The Rodney L. White Center for Financial Research

Competition Among Ratings Agencies and Information Disclosure

Neil A. Doherty
Anastasia V. Kartasheva
Richard D. Phillips

04-09

The Wharton School
University of Pennsylvania

Competition among rating agencies and information disclosure*

Neil A. Doherty† Anastasia V. Kartasheva‡ Richard D. Phillips§

February 13, 2009

Abstract

The paper proposes an explanation for why a rating agency chooses to pool different credit risks in one rating class, and analyzes how information disclosure depends on the value of information to the market. We show that an optimal disclosure policy of a monopoly rating agency is to pool companies or issuers in multiple rating classes and to have partial market coverage. It provides an opportunity for market entry. We then describe the potential market and the strategy of the entrant. We find that entry of an identical rating agency results in asymmetric rating scales. It justifies why some companies obtain multiple ratings and suggests that similar ratings from different agencies may mean different credit risks. We use Standard and Poor's entry in to the market for insurance ratings - a market that was previously covered by the monopolist agency the A.M. Best Company - to empirically test the qualitative predictions of the model regarding the impact of competition on the information content of ratings.

Keywords: rating agency, entry, competition, precision and disclosure of information, insurance.

JEL Codes: D8, G22, G28, L1, L43

*We are grateful to Franklin Allen, Jacques Cremér, Kose John, Gregory Nini, Olivia Mitchell, Patrick Rey, Jean-Charles Rochet, Anjolein Schmeits, Kent Smetters, Ajay Subramanian and the seminar participants at Wharton, Toulouse, IIOC 2007, EGRIE 2007, LMU Munchen, NBER Insurance Group, ARIA, FIRS Alaska for valuable comments. We gratefully acknowledge the support of Rodney L. White Center for Financial Research at Wharton.

†Insurance and Risk Management Department, Wharton School of the University of Pennsylavania, doherty@wharton.upenn.edu

‡*Corresponding author.* Insurance and Risk Management Department, Wharton School of the University of Pennsylavania, karta@wharton.upenn.edu

§Department of Risk Management and Insurance, Robinson College of Business, Georgia State University. rphillips@gsu.edu.

1 Introduction

Most credit rating agencies (CRAs) pool quantitative and qualitative information in letter grades. It results in clustering of companies and issuers in fairly broad rating classes. Nowadays, major CRAs use between 10 and 20 grades. The first objective of the paper is to examine the information disclosure by a monopoly CRA. The second objective is to analyze the entry strategy of a new CRA. Information pooling generates an opportunity for a new agency to cherry pick. If a company or an issuer decides to pay for a rating of a new CRA, it suggests either that the entrant uses a different rating scale or that its rating contains additional information. The third objective is to provide evidence of the effect of entry on the information content of the ratings of insurance companies. The insurance industry provides a unique natural experiment as it was served by the monopoly rating agency, the A.M. Best Company, for almost 100 years until experienced the entry of Standard & Poor's in the end of 1980s.

Pooling can be a profit maximizing strategy for a monopoly credit rating agency (Lizzeri (1999)). When all parties are risk neutral, Lizzeri shows that a monopoly CRA's optimal disclosure strategy is to pool all companies in one rate class. Surprisingly, all sellers pay to be rated in spite of the fact that a seller does not have to obtain a rating and *de facto* the CRA discloses no information. The logic of this result is as follows. The profit of a CRA is a product of the market coverage and a uniform rating fee. The fee cannot exceed the willingness to pay of the lowest quality rated seller. By pooling this seller with better quality sellers into one rating, a CRA can charge a higher fee without reducing demand for its services.

These results contrast with practice. For example, as a monopolist CRA in the insurance industry for many decades, A.M. Best had multiple rating categories and did not have complete market coverage. Moreover, in most settings, multiple CRAs are in competition. In 2000, A.M. Best covered 80.5% of insurance companies while S&P's coverage was 27.5%. We present a model that explains these phenomena and addresses the impact of new entry.

A distinguishing feature of our model is that buyers value the precision of information contained in ratings. Because rating-based guidelines are widely used in the conduct of business activities, buyers are ready to pay a higher price for a good with less ambiguous quality[1]. For their part, sellers have private information about the quality of a good which they cannot communicate credibly to buyers. A CRA can learn seller's quality but has discretion in how this information communicated to buyers. The evaluation of a seller is reached in two stages. In the first stage, a CRA chooses and commits to a disclosure policy and the fee for its services behind the veil of ignorance. In the second stage, privately informed sellers decide whether to demand a rating from the CRA. Once the CRA evaluates sellers who demanded a rating, buyers form a belief about each seller based on its decision to be rated and possibly, a rating.

[1] For example, in a survey of 200 plan sponsors and investment managers in the U.S. and Europe (Cantor, Gwilym and Thomas (2007)), 60% of fund managers and 47% of plan sponsors report that CRA should put more emphasis on accuracy of ratings.

We derive four main results. First, when buyers care about the precision of information, Lizzeri's single rating result holds only as a special case. The optimal rating scale derived from the model resembles the interval disclosure rule employed by the major credit agencies. Pooling the lowest rated seller with better types has two countervailing effects: the expected quality in the eyes of buyers increases while the precision of information goes down. For low value of information precision the first effect dominates and full pooling is optimal for the CRA. As the value of precision increases, the trade off between the two effects defines the boundary of the lowest rating. The model results in multiple equilibrium disclosure policies for rated sellers of higher quality. However, in all equilibria the payoff of rated sellers is non-decreasing in quality. Also we show that as the value of information precision goes to infinity, the optimal disclosure of a monopoly CRA converges to full disclosure. Hence, the CRA's incentives to provide information cannot be fully attributed to competitive market structure.

Our second result is that the optimal disclosure policy of a CRA implies partial market coverage. The reason is that the presence of unrated companies widens the gap between the prices rated and unrated sellers receive on their products or securities. As a result it permits the CRA to charge a higher fee.

The next two results are related to the optimal entry strategy of a new CRA to a market previously served by the incumbent. Our third result is that the entrant will design a rating scale more stringent relative to that of the incumbent. A seller purchases a second rating from the entrant only if this enables it to increase the price charged to buyers. The incumbent's rating states the seller's quality belongs to some interval. Then a seller purchases a second rating only if it either increases the expected quality by pooling it with better quality types or if it improves information precision by reducing the diversity of the group. In both cases, an entrant will require higher standards to be met to provide a similar rating to that of the incumbent. It also follows from our model that sellers are more likely to demand a second rating in markets where precision is of greater value.

The final result deals with the demand for ratings of the entrant. In each rating grade of the incumbent, sellers of higher than average quality are disadvantaged by pooling. We show that an optimal entry strategy of a new CRA is to target these sellers. Hence, the number of ratings each seller obtains depends on its quality. However, there is no congruency between the number of ratings and the quality of the seller. High and low quality sellers can be rated by both agencies while the intermediate quality seller obtains only one rating.

We test the predictions on the entry strategy of a new CRA using data on the U.S. property-liability insurance market. The insurance industry provides an ideal natural experiment to study entry for two reasons. First, unlike the market for bond ratings, there are no regulatory barriers to enter the market for insurance ratings. Second, until recently the market for insurance ratings has largely been dominated by a single monopoly agency - the A.M. Best Company. Standard & Poor's made its initial foray into the insurance ratings market in the late 1980's and dramatically increased the number of ratings it provided to insurers during the 1990's[2].

Insurance ratings measure an insurer's financial strength and ability to meet its ongoing insurance policy and contract obligations. Prior research has shown that rating affect an insurer's cost of capital. Epermanis and Harrington (2006) provide evidence that insurance buyers are very sensitive to insurers' ratings with lower rated insurer receiving lower prices for their policies in the marketplace[3]. Since policy liabilities are the primary source of capital for insurers, lower prices imply higher costs of capital. Also the importance of a rating for an insurer depends on the type of insurers buyer. Corporate insurance buyers usually require that insurers are highly rated as their insurance policies are very detailed and tailored to company's profile. These policies are mostly sold through insurance agents and brokers who often will not recommend an insurer with an A.M. Best rating below A- (for example, see Bradford (2003)). Personal automobile and homeowners insurance, on the other hand, is protected by state guarantee funds and is sold to less sophisticated customers. As a result, prices and demand are less sensitive to the financial strength of an insurer.

We employ two methodologies to examine the entry strategies of a new CRA. First, we use a hazard model to estimate a one-year probability of insolvency using publicly available data for all U.S. property-liability insurers. We show that S&P applied more stringent standards compared to A.M. Best for an insurer to achieve a similar rating.

Comparing the probabilities of default for insurers that receive a rating from A.M. Best and S&P ignores the possibility that firms will decide strategically whether to request a second rating from the entrant. This strategic behavior potentially biases our results. The second empirical test is designed to investigate the differences in rating opinions across the incumbent and the entrant using the Heckman-style sample selection methodology (Heckman (1979)). It allows us to correct for the strategic decision-making of firms and to decompose the sources of rating differences in two components: stringency differences across the two agencies and the insurer's financial quality. We find that higher-than-average quality insurers in each rating category chose to receive a second rating from S&P and that S&P required higher standards for an insurer to achieve a similar rating. Both results are consistent with our theory.

The plan of the paper is as follows. The next section discusses related literature. Section 3 presents the model and examines the disclosure policy of the monopoly. Section 4 analyzes which market segments are profitable for an entrant. We discuss the institutional details of insurance ratings and describe the data in Section 5. Our empirical analysis is presented in Section 6 and the conclusion follows. All proofs and tables are provided in the Appendix.

[2] For example, in 1992, S&P issued full rating opinions on only 25 property-casualty insurers and this number increased to over 250 insurers by the end of the decade. By year 2000, S&P was the second largest insurance rating agency and now rates over 800 companies representing more than 45 percent of the industry's assets.

[3] Epermanis and Harrington (2006) report that the rating downgrade of a property-liability insurer rated A or higher by A.M. Best results in an estimated abnormal premium growth of about -5%; for A- rated insurers, the estimated impact of a downgrade is -12%.

2 Related Literature

The paper belongs to the growing literature on the incentives of information intermediaries to manipulate information disclosed to interested parties. Since Akerlof's (1970) "lemon markets" paper, it is recognized that information intermediaries may play crucial role for markets under adverse selection[4] (see Biglaiser (1993)). Boot, Milbourn and Schmeits (2005) show that intermediaries can help to coordinate on a desired equilibrium. However, if an intermediary cannot perfectly assess the quality of the good and/or it has discretion about how the results of the assessment are communicated to buyers, the incentive problems may reduce the amount and the precision of information disclosed to the market.

The theory we develop in this paper builds on Lizzeri (1999) who studies optimal disclosure policies of an intermediary who can learn perfectly the information about the quality of the seller and communicate it to the buyer. Lizzeri shows a unique equilibrium in which all types of sellers pay to be rated. However, the intermediary does not disclose any information except that the seller has obtained a rating. Risk neutrality is essential for this result. It implies that the buyer is ready to pay the same price regardless of whether the quality is known for sure or is uncertain. In other words, the buyer does not value the precision of information disclosed by the intermediary. We change this assumption and assume that buyers care about the quality of information contained in the rating. In this respect our analysis is related to the literature on information quality and ambiguity aversion (Veronesi (2000), Epstein and Schneider (2008)).

There are other explanations for why an information intermediary may manipulate information. Manipulation can also occur due to collusion between the intermediary and the seller. Strausz (2005) shows that the threat of collusion makes honest certification a natural monopoly. Peyrache and Quesada (2005) argue that mandatory certification makes intermediaries more prone to collusion by increasing participation of low quality sellers. Mathis, McAndrews and Rochet (2008) show that reputation is sufficient to discipline CRAs only when a large fraction of CRAs income comes from rating simple assets. Benabou and Laroque (1992) analyze incentives of an intermediary to manipulate information when it also acts as speculator on the market.

When intermediaries compete for clients, and are not certain about their ability as experts, reputation concerns may lead to misreporting of information. Scharfstein and Stein (2000) and Ottaviani and Sorensen (2006a, 2006b, 2006c) study the impact of reputation concerns on the reports of analysts. These papers consider cheap talk models (Crawford and Sobel (1982)) in which intermediaries are concerned with establishing a reputation of being well informed. In order to signal its ability to provide information with high precision, the intermediary biases

[4] Grossman and Hart (1980), Grossman (1981) and Milgrom (1981) analyze direct communication between the buyer and the seller. In their framework a seller can disclose any information but must include the true type in its report. Then the conjecture of the buyer is that the true type is the most pessimistic element of the report, and full disclosure obtains. The key distinction of our approach is that a rating agency provides information about multiple sellers. It permits clustering different sellers in one rating class, and the unraveling need not happen.

its private observation in favor of prior belief. Mariano (2006) addresses a similar issue in the context of rating agencies. Skreta and Veldkamp (2008) study how higher complexity of rated assets affects incentives for rating shopping. They show that ability of sellers to compare ratings from different CRAs before ratings are disclosed to the market leads to rating shopping and ultimately inflates ratings.

In spite the fact that most information intermediaries function in oligopolistic markets, there is little research on the impact of competition on the disclosure of information[5]. Lizzeri (1999) shows that competition leads to full disclosure and zero fees for certification. In this paper we study the impact of entry into a previously monopolistic market for ratings.

3 Ratings of a Monopoly Credit Rating Agency

3.1 Model

A credit rating agency (CRA) provides information services that can lessen the information asymmetry between buyers and sellers. Sellers have private information about their quality v. Higher v corresponds to higher quality. CRA and buyers share a common prior about the quality of a seller. For simplicity, we assume that v is distributed uniformly on $[0, 1]$.

A seller cannot credibly communicate its quality to buyers. A rating agency offers an evaluation service for a fee t and can perfectly observe the type v of a seller. The fee is flat for all sellers purchasing a rating, and a CRA cannot screen companies by demanding higher fee for more favorable rating[6][7].

Having evaluated a company, a CRA strategically communicates results. The disclosure policy of the agency defines how rated sellers' quality types are communicated to buyers. For example, under full disclosure A CRA communicates the observed quality v. In general, a disclosure policy is a measurable function from the set of signals $[0, 1]$ into the set of Borel probability distributions on real numbers. An optimal disclosure policy in our model (Proposition 3) is similar to the discrete system of ratings employed by the major rating agencies. Under this system an agency partitions the set of realization of v in subintervals, and discloses that its estimate of quality belongs to a subinterval.

There is a unit mass of identical buyers. A buyer purchases at most one unit of a good from one seller. The buyer's willingness to pay for the good depends on the expected quality

[5] A few exceptions include Lerner and Tirole (2006) who study competition in standard setting; Farhi, Lerner and Tirole (2008) analyze the interaction between rating shopping behavior and transparency of certification; Morrison and White (2005) study banks decisions to apply to regulators with different perceived abilities.

[6] A rated seller does not value an option to withhold its rating. Faure-Grimaud, Peyrache and Quesada (2005) show that firms may have incentives to hide their ratings only if they are sufficiently uncertain about their quality. In our setting firms have perfect information about their quality, and thus will not apply for rating unless it increases their reservation price.

[7] In practice, CRA fees depend on the type of provided service but do not depend on the assigned rating (Cantor and Parker (1994))

and the precision of information about quality. Precision is measured by the variance of quality conditional on information available to buyers. For example, under full disclosure the variance of quality of a rated seller is zero, and precision is the highest. To model demand for precision we assume that buyers have mean-variance preferences. Given information I available to buyers, their valuation of a good is equal to

$$u(I) \equiv E[v|\, I] - aVar[v|\, I],$$

where $E[v|\, I]$ is the expected quality, and $Var[v|\, I]$ is the variance of quality. $a > 0$ measures the marginal value of information precision to buyers. Buyers are price takers, and $u(I)$ is the price paid for the good. Under the prior distribution, buyers valuation is equal to

$$u_0 = \frac{1}{2} - \frac{1}{12}a.$$

If the marginal value of information is low, $0 < a < 6$, the reservation price u_0 is positive. In this case, providing new information is not essential for functioning of the market. When $a > 6$, a buyer does not purchase a good unless it has some additional information about a seller. When $a = 0$ this model is equivalent to Lizzeri (1999).

Obtaining ratings is voluntary to sellers. The decision to be rated is based on the cost of rating and its effect on the buyer's valuation. The information impact of a rating depends on the disclosure rule employed by the agency and on the set of rated types. The expected payoff to a seller of type v depends' upon the buyers' valuation and is equal to

$$\begin{array}{l} u_R(v) - t, \text{ if it is rated,} \\ u_N(v), \text{ if it is not rated,} \end{array}$$

where $u_R(v)$ and $u_N(v)$ are the expected payoffs of type v with and without a rating, respectively. Denote δ the mass of sellers demanding a rating. Then the payoff of the rating agency is equal to

$$V = \delta t.$$

The game consists of three stages.

