Competition among rating agencies and information disclosure*

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

September 22, 2008

Abstract

The paper analyzes why a rating agency pools different credit risks in one credit grade, and how information disclosure depends on the value of information to the market. We build a model to analyze the optimal disclosure policy of a monopoly rating agency depending on the value of information to investors, and then describe the potential market and the strategy of the entrant. We find that entry of a symmetric rating agencies 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 empirically test the qualitative predictions of the model. Standard and Poor's entry to the insurance market that was previously covered by a monopoly agency, A.M. Best, is used as a natural experiment to study the impact of competition on the information content of ratings.

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

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

*We are gratefully to Gregory Nini, Anjolein Schmeits, and the seminar participants at University of 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

Following recent failures of rating agencies to provide timely accurate information (Enron, Worldcom, Global Crossing, AB&B, AT&T, to name the big ones), the industry has been criticized for the significant concentration of power in a small number of rating agencies. In response to these criticisms, the U.S. House of Representatives passed the "Credit Rating Agency Reform Act of 2006" that aimed to substantially simplify the process of obtaining a NRSRO (Nationally Recognized Statistical Rating Organization) status and, ultimately, promote competition. From June 2007 SEC has adopted rules to implement provisions of the Act. In particular, rating agencies who have issued ratings for 3 years and satisfy certain requirements[1] have the option of registering as NRSROs. This policy has substantially diminished the regulatory barriers to entry. However, till now there is no agreement about the impact of competition on the quality of information provided by credit raters.

The current paper addresses information quality, specifically incentives for information disclosure. Credit quality assessment implies that a rating agency produces an estimate of probability of default[2]. However, with the exception of KMV^{TM}, major rating agencies do not disclose a numerical estimate of credit quality[3]. Instead, they assign letter grades, and companies or issuers with similar credit risk characteristics have the same grade. Pooling reduces the amount of information disclosed to the market.

The first objective of our paper is to analyze the incentives of a monopoly rating agency to pool information. We find that the value of information to the ultimate users (investors, policyholders, etc.) is crucial in determining the agency's disclosure policy. The second objective is to characterize the optimal entry strategy for a new rating agency. We find that a new rating agency enters by targeting the companies of the highest credit quality in each rating category. We empirically test the qualitative predictions of the model on the entry strategy using the data on ratings of creditworthiness of insurance companies.

The model has three groups of players - sellers (debt issuers, corporations), rating agencies, and buyers (investors, consumers, policyholders). We assume that purchasing a rating is voluntary to a seller, and a rating agency cannot charge a fee contingent on the rating assigned. The optimal rating system of the agency is designed to trade-off the ability of high quality sellers to signal their quality by purchasing a rating and the benefits for the low quality companies to be pooled with better companies. This trade-off determines the rating agency's disclosure policy

[1] These requirements are listed in Section 15E of the "Credit Rating Agency Reform Act of 2006". An application must include information on credit rating performance measurement statistics, procedures and methodologies used to assign ratings, organizational structure, the list of the 20 largest issuers and subscribers that use services of the applicant (on confidential basis). An applicant must also submit a written certification from at least 10 qualified institutional buyers.

[2] In general the output of the credit risk model is multidimensional. Other important outputs are expected loss given default, rating stability, etc.

[3] KMV LLC offered Merton structured default probabilities in mid-1990s. It was acquired by Moody's in 2002.

to pool companies into rating categories and the optimal coverage of the market. The optimal rating scale derived from the model resembles the interval disclosure rule employed by the major credit agencies. We study how the market coverage and information precision depend on the value of information to buyers.

In the model, a rating agency decides how credit risk information is disclosed to the market. Lizzeri (1999) establishes a striking result; a monopoly rating agency's optimal rating scale is to pool all companies in one rating. Surprisingly, all companies pay to be rated in spite the fact that de facto the agency discloses no information. The crucial assumption that leads to no disclosure result is that all parties are risk neutral. It implies that no party values the precision of information contained in rating. We relax this assumption and show that when precision of information has value to end users of ratings (investors, buyers of a product, etc.), the no disclosure result no longer holds. Also we study how the marginal value for precision of information affects the optimal rating scale. As the value of information increases ratings become more precise.

