(x)$ only in terms of E_i, $\tilde{E}$, and $\hat{E}$. The derivative of $\tilde{E}$ with respect to x is

$$\sum_i \hat{\epsilon}_i \frac{\partial E_i}{\partial x} = \sum_i \hat{\epsilon}_i \frac{E_i - x}{\lambda_i(1-F(x)) + \mu_i} f(x)\lambda_i.$$

and analogously it can be shown that

$$\frac{\partial \hat{E}}{\partial x} = \frac{\hat{E} - x}{1 - F(x) + \tilde{\mu}} f(x).$$

Using these two expressions in (23), we can write

$$\begin{aligned} L'(x) =& f(x)\left(\hat{E} - \tilde{E} + \frac{1-F(x)+\tilde{\mu}}{f} \left(\sum_i \hat{\epsilon}_i \frac{E_i - x}{\lambda_i(1-F(x)) + \mu_i}\lambda_i f - \frac{\hat{E}-x}{1-F(x)+\tilde{\mu}} f \right) \right) \\ =& f(x)\left(\hat{E} - \tilde{E} + (1-F(x)+\tilde{\mu}) \sum_i \hat{\epsilon}_i \frac{E_i - x}{\lambda_i(1-F(x)) + \mu_i}\lambda_i - (\hat{E} - x) \right). \end{aligned}$$

From the definitions of E_i and $\hat{E}$ we derive $\mu_i = \lambda_i(1-F(x))\frac{E_i - E_0}{\bar{t}-E_i}$ and $\tilde{\mu} = (1 -$

$F(x))\frac{\hat{E}-E_0}{\bar{t}-\hat{E}}$, which leads to

$$
\begin{aligned}
L'(x) =& f(x)\left(x-\tilde{E}+(1-F(x)+(1-F(x))\frac{\hat{E}-E_0}{\bar{t}-\hat{E}})\sum_i \hat{\epsilon}_i \frac{E_i-x}{\lambda_i(1-F(x))+\lambda_i(1-F(x))\frac{E_i-E_0}{\bar{t}-E_i}}\right)\\
=& f(x)\left(x-\tilde{E}+(1+\frac{\hat{E}-E_0}{\bar{t}-\hat{E}})\sum_i \hat{\epsilon}_i \frac{(E_i-x)(\bar{t}-E_i)}{\bar{t}-E_0}\right)\\
=& f(x)\left(x-\tilde{E}+\sum_i \hat{\epsilon}_i \frac{(E_i-x)(\bar{t}-E_i)}{\bar{t}-\hat{E}}\right).
\end{aligned}
$$

Recall that $W'(0)=0$, which implies that $L'(0)=\Pi'(0)$. Thus, the sign of $L'(0)$ determines the sign of the profit maximizing cutoff. To determine the sign of $L'(0)$, we set $x=0$ in the above expression and

$$
\begin{aligned}
L'(0) =& f(0)\left(-\tilde{E}+\frac{\sum_i \hat{\epsilon}_i E_i(\bar{t}-E_i)}{\bar{t}-\hat{E}}\right)\\
=& f(0)\left(-\tilde{E}+\frac{\bar{t}\tilde{E}-\sum_i \hat{\epsilon}_i E_i^2-\tilde{E}\hat{E}+\tilde{E}\hat{E}}{\bar{t}-\hat{E}}\right)\\
=& f(0)\left(-\tilde{E}+\frac{\tilde{E}\hat{E}-\sum_i \hat{\epsilon}_i E_i^2}{\bar{t}-\hat{E}}+\frac{(\bar{t}-\hat{E})\tilde{E}}{\bar{t}-\hat{E}}\right)\\
=& f(0)\left(\underbrace{-\tilde{E}}_{marginal\ effect}+\underbrace{\frac{\tilde{E}}{\bar{t}-\hat{E}}\left(\hat{E}-\frac{\sum_i \hat{\epsilon}_i E_i^2}{\tilde{E}}\right)+\tilde{E}}_{inframarginal\ effect}\right).
\end{aligned}
$$

This expression gives us another way to write the inframarginal effect of a change of the threshold at $x=0$ on the profit Π. We can simplify $L'(0)$ to

$$
\begin{aligned}
L'(0) =& \frac{f(0)\tilde{E}}{\bar{t}-\hat{E}}\left(\hat{E}-\frac{\sum_i \hat{\epsilon}_i E_i^2}{\tilde{E}}\right)\\
=& \frac{f(0)\tilde{E}}{\bar{t}-\hat{E}}\left(\hat{E}-\tilde{E}-\frac{(\sum_i \hat{\epsilon}_i E_i)^2-\sum_i \hat{\epsilon}_i E_i^2}{\tilde{E}}\right).
\end{aligned}
$$