1. Sellers learn their types v. A rating agency designs its disclosure policy, and sets a fee.

2. Sellers observe the disclosure policy of the rating agency and a fee. They decide whether to purchase a rating. The participating sellers are evaluated, and the results are disclosed to buyers according to the disclosure policy of the CRA.

3. The buyers observe the disclosure policy and the rating if the seller is rated. They decide whether to purchase the credit sensitive product. Sellers receive a payoff which depends on rating status.

We study sequential equilibria of the game. Strategies of all players must be optimal at every stage of the game given the beliefs about other players information. Beliefs must be consistent with the Bayes rule whenever possible.

3.2 Full disclosure and the benefits of information pooling

Analysis of full disclosure can be useful to highlight the CRA's benefits from pooling information. Under full disclosure a rating is a perfect signal about seller's type to buyers. If a seller $\widehat{v} \in [0,1)$ decides to purchase a rating, better sellers $v > \widehat{v}$ do the same, as their payoff when rated is increasing in quality. Then non-rated sellers must have quality below $\widehat{v}$. Also if a seller $\widehat{\widehat{v}}$ is not rated, rating does not benefit sellers with lower quality $v < \widehat{\widehat{v}}$. This intuition implies that given a fee charged for ratings, there is a seller type indifferent between purchasing a rating or operating without a rating[8]. Demand for ratings will come from higher seller types. The optimal fee for the rating agency and the resulting coverage of the market are derived from the trade off between the marginal benefit of charging higher fee and the marginal cost of the reduced demand for ratings.

Proposition 1 *Suppose that a monopoly rating agency commits to full disclosure, and the fee for the rating services is less than $\frac{1}{2}+\frac{1}{12}a$. Then the unique sequential equilibrium of the subgame has a threshold structure: There is a type $v_F \in [0,1]$ such that all types above v_F purchase a rating, and no type below v_F is rated.*

Is full disclosure optimal for the CRA? Suppose that, instead of reporting the type v_F, the rating agency announces that this type is from an interval $[v_F, v_F + \Delta]$, $\Delta > 0$. Thus v_F is pooled with better types and the rating agency may be able to charge a higher fee without reducing demand for ratings. It is the case when the valuation of pooled types is higher than the valuation of the lowest rated type,

$$v_F + \frac{1}{2}\Delta - \frac{1}{12}a\Delta^2 > v_F.$$

If the marginal value of information is zero, $a = 0$, Lizzeri's (1999) result obtains: all types should be pooled and assigned the same rating grade. When precision of information matters, $a > 0$, pooling imposes a cost in lost precision, $\frac{1}{12}a\Delta^2$. This intuition suggests that the optimal disclosure policy trades off the benefits of pooling due to higher fees and the cost of pooling due to reduced precision.

3.3 Optimal disclosure

We now analyze a profit maximizing disclosure policy of a monopoly rating agency. A rating agency may fully disclose the rated seller's type, or may pool a seller with any other sellers and disclose that it belongs to a particular group. Formally, a disclosure policy is a correspondence

[8] For this to hold, the fee should not exceed the gain of a rating to the highest type $v = 1$ when it is the only rated type, $1 - (\frac{1}{2} - \frac{1}{12}a) = \frac{1}{2} + \frac{1}{12}a$.

$s : [0,1] \to [0,1]$. The expected quality $\mu(s(v))$ and the variance $\sigma^2(s(v))$ of type v rated $s(v)$ depend the set of types that obtain the same rating,

$$\begin{aligned} \mu(s(v)) &= E[v' : s(v') = s(v)], \\ \sigma^2(s(v)) &= Var[v' : s(v') = s(v)], \end{aligned}$$

and results in buyers' valuation of seller type v equal to

$$u(s(v)) = \mu(s(v)) - a\sigma^2(s(v)).$$

Denote $V_R(s)$ the set of rated seller types, $V_R(s) \subset [0,1]$, and $V_N(s)$ the set of non-rated types, $V_N(s) = [0,1] \backslash V_R(s)$. Sellers purchase a rating only if it has a positive return,

$$u(s(v)) - t \geq \max\{u(V_N(s)), 0\} \text{ for all } v \in V_R(s). \tag{1}$$

The right hand side of the inequality reflects that a non-rated seller will trade only if its valuation without a rating is positive. A disclosure policy $s(\cdot)$ generates demand

$$\delta(s) = \int_{V_R(s)} dF(v).$$

A strategy of the rating agency is a disclosure policy $s(\cdot)$ and a fee for the rating t. A strategy of each seller type is its decision to be rated. We restrict attention to pure strategies, and study sequential equilibria of this game. In equilibrium, the following two conditions must be met. First, the disclosure policy is optimal for the rating agency,

$$(s(\cdot), t) \in \arg\max_{\widetilde{s}, \widetilde{t}} \delta(\widetilde{s})\widetilde{t}.$$

Second, the decision to obtain a rating is optimal for a seller. That is, for any $(s(\cdot), t)$ and strategies of sellers $[0,1] \backslash v$, seller type v is rated if and only if (1) holds for this seller.

To analyze the optimal disclosure policy, we proceed in two steps. In the next proposition we describe the structure of an optimal disclosure policy. Then we apply this result to characterize the policy and analyze how it depends on the marginal value of information a.

Proposition 2 *An optimal disclosure policy of a monopoly rating agency has the following structure. There is a type $v_M \in [0,1]$ such that all types $v \geq v_M$ are rated, and no type $v < v_M$ is rated. Types $[v_M, v_M + b_M]$, $b_M \geq 0$, $v_M + b_M \leq 1$ are assigned the same rating. The fee charged for the rating equals to the value of the rating to the lowest rated type, $t_M = u([v_M, v_M + b_M]) - \max\{u([0, v_M]), 0\}$.*

The rating agency faces demand $\delta_M = 1 - v_M$, and earns profits

$$(1 - v_M)t_M,$$

Denote $u(N)$ and $u(L)$ the valuation on non-rated types $N = [0, v_M]$ and the lowest rated types $L = [v_M, v_M + b_M]$, respectively.

$$\begin{aligned} u(L) &= v_M + \frac{1}{2}b_M - \frac{1}{12}ab_M^2, \\ u(N) &= \max(\frac{1}{2}v_M - \frac{1}{12}av_M^2, 0). \end{aligned} \tag{2}$$

Then

$$t_M = u(L) - \max\{u(N), 0\}.$$

If a seller cannot trade without a rating, $u(N) < 0$, the fee is equal to the valuation of types in the lowest grade L. When $u_N > 0$, CRA charges the difference between the valuations of lowest grade sellers and non-rated sellers, $u(L) - u(N)$.

In equilibrium, non-rated sellers N must be better off without a rating. If a seller $v \in N$ deviates and purchases a rating, the rating agency announces that seller's quality is from the interval N. Then the deviation is not profitable and purchasing a rating cannot increase the reservation price charged by these sellers.

An optimal disclosure policy of the rating agency solves

$$\max_{(v_M, b_M)} (1 - v_M)(u(L) - \max\{u(N), 0\}).$$

In the next proposition we summarize the solution to this problem.

Proposition 3 *The optimal monopoly rating system is summarized in the following table.*

a	v_M	b_M	t_M	π_M	$\max\{u_N, 0\}$
$0 \le a \le 2$	0	1	$\frac{1}{2} - \frac{1}{12}a$	$\frac{1}{2} - \frac{1}{12}a$	0
$2 \le a \le 6$	$\frac{3}{4} - \frac{3}{2a}$	$\frac{1}{4} + \frac{3}{2a}$	$\frac{1}{4} + \frac{1}{24}a$	$\frac{(a+6)^2}{96a}$	$\frac{3(10-a)(a-2)}{64a}$
$6 \le a \le \frac{21}{2}$	$\frac{2}{3} - \frac{1}{a}$	$\frac{3}{a}$	$\frac{(a+3)^2}{27a}$	$\frac{(a+3)^3}{81a^2}$	$\frac{(2a-3)(21-2a)}{108a}$
$\frac{21}{2} \le a \le \frac{51}{4}$	$\frac{6}{a}$	$\frac{3}{a}$	$\frac{27}{4a}$	$\frac{27(a-6)}{4a^2}$	0
$a \ge \frac{51}{4}$	$\frac{1}{2} - \frac{3}{8a}$	$\frac{3}{a}$	$\frac{3}{8a} + \frac{1}{2}$	$\frac{(4a+3)^2}{64a^2}$	0

When the marginal value of information is relatively low, $a \le 2$, all seller types are rated and are pooled in the same rating grade. As the value of information increases, the rating becomes more precise, i.e. b_M decreases. The market coverage is decreasing in a when $u(N) > 0$, increasing in a when $u(N) = 0$, and decreases in a when $u(N) < 0$. The profit of the rating agency is non-monotone in the value of information a. It is decreasing when $u(N) > 0$, increasing when $u(N) = 0$ and decreasing when $u(N) < 0$. Profit is the highest when the value of information is the lowest, $a = 0$. As $a \to +\infty$, the profit converges to $\frac{1}{4}$.

What is the measure of types pooled in the lowest rating? From (2), pooling db_M sellers in one rating increases the expected quality of $u(L)$ by $\frac{1}{2}db_M$ and reduces the precision of the rating by $(-\frac{1}{6}ab_M)db_M$. For low values of a the increase in expected quality from pooling outweights the precision cost, and that leads to extensive pooling. For higher values of a, the interior solution obtains when the marginal increase in quality is equal to the marginal cost of reduced precision, resulting in $b_M = \frac{3}{a}$. As the value of precision increases, the measure of types pooled in the lowest rating goes to zero.

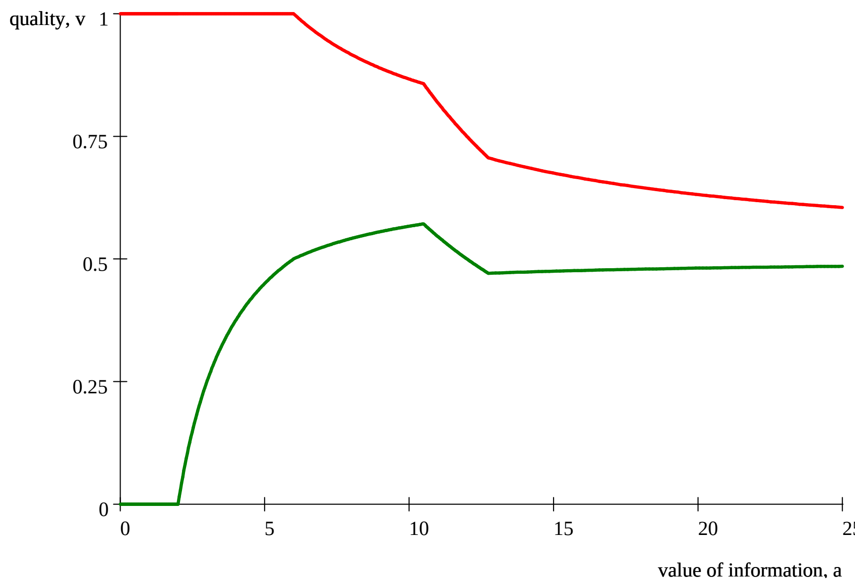
Marginal value of information entails five disclosure policy regimes. When a is low, $0 \leq a \leq 2$, the optimal disclosure policy of the rating agency is to pool all sellers in the same rating grade. It shows that Lizzeri's "no disclosure" result is more general. It also holds when buyers have relatively low value for information precision of the rating.

For moderate information values, $2 \leq a \leq 6$, a monopoly rating agency has partial coverage of the market, $v_M > 0$, but all rated sellers are still pooled in the same rating grade. This regime resembles a minimum standard setting. Reducing the coverage of the market is beneficial for the rating agency because it widens the difference between the valuation of rated and non-rated sellers, and allows the CRA to charge a higher fee for the rating. At the same time, the value of precision is too low to benefit from increasing precision, so all rated sellers are pooled in order to increase the expected quality of the lowest rated type.

As the value of information increases, $a \geq 6$, providing precision becomes more valuable than increasing expected quality by pooling. The distinction between the last three regimes for $a \geq 6$ is the ability of non-rated sellers to trade. Higher demands for precision imply that rating becomes essential for trade, and the agency expands the coverage of the market for $\frac{21}{2} \leq a \leq \frac{51}{4}$. However, when the payoff of non-rated sellers is negative, $u(N) < 0$, precision becomes secondary to improving the pool of rated companies. The coverage is increasing for $a \geq \frac{51}{4}$.

The profit of the rating agency is non-monotone in the value of information. For relatively low values, the rating agency can benefit from its unique ability to screen sellers and selectively disclose the results. However, as the value of information increases, the optimal rating system requires finer information disclosure and reduces the ability of the agency to increase the fee by pooling types in one rating.

Figure 1 shows the boundaries for rating L as a function of a. Types located below the lower curve are not rated. Types located between the lower and the upper curves are pooled in the lowest rating grade L. Like under full disclosure, the coverage of the market is non-monotone in a and depends on the ability of non-rated companies to trade. It is decreasing in a for low information values because the rating agency has incentive to widen the gap between valuations of rated and non-rated sellers.



Optimal rating scale of a monopoly rating agency

Unfortunately, Proposition 2 does not allow to pin down a unique rating scale for types $[v_M + b_M, 1]$, when the value of information is relatively high. A CRA is indifferent among all disclosure policies that induce participation of these sellers. In the next proposition we derive a necessary condition for an optimal interval disclosure policy to be an equilibrium, and show how the size of a rating interval changes with the marginal value of information.

Proposition 4 *Any system of intervals* $(R_1, ..., R_N)$, $N \leq +\infty$ *that satisfies* $b_{k+1} \leq b_k + \frac{6}{a}$ *is consistent with the optimal disclosure policy. As the value of precision increases, the measure of types pooled in the same rating interval decreases.*

Disclosure policies are not equivalent from the seller's perspective. In each pooling interval the types at the bottom of the interval benefit from pooling at a cost of types on the top of an interval. It is immediate to show the following.

Proposition 5 *In each pooling interval, the measure of types that prefer full disclosure is greater than the measure of types that prefer pooling, and the difference is increasing in the value of precision* a.

In the next section we study entry under the assumption that the incumbent CRA uses interval disclosure. Our motivation is twofold. First, it is an equilibrium that is consistent with industry practice. Second, as we will show it allows entry in multiple segments of the market.

Indeed, if there are segments of the market where the incumbent makes seller's type perfectly known to the buyer, a new rating agency has no benefit to enter these segments[9].

4 Entry of a New Credit Rating Agency

We analyze the entry strategy of a new agency on the following time line. After the ratings has been purchased from the incumbent, but before the transaction between buyers and sellers, a new agency offers an additional rating for a fee. If a new rating agency attracts any sellers, they are rated by the entrant. Then buyers form their valuations based on all available sellers' ratings (i.e. from the incumbent and the entrant) and trade takes place.

In this setup, the incumbent has no possibility to adjust its disclosure policy. Our motivation for this assumption is that sellers, buyers and the incumbent rating agency exhibit inertia in designing and understanding rating standards and the industry structure cannot change overnight. The section analyzes how a new rating agency should structure its disclosure policy to create demand for its services.

A seller will pay for an additional rating only if it increases its value in the eyes of the buyer. This occurs either when the second rating allows the seller to signal higher quality or when it improves the precision of information. If a seller is rated by the incumbent and the entrant, it must be that two ratings are better than one,

$$u(R_m, R_e; v) - t_m - t_e \geq u(R_m; v) - t_m,$$

where $u(R_m, R_e; v)$ and $u(R_m; v)$ respectively are the payoffs of seller v rated by both agencies and rated only by the incumbent and t_m and t_e are the fees for ratings by two rating agencies. If a seller is rated only by the entrant, then

$$u(R_e; v) - t_e \geq \max\{u(V_N), 0\}.$$

In the next proposition we characterize the demand for the entrant's rating.

Proposition 6 *Suppose that the incumbent CRA employs an interval disclosure rating system. An entrant can always design a rating system that is attractive to sellers at the top of each rating interval of the incumbent. The rating standards of the entrant are more stringent than that of the incumbent.*

Demand for a new rating comes from types rated below their true valuation by the incumbent. These are the types on the top of each pooling interval. It implies that a seller with two ratings, from the entrant and the incumbent, is not necessarily better than a seller rated only by the incumbent.

[9] Note that if a rating agency's evaluation technology is imperfect, entry can be beneficial even in the case of full disclosure.

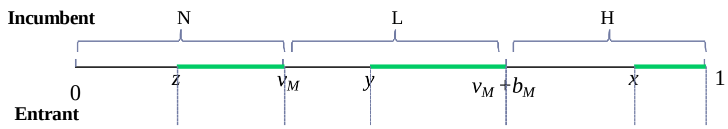


Figure 1: Demand for entrant's ratings

In general, the entry strategy of a new agency depends on the rating scheme of the incumbent and on the marginal value of information to sellers. The entrant can target multiple groups, including unrated sellers. The highest quality rated sellers are willing to pay the highest price to refine the information about their quality. However, charging high fee to this group reduces the demand from the other sellers. An optimal entry strategy is a trade-off between the fee charged by the entrant and the coverage of the market.