We use the basic model to analyze the entry strategy of a new rating agency to the market that is already served by the incumbent. We show that a new agency enters by targeting high quality companies in each rating category. Interestingly, the entrant competes by differentiating its rating scale such that companies may be rated differently by the incumbent and entrant. Also the number of ratings each firm obtains depends on the credit quality of a rated company. However, there is no congruency between the number of ratings and the quality of the company. High and low companies can be rated by both agencies while the intermediate risk company obtained only one rating.

The theoretical model yields two major predictions regarding the entry strategy of the new agency. First, the entrant targets the highest quality companies in each rating category. Second, the entrant employs more stringent standards relative to the rating scale of the incumbent. We test these predictions using data on the U.S. property-liability insurance market. The insurance industry provides an ideal natural experiment to test our hypotheses for two reasons. 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. For example, in 1992, S&P issued full rating opinions on only 23 property-casualty insurers and this number increased to over 340 insurers by the end of the decade. By year 2004, S&P was the second largest insurance rating agency and now rates almost 800 companies representing more than 90 percent of the industry's assets.

We examine the strategies employed as the new entrant came into the insurance ratings business. We employ two methodologies. 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 and then use the results to compare the standards that were necessary for a firm to receive similar ratings from both the incumbent and the entrant agency.

Although the results of the hazard model are interesting and useful, the analysis suffers from two limitations. First, the hazard model relies completely on publicly available information to determine the one-year probabilities of default. Therefore, we are unable to control for any private information that might be learned through the ratings process itself. Second, comparing the probabilities of default for insurers that receive a rating from both Best's and from S&P ignores the possibility that firms will decide strategically whether to request a second rating from the new entrant. This second problem potentially biases our results. The second empirical test is designed to control for both limitations of the hazard model and we investigate differences in rating opinions across the incumbent and the entrant using a Heckman-style sample selection methodology.

By way of preview, we find that higher than average quality insurers in each rating category were attracted to receive a second rating from S&P and that S&P required higher standards in order for an insurer to achieve a similar rating. Both results are consistent with our theory. The remainder of this section reviews the prior literature. Section 2 presents the model. Section 3 examines the strategy for the monopoly. Section 4 analyses which market segments are profitable for an entrant. We present our empirical analysis in Section 5 and the conclusion follows.

Related Literature

The paper belongs to the growing literature on 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 (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. This disclosure policy is equivalent to assigning a unique rating to all types of debt issuers. The intuition goes as follows. Since the benefit of a rating is higher for better types of sellers, the coverage of the market is determined by the lowest rated type. To increase the willingness of this type to pay, the intermediary pools it will all better types. 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 (2003) 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 poor types.

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 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 its private observation in favor of prior belief. Mariano (2006) addresses a similar issue in the context of rating agencies.

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. Lizzeri (1999) shows that competition leads to full disclosure and zero fees for certification. In this paper we study the impact of new entry into a previously monopolistic market for ratings.

2 The model

There are three groups of agents: sellers and buyers of a credit sensitive product, and rating agencies. Sellers have private information about their quality v. Higher v corresponds to higher quality. Rating agencies and buyers share a common prior about the quality of a seller. For simplicity, we assume that v is distributed uniformly on $[0, 1]$.

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 prior on quality and the accuracy of information about quality. To model demand for accuracy 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 $E[I]$ is the expected quality, and $Var[I]$ is the variance of quality. $a > 0$ measures the marginal value of information accuracy to buyers. Buyers are price takers, and $u(I)$ is the price they pay 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).

We assume that a seller cannot credibly communicate its financial strength to buyers. A rating agency offers an evaluation service for a fee and can perfectly observe the type v of a seller. We assume that the fee is the same for all sellers purchasing a rating, and the rating agency cannot screen companies by demanding higher fee for more favorable rating. A rated company does value an option to withhold its rating[4].