□

Proof of Proposition 7. The expression for ϵ_N can be derived by observing that

$$
O_{N-1}=\sum_{i=N}^{N}\epsilon_i(E_i-y_{N-1})=\epsilon_n(y_N-y_{N-1}).
$$

The expression for ϵ_j for $1 < j < N$ can be obtained by first observing that

$$O_{j-1} - O_j = \sum_{i=j}^{N} \epsilon_i (E_i - E_{j-1}) - \sum_{i=j+1}^{N} \epsilon_i (E_i - E_j) = \sum_{i=j}^{N} \epsilon_i (E_j - E_{j-1}),$$

where the first equality follows from (8) and the second equality can be obtained by rearranging the sums. Dividing by $E_j - E_{j-1}$ yields

$$\frac{O_{j-1} - O_j}{E_j - E_{j-1}} = \sum_{i=j}^{N} \epsilon_i,$$

and taking differences

$$\frac{O_j - O_{j+1}}{y_{j+1} - y_j} - \frac{O_{j-1} - O_j}{y_j - y_{j-1}} = \sum_{i=j+1}^{N} \epsilon_i - \sum_{i=j}^{N} \epsilon_i = \epsilon_j,$$

that is, the expression for ϵ_j for $1 < j < N$ in the proposition. The expression for ϵ_1 simply follows from that fact that probabilities add up to one. □

Proof of Proposition 8. We want to analyze the effect of a change of $\bar{t}_2$, while keeping $z_i = \gamma_i \bar{t}_2$ constant. For this purpose the variance $\sigma_i(x)$ can be rewritten as

$$\begin{aligned}
\sigma_i(x) =& \frac{\lambda_i \int_x^{\bar{t}_1} t^2 dF(t) + \mu_i \bar{t}_1^2 + \gamma_i \bar{t}_2^2}{\lambda_i(1 - F(x)) + \mu_i + \gamma_i} - \left(\frac{\lambda_i \int_x^{\bar{t}_1} t dF(t) + \mu_i \bar{t}_1 + \gamma_i \bar{t}_2}{\lambda_i(1 - F(x)) + \mu_i + \gamma_i} \right)^2 \\
=& \frac{(\lambda_i(1 - F(x)) + \mu_i + \gamma_i)(\lambda_i \int_x^{\bar{t}_1} t^2 dF(t) + \mu_i \bar{t}_1^2 + \gamma_i \bar{t}_2^2) - (\lambda_i \int_x^{\bar{t}_1} t dF(t) + \mu_i \bar{t}_1 + \gamma_i \bar{t}_2)^2}{(\lambda_i(1 - F(x)) + \mu_i + \gamma_i)^2} \\
=& \frac{(\lambda_i(1 - F(x)) + \mu_i + z_i/\bar{t}_2)(\lambda_i \int_x^{\bar{t}_1} t^2 dF(t) + \mu_i \bar{t}_1^2 + z_i \bar{t}_2) - (\lambda_i \int_x^{\bar{t}_1} t dF(t) + \mu_i \bar{t}_1 + z_i)^2}{(\lambda_i(1 - F(x)) + \mu_i + z_i/\bar{t}_2)^2}
\end{aligned}$$

For $\bar{t}_2 \to \infty$ we get that the variance $\sigma_i(x)$ goes to infinity

$$\lim_{\bar{t}_2 \to \infty} \sigma_i(x) = \infty$$

and the expected type $Q_i(x)$ converges to

$$\lim_{\bar{t}_2 \to \infty} Q_i(x) = \frac{\lambda_i \int_x^{\bar{t}_1} t dF(t) + \mu_i \bar{t}_1 + z_i}{\lambda_i(1 - F(x)) + \mu_i} < \infty.$$

It follows that for $a > 0$, the utility in state i, $Q_i(x) - a\sigma_i(x)$, becomes negative if $\bar{t}_2$

is large enough. This implies that buyers are not willing to pay a positive price for a rated firm if the variance of types in one rating class is too high. Thus, for every cutoff x there is a $\bar{t}_2$ large enough such that the rating agency is better off pooling $t \in [x, \bar{t}_1]$ and rating $\bar{t}_2$ in a separate class than pooling all types above x.

A further condition that needs to be satisfied is that the agency does not prefer to charge a rating fee of $\bar{t}_2$ and to rate only firms with type $\bar{t}_2$ which yields profits of $\bar{t}_2\tilde{\gamma}$. Since we keep $\gamma_i\bar{t}_2$ constant when we increase $\bar{t}_2$, the profit of only rating types $\bar{t}_2$ stays the same. A sufficient condition for the agency to prefer to rate $[x, \bar{t}_1]$ and $\bar{t}_2$ is that the profit from pooling $t \in [x, \bar{t}_1]$ and not rating $\bar{t}_2$ is larger than the profit from only rating $\bar{t}_2$. □