Proposition 6 is based on a "cherry picking" reasoning and does not rest on any distributional assumption over seller types. The particular incumbent rating categories in which the entrant chooses to harvest these cherries depends on how the incumbent formed its rate categories in the first place. The size, number and thresholds of rate categories depend on the distributional assumption (in our case a uniform distribution) and the value of information a. Not surprisingly, then, it turns out that where the entrant chooses to compete is sensitive to the distributional assumption and, for this reason, any empirical prediction will be suspect. Nevertheless, we can gain some insight into the qualitative nature of the entry strategy by following a particular case.

In the rest of this section we will revert to the uniform distribution and examine the entrant's strategy when the incumbent assigns two ratings[10], $H = [v_M + b_M, 1]$ and $L = [v_M, v_M + b_M]$. Then a new CRA can enter in three segments, non-rated sellers $N = [0, v_M]$ and rated sellers L and H. Denote x, y and z the lowest types rated by the entrant in each interval N, L and H, with $v_M + b_M \leq x \leq 1$, $v_M \leq y \leq v_M + b_M$ and $0 \leq z \leq v_M$ (see Figure 1). When $a \leq 6$ and all types are pooled to the same rating grade by the incumbent, x is set to zero. The entrant charges a fee t_e for its rating. An optimal entry strategy of a new rating agency solves

$$\max_{(x,y,z,t_e)} ((1-x) + (v_M + b_M - y) + (v_M - z))t_e \tag{3}$$
$$u(R_m, R_e; v) - t_m - t_e \geq u(R; v) - t_m, \; v \in V_R,$$
$$u(R_e; v) - t_m \geq \max\{u(V_N), 0\}, \; v \in V_N.$$

The total demand for the services of the entrant is the sum of coverage in each rating category. The entrant has a leeway to decide which categories of sellers to target. The difference between this problem and the problem of the monopolist CRA is that the potential market of the entrant

[10] This policy is optimal when $u(H) \geq u(L)$, or $a \leq 23\frac{1}{4}$. For higher value of information a rating agency needs more rating categories.

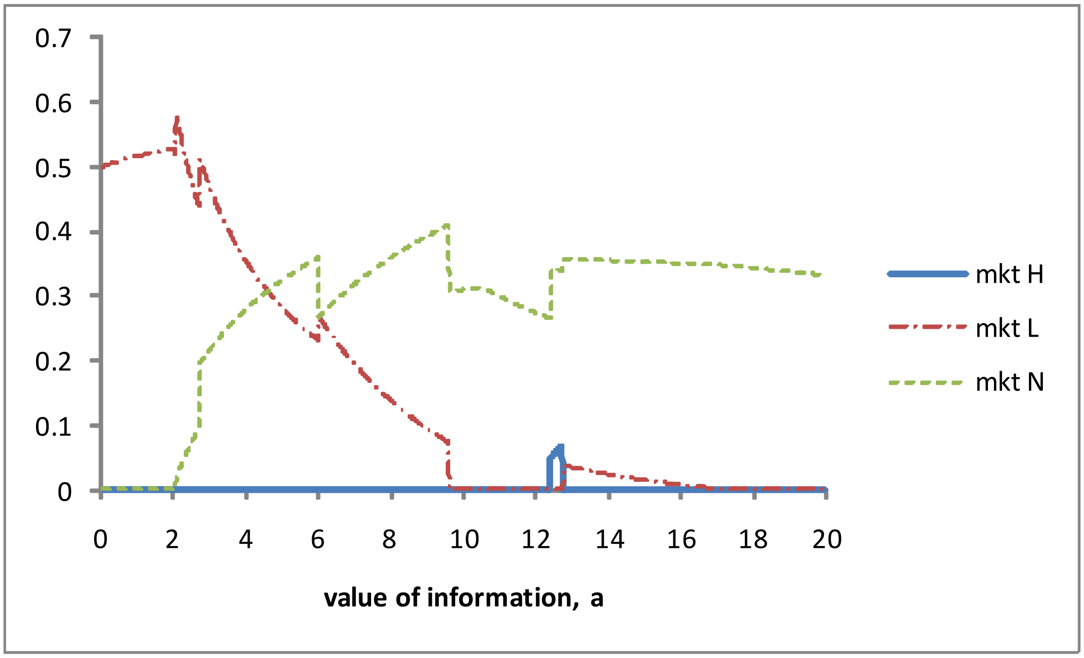


Figure 2: Entrant's coverage of highly rated, low-rated and non-rated segments of the market

consists of discrete segments in each rating category of the incumbent. Thus the entrant has two decisions to make. First it selects the segments of the market to be rated; second, it specifies the stringency of the rating in each targeted segment. Optimal entry strategy does not have a closed form solution for $a > 2$, but the problem can be solved numerically. The technical details of finding the solution to the optimization problem are summarized in supplementary materials available from authors on request.

Proposition 7 *The optimal strategy of the entrant is to target sellers rated N and L when the value of information is low. As the value of information increases and non-rated companies cannot trade, the entrant specializes on providing ratings to sellers not rated by the incumbent.*

The solution is summarized in Figure 2. The curves show the optimal market coverage of the entrant for different values of information a. When the value of information is relatively low, the entry occurs in two segments of the market, companies rated L by the incumbent and non-rated companies. The incumbent pools a large number of companies in a single rating. Designing ratings that allow better quality sellers to differentiate from the lower quality sellers is the best entry strategy. As the value of information increases, the coverage of sellers rated L by the incumbent goes down, while the coverage of non-rated companies goes up. Recall that as a increases, sellers cannot trade without a rating. However, the incumbent rating agency does not provide coverage to all companies because increasing the gap between the payoff of rated

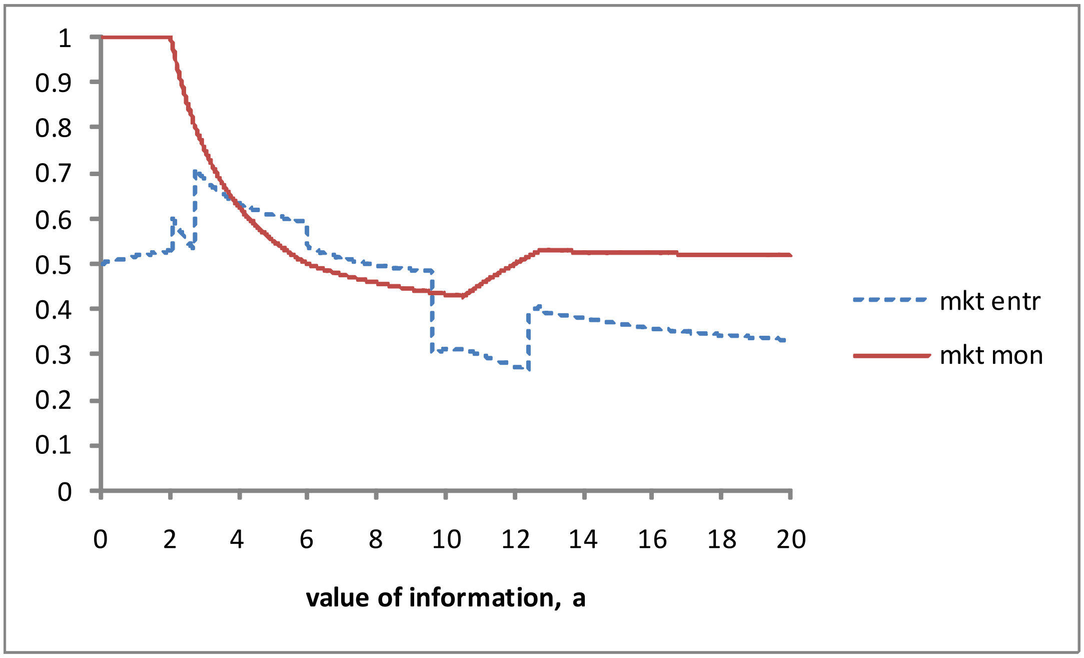


Figure 3: Market coverage by the incumbent and the entrant

and non-rated companies increases the value of the rating, and ultimately, the fee. The market of non-rated sellers provides the highest value for the entrant. At the same time the other two segments, L and H, become less profitable and there coverage decreases.

There is almost no entry in the highest segment of the market. The reason is that the incumbent's rating is the most opaque about the intermediate and low quality companies, making these segments more attractive for the entrant.

In Figure 3 we show the market coverage of the entrant and the incumbent agency. When the value of information is low, there is an interval of values of a where the coverage of the entrant is higher that that of the incumbent. As the value of information increases, the coverage of the entrant is about one third of the market.

When the value of information is low and a rating is not essential for trade, the profit of the entrant is low relative to the incumbent's. However, it substantially increases once sellers are unable to trade without the incumbent's rating (Figure 4). In this case the services of the entrant to non-rated sellers become necessary for trade, and create a niche for the new agency. The entrant's strategy of targeting the non-rated segment of the market echoes the results obtained in Shaked and Sutton (1982). They study the price competition among firms that can differentiate the quality of their goods, and show that firms choose distinct qualities in equilibrium. The reason is that offering distinctive products relaxes price competition of firms. Our setup is different in many respects, most importantly, in that the entrant offers its services

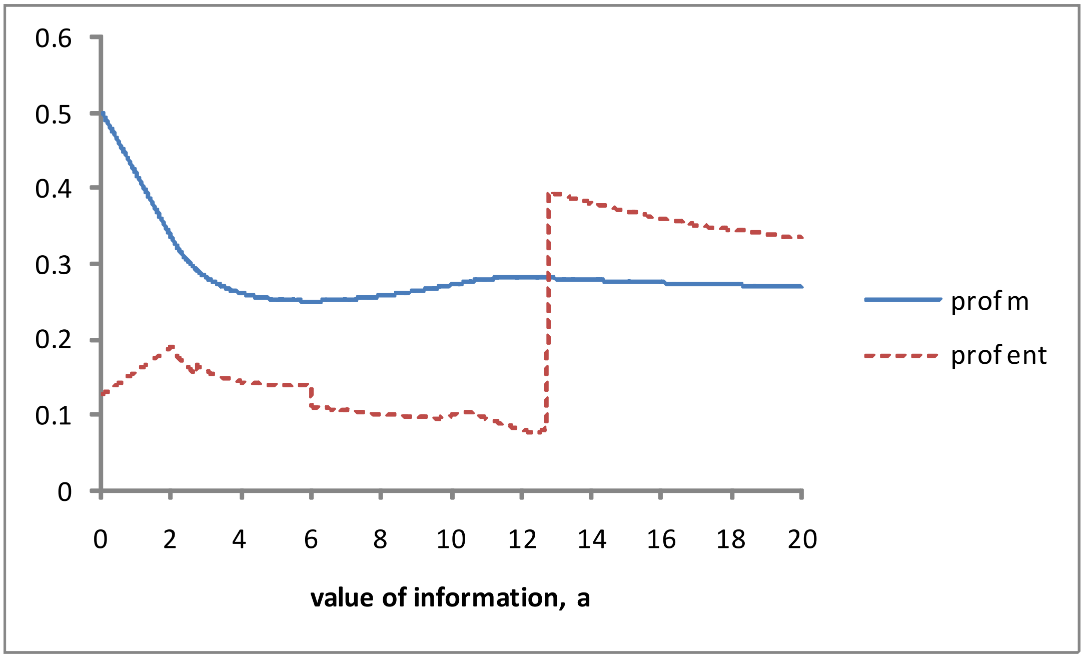


Figure 4: Profits of the incumbent and the entrant

to sellers already rated by the incumbent. For low information values the market of the entrant is rather narrow because all seller types, rated or not, can trade. However, for higher information values providing rating is essential for non-rated sellers. It creates a niche where the entrant becomes a monopolist, and increases its profits.

5 Institutional Setting and Data

In the remainder of the paper we test the primary predictions of our theoretical model that (i) the entrant CRA will have higher standards, on average, in order for a firm to receive a rating similar to the one they received from the incumbent CRA; (ii) the entrant CRA will obtain the greatest demand for its services from higher-than-average quality insurers within a rating class of the incumbent CRA; and (iii) insurers for which market participants have a more difficult time assessing the true financial strength of the firm will be more likely to seek an additional rating. To do so we take advantage of a natural experiment which began in the late 1980's and continued through the 1990's when the well known bond rating agency, Standard & Poor's (S&P), entered the market for insurance ratings[11].

[11] Although our theoretical resutls could be useful to analyze entry in bond rating industry, this market has been consistently dominated by a few number of incumbents. As noted by Professor White of NYU, "A striking fact about the structure of the industry in the U.S. is its persistent fewness of incumbents. There have never been

In this section we present the institutional setting for our tests and discuss our data sources. We then discuss our empirical tests and results based upon making univariate comparisons between the stringency standards employed by S&P and A.M. Best. We conclude the empirical section by presenting an econometric analysis of the differences in the ratings assigned to insurers that opted to request a rating from both agencies.

5.1 Insurance ratings and Standard and Poor's entry

Before the end of 1980s, the market for insurance ratings was largely dominated by the A.M. Best Company. Incorporated in 1899, A.M. Best publishes "financial strength ratings" on the majority of U.S. insurers and, for most of their history, they were the only agency doing so. The monopoly position Best's enjoyed, however, began to erode after A.M. Best was criticized following the liability insurance crisis of the mid-1980's and several natural catastrophes in the early 1990's bankrupted numerous insurers.[12] The most aggressive agency to enter the market was Standard & Poor's (S&P) who entered in three phases[13]. As shown in Table 1 and graphically in Figure T1, S&P's initial entry occurred in the late 1980's when it announced it would offer to publish "claims paying ability" ratings on property-liability insurers using methods similar to those employed by A.M. Best and also similar to the same methodology they use to issue ratings on corporate debt. Namely, A.M. Best and S&P both combine information publicly available together with proprietary information they learn by conducting interviews with the management of the insurer to determine their ratings. In addition, and consistent with our theoretical model, A.M. Best and S&P both require that a company has to request a rating and it pays a fee to initiate the ratings process.

Phase two of S&P's entry began in 1991 when they introduced their "qualified solvency rating" service to complement their traditional ratings. The methodology S&P uses to determine a qualified rating for an insurer differs in at least three important ways from the traditional manner. First, qualified ratings are solely based upon publicly available data. Thus, unlike the traditional method, S&P analysts do not interview the management of an insurance company prior to issuing the qualified rating. Second, insurance companies are not required to request the rating nor pay a fee to receive the qualified rating. Finally, when the system of qualified ratings

more than five general-purpose bond rating firms; currently there are only three. Network effects - users' desires for consistency of rating categories across issuers - are surely part of the explanation. But, for the past 25 years, regulatory restrictions (by the Securities and Exchange Commission) on who can be a "nationally recognized statistical rating organization" (NRSRO) have surely also played a role." See White (2002).

[12] Winter (1991) provides an overview of the crisis in U.S. liability insurance markets that occurred between 1984-1986. Lewis and Murdock (1996) describe the state of U.S. property insurance markets following several large natural disasters that occurred in the early 1990's including Hurricanes Andrew (in Florida) and Iniki (in Hawaii) in 1992 and the Northridge Earthquake in California in 1994.

[13] The three phases has been documented in financial press. See (i) PR Newswire, 27 August 1987, "S&P launches insurance rating service"; (ii) Risk Management, June 1991, "S&P launches service to rate solvency"; (iii) National Underwriter, 25 September 1995, "S&P expanding P-C range".

was introduced, S&P maintained a policy that no insurer could receive above a BBB rating – regardless of the characteristics of the company or regardless of the rating that company may have received from A.M. Best. The introduction of qualified ratings allowed S&P to dramatically expand their coverage of industry participants although most of the insurers they rated were done so under the qualified system and not in the traditional system that generated revenue.

The final phase of S&P's entry occurred in late 1994 when S&P decided to relax the rule which said that insurers could not receive a qualified rating above BBB. As shown in the figure, this decision led to a significant increase in the demand for its services over the next couple of years. Thus, by the end of 1990s, S&P provided full rating opinions to about 30-50% of insurers measured by the asset size.[14]

5.2 Data

We gathered two sets of data: the first documents the financial quality, business strategy, and organizational characteristics of U.S. insurance companies while the second data set contains the financial strength/claims paying ability ratings assigned to insurance companies by A.M. Best and by S&P, respectively. The data documenting the financial quality and other relevant characteristics of the companies comes from the annual regulatory statements of all property-liability insurers maintained in electronic form by the National Association of Insurance Commissioners (NAIC). We include all firms that meet our data requirements (discussed below). The ratings information for A.M. Best comes from Best's annual *Key Ratings Guide* (various years). We obtained the S&P ratings directly from S&P via a custom data order. Our data spans the years of S&P's entry into this market - 1989-2000.