The disclosure policy of the agency defines how the estimates of quality are communicated to buyers. One particular case is full disclosure, where the rating agency 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. The optimal disclosure policy in our model (Proposition 4) 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.

We assume that 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 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 agency.

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.

[4] 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.

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 Monopoly Rating Agency

3.1 Preliminary results: Full disclosure

This section describes the demand for ratings and the profits of the rating agency under full disclosure. Though this system is not optimal for the rating agency, the analysis can be useful to highlight the agency's gains from pooling different risk types in the same rating grade.

Proposition 1 *Suppose that a monopoly rating agency commits to full disclosure, and the fee for the rating services is such that* $t < \frac{1}{2} + \frac{1}{12}a$. *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.*

Proof. Let's consider some type $v \in [0,1]$ and assume that all types above v and no types below v are rated. Under full disclosure a rated seller of type v is paid by buyers

$$u_R(v) \equiv v.$$

If seller type v is not rated, it is pooled with types $[0, v]$. The reservation price of non-rated sellers is then equal to

$$u_N(v) \equiv \frac{1}{2}v - \frac{1}{12}av^2,$$

where $\frac{1}{2}v$ is the expected quality and $\frac{1}{12}v^2$ is the variance of quality of non-rated types $[0, v]$. If this price is negative, the non-rated sellers do not trade.

A necessary condition that seller type v purchases a rating is that it increases its reservation price net of the rating fee,

$$u_R(v) - t \geq \max\{u_N(v), 0\}. \tag{1}$$

Note that as v increases, the difference between the two prices increases,

$$\frac{d}{dv}(u_R(v) - u_N(v)) = \frac{1}{2} + \frac{1}{6}av > 0, \tag{2}$$

and

$$u_R(v) - u_N(v) = \begin{cases} 0 \text{ if } v = 0, \\ \frac{1}{2} + \frac{1}{12}a > 0 \text{ if } v = 1. \end{cases}$$

If the fee t is below $\frac{1}{2} + \frac{1}{12}a$, then there is a seller type $v_F \in (0,1)$ for which the participation constraint (1) is binding. (2) implies that all types above v_F strictly prefer to obtain a rating, and no type below v_F obtains a rating. The buyers' beliefs that the company's quality is above

v_F if it is rated and is below v_F if it is not rated, are consistent with the equilibrium. The proof of the uniqueness is shown in the Appendix. ■

Under full disclosure, the fee charged by the rating agency is $t = u_R(v_F) - \max\{u_N(v_F), 0\}$. It is equal to the amount that the lowest rated seller is willing to pay for rating. The demand for ratings is $\delta_F = 1 - v_F$. So the profit of the rating agency is

$$\max_{v_F}(1 - v_F)(u_R(v_F) - \max\{u_N(v_F), 0\}).$$

Since increasing the fee reduces the demand for ratings, the optimal fee for the rating agency and the resulting coverage of the market, $\delta_F = 1 - v_F$, 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 2 *Under full disclosure, the optimal market coverage,*

$$\delta_F = \begin{cases} \frac{2(a+3)-\sqrt{a^2+6a+36}}{3a}, \ a \leq 10, \\ \frac{a-6}{a}, \ 10 \leq a \leq 12, \\ \frac{1}{2}, \ a \geq 12, \end{cases}$$

is decreasing in the value of information for $a < 10$ and increasing for $10 < a < 12$, and is independent of a for $a > 12$. The profit of the rating agency is

$$\pi_F = \begin{cases} \frac{(a+12)(a+3)(a-6)+(6a+a^2+36)^{\frac{3}{2}}}{162a^2}, \ a \leq 10, \\ \frac{6(a-6)}{a^2}, \ 10 \leq a \leq 12, \\ \frac{1}{4}, \ a \geq 12. \end{cases}$$

Profit is increasing in the marginal value of information a.

Different regimes for the optimal coverage are driven by the ability of non-rated sellers to trade. When the marginal value of information is relatively low, obtaining a rating is not essential for trade. However, rated sellers can charge v instead of selling at the average valuation of non-rated types, u_N. The fee charged by the rating agency is equal to the gain of a rating to the lowest rated type, and it increases as the gap between v and u_N becomes higher. Due to this effect the rating agency benefits from reducing the market coverage.