6 Econometric Analysis and Results

6.1 Comparing Rating Stringency

Empirical strategy. In our first set of tests we seek to empirically compare the stringency of the ratings assigned by the incumbent firm (A.M. Best) relative to the entrant (S&P). To do so we need a statistic that summarizes the financial quality of each insurer in our data. Consistent with the ratings literature and with A.M. Best's and S&P's own stated objectives, the metric we use to proxy for financial quality is the insurer's one-year probability of default estimated using a discrete-time hazard model (Shumway (2001)). We chose the hazard model as it has at least three primary advantages over more traditional static models (e.g., Altman Z-scores; Altman (1968)). First, hazard models allow for time-varying covariates that explicitly recognize

[14] Like A.M Best, Weiss Research provides ratings on almost every insurer that operates in the U.S. marketplace. However, the process Weiss uses in the assignment of their ratings is fundamentally different than the process used by Bests and S&P. We consider S&P to be the more influential new entrant into the market for property-liability insurance ratings given their established reputation in the bond rating market.

that the financial health of some firms will deteriorate over time even though the firm may not declare bankruptcy for many years.[15] Second, hazard models allow us to exploit all available information about the condition of the firm rather than just the last year's observations. And finally, the hazard model allows us to fully exploit our large panel data set of insurers.

Shumway (2001) shows that the likelihood function of a discrete time hazard model is identical to the likelihood function for a multiperiod logit model. Thus, estimating the model is equivalent to estimating the traditional static logistic model except the coding of the dependent variable is slightly different. Specifically, the dependent variable for the hazard model, y_{it}, is a binary indicator set equal to 1 if firm i is declared bankrupt in year $t+1$ and equals 0 otherwise. In other words, the dependent variable equals 0 for each year the firm does not exit the system and each bankrupt firm contributes only one failure observation, i.e., $y_{it} = 1$, in the last year the firm has data. Time varying covariates are easily incorporated by using each firm's annual data estimated from a discrete-time hazard model. Consistent with the insurer solvency prediction literature, we classify the year of insolvency as the year that the first formal regulatory action is taken against a troubled insurer (e.g., Cummins, Harrington and Klein (1995)). We identify the year of first regulatory action through a variety of sources including the NAIC's Report on Receiverships (various years), the Status of Single-State and Multi-State Insolvencies (various years), and from the list of insolvent insurers provided in a report by A.M. Best Company which lists all property-liability insurers that failed from 1969-2001 (A.M. Best, 2002). From these sources we identify 300 property-liability insurers that failed between 1990 and 2001.

The explanatory variables we use to construct the probability of default are the nineteen balance sheet and income statement ratios that make up the NAIC's Financial Analysis and Surveillance Tracking (FAST) system. Grace, Harrington and Klein (1995) conclude that there are diminishing marginal returns to incorporating additional balance sheet and income statement ratios not already included in the FAST system except to include a control for firm size[16] and an organizational form indicator variable to control for whether the insurer belongs to a mutual or reciprocal group of insurers.

The probability of default is estimated for all insurers for which we have data to calculate the FAST ratios. Thus, not only do we include insurers rated by A.M. Best or S&P, but we also include insurer firm-year observations that do not receive ratings from either of these two agencies. The only insurers we delete from the analysis are those with insufficient data needed to calculate the nineteen FAST variables or those who do not have data available in the year prior to their first event year. In an effort to include as many insolvent observations in the analysis, we include insurers who report data two years prior to their first event year but who do not report in the year prior to their first event year. We delete any bankrupt firms for which we

[15] Shumway (2001) shows that ignoring this information creates a selection bias which leads to inconsistent parameter estimates.

[16] Firm size is equal to the natural logarithm of the real assets of the firm where the price deflator is the Consumer Price Index.

were unable to locate data within two years of their first event year. The final data set contains 24,236 solvent firm-year observations and 217 insolvent firm-year observations.

The second decision we need in order to evaluate the standards necessary to obtain a comparable rating across the two CRAs is to define a mapping between the rating systems used by the two agencies. Unfortunately a single one-to-one mapping does not exist and prior research investigating insurance ratings across agencies have used different definitions.[17] For this study we reviewed the verbal descriptions each agency ascribed to their individual rating categories as they existed at the time S&P began to enter the market for insurance ratings (A.M. Best 1987; Standard & Poor's 1992). Identifying the upper-end of each agency's ratings scale is straight forward so we map S&P's AAA rating into A.M. Best's A++/A+ category.[18] Mid-way through the each agency's rating scale, we map S&P's BBB rating into Best's B++/B+ category as both agencies suggest these categories define the threshold between "Investment vs. Non-investment grade insurers" (using S&P's language) and between "Secure vs. Vulnerable" insurers (using A.M. Best's language). For the remaining ratings, we map all categories in between the upper and lower anchor points to each other (S&P's AA and A categories map into Best's A and A- categories) and we map each agency's categories below the investment grade/secure threshold to each other. Table 2 summarizes the mapping used in the study and also presents the rating descriptions each agency uses to verbally describe the categories taken from their respective ratings manuals . Also shown in Table 2 are numerical values that we assigned to each rating category that will be used later to facilitate comparisons across agencies and over time.

Results. Table 3 presents summary statistics comparing the solvent and insolvent samples of the data set used to estimate the hazard model while the estimated discrete-time hazard model is shown in Panel A of Table 4. The regression results are consistent with many of the inferences obtained after reviewing the summary statistics. For example, the estimated coefficients suggest highly levered firms, rapidly growing firms, and firms that rely more heavily upon reinsurance to support their capital positions are associated with higher failure rates. Larger firms and insurers that are part of mutual organizations are relatively less likely to default. Finally, firms that have high cash outflows relative to inflows and who experience adverse reserve development are more likely to fail. Overall the explanatory power of the model is reasonable as the pseudo R^2 statistic is 26 percent.

[17] For example, Pottier and Sommer (1999) use a four category system to map the individual ratings assigned by each agency. Both the Governemnt Accounting Office (1994) and Doherty and Phillips (2002) use a five level system but the individual ratings assigned to the five categories differ slightly. In work not shown here, we compared the results of the five-level and four-level categorization systems with the results shown in this paper. Some of the details are different, but the primary conclusions are similar regardless of which categorization is used.

[18] In the late 1980's, the highest rating category in the A.M. Best rating system was A+. In 1992, Bests announced the expansion of its rating scale to add finer distinctions among certain ratings. One rating that was split was the A+ category which was split into A++ and A+. Another category that was split was B+ which was split into B++ and B+. See A.M. Best (1992).

Panel B of Table 4 shows summary statistics of the estimated one-year probabilities of default for solvent and insolvent insurers. The average/median probability of default for the healthy firms is 0.8/0.2 percent while the average/median statistics for the firms in the year before they become bankrupt is 9.2/4.4 percent. Thus, the average estimated one-year probability of default for firms that become bankrupt in the next year is over 10 times larger than the average probability for healthy firms. Clearly the model does a reasonable job assigning high default probabilities to firms that ultimately fail and low probabilities to healthy firms. We also note that A.M. Best reports the average annual probability of default for property-liability insurers from 1991-2002 was 0.95 percent - a result very consistent with the probabilities produced by our model (A.M. Best, 2004).

We now consider differences in the average financial quality of the insurers rated by each agency. To do so we calculate the average and median probability of default for the insurers that requested a full rating by A.M. Best or by S&P over the time period of this study. The results are shown in Table 5. Except for the 1989, when S&P only rated three insurers, the average and median probability of default statistics for S&P rated firms that requested a full rating are always lower through 1998 and these difference are always statistically significant at a one percent p-value.[19] To give economic significance, the analysis suggests that from 1989 through 1998, the one-year default probability for the average insurer that requested a rating from S&P was 56% lower than the probability of default for the average firm that requested a rating from A.M. Best. These results are consistent with our hypothesis that insurers of higher than average quality will have the greatest demand to request a rating from the new entrant.

It is also interesting to note two additional facts displayed in Table 5. First, differences in average or median quality appear to converge over time as S&P's presence in this market grew more substantial. This result is most easily seen in the t-statistics which generally decline over time. Second, unlike the comparison between A.M. Best rated firms and insurers requesting a full rating from S&P, the average/median probability of default statistics for insurers that requested an A.M. Best rating relative to those that were assigned a qualified rating by S&P are never statistically significant. Although interesting, this latter result is neither consistent nor inconsistent with our model as we do not theoretically consider the rationale for an information intermediary to assign qualified ratings.

Table 6 and the accompanying chart, Figure T6, shows the average rating issued by each agency over the time period of our study where the numerical scores assigned to each category were shown in Table 3. The results stand in stark contrast to those reported in Table 5 as the average rating issued by A.M. Best over this time period was very similar to the average rating S&P issued on a full unqualified basis even though the probability of default for the average

[19] Table 5 only shows the statistics for the null hypothesis of differences in means. The statistics for the null hypothesis of differences in medians provide similar results are thus are not shown. The results are available from the authors upon request.

insurer that requested a rating from S&P was significantly lower . This result is consistent with our hypothesis that the entrant CRA will have higher standards, on average, in order for a firm to receive a rating similar to the one they received from the incumbent CRA. By contrast, the average qualified rating issued by S&P was always dramatically lower than the average rating issued by A.M. Best even though the average default probabilities for S&P qualified vs. A.M. Best insurers were essentially identical. Even after 1994 when S&P removed the restriction that firms could not receive a rating above BBB on a qualified basis, the average qualified rating is always significantly less than the average rating given by S&P on an unqualified basis using the traditional methodology.

The final univariate comparison we make is shown in Table 7 where we compare the ratings assigned to insurers that requested a rating from both agencies. The data in Panel A includes all firm-year observations over the years 1989 - 2000 that received a rating from A.M. Best and a full rating from S&P. The data in Panel B includes all firm-year observations over the years 1991 - 2000 that received a rating from A.M. Best and a qualified rating from S&P. Each cell of the matrix, c_{ij}, equals the number of firm-year observations that receive rating i from A.M. Best and rating j from S&P where $i, j \epsilon$ {Superior, Excellent, Good, Marginal}.

The results in Panel A suggest S&P generally agreed with the rating that A.M. Best assigned to the same insurer as the two agencies issued the identical rating for 56.6 percent of the firm-year observations. Panel A also suggests that, when S&P's opinion differed from A.M. Best, it was generally more pessimistic as 39.5 percent of the firm-year observations represent cases where S&P assigned a lower rating to the insurer than did A.M. Best. S&P assigned a higher rating to an insurer than did A.M. Best in only 3.9 percent of the firm-year observations. Consistent with our earlier comments, the results shown in Panel B where S&P was assigning a qualified rating, are quite different as S&P issued the identical rating as A.M. Best in only 19.8 percent of the firm-year observations and they issued a lower rating in 78.9 percent of the observations. It is also interesting to note that it was quite routine for S&P's rating to be two or three grades lower than A.M. Best's on a qualified basis yet this almost never occurred when S&P was assigning the rating using their full methodology.

6.2 Investigating the Differences in Ratings

The univariate results suggest insurers that opted to receive a full unqualified rating from S&P were, on average, of higher financial quality yet the average rating they received was either the same or lower than the rating they received from A.M. Best. Although consistent with our theory, this analysis suffers from two shortcomings that preclude firm conclusions. First, we have only been able to compare the two rating systems for firms that received a rating from both agencies. Thus, we have not ruled out the possibility that differences in the assigned ratings may be distorted by a potential selection bias between insurers that chose to be rated by the new entrant and those that did not. Second, the hazard model used to calculate the one-year probabilities of default was estimated using publicly available information only. Presumably one of the advantages of a rating system is the ability of the CRA to learn private information. Thus, in this section, we present a methodology where we investigate the determinants of differences in the ratings assigned by the incumbent versus the entrant CRA while controlling for the private information the agencies learn through the rating process and the strategic incentives of the insurer.

Empirical Strategy. Assume the following model is used by the incumbent rating agency to determine the rating for a particular firm:

$$r_{if} = \alpha_i + \boldsymbol{\beta}_i' \mathbf{X}_f + \varepsilon_{if} \tag{4}$$

where

r_{if} = rating issued firm f by the incumbent agency

α_i = constant term for the incumbent agency

β_i = vector of coefficients summarizing the incumbent agency's rating technology

X_f = vector of observable information for firm f

ε_{if} =error term of the incumbent agency's rating of firm f

In addition, assume the new entrant has a model of similar structure. We want to explain differences between the new entrant's ratings and the incumbent's, that is

$$r_{ef} - r_{if} = (\alpha_e - \alpha_i) + (\boldsymbol{\beta}_e - \boldsymbol{\beta}_i)' \mathbf{X}_f + (\varepsilon_{ef} - \varepsilon_{if}), \tag{5}$$

where all variables subscribed with an e denote the new entrant agency. As discussed by Cantor and Packer (1997), estimating equation (5) directly using OLS will lead to biased results if the decision to seek a second rating is correlated with the ratings assigned by that agency. This selection bias makes it impossible to understand if the estimated differences between the two rating systems are due to actual differences between the rating scales or because the sample of firms that choose to get a rating from the new entrant have a common set of characteristics. In particular, the theory presented in this paper suggests that firms which elect to receive a full rating from S&P will be those that are of higher than average financial quality in their rating category and have some belief that they are likely to obtain a favorable outcome from the new entrant. Thus, the average rating difference that we see may underestimate the true difference in standards across the two rating systems.

We employ a standard Heckman two-stage regression methodology to control for this potential sample selection bias (Heckman (1979)). The Heckman methodology is ideal in this setting because it not only allows us to control for the possibility of selection bias, but we can also incorporate private information garnered in the ratings process by including the insurer's A.M. Best rating as explanatory variables. The empirical procedure is as follows: we first estimate a Probit regression that models the insurer's decision to request a second rating by S&P; second, we use the results of the Probit regression to estimate an inverse Mill's ratio, that, when included in the ratings difference model, will control for the selection bias and lead to unbiased estimates. Thus, in the second stage we estimate

$$r_{ef} - r_{if} = \alpha + \gamma IMR_f + n_f,$$

where the estimated constant term measures the mean difference in ratings standards across the two agencies and the coefficient on the inverse Mill's ratio (IMR) captures the sample selection effect.[20] We hypothesize α will be negative consistent with our theory that the new entrant, on average, will employ higher rating standards than the incumbent firm. In addition, we predict the estimated coefficient γ will be positive consistent with the hypothesis that insurers who believe they will receive a favorable rating from S&P will self-select to receive that rating.

We include the following variables to explain demand for a second rating. First, theory suggests that higher quality insurers within each rating category of the incumbent have a stronger demand to receive a second rating from S&P. To test this hypothesis, we construct a variable for each insurer equal to the difference between the median probability of default of insurers within its A.M. Best's rating category and insurer's own estimated default probability. We expect a positive estimated coefficient on this variable.

Second, we include a variable where we interact the "higher quality than the average insurer" variable with a time trend variable to control for differences in demand over time. We expect insurers with the highest demand will choose a rating from the new entrant first and this effect will dissipate over time. Thus, we expect a positive coefficient on the "higher than average insurer" variable and a negative coefficient on the interacted variable.

Third, we include indicator variables for each rating category assigned by A. M. Best to the insurer to control for any differences in the demand for a second rating based upon insurers' credit quality and to control for the private information that Best's learns during the rating process.

Fourth, the model predicts that rating agencies have incentives to reveal more information to the market when the buyers value information higher. To test this hypothesis we include a variable equal to the percentage of the insurer's premiums in retail lines of insurance[21]. Since state guaranty funds provide greater protection to the retail policyholders of insurers that become bankrupt, we expect retail policyholders should place less value on information and therefore find little value in the additional information revealed by obtaining the second rating.

[20] Note, differences in rating standards can be due to a shift in the cardinal ranking across the two systems (i.e., differences in the intercept terms) or due to different weightings employed by the two agencies (i.e., differences in the beta coefficients). We are unaware of any theory that can guide us in selecting exogenous variables that might explain why two agencies might place different weighting on rating factors. Therefore, like Cantor and Pakcer (1997), we only include an intercept term and a control for sample selection bias in the second stage rating difference regressions.

[21] The lines of insurance we considered to be retail lines included personal automobile insurance (both liability and property damage), homeowners insurance and farmowners insurance.

We include two variables to control for differences in firm complexity: a size variable (equal the natural logarithm of the firm's assets), and a variable measuring the geographical concentration of the firm's premium writings (a Herfindahl index of the premiums written across each state in which the insurer operates). We expect a positive coefficient on the firm size variable and a negative coefficient on the Herfindahl index consistent the hypothesis larger and more geographically dispersed insurers are more complex.

Finally, we include an indicator variable for whether an insurer is a mutual or reciprocal. We have two competing hypotheses this variable. First, the managerial discretion literature predicts that mutual insurers should underwrite less risky lines of insurance and have more transparent business models than stock insurers due to the reduced ability of a diffuse set of owner/policyholders to monitor management (Mayers and Smith (1987)). This model suggests that mutuals should have a lower demand for a second rating. An alternative is that stock insurers have lower demand for an additional rating because their financial quality is already conveyed in share price.