As the marginal value of information increases, it becomes impossible to trade without a rating. For intermediate values, $10 < a < 12$, non-rated companies have zero valuation, and the effect on the market coverage is reversed. Market coverage increases as information becomes more valuable. For high values of information, $a > 12$, non-rated sellers cannot trade, $u_N < 0$. The fee charged by the rating agency is the valuation of the lowest rated type, v_F. The optimal market coverage does not depend on the value of information, and it is derived from a trade off between the gain of charging higher fee to better types and the cost of reduced market coverage.

Is full disclosure optimal? 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. When the valuation of pooled types is higher than the valuation of the lowest rated type, then

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

If the marginal value of information is zero, $a = 0$, one obtains Lizzeri's (1999) result that all types should be pooled and assigned the same rating grade. When precision of information matters, $a > 0$, pooling imposes a cost in lost accuracy. 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 informativeness of ratings.

3.2 Optimal disclosure of a monopoly rating agency

In this section, we 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 the seller with any other sellers and disclose that it belongs to a particular group. Formally, a disclosure policy is a correspondence $s : [0, 1] \to [0, 1]$. 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 result 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,

$$u(s(v)) - t \geq \max\{u(V_N(s)), 0\} \text{ for all } v \in V_R(s). \tag{3}$$

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

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

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 (3) 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 3 *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. There are multiple optimal disclosure policies for types $v > v_M + b_M$. However, the valuation $\mu(s(v)) - a\sigma^2(s(v))$ of types $[v_M + b_M, 1]$ under these policies is at least $\mu([v_M, v_M + b_M]) - a\sigma^2([v_M, v_M + b_M])$. The fee charged for the rating equals to the value of the rating to the lowest rated types, $t = u([v_M, v_M + b_M]) - \max\{u([0, v_M]), 0\}$. As the marginal value of information increases, $a \to +\infty$, an optimal disclosure policy converges to full disclosure.*

Pooling the lowest rated type with higher neighboring types increases its expected quality and costs the least in terms of lost precision. For higher types there are multiple disclosure policies. The reason is that the profit of the rating agency does not depend on rating schedule of this group as long as ratings have value to sellers. Pooling neighboring types is one example of rating scale for higher types.

In the remaining of the section we derive an optimal disclosure of a monopoly rating agency. The payoff of the lowest rated types $v \in L = [v_M, v_M + b_M]$ is

$$u(L) = v_M + \frac{1}{2}b_M - \frac{1}{12}ab_M^2, \tag{4}$$

The payoff of non-rated sellers $N = [0, v_M]$ is

$$u(N) = \max(\frac{1}{2}v_M - \frac{1}{12}av_M^2, 0).$$

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

$$(1 - v_M)t_M,$$

where t_M is the uniform fee charged for the rating. These sellers purchase a rating only if

$$u(s(v)) - t_M \geq \max\{u(N), 0\} \text{ for } v \in [v_M, 1]. \tag{5}$$

Proposition 3 implies that the fee charged by the rating agency is determined by the willingness of pooled types $v \in [v_M, v_M + b_M]$ to pay for the rating, and (5) reduces to

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

If a seller cannot trade without a rating, $u(N) < 0$, the fee is equal to the valuation for pooled types $v \in [v_M, v_M + b_M]$. When $u_N > 0$, the fee equals to the difference between the valuations of pooled companies $[v_M, v_M + b_M]$ and non-rated companies $[0, v_M]$.

In equilibrium, non-rated sellers $v \in [0, v_M]$ must be better off without a rating. If a seller $v \in [0, v_M]$ deviates and purchases a rating, the rating agency announces that seller's quality is from the interval $[0, v_M]$. 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 4 *Under assumptions (i)-(iii) the optimal monopoly rating system is summarized in the following table.*

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

When the value of information is relatively low, $a \leq 2$, 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 is decreasing. The coverage of the market is decreasing in a when the non-rated sellers can trade, $u(N) > 0$, 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 $\frac{1}{4}$.

The marginal value of information affects the design of the optimal rating system. 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. Recall that the valuation by buyers is composed of two components, the expected quality of the seller and the precision of information about quality. Since the value information precision is low, rating agency can increase the expected quality of the lowest rated seller by pooling it with the highest types.

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. 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 companies, and allows the agency 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. As a result, the price that can be charged when a subset $[v_M, v_M + b_M]$, $v_M + b_M < 1$, is pooled in one rating grade is higher than the price that can be charged when all rated sellers $[v_M, 1]$ are pooled in one rating, even though the later has higher expected quality.

The distinction between the last three regions 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 $a \geq \frac{51}{4}$.

How the rating agency decides the precision of $u(L)$? From (4), 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 impact on precision is dominated by the increase in expected quality from pooling, and that leads to extensive pooling. Note that the cost of precision is zero when $a = 0$. 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 one rating goes to zero and rating L effectively means that the type is v_M when $a \to +\infty$.

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 same 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.

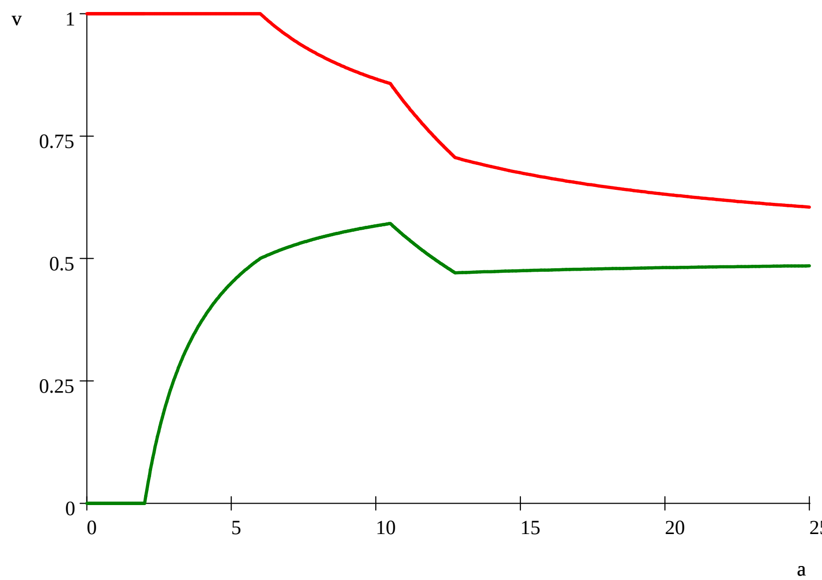

Figure 1: Optimal rating scale of a monopoly rating agency

In the next Corollary we compare the optimal rating scale to the full disclosure rating scale.

Corollary 1 *An optimal rating system derived in Proposition 4 provides higher profit to the rating agency than full disclosure rating system.*

Unfortunately, Proposition 3 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. Full disclosure and interval disclosure, where the neighboring types are pooled in the same rating, are two examples, but there are many others. In the next proposition we derive a necessary condition for an optimal interval disclosure policy, and show how the size of a rating interval changes with the marginal value of information.

Proposition 5 *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 6 *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 focus on the interval disclosure equilibrium where a rating agency pools higher types $H = [v_M + b_M, 1]$ in one rating[5]. Our motivation is twofold. First, it is an equilibrium that is consistent with industry practice. Second, 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[6].

4 Entry Strategy of a New Rating Agency

In this section, 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 on the eyes of the buyer. This occurs either when the second rating allows the seller to signal higher quality, or when it improves the accuracy of information of the buyers. 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 7 *An entrant can always design a rating system that attracts the best companies within each rating interval of the incumbent. The rating standard of the entrant is more stringent than that of the incumbent.*

[5] 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.

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

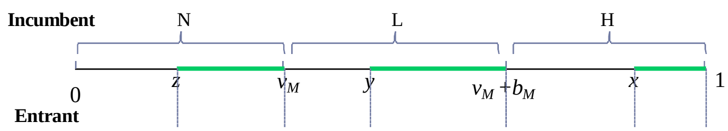

Figure 1: Demand for entrant's ratings

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.