The data for the rating difference tests includes any insurer that received a rating from A.M. Best over the time period 1989 - 2000. We estimate the first stage probit regression using all firm-year observations and the second stage OLS regressions only for insurers that receive full ratings from S&P. There are 14,956 insurer-year observations in the Best's sample of which 1,522 also obtained a rating S&P on a fully unqualified basis.

Results[22]. Table 8 displays summary statistics for all the variables used in the ratings differences tests. The most important statistic in this table is the average difference in rating assigned to an insurer by S&P relative to A.M. Best which, over the time period of this study was 0.35 notches lower. The results in Table 8 also suggest that insurers that requested a full rating were less likely to be a member of a mutual group of insurers than the average Best insurer, were much larger than the average Best insurer (based on assets under management), and conducted less business in retail lines of insurance. All of these latter results are consistent with prior hypotheses.

The first-stage probit regression results are shown in Table 9. The most important results concern the test that higher than average quality insurers in each rating category of the incumbent are more likely to demand a second rating. The variable measuring the difference between the median insurer's probability of default in the Best's rating category and the estimated default probability for the insurer, is both positive and significant as predicted. At the same time, the interaction of this variable with the time indicator is negative. This suggests that the importance of S&P's entry strategy to target the best companies in each rating category declined over time. The results for the separate indicator variables for Best's rating classes reveal that

[22] We only show results in the paper based upon insurers that requested a full rating from S&P. The results for insurers that requested an unqualified rating are similar to those reported here and are available upon request. We chose not to display results on the unqualified insurer data sample as our theoretical model assumes a full ratings process.

the likelihood of requesting a second rating from S&P increases as the quality of the insurer (as rated by Best's) also increases. For example, focusing on the marginal effects of the model, a firm rating Superior by A. M. Best full rating category, is 23.9% more likely to request a full rating from S&P than an insurer who Best's rates as Marginal.

Consistent with our theory, large insurers and insurers with more complex businesses are all more likely to seek an additional rating. Finally, mutual insurers are less likely to request a full rating from S&P consistent with the managerial discretion hypothesis. We find little support for the hypothesis that insurers writing more retail lines of insurance have lower demand for a second rating.

The results from the second stage OLS regressions are shown in Table 10. Each model shown in Table 10 corresponds to the estimated Probit regression model shown in Table 9 and contains the inverse Mill's term corresponding to that same model. Our first conclusion is that we find evidence of a selection effect as the estimated coefficient on the inverse Mill's ratio is always positive and significantly different from zero. Thus, we find strong evidence insurers that seek a second rating expect, on average, to receive a more favorable rating from S&P even after we control for publicly available information to market participants and for the private information revealed to A. M. Best. For example, based on the results shown in Model 2, insurers strategically choosing a second rating expect to receive, on average, a 0.88 higher rating from S&P than they received from A. M. Best. These results support the contention that better than average insurers within any pre-existing Best's rating category sought to differentiate themselves through a second rating.

The second conclusion we draw is that S&P maintained significantly higher standards relative to Best's as the intercept term in each model is negative and significantly different than zero. Focusing again on Model 2, we see that the estimated mean difference in rating standards is 1.23 grades lower on the S&P scale than under the A.M. Best's rating system. This result provides strong evidence consistent with our theory that, conditional upon the rating provided by the incumbent agency, insurers sought to differentiate themselves by seeking an additional rating from a new entrant agency that required higher standards in order to maintain the same rating.

7 Conclusion

This paper analyzes optimal information disclosure of a monopoly rating agency depending on the marginal value of information to buyers. The paper also characterizes the optimal entry strategy of a rating agency to a market previously dominated by an incumbent. The qualitative results of the paper are that the information disclosure of the rating agency significantly depends on the value of information to its end users. As the value of information increases, the ratings become more precise. The entry strategy of a new agency is to target the companies of the highest financial quality in each rating class. This policy benefits companies who voluntarily obtain the second rating. However, it decreases the payoff to companies on the bottom side of each rating class. These results are supported by our empirical analysis of the insurance industry.

References

[1] Altman, Edward I., 1968, "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", *Journal of Finance,* 23: 589-609.

[2] A.M. Best, 2004, "Best's Insolvency Study: Property/Casualty U.S. Insurers 1969-2002 (Oldwick, NJ: A.M. Best Company, Inc).

[3] Benabou, Roland and Guy Laroque, 1992, "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility", *Quarterly Journal of Economics*, 107(3): 921-958.

[4] Boot, Arnoud W.A., Todd T. Milbourn, Anjolein Schmeits, 2006, "Credit Ratings as Coordination Mechanisms", *Review of Financial Studies*, 19(1): 81-118.

[5] Bradford, Michael, 2003, "Big Changes at Kemper Prompt Buyer Concerns", *Business Insurance*, January 6.

[6] Cantor, Richard, and Frank Packer, 1997, "Differences of Opinion and Selection Bias in the Credit Rating Agency," *Journal of Banking and Finance*, 21: 1395-1417.

[7] Cummins, J. David, Martin F. Grace, Richard D. Phillips, 1999, "Regulatory Solvency Prediction in Property-Liability Insurance: Risk-Based Capital, Audit Ratios, and Cash Flow Simulation", *Journal of Risk and Insurance*, 66(3): 417-458.

[8] Cummins, J. David, Scott E. Harrington and Robert W. Klein, 1995, "Insolvency Experience, Risk-Based Capital, and Prompt Corrective Action in Property-Liability Insurance", *Journal of Banking and Finance*, 19: 511-527.

[9] Doherty, Neil A. and Richard D. Phillips, 2002, "Keeping up with the Joneses: Changing Rating Standards and the Buildup of Capital by U.S. Property-Liability Insurers", *Journal of Financial Services Research,* 21: 55–78.

[10] Dolton, P.J. and G.H. Makepeace, 1987, "Interpreting Sample Selection Effects", *Economics Letters* 24, 373-9.

[11] Epermanis, Karen and Scott E. Harrington, 2005, "Market Discipline in Property/Casualty Insurance: Evidence from Premium Growth Surrounding Changes in Financial Strength Ratings", *Journal of Money, Credit and Banking*, 38(6): 1515-1544.

[12] Epstein, Larry G., Martin Schneider, 2008, "Ambiguity, Information Quality, and Asset Pricing", *Journal of Finance,* 63(1):197-228.

[13] Farhi, Emmanuel, Josh Lerner and Jean Tirole, 2008, "Fear of Rejection? Tiered Certification and Transparency", working paper Harvard and Toulouse.

[14] Grace, Martin, Scott E. Harrington and Robert W. Klein, 1995, *An Analysis of the FAST Solvency Monitoring System*, presented to the NAIC's Financial Analysis Working Group. (Kansas City: NAIC)

[15] Grossman, Sanford J., 1981, "The Informational Role of Warranties and Private Disclosure about Product Quality", *Journal of Law and Economics*, 24(3): 461-483.

[16] Grossman, Sanford J. and Olivier Hart, 1980, "Disclosure Laws and Takeover Bids", *Journal of Finance*, 35: 323-334.

[17] Heckman, James J., 1979, "Sample Selection Bias as a Specification Error", *Econometrica*, 47(1): 153-161.

[18] Lerner, Josh and Jean Tirole, 2006, "A Model of Forum Shopping", *American Economic Review*, 96(4): 1091-1113.

[19] Lewis, Christopher M. and Kevin C. Murdock, 1996, "The Role of Government Contracts in Discretionary Reinsurance Markets for Natural Disasters", *Journal of Risk and Insurance*, 63(4):567-597.

[20] Lizzeri, Alessandro, 1999, "Information Revelation and Certification Intermediaries", *RAND Journal of Economics*, 30(2): 214-231.

[21] Mathis, Jerome, James McAndrews and Jean-Charles Rochet, 2008, "Rating the Raters: Are Reputational Concerns Powerful Enough to Discipline Rating Agencies?", working paper Toulouse.

[22] Mayers, D., and C. W. Smith, Jr., 1987, "Corporate Insurance and the Underinvestment Problem", *Journal of Risk and Insurance*, 54(1): 45-54.

[23] Milgrom, Paul R., 1981, "Good News and Bad News: Representation Theorems and Applications", *Bell Journal of Economics*, 12(2): 380-391.

[24] Millon, M.H. and A.V. Thakor, 1985, "Moral Hazard and Information Sharing: A Model of Financial Information Gathering Agencies", *Journal of Finance*, December, 1403-1422.

[25] Morrison, Andrew and Lucy White, 2005, "Crises and Capital Requirements in Banking", *American Economic Review*, 95(5): 1548 - 1572.

[26] Ottaviani, Marco and Peter Norman Sorensen, 2006a, "Professional Advice", *Journal of Economic Theory*, 126: 120-142.

[27] Ottaviani, Marco and Peter Norman Sorensen, 2006b, "Reputational Cheap Talk", *RAND Journal of Economics*, 37(1): 155-175.

[28] Ottaviani, Marco and Peter Norman Sorensen, 2006c, "The Strategy of Professional Forecasting", *Journal of Financial Economics*, 81: 441-466.

[29] Peyrache, Eloic and Lucia Quesada, 2005, "Intermediaries, Credibility and Incentives to Collude", working paper, HEC Paris and Universidad Torcuato Di Tella.

[30] Pottier, Stephen W., and David W. Sommer, 1999, "Property-Liability Insurer Financial Strength Ratings: Differences Across Rating Agencies," *Journal of Risk and Insurance* 66(4): 621-642.

[31] Ramakrishnan, R.T.S., Thakor, A., 1984, "Information reliability and a theory of financial intermediation", *Review of Economic Studies,* 415–432.

[32] Scharfstein, David, and Jeremy Stein, 1990, "Herd Behavior and Investment", *American Economic Review*, 80: 465-479.

[33] Shaked, Avner, and John Sutton, 1982, "Relaxing Price Competition through Product Differentiation", *Review of Economic Studies*, 49(1):3-13.

[34] Shumway, Tyler, 2001, "Forecasting Bankruptcy More Accurately: A Simple Hazard Model," *Journal of Business* 74(1): 101-124.

[35] Skreta, Vasiliki and Laura Veldkamp, 2008, "Rating Shopping and Asset Complexity: A Theory of Ratings Inflation", working paper, NYU Stern.

[36] Standard and Poor's Insurance Company Rating Guide, 1995 Edition, McGraw-Hill.

[37] Strausz, Roland, 2005, "Honest Certification and the Threat of Capture", *International Journal of Industrial Organization*, 23:45-62.

[38] Veronesi, Pietro, 2000, "How does Information Quality Affect Stock Returns?", *Journal of Finance*, 55: 807-837.

[39] White, Lawrence, J., 2002, "The Credit Rating Industry: An Industrial Organization Analysis," in R.M. Levich, G. Majnoni, and C.M. Reinhart, eds. Ratings, Rating Agencies and the Global Financial System, Kluwer, 2002.

[40] Winter, Ralph A., 1991, "The Liability Insurance Market", *Journal of Economic Perspectives*, 5(3):115-136.

Appendix: Proofs

Proof of Proposition 1. We first prove uniqueness of equilibrium under full disclosure. Consider a set of non-rated types V_N. Then the valuation of these types in the eyes of buyers is

$$E(v|V_N) - aVar(v|V_N),$$

where

$$E(v|V_N) = \frac{1}{|V_N|}\int_{V_N} vdv,$$

$$Var(v|V_N) = \frac{1}{|V_N|}\int_{V_N}(v - E(v|V_N))^2 dv.$$

If $V_N = [0,1]$, then the valuation is equal to $max(0, \frac{1}{2} - \frac{1}{12}a)$. If type $v = 1$ is the only one rated, it is paid 1. When $t < 1 - max(0, \frac{1}{2} - \frac{1}{12}a) = \max(1, \frac{1}{2} + \frac{1}{12}a)$, among the non-rated types there are types that prefer to be rated. Denote v_r any of these types. Then all types above v_r prefer to be rated. Also the benefits of rating are decreasing for types below v_r, and there is a type $v_F < v_r$ that is indifferent between being rated or not.

We now derive the optimal market coverage v_F. Consider two cases, $u_N > 0$ and $u_N < 0$. $u_N > 0$ is equivalent to

$$\frac{1}{2} - \frac{1}{12}av_F > 0. \tag{6}$$

If $u_N > 0$, the agency charges the fee $t = v_F - u_N$, and the problem of the rating agency writes

$$(1 - v_F)(\frac{1}{2}v_F + \frac{1}{12}av_F^2)$$
$$\text{subject to (6).}$$

Denote $\lambda \geq 0$ the Lagrangian multiplier of (6). Suppose first that $\lambda > 0$. Then $v_F = \frac{6}{a}$, and

$$\lambda = \frac{12}{a}(\frac{3}{2} - \frac{15}{a}),$$

so $\lambda > 0$ when $a > 10$. In this case the profit of the agency is $\frac{6(a-6)}{a^2}$. Now suppose that (6) is not binding, $\lambda \geq 0$. Then $v_F = \frac{a-6+\sqrt{a^2+6a+36}}{3a}$, and (6) is satisfied when $a < 10$. The profit is $\frac{a^3+9a^2-54a-216+(6a+a^2+36)^{\frac{3}{2}}}{162a^2}$.

Consider the case $u_N < 0$. In this case the agency charges the fee $t = v_F$, and the problem of the rating agency writes

$$(1 - v_F)v_F$$
$$\text{subject to } -\frac{1}{2} + \frac{1}{12}av_F > 0.$$

Denote $\lambda \geq 0$ the Lagrangian multiplier of the constraint. If $\lambda > 0$, then $v_F = \frac{6}{a}$ and $\lambda = \frac{12}{a}(\frac{12}{a} - 1)$, implying $a < 12$. The profit in this case is $\frac{6(a-6)}{a^2}$. Now assume that the constraint is not binding. Then $v_F = \frac{1}{2}$, the profit is $\frac{1}{4}$, and the constraint is satisfied when $a > 12$.

To find the optimal v_F, compare the solutions in cases $u_N > 0$ and $u_N < 0$ for different values of a. When $a < 10$, the global solution is $v_F = \frac{a-6+\sqrt{a^2+6a+36}}{3a}$, resulting in profit $\frac{(a+12)(a+3)(a-6)+(6a+a^2+36)^{\frac{3}{2}}}{162a^2}$. When $10 < a < 12$, solutions in the two cases are the same, $v_F = \frac{6}{a}$, and the profit is $\frac{6(a-6)}{a^2}$. Finally, when $a > 12$, the global solution is $v_F = \frac{1}{2}$ and the profit is $\frac{1}{4}$. ■

Proof of Proposition 2. Define a disclosure policy correspondence

$$s(v) = \{[s_i(v), s_{i+1}(v)]_{i=1}^{K(v)}\} \text{ for all } v \in [0,1],$$

where $v \in s(v)$, $s_i(v) \leq s_{i+1}(v)$, $s_1(v) \geq 0$, $s_{K(v)}(v) \leq 1$. Under this policy a rating assigned to type v is a union of intervals and points with elements from $[0,1]$ containing v. For example, $v = \frac{1}{5}$ and $s(\frac{1}{5}) = \{[\frac{1}{6}, \frac{2}{5}], [\frac{1}{2}, \frac{2}{3}], \frac{9}{10}\}$ means that if type $v = \frac{1}{5}$ is rated, the rating agency discloses that the seller's type can be in intervals $[\frac{1}{6}, \frac{2}{5}]$ or $[\frac{1}{2}, \frac{2}{3}]$, or equal to $\frac{9}{10}$. Note that $\frac{1}{5} \in s(\frac{1}{5})$ and $K(\frac{1}{5}) = 3$. The disclosure policy can also be an infinite sequence of elements, for example, $v = \frac{5}{9}$ and $s(v) = \{(\frac{i}{i+1}, \frac{2i+1}{2i+3})_{i=1}^{\infty}\}$. Clearly, any disclosure policy can be characterized as $s(\cdot)$.

Denote $\underline{v}(s)$ the lowest rated type under disclosure policy s. Assume that all types $v > \underline{v}$ are rated and obtain a valuation of at least $u(s(\underline{v}))$. Then the participation constraint of type $\underline{v}$ determines the fee $t = u(s(\underline{v})) - \max\{u([0,\underline{v}]), 0\}$, where $u([0,\underline{v}])$ is the valuation of non-rated types $[0,\underline{v}]$. The lowest payoff of type $\underline{v}$ equals to $\underline{v}$ and obtains when the signal $s(\underline{v})$ contains only $\underline{v}$. Thus the fee can be increased only if $s(\underline{v})$ contains some higher types $v > \underline{v}$.

Suppose that $s(\underline{v})$ consists of a system of disjoint intervals, $s(\underline{v}) = \{[\underline{v}, v_1], ...[v_{K(v)-1}, v_{K(v)}]\}$ with $v_{2j-1} < v_{2j}$ for all j such that $2j < K(\underline{v}) \leq \infty$. Then

$$\mu(s(\underline{v})) = \frac{\int_{\underline{v}}^{v_1} v dF(v) + \int_{v_2}^{v_3} v dF(v) + ... + \int_{v_{K(\underline{v})-1}}^{v_{K(\underline{v})}} v dF(v)}{F(v_1) - F(\underline{v}) + F(v_3) - F(v_2) + ... + F(v_{K(\underline{v})}) - F(v_{K(\underline{v})-1})},$$

$$\sigma^2(s(\underline{v})) = \frac{\int_{\underline{v}}^{v_1} (v - \mu(s(\underline{v})))^2 dF(v) + \int_{v_2}^{v_3} (v - \mu(s(\underline{v})))^2 dF(v) + ... + \int_{v_{K(\underline{v})-1}}^{v_{K(\underline{v})}} (v - \mu(s(\underline{v})))^2 dF(v)}{F(v_1) - F(\underline{v}) + F(v_3) - F(v_2) + ... + F(v_{K(\underline{v})}) - F(v_{K(\underline{v})-1})}.$$

Define $\widehat{s}(\underline{v}) = [\underline{v}, \widehat{v}]$ where $\widehat{v}$ is such that

$$\frac{1}{F(\widehat{v}) - F(\underline{v})} \int_{\underline{v}}^{\widehat{v}} v dF(v) = \mu(s(\underline{v})).$$

That is, under disclosure policy $\widehat{s}$ type $\underline{v}$ is pooled with types $[\underline{v}, \widehat{v}]$ in the same signal, and $\widehat{v}$ is such that $\mu(\widehat{s}(\underline{v})) = \mu(s(\underline{v}))$. Then it is easy to show that the signal $\widehat{s}(\underline{v}) = [\underline{v}, \widehat{v}]$ results in lower variance $\sigma^2(\widehat{s}(\underline{v}))$ than $\sigma^2(s(\underline{v}))$. Since increasing σ^2 for fixed μ reduces the fee t, we conclude that it is optimal to pool type $\underline{v}$ with neighboring types $(\underline{v}, \widehat{v}]$, $\underline{v} \leq \widehat{v}$.

It remains to prove that under an optimal disclosure policy all types $v > \widehat{v}$ are rated and obtain a valuation of at least $\mu(s(\underline{v})) - \sigma^2(s(\underline{v}))$. Suppose that there is a rated type $\widehat{\widehat{v}} > \widehat{v}$ such that $\mu(s(\widehat{\widehat{v}})) - a\sigma^2(s(\widehat{\widehat{v}})) < \mu(s(\underline{v})) - a\sigma^2(s(\underline{v}))$. It means that $\widehat{\widehat{v}}$ is either pooled with lower types resulting in lower $\mu(s(\widehat{\widehat{v}}))$, or with types that are higher but are very distinct resulting in high $\sigma^2(s(\widehat{\widehat{v}}))$, or both. At the same time $\widehat{\widehat{v}}$ can always obtain a payoff $\widehat{\widehat{v}} > \mu(s(\underline{v})) - a\sigma^2(s(\underline{v}))$ if the type is fully disclosed, and $s(\widehat{\widehat{v}}) = \widehat{\widehat{v}}$. As long as $\widehat{\widehat{v}}$ obtains a rating, its valuation has no impact on the fee charged by the agency. Thus the rating agency cannot benefit by designing a disclosure policy resulting in payoff of $\widehat{\widehat{v}}$ lower than $\mu(s(\underline{v})) - a\sigma^2(s(\underline{v}))$. For the same reason the rating agency cannot benefit from excluding types $v > \widehat{v}$ from the rating. We conclude that under optimal rating system all types $v > \underline{v}$ are rated, types $[\underline{v}, \widehat{v}]$, $\underline{v} \leq \widehat{v} \leq 1$ are pooled in one rating, and types $v > \widehat{v}$ obtain a valuation of at least $\mu(s(\underline{v})) - a\sigma^2(s(\underline{v}))$.

There are multiple disclosure policies for types $v > \widehat{v}$ that are compatible with the last condition. In particular, any interval disclosure system that results in valuation of at least $\mu(s(\underline{v})) - a\sigma^2(s(\underline{v}))$ is an equilibrium.

As the value of information increases, $a \to +\infty$, pooling has infinite costs. Consequently, the support of each rating contains only one type and has zero variance. It implies that as the value of information tends to infinity, the optimal disclosure policy converges to full disclosure.

■

Proof of Proposition 3. Let's consider a rating system when the agency pools companies $[v_M, v_M + b_M]$ in one rating. We distinguish between two cases, $u_N > 0$ and $u_N < 0$.

Consider a rating system with $u_N > 0$. The problem of the rating agency writes

$$\begin{aligned} \max_{(b_M, v_M)} (1 - v_M)(u_R(v_M, v_M + b_M) - u_N) &= (1 - v_M)(\frac{1}{2}v_M + \frac{1}{12}av_M^2 + \frac{1}{2}b_M - \frac{1}{12}ab_M^2) \\ \frac{1}{2} - \frac{1}{12}av_M &\geq 0, \qquad (7) \\ 1 - b_M - v_M &\geq 0. \qquad (8) \end{aligned}$$

Constraint (7) is equivalent to $u_N > 0$, and (8) is a feasibility condition. Denote $\lambda \geq 0$ and $\mu \geq 0$ the Lagrangian multipliers of these constraints. The first order conditions of the problem are

$$\begin{aligned} b_M: &\quad (1 - v_M)(\tfrac{1}{2} - \tfrac{1}{6}ab) - \mu = 0, \\ v_M: &\quad -\tfrac{1}{4}av_M^2 + \tfrac{1}{6}(a - 6)v_M + \tfrac{1}{2} - \tfrac{1}{2}b_M + \tfrac{1}{12}ab_M^2 - \tfrac{1}{12}a\lambda - \mu = 0. \end{aligned}$$

Suppose that $\lambda > 0$ and $\mu > 0$. Then $v_M = \frac{6}{a}$ and $b_M = 1 - \frac{6}{a}$. It implies that $\mu = \frac{(a-6)(9-a)}{6a}$ and $\lambda = \frac{3(a-10)}{a}$. $\mu > 0$ when $6 < a < 9$, and $\lambda > 0$ when $a > 10$. A contradiction.

Suppose that $\lambda > 0$ and $\mu = 0$. Then $b_M = \frac{3}{a}$ and $v_M = \frac{6}{a}$. It implies that $\lambda = \frac{9(2a-21)}{a^2}$, and $\lambda > 0$ when $a > \frac{21}{2}$. $\mu = 0$ implies that (8) must be satisfied, $\frac{6}{a} + \frac{3}{a} < 1$, or $a > 9$. Then this case is possible when $a > \frac{21}{2}$. The profit of the rating agency in this case is $\frac{27(a-6)}{4a^2}$.

Suppose that $\lambda = 0$ and $\mu > 0$. Then $b_M = 1 - v_M$, and $\mu = (1 - v_M)(\frac{1}{2} - \frac{1}{6}a(1 - v_M))$. The first order condition with respect to v_M writes $\frac{1}{4}a - \frac{1}{3}av_M - \frac{1}{2} = 0$, and $v_M = \frac{3}{4} - \frac{3}{2a}$. $v_M > 0$ when $a > 2$. $\mu = \frac{36-a^2}{96a}$, and $\mu > 0$ when $a < 6$. $\lambda = 0$ implies that (8) must be satisfied, $\frac{3}{4} - \frac{3}{2a} \leq \frac{6}{a}$, or $a < 10$. Then this case is possible when $2 < a < 6$. The profit of the rating agency in this case is $\frac{(a+6)^2}{96a}$. If $a < 2$, then $v_M = 0$, and $b_M = 1$. The profit of the rating agency in this case is $\frac{1}{2} - \frac{1}{12}a$.

Suppose that $\lambda = \mu = 0$. Then $b_M = \frac{3}{a}$, and the first order condition with respect to v_M writes $-\frac{1}{4}av_M^2 + \left(\frac{1}{6}a - 1\right)v_M + \frac{1}{2} - \frac{3}{4a} = 0$ implying that $v_M = \frac{2}{3} - \frac{1}{a}$. $\lambda = 0$ implies that $\frac{2}{3} - \frac{1}{a} \leq \frac{6}{a}$, or $a \leq \frac{21}{2}$. $\mu = 0$ implies that $\frac{3}{a} + \frac{2}{3} - \frac{1}{a} \leq 1$, or $a \geq 6$. Then this case is possible when $6 \leq a \leq \frac{21}{2}$. The profit of the agency in this case is $\frac{(a+3)^3}{81a^2}$.

The next table summarizes the case $u_N > 0$.

a	v_M	b_M	t_m	π_M
$0 \leq a \leq 2$	0	1	$\frac{1}{2} - \frac{1}{12}a$	$\frac{1}{2} - \frac{1}{12}a$
$2 \leq a \leq 6$	$\frac{3}{4} - \frac{3}{2a}$	$\frac{1}{4} + \frac{3}{2a}$	$\frac{1}{4} + \frac{1}{24}a$	$\frac{(a+6)^2}{96a}$
$6 \leq a \leq \frac{21}{2}$	$\frac{2}{3} - \frac{1}{a}$	$\frac{3}{a}$	$\frac{(a+3)^2}{27a}$	$\frac{(a+3)^3}{81a^2}$
$a \geq \frac{21}{2}$	$\frac{6}{a}$	$\frac{3}{a}$	$\frac{27}{4a}$	$\frac{27(a-6)}{4a^2}$

Consider an alternative case with $u_N < 0$. The problem of the rating agency in this case writes

$$\begin{aligned} \max_{(b_M, v_M)} (1 - v_M)u_R(v_M, v_M + b_M) &= (1 - v_M)(v_M + \frac{1}{2}b_M - \frac{1}{12}ab_M^2) \\ -\frac{1}{2} + \frac{1}{12}av_M &\geq 0 \text{ and (8).} \end{aligned}$$

Again, denote $\lambda \geq 0$ and $\mu \geq 0$ the Lagrangian multipliers of the constraints. The first order conditions of this problem write

$$\begin{aligned} b_M: &\quad (1 - v_M)(\tfrac{1}{2} - \tfrac{1}{6}ab_M) - \mu = 0, \\ v_M: &\quad 1 - 2v_M - \tfrac{1}{2}b_M + \tfrac{1}{12}ab_M^2 + \tfrac{1}{12}a\lambda - \mu = 0. \end{aligned}$$

Suppose that $\lambda > 0$ and $\mu > 0$. Then $v_M = \frac{6}{a}$ and $b_M = 1 - \frac{6}{a}$. It implies that $\lambda = -\frac{3(a^2-12a+12)}{a^2}$, and $\lambda > 0$ when $6 - 2\sqrt{6} < a < 6 + 2\sqrt{6}$. $\mu = \frac{(9-a)(a-6)}{6a}$, and $\mu > 0$ when $6 < a < 9$. Then this case is possible when $6 < a < 9$. The profit of the rating agency is $\frac{(a-6)(18-a)}{12a}$.

Suppose that $\lambda > 0$ and $\mu = 0$. Then $v_M = \frac{6}{a}$ and $b_M = \frac{3}{a}$. It implies that $\lambda = \frac{3(51-4a)}{a^2}$, and $\lambda > 0$ when $a < \frac{51}{4}$. $\mu = 0$ implies that (8) must be satisfied, $\frac{6}{a} + \frac{3}{a} \leq 1$, or $a \geq 9$. So this case is possible when $9 \leq a < \frac{51}{4}$. The profit of the rating agency in this case is $\frac{27(a-6)}{4a^2}$.

Suppose that $\lambda = 0$ and $\mu > 0$. Then $b_M = 1 - v_M$, and $\mu = (1 - v_M)(\frac{1}{2} - \frac{1}{6}a(1 - v_M))$. The first order condition with respect to v_M becomes $av_M^2 - 2(a+2)v_M + a = 0$, implying that $v_M = \frac{a+2-2\sqrt{a+1}}{a}$. Then $\mu = \frac{7\sqrt{a+1}-2a-7}{3a}$, and $\mu > 0$ when $4a^2 + 21a + 42 < 0$. A contradiction.

Suppose that $\lambda = \mu = 0$. Then $b_M = \frac{3}{a}$ and $v_M = \frac{1}{2} - \frac{3}{8a}$. $\lambda = 0$ implies that $\frac{1}{2} - \frac{3}{8a} \geq \frac{6}{a}$ must be satisfied, or $a \geq \frac{51}{4}$. $\mu = 0$ implies that $\frac{1}{2} - \frac{3}{8a} + \frac{3}{a} \leq 1$ must be satisfied, or $a \geq \frac{21}{4}$. So this case is possible when $a \geq \frac{51}{4}$. The profit of the rating agency in this case is $\frac{(4a+3)^2}{64a^2}$.

The next table summarizes the case of $u_N < 0$.

a	v_M	b_M	t_M	π_M
$6 \leq a \leq 9$	$\frac{6}{a}$	$1 - \frac{6}{a}$	$\frac{3}{2} - \frac{1}{12}a$	$\frac{(a-6)(18-a)}{12a}$
$9 \leq a \leq \frac{51}{4}$	$\frac{6}{a}$	$\frac{3}{a}$	$\frac{27}{4a}$	$\frac{27(a-6)}{4a^2}$
$a \geq \frac{51}{4}$	$\frac{1}{2} - \frac{3}{8a}$	$\frac{3}{a}$	$\frac{3}{8a} + \frac{1}{2}$	$\frac{(4a+3)^2}{64a^2}$

The global solution to the problem can be found by comparing the profit of the rating agency under the two alternative rating systems. The next table summarizes the global solution.

a	v_M	b_M	t_M	π_M
$0 \leq a \leq 2$	0	1	$\frac{1}{2} - \frac{1}{12}a$	$\frac{1}{2} - \frac{1}{12}a$
$2 \leq a \leq 6$	$\frac{3}{4} - \frac{3}{2a}$	$\frac{1}{4} + \frac{3}{2a}$	$\frac{1}{4} + \frac{1}{24}a$	$\frac{(a+6)^2}{96a}$
$6 \leq a \leq \frac{21}{2}$	$\frac{2}{3} - \frac{1}{a}$	$\frac{3}{a}$	$\frac{(a+3)^2}{27a}$	$\frac{(a+3)^3}{81a^2}$
$\frac{21}{2} \leq a \leq \frac{51}{4}$	$\frac{6}{a}$	$\frac{3}{a}$	$\frac{27}{4a}$	$\frac{27(a-6)}{4a^2}$
$a \geq \frac{51}{4}$	$\frac{1}{2} - \frac{3}{8a}$	$\frac{3}{a}$	$\frac{3}{8a} + \frac{1}{2}$	$\frac{(4a+3)^2}{64a^2}$

It completes the proof. ■

Proof of Proposition 4. Consider an interval disclosure policy,

$$\{[v_M, v_M + b_M], [v_M + b_M, v_M + b_M + b_1], ..., [v_M + b_M + \sum_{i=1}^{N-1} b_i, v_M + b_M + \sum_{i=1}^{N} b_i]\},$$

where v_M and b_M are derived in Proposition 3 and the valuation of all rated types is at least $u([v_M + b_M, v_M + b_M + b_1])$, that is, $u_k([v_M + b_M + \sum_{i=1}^{k-1} b_i, v_M + b_M + \sum_{i=1}^{k} b_i]) \geq u([v_M + b_M, v_M + b_M + b_1])$.A disclosure policy where the valuation of a seller is non-decreasing in rating must satisfy $u_{k+1}([v_M + b_M + \sum_{i=1}^{k} b_i, v_M + b_M + \sum_{i=1}^{k+1} b_i]) \geq u_k([v_M + b_M + \sum_{i=1}^{k-1} b_i, v_M + b_M + \sum_{i=1}^{k} b_i])$, or

$$\begin{aligned} v_M + b_M + \sum_{i=1}^{k-1} b_i + b_k + \frac{1}{2} b_{k+1} - \frac{1}{12} a b_{k+1}^2 \quad &\geq \quad v_M + b_M + \sum_{i=1}^{k-1} b_i + \frac{1}{2} b_k - \frac{1}{12} a b_k, \\ &\Leftrightarrow \quad b_{k+1} \leq b_k + \frac{6}{a}. \end{aligned}$$

It implies that $b_{k+1} - b_k \leq \frac{6}{a}$, and the size of an interval decreases as a increases. ■

Proof of Proposition 5. In each rating interval k denote $\widehat{v}_k$ the type that is indifferent between full disclosure and pooling,

$$\widehat{v}_k = v_M + b_M + \sum_{i=1}^{k-1} b_i + \frac{1}{2} b_k - \frac{1}{12} a b_k^2.$$

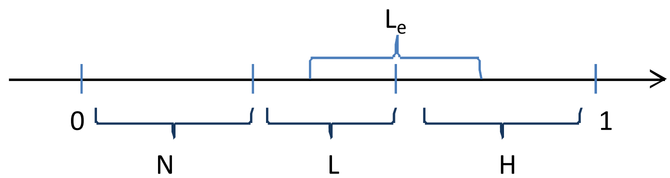


Figure 5: Ratings of the entrant must be nested in ratings of the incumbent

Types $v \in [v_M + b_M + \sum_{i=1}^{k-1} b_i, \widehat{v}_k]$ prefer pooling and types $v \in [\widehat{v}_k, v_M + b_M + \sum_{i=1}^{k-1} b_i + b_k]$ prefer full disclosure. The measures of types in each group are, respectively, $\frac{1}{2}b_k - \frac{1}{12}ab_k^2$ and $\frac{1}{2}b_k + \frac{1}{12}ab_k^2$. Hence, for any $a > 0$ the measure of types that prefer full disclosure is higher than the measure of types that prefer pooling. ■

Proof of Proposition 6. In each interval N, L, H types at the top of the interval obtain the valuation below their true types (Proposition 5). These are the segments of the market where the entrant can increase the seller's valuation. For example, if an entrant fully discloses the true type, all types $v > u(R_m)$, $v \in R_m$ obtain a second rating. However, like an incumbent, the entrant has incentives to pool the lowest rated type in each rating category (Proposition 2).