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 at most three groups - sellers with high rating, sellers with low rating, and sellers with no rating by the incumbent. The top 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.

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

$$\max_{(x,y,z,t_e)} ((1-x) + (v_M + b_M - y) + (v_M - z))t_e \tag{6}$$
$$u(R_m, R_e; v) - t_m - t_e \geq u(R; v) - t_m, \ v \in V_R,$$
$$u(R_e; v) - t_m \geq \max\{u(V_N), 0\}, \ v \in V_N.$$

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 is that the potential market of the entrant 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. We devote the technical details to the appendix.

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.

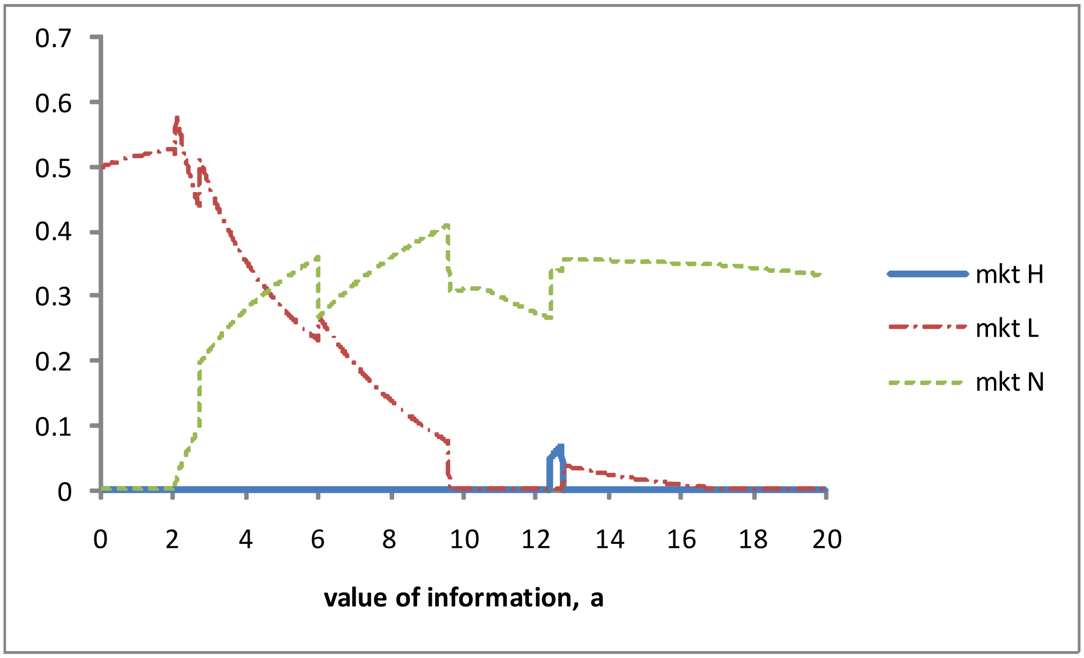

Figure 2: Entrant's coverage of segments H, L and N

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 becasue increasing the gap between the payoff of rated 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.