Ratings of the entrant are nested in ratings of the incumbent. Indeed, if an entrant assigns the same rating to L_e to companies rated L and H by the incumbent (see Figure 5), the entrant's rating will reduce the valuation of H sellers, and will never be requested. Thus it must be that $L_e \subset L$, and ratings of the entrant are more stringent, or precise, than ratings of the incumbent. ■

Optimal entry strategy. Optimal entry strategy solves problem (3). For each regime of the incumbent, the entrant has to decide (i) the segments of the market to enter, and (ii) the stringency of the rating in each segment. In the Supplementary material available from the authors we derive the first order conditions of the optimal entry strategy for each regime of the incumbent and solve numerically for the optimal entry strategy. ■

Table 1
Coverage of the U.S. Property-Liability Insurance Industry
By A.M. Best vs. Standard & Poor's 1989 - 2000

Table 1 displays the number of property-liabilty insurers operating in the United States, the number of firms rated by A.M. Best Company, and the number of firms that either requested a full rating or were assigned a qualified rating from Standard & Poor's over the years 1989 - 2000. The table also displays the total assets of the industry and the total assets of the firms rated by A.M. Best and Standard & Poor's. The numbers in parantheses show the percentage of firms (or assets) of the industry receiving a rating from either firm.

	Number of Companies				Total Assets ($ billions)			
Year	NAIC	A.M. Best	S&P Full	S&P Qualified	NAIC	A.M. Best	S&P Full	S&P Qualified
1989	1904	1110 (58.3%)	3 (0.2%)		534.6	491.6 (91.9%)	0.6 (0.1%)	
1990	1898	1176 (62.0%)	12 (0.6%)		575.2	520.2 (90.4%)	42.1 (7.3%)	
1991	1973	1266 (64.2%)	25 (1.3%)	610 (30.9%)	618.0	567.9 (91.9%)	44.9 (7.3%)	175.6 (28.4%)
1992	2031	1370 (67.5%)	26 (1.3%)	708 (34.9%)	688.6	626.7 (91.0%)	69.6 (10.1%)	204.7 (29.7%)
1993	2068	1444 (69.8%)	30 (1.5%)	751 (36.3%)	703.3	640.9 (91.1%)	51.3 (7.3%)	232.3 (33.0%)
1994	2068	1517 (73.4%)	29 (1.4%)	495 (23.9%)	731.5	670.9 (91.7%)	53.9 (7.4%)	115.5 (15.8%)
1995	2097	1563 (74.5%)	36 (1.7%)	723 (34.5%)	786.8	726.9 (92.4%)	30.2 (3.8%)	194.9 (24.8%)
1996	2127	1603 (75.4%)	281 (13.2%)	775 (36.4%)	842.6	786.4 (93.3%)	411.4 (48.8%)	185.1 (22.0%)
1997	2115	1617 (76.5%)	314 (14.8%)	762 (36.0%)	916.1	865.9 (94.5%)	447.7 (48.9%)	197.9 (21.6%)
1998	2136	1660 (77.7%)	354 (16.6%)	751 (35.2%)	980.7	928.8 (94.7%)	495.3 (50.5%)	202.2 (20.6%)
1999	2072	1647 (79.5%)	297 (14.3%)	706 (34.1%)	970.6	920.5 (94.8%)	283.4 (29.2%)	196.3 (20.2%)
2000	1964	1582 (80.5%)	258 (13.1%)	682 (34.7%)	942.2	892.0 (94.7%)	258.7 (27.5%)	172.6 (18.3%)

Figure T1
U.S. Property-Liability Industry Assets Covered by
A.M. Best vs. Standard & Poor's 1989 - 2000

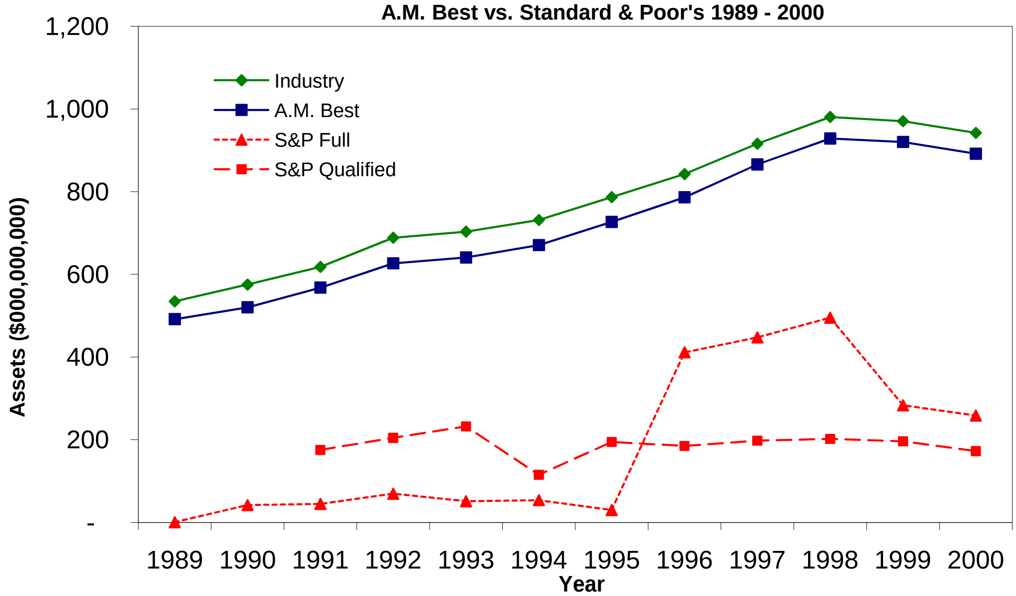


Table 2
Insurer Rating Categories: A.M Best vs. Standard & Poor's

Table 2 displays the mapping used in this study to compare financial strength rating categories across A.M. Best and Standard and Poor's. The terms shown in the column labeled "Description" are the descriptive words that each agency uses in their rating manuals to characterize a specific rating. The terms shown in the column labeled "Category" are the two broad terms each agency uses to group the individual ratings.

	A.M. Best			Standard & Poor's		
Number	**Description**	**Rating**	**Category**	**Description**	**Rating**	**Category**
3	Superior	A++,A+	Secure	Extremely Strong	AAA	Investment Grade
2	Excellent	A, A-	Secure	Very Strong	AA, A	Investment Grade
1	Good	B++,B+	Secure	Good	BBB	Investment Grade
0	Marginal	B and below	Vulnerable	Marginal	BB and below	Non-Investment Grade

Table 3: FAST Ratio and Control Variable Summary Statistics: Solvent versus Insolvent Insurers 1989 - 2000

The table displays summary statistics of the variables used to estimate the one year default probabilities using the discrete-time hazard model. The statistics are shown separately for the solvent insurers and the insolvent insurer samples. All insurers are included in the analysis except insurers that have insufficient data or those that fail for which data is not available either one year or two years prior to the first regulatory action being taken against the firm. There are 217 firm-year observations in the insolvent sample and 24,236 in the solvent sample.

	Solvent Insurers		Insolvent Insurers		Test Statistic
FAST Ratios and other Control Variables	μ_{sol}	σ_{sol}	μ_{ins}	σ_{ins}	$H_0: \mu_{sol} = \mu_{ins}$
Kenney Ratio: Net Premiums Written to Equity Capital	1.13	0.85	1.87	1.12	9.654
Insurance Reserves to Equity Capital	1.03	0.94	1.64	1.24	7.229
1 Yr. Growth in Net Premiums Written (%)	11.85	41.58	11.18	60.92	0.160
1 Yr. Growth in Gross Premiums Written (%)	11.90	37.48	10.86	52.45	0.292
Aid to Equity Capital due to Reinsurance	2.05	4.34	6.04	7.50	7.838
Investment Yield (%)	5.71	1.38	5.42	1.54	2.778
1 Yr. Growth in Equity Capital (%)	8.81	16.33	-8.74	19.84	12.990
Adverse Reserve Development to Equity Capital (%)	-2.72	10.84	4.18	11.73	8.635
Gross Expenses to Gross Premiums Written	0.58	0.76	0.55	0.64	0.839
1 yr. Change in Gross Expenses (%)	0.05	0.46	0.08	0.57	0.877
1 yr. Change in Liquid Assets (%)	1.17	2.67	0.35	1.78	6.694
Investments in Affiliates to Equity Capital	0.58	1.32	0.95	1.75	3.105
Receivables from Affiliates to Equity Capital	0.02	0.04	0.04	0.05	5.237
Misc. Recoverables to Equity Capital	0.03	0.05	0.07	0.08	6.735
Non-investment Grade Bonds to Equity Capital	0.69	2.51	0.76	2.70	0.353
Other Invested Assets to Equity Capital	0.01	0.03	0.02	0.04	3.567
Dummy = 1 if insurer has a large single agent	0.12	0.33	0.22	0.42	3.414
Dummy = 1 if insurer has a large single agent they control	0.08	0.28	0.12	0.32	1.437
Losses, Exp's, Div's and Taxes Paid to Premiums Collected	1.29	0.73	1.60	0.84	5.364
Total Assets (000000's in 2000 $)	435.88	2218.95	134.63	695.43	6.109
Ind. = 1 if Insurer is Part of a Mutual Group	0.26	0.44	0.08	0.28	9.070

Table 4: Discrete-Time Hazard Bankruptcy Model Regression Results

Panel A:

Table displays the results of the discrete-time hazard regression model. The dependent variable $y_{it} = 1$ for each insurer that has a formal regulatory action taken against the insurer in year t+1. Otherwise $y_{it} = 0$ for all other observations. All U.S. property-liability insurers from 1989 - 2000 are included assuming they have adequate data. There are 24,236 healthy firm-year observations and 217 insolvent company observations.

Independent Variable	Coefficient Estimate	Standard Error	χ^2 Statistic	
Intercept	-0.2458	0.9069	0.0734	
Kenney Ratio: Net Premiums Written to Equity Capital	0.6619	0.0995	44.2155	***
Insurance Reserves to Equity Capital	0.2637	0.0910	8.4058	***
1 Yr. Growth in Net Premiums Written (%)	0.0019	0.0018	1.1025	
1 Yr. Growth in Gross Premiums Written (%)	0.0007	0.0022	0.0969	
Aid to Equity Capital due to Reinsurance	0.0453	0.0114	15.9287	***
Investment Yield (%)	-0.0538	0.0534	1.0180	
1 Yr. Growth in Equity Capital (%)	-0.0469	0.0056	69.7371	***
Adverse Reserve Development to Equity Capital (%)	0.0282	0.0069	16.9760	***
Gross Expenses to Gross Premiums Written	-0.1024	0.1112	0.8489	
1 yr. Change in Gross Expenses (%)	0.1509	0.1580	0.9120	
1 yr. Change in Liquid Assets (%)	-0.0821	0.0421	3.7955	*
Investments in Affiliates to Equity Capital	0.1527	0.0481	10.0745	***
Receivables from Affiliates to Equity Capital	3.3804	1.4911	5.1395	**
Misc. Recoverables to Equity Capital	1.3329	1.0189	1.7112	
Non-investment Grade Bonds to Equity Capital	0.0144	0.0285	0.2559	
Other Invested Assets to Equity Capital	4.6585	1.9498	5.7083	**
Dummy = 1 if insurer has a large single agent	0.4375	0.2001	4.7827	**
Dummy = 1 if insurer has a large single agent they control	-0.2045	0.2580	0.6285	
Losses, Exp's, Div's and Taxes Paid to Premiums Collected	0.5949	0.1028	33.4640	***
Ln(Total Assets in $2000)	-0.3856	0.0519	55.2763	***
Ind. = 1 if Insurer is Part of a Mutual Group	-1.1267	0.2653	18.0353	***
Log Likelihood Function Value	-876.46			
Pseudo R^2	26.21%			

*** - significant at the 1 percent level; ** - significant at the 5 percent level; * - significant at the 10 percent level

The pseudo R^2 equals 1 minus the ratio of the log likehood function value divided by the log likelihood function value where all coefficients are constrained to be zero (see Greene 1997 p. 891).

Panel B:

Table displays summary statistics of the predicted one-year probability of default for solvent firm-year observations and for bankrupt firm-year observations.

Firm Type	Num	Ave.	Median	Standard Deviation	1st Percentile	99th Percentile
Solvent	24,236	0.81%	0.20%	2.50%	0.01%	11.48%
Insolvent	217	9.22%	4.36%	12.55%	0.08%	68.69%

Panel C:

The classification table displays the number of actual insolvent and solvent firm-year observations versus the number of predicted insolvent and solvent firm-year observations using the estimated hazard model. Observations were predicted to be insolvent (solvent) that had one-year probabilities of default estimated to be greater than the population average over this time period (0.89 percent).

	Predicted Firm Type		
Firm Type	Insolvent	Solvent	Totals
Insolvent	177	40	217
Solvent	4,240	19,996	24,236

Table 5
Summary Statistics One-Year Probability of Default for U.S. Property-Liability Insurers Rated By A.M Best vs. Standard & Poor's: 1989 - 2000

Table 5 displays the average and median probability of default of the firms that receive ratings by A.M. Best and either a full or qualified rating from Standard & Poor's. The T-test column reports the results of the null hypothesis of equal means for the proability of default for A.M. Best rated insurers versus S&P rated insurers assuming unequal variances. The chart below displays the average probability of default time series for each agency over time period of this study.