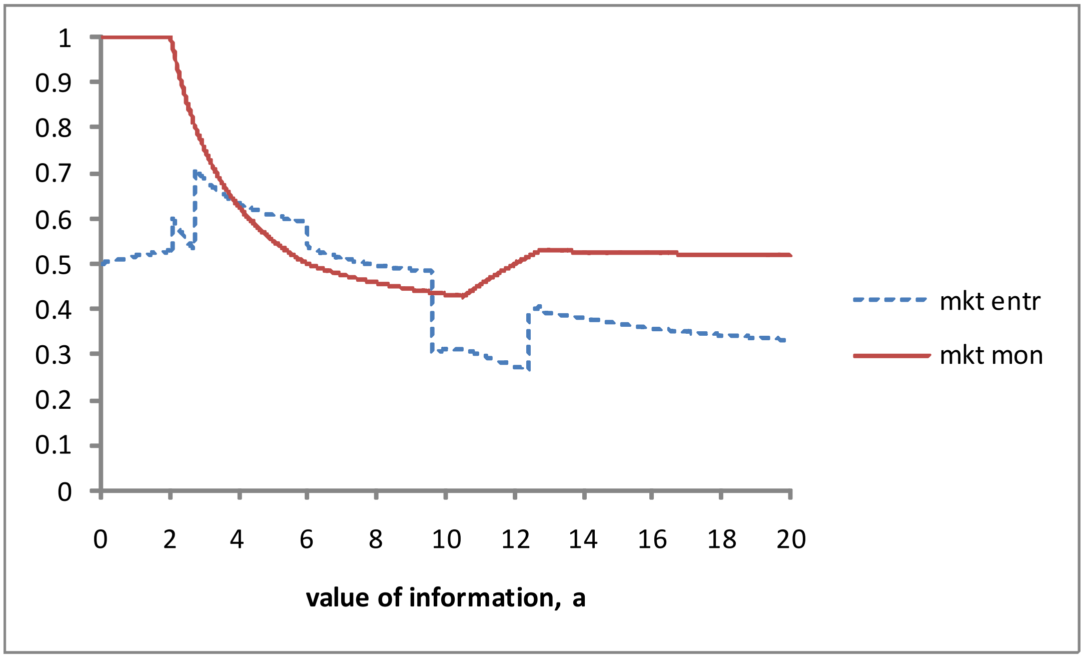

Figure 3: Market coverage by the incumbent and the entrant

The idea that the entrant is better off targeting the non-rated segment of the market echoes the result 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 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.

In the next proposition we summarize the properties of the entrant's optimal strategy.

Proposition 8 *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 general structure of the entry strategy derived in Proposition 7 does not depend on the probability distribution of types. If a pool of sellers obtain the same rating, sellers with types above average in the pool benefit from obtaining a second rating. Results summarized in Proposition 8 are a special case of uniform distribution. The decision to enter in each rating

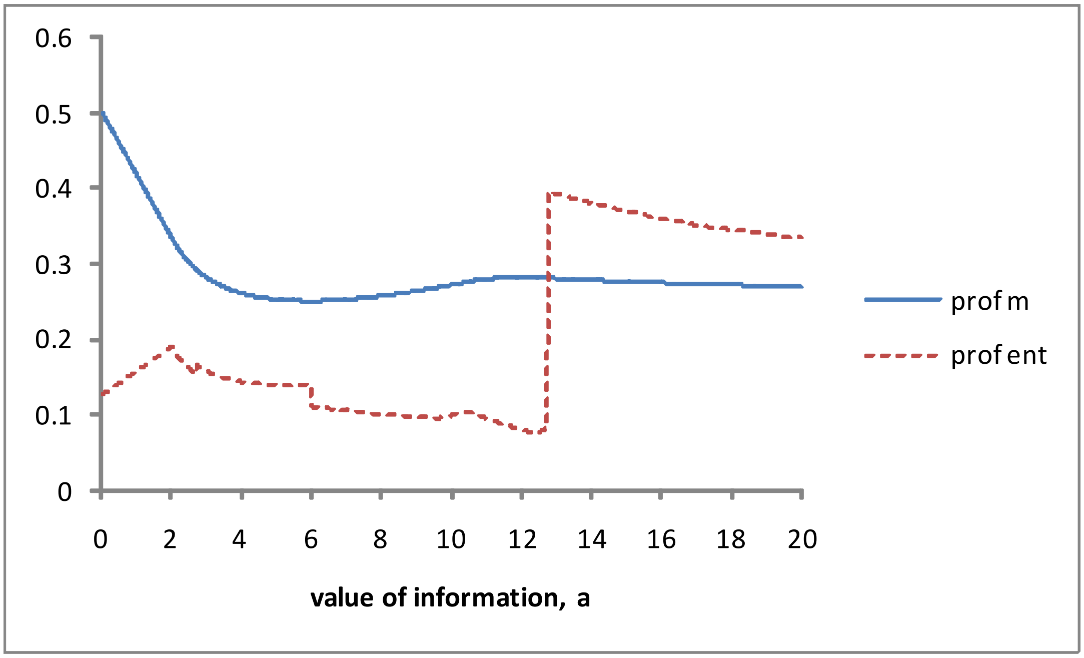

Figure 4: Profits of the incumbent and the entrant

interval is determined by the profitability of offering a rating for this pool of sellers relative to others, and ultimately, the probability distribution.