	A.M. Best			Standard & Poor's Full Rating			Standard & Poor's Qualified Rating			Test Statistics H_0: $\mu_{Best} = \mu_{S\&P}$		
Year	Num	Mean	Median	Num	Mean	Median	Num	Mean	Median	Full		Qualified
1989	1110	0.39%	0.14%	3	0.18%	0.24%				-		
1990	1175	0.59%	0.26%	12	0.32%	0.11%				1.991	**	
1991	1261	0.50%	0.14%	25	0.13%	0.11%	610	0.48%	0.14%	6.500	***	0.210
1992	1352	0.63%	0.15%	26	0.22%	0.12%	708	0.59%	0.13%	4.585	***	0.224
1993	1437	0.47%	0.12%	30	0.11%	0.05%	751	0.59%	0.11%	5.668	***	0.199
1994	1515	0.46%	0.14%	29	0.21%	0.13%	495	0.46%	0.13%	4.335	***	0.110
1995	1551	0.42%	0.10%	36	0.20%	0.04%	723	0.37%	0.10%	2.272	**	0.073
1996	1577	0.87%	0.18%	281	0.35%	0.18%	775	0.73%	0.17%	5.638	***	0.177
1997	1598	0.51%	0.14%	314	0.32%	0.14%	762	0.59%	0.13%	3.248	***	0.068
1998	1620	0.68%	0.16%	354	0.43%	0.15%	751	0.94%	0.15%	2.773	***	0.091
1999	1620	0.79%	0.19%	297	1.21%	0.21%	706	0.87%	0.19%	2.128	**	0.151
2000	1570	0.68%	0.20%	258	0.94%	0.25%	682	0.80%	0.20%	1.360	*	0.092

***, **, or * - significant at the 1, 5, or 10 percent level, respectively

Figure T5
Average One-Year Probability of Default for U.S. Property-Liability Insurers Rated By A.M Best and Standard & Poor's 1989-2000

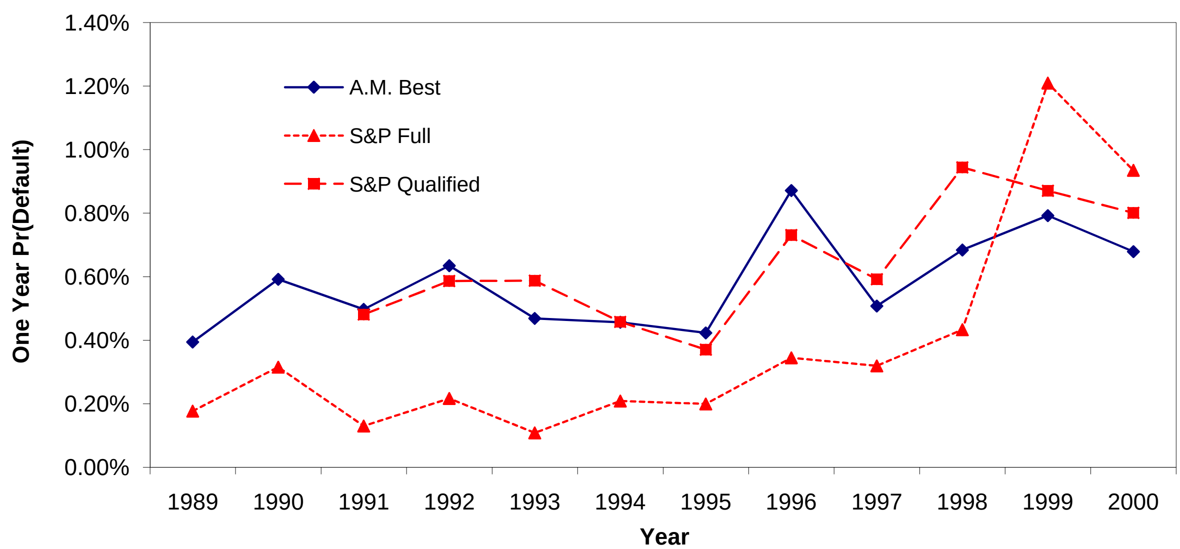


Table 6
A.M. Best vs. S&P Full vs. S&P Qualified Full Ratings: 1989 - 2000

Table displays summary statistics of the ratings assigned by A.M. Best and S&P to U.S. property-liability insurers over the years 1989 - 2000. The final two columns display t-statistics reporting the difference in means test assuming unequal variances

	A.M. Best					Standard & Poor's Full Ratings					Standard & Poor's Qualified Ratings					Difference in Means Test Statsitics S&P - A.M. Best			
Year	Num	μ_{full}	σ_{full}	Min	Max	Num	μ_{full}	σ_{full}	Min	Max	Num	μ_{qual}	σ_{qual}	Min	Max	Full		Qualified	
1989	1110	2.16	0.82	1	3	3	1.33	1.15	0	2						-0.83			
1990	1176	2.13	0.86	1	3	12	2.17	0.72	1	3						0.04			
1991	1266	2.04	0.87	1	3	25	2.20	0.58	1	3	610	0.25	0.43	0	1	0.16	*	-1.79	***
1992	1370	2.02	0.86	1	3	26	2.12	0.65	0	3	708	0.44	0.50	0	1	0.09		-1.59	***
1993	1444	2.00	0.85	1	3	30	1.73	0.91	0	3	751	0.43	0.49	0	1	-0.26	*	-1.57	***
1994	1517	1.92	0.86	1	3	29	1.55	0.83	0	2	495	0.49	0.50	0	1	-0.37	***	-1.43	***
1995	1563	1.89	0.84	1	3	36	2.00	0.63	0	3	723	0.89	0.82	0	2	0.11		-1.00	***
1996	1603	1.90	0.84	1	3	281	2.11	0.45	1	3	775	0.94	0.83	0	3	0.21	***	-0.96	***
1997	1617	1.92	0.81	1	3	314	2.08	0.45	1	3	762	0.90	0.76	0	3	0.15	***	-1.02	***
1998	1660	1.97	0.81	1	3	354	2.09	0.42	1	3	751	0.97	0.75	0	3	0.12	***	-1.00	***
1999	1647	1.95	0.84	1	3	297	1.84	0.54	0	3	706	1.01	0.74	0	3	-0.11	***	-0.94	***
2000	1582	1.93	0.85	1	3	258	1.82	0.54	0	3	682	0.96	0.73	0	3	-0.11	***	-0.97	***

***, **, or * - significant at the 1, 5, or 10 percent level, respectively

Figure T6
Average Rating for U.S. Property-Liability Insurers Rated By
A.M Best and Standard & Poor's 1989-2000

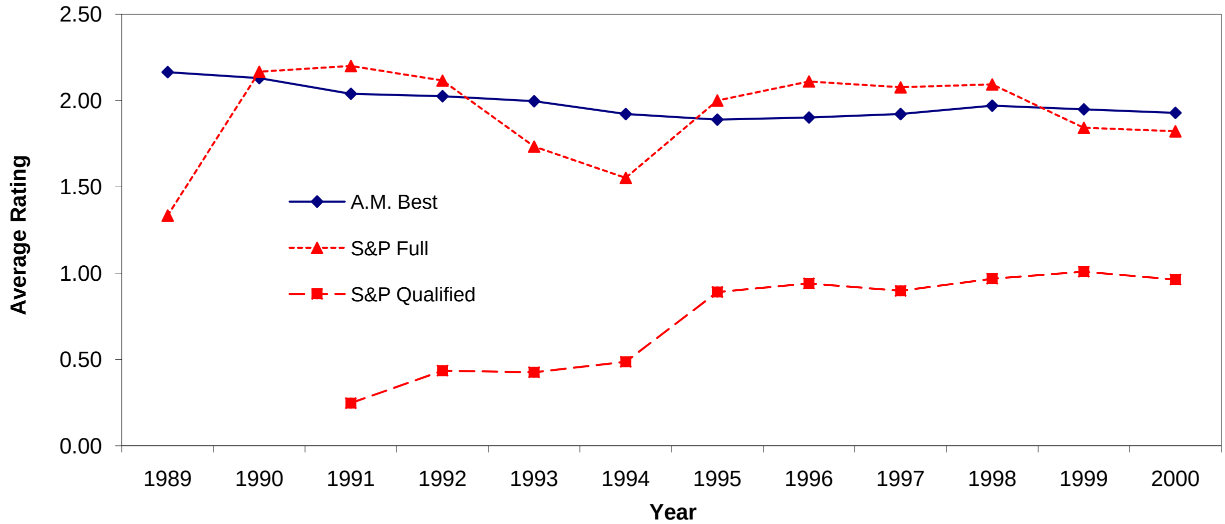


Table 7
Comparing A.M. Best and Standard & Poor's
Ratings for Common Insurers: 1989 - 2000

Table compares the ratings assigned by A.M Best and S&P for insurers that receive a rating from both firms. The data in Panel A includes all firm-year observations over the years 1989 - 2000 that received a rating from A.M Best and a full rating from S&P. The data in Panel B includes all firm-year observations over the years 1991 - 2000 that received a rating from A.M Best and a qualified rating from S&P. Each cell of the matrix, c_{ij}, equals the number of firm-year observations that receive rating i from A.M. Best and rating j from S&P where i,j ε {Superior, Excellent, Good, Marginal}.

Panel A

	S&P Full Rating			
A.M. Best Rating	Marginal	Good	Excellent	Superior
Marginal	23	5	6	0
Good	12	31	38	0
Excellent	1	63	706	15
Superior	0	0	567	162

Total number of firm-year observations: 1629

S&P gives the same rating as A.M. Best:	56.60%
S&P rates higher than A.M. Best:	3.93%
S&P rates lower than A.M. Best:	39.47%

Panel B

	S&P Qualified Rating			
A.M. Best	Marginal	Good	Excellent	Superior
Marginal	365	66	1	0
Good	605	295	15	0
Excellent	1044	1640	582	0
Superior	428	686	568	7

Total number of firm-year observations: 6302

S&P gives the same rating as A.M. Best:	19.82%
S&P rates higher than A.M. Best:	1.30%
S&P rates lower than A.M. Best:	78.88%

Table 8
Summary Statistics of Insurer's Receiving Full vs. Qualified Ratings from Standard & Poor's: 1989 - 2000

The sample includes insurer-year observations for all firms that receive an A.M. Best rating and any firm that receives either a full or a qualified rating from S&P over the years 1989 - 2000. The variable labeled "[Median Pr(Def. | A.M. Best) - Insurer Pr(Def.)]" equals the difference between the median probability of default for all insurers rated by A.M. Best conditional upon their A.M. Best rating category minus the specific insurer's estimated default probability.

	A.M. Best Only		Standard & Poor's Qualified Rating		Standard & Poor's Full Rating		Test Statistics		
	μ_a	σ_a	μ_q	σ_q	μ_f	σ_f	$H_0: \mu_a = \mu_q$	$H_0: \mu_a = \mu_f$	$H_0: \mu_q = \mu_f$
S&P Rating	-		0.817	0.736	1.989	0.506	-	-	72.6 ***
A.M. Best Rating	1.916	0.847	2.004	0.817	2.339	0.684	6.1 ***	21.0 ***	16.3 ***
S&P Rating - A.M. Best Rating	-		-1.187	0.865	-0.350	0.564	-	-	45.6 ***
Ind. = 1 for Marginal A.M. Best Rating	0.076	0.266	0.064	0.246	0.024	0.153	2.7 ***	10.5 ***	8.0 ***
Ind. = 1 for Good A.M. Best Rating	0.175	0.380	0.138	0.345	0.050	0.218	5.8 ***	17.5 ***	12.3 ***
Ind. = 1 for Excellent A.M. Best Rating	0.505	0.500	0.526	0.499	0.489	0.500	2.4 ***	1.2	2.6 ***
Ind. = 1 for Superior A.M. Best Rating	0.243	0.429	0.271	0.445	0.437	0.496	3.7 ***	14.1 ***	11.8 ***
[Median Pr(Def. | A.M. Best) - Insurer Pr(Def.)]	-0.003	0.011	-0.002	0.009	-0.004	0.011	4.8 ***	1.6 *	4.4 ***
Ind. = 1 if Insurer is Part of a Mutual Group	0.274	0.446	0.374	0.484	0.171	0.377	12.3 ***	9.4 ***	17.6 ***
Total Assets (000000's in 2000 $)	623.8	2833.5	334.4	1,502.3	1,860.4	5,963.9	7.6 ***	7.9 ***	9.8 ***
% Net Premiums Written in Retail Lines of Insurance	0.358	0.353	0.392	0.372	0.288	0.311	5.3 ***	7.8 ***	11.1 ***
Ind. = 1 if year = 1990	0.143	0.350	0.000	0.000	0.007	0.081	35.1 ***	29.8 ***	3.2 ***
Ind. = 1 if year = 1991	0.081	0.273	0.083	0.276	0.016	0.125	0.3	14.4 ***	13.9 ***
Ind. = 1 if year = 1992	0.086	0.281	0.096	0.295	0.015	0.123	2.0 **	15.6 ***	16.3 ***
Ind. = 1 if year = 1993	0.087	0.281	0.108	0.310	0.015	0.123	4.1 ***	15.6 ***	18.1 ***
Ind. = 1 if year = 1994	0.125	0.330	0.069	0.254	0.017	0.128	11.0 ***	21.3 ***	11.3 ***
Ind. = 1 if year = 1995	0.108	0.311	0.101	0.302	0.023	0.151	1.3 *	16.0 ***	14.2 ***
Ind. = 1 if year = 1996	0.073	0.260	0.112	0.316	0.174	0.380	7.7 ***	9.9 ***	5.9 ***
Ind. = 1 if year = 1997	0.067	0.251	0.108	0.311	0.185	0.388	8.2 ***	11.3 ***	7.1 ***
Ind. = 1 if year = 1998	0.068	0.252	0.109	0.311	0.212	0.409	8.1 ***	13.1 ***	9.1 ***
Ind. = 1 if year = 1999	0.078	0.269	0.110	0.313	0.174	0.379	6.1 ***	9.3 ***	6.0 ***
Ind. = 1 if year = 200C	0.083	0.276	0.103	0.304	0.162	0.369	4.0 ***	7.9 ***	5.8 ***

*** - significant at the 1 percent level; ** - significant at the 5 percent level; * - significant at the 10 percent level. The number of firm-year observations that received a rating from A.M. Best was 7398. The number of firm-year observations that received both an A.M. Best rating and a qualified/full rating by Standard &

Table 9
Probit Regression Results Predicting Whether Insurer Received a Full Rating from Standard & Poor's: 1989 - 2000

Table displays Probit regression results where the dependent variable equals 1 when insurer i was assigned a full rating by Standard & Poor's in year t and 0 otherwise. Panel A displays the estimated coefficients on the independent variables. Panel B displays the marginal effects. The model also contains year indicator variables but the results are suppressed to save space. The full results with the estimated coefficients for the year indicator variables are avialable upon request.

Panel A: Regression Results	**Did Insurer Request Full Rating from S&P?**		
Independent Variable	**Model 1**	**Model 2**	**Model 3**
Intercept	-1.020 ***	-1.741 ***	-4.648 ***
	(0.040)	(0.099)	(0.232)
[Median Pr(Def. | A.M. Best) - Insurer Pr(Def.)]	47.130 ***	47.998 ***	40.442 ***
	(9.542)	(10.681)	(11.023)
[Median Pr(Def. | A.M. Best) - Insurer Pr(Def.)] x (Year - 1988)	-4.832 ***	-5.581 ***	-4.568 ***
	(0.914)	(1.020)	(1.053)
Ind. = 1 for Good A.M. Best Rating		0.034	-0.044
		(0.104)	(0.114)
Ind. = 1 for Excellent A.M. Best Rating		0.681 ***	0.357 ***
		(0.094)	(0.103)
Ind. = 1 for Superior A.M. Best Rating		1.099 ***	0.613 ***
		(0.095)	(0.106)
Ln(Assets in $2000)			0.192 ***
			(0.011)
Ind. = 1 if Insurer is Part of a Mutual Group			-0.460 ***
			(0.043)
% Net Premiums Written in Retail Lines of Insurance			-0.086
			(0.052)
State of Business of Herfindahl			-0.450 ***
			(0.053)
Log-likehood function value	-4186.7	-3957.6	-3549.9
Pseudo R^2	13.99%	18.69%	27.07%

Panel B: Estimated Marginal Effects			
[Median Pr(Def. | A.M. Best) - Insurer Pr(Def.)]	5.933 ***	5.325 ***	3.339 ***
[Median Pr(Def. | A.M. Best) - Insurer Pr(Def.)] x (Year - 1988)	-0.608 ***	-0.619 ***	-0.377 ***
Ind. = 1 for Good A.M. Best Rating		0.004	-0.004
Ind. = 1 for Excellent A.M. Best Rating		0.076 ***	0.029 ***
Ind. = 1 for Superior A.M. Best Rating		0.122 ***	0.051 ***
Ln(Assets in $2000)			0.016 ***
Ind. = 1 if Insurer is Part of a Mutual Group			-0.038 ***
% Net Premiums Written in Retail Lines of Insurance			-0.007
State of Business of Herfindahl			-0.037 ***

*** - significant at the 1 percent level; ** - significant at the 5 percent level; * - significant at the 10 percent level
The pseudo R^2 equals 1 minus the ratio of the log likehood function value divided by the log likelihood function value where all coefficients are constrained to be zero (see Greene 1997 p. 891).
Standard errors are shown in parentheses.

Table 10
Regression Results Explaining Rating Difference
Between Standard & Poor's and A.M. Best Ratings: 1989 - 2000

Table displays OLS regression results where the dependent variable is the difference between the full rating assigned by Standard & Poor's minus the rating assigned by A.M. Best in year t. The inverse Mill's terms in each model were calculated using the results of each Probit regression shown in Table 10, respectively.

	Did Insurer Request Full Rating from S&P?		
Independent Variable	**Model 1**	**Model 2**	**Model 3**
Intercept	-0.5826 ***	-1.2307 ***	-0.7177 ***
	(0.075)	(0.054)	(0.042)
Inverse Mills Ratio	0.1527 ***	0.6118 ***	0.2838 ***
	(0.048)	(0.035)	(0.030)
R^2	0.66%	19.93%	5.73%
Expected change in rating due to insurer strategic choice	0.233	0.881	0.368
Expected change in ratings due difference in S&P vs. A.M. Best standards	-0.583	-1.231	-0.718
Average Rating Difference S&P Rating - A.M. Best Rating	-0.350	-0.350	-0.350

The Rodney L. White Center for Financial Research
The Wharton School
University of Pennsylvania
3254 Steinberg Hall-Dietrich Hall
3620 Locust Walk
Philadelphia, PA 19104-6367

(215) 898-7616
(215) 573-8084 Fax
http://finance.wharton.upenn.edu/~rlwctr

The Rodney L. White Center for Financial Research is one of the oldest financial research centers in the country. It was founded in 1969 through a grant from Oppenheimer & Company in honor of its late partner, Rodney L. White. The Center receives support from its endowment and from annual contributions from its Members.

The Center sponsors a wide range of financial research. It publishes a working paper series and a reprint series. It holds an annual seminar, which for the last several years has focused on household financial decision making.

The Members of the Center gain the opportunity to participate in innovative research to break new ground in the field of finance. Through their membership, they also gain access to the Wharton School's faculty and enjoy other special benefits.