In the next section we empirically test the predictions about the optimal entry strategy.

5 Empirical Analysis

Our theoretical model yields several empirical predictions regarding the optimal strategies of a profit-maximizing entrant. We seek to test these predictions taking advantage of data on the U.S. property-liability insurance industry during the years 1992 - 2000. The insurance industry during the time period of the late 1980's and through the decade of the 1990's is uniquely suited to test our hypotheses as Standard & Poor's, the well-respected bond rating agency, invested significant resources to expand their influence and entered the market for insurance ratings. Prior to this time period, the market for insurance ratings was largely dominated by the A.M. Best Company. Incorporated in 1899, A.M. Best has published ratings on virtually all U.S. insurers and, for a majority of their history, they were the only agency doing so. The monopoly position Best's enjoyed, however, began to erode after Best's was criticized following the liability insurance crisis of the mid 1980's and after several natural catastrophes in the early 1990's that bankrupted numerous insurers. The most aggressive agency to enter the market was Standard & Poor's (S&P) who began publishing ratings on property-liability insurers in 1983 and then expanded coverage once in 1987 and then again in 1991 (Standard & Poor's 1987; A.M. Best 1992). Today, S&P provides ratings on insurers that represent in excess of 80 percent of the assets of the industry - more than any other new entrant except Weiss Research.[7]

5.1 Hypothesis Development

We test four distinct hypotheses that are derived either directly from the theoretical model presented in this paper or from the prior literature. The first hypothesis comes from Proposition 7 where we predict

Hypothesis 1 *The new entrant agency will find the greatest demand for its services from the high quality insurers seeking to differentiate themselves from other insurers that have a rating from the incumbent similar to their own.*

More specifically, we predict that higher-than -average quality insurers within a rating class who are bundled together with lower-than-average insurers in that same class will self-select and seek to differentiate themselves by obtaining a new rating.[8]

In a related hypothesis, taken from Propositions 4 and 7, we predict

Hypothesis 2 *The new entrant rating agency will require higher standards, on average, in order for a firm to receive a rating similar to the one they received from the incumbent agency.*

Thus firms that seek a rating from the new agency may not receive higher ratings but instead, for each rating class, the newly rated insurers should have higher average financial quality.

The next hypothesis is related both to the amount of information that is available to market participants regarding the financial quality of the insurer and how valuable that information is to market participants. First, consistent with Ramakrishnan and Thakor (1984) and Millon and Thakor (1985), we expect

Hypothesis 3 *More opaque insurers or 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.*

5.2 Methodology

We conduct several tests using two different econometric methodologies to investigate the hypotheses stated above. In the 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). More

[7] 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.

[8] Our's is not the first to investigate this hypothesis as Cantor and Packer (1997) empirically investigate a similar hypothesis using data on firms that obtain ratings from agencies other than the dominant bond rating firms Moody's and S&P. Unlike Cantor and Packer, however, we find strong evidence of self-selection using data from the insurance market.

stringent ratings standards are said to exist when the average/median probability of default for insurers in a particular rating class is lower for one agency than the other. The tests are designed to analyze the following questions. How do the ratings assigned by S&P compare to the ratings of A.M. Best for firms that are jointly rated by both companies? What is the average financial quality of the insurers in each rating category across the two agencies? Answering these questions requires us to develop a summary statistic of the financial quality of the insurers and then to use that statistic to compare the ratings systems of the two agencies. The benchmark we use is the one-year probability of default for each firm in our data set. We argue the one-year probability of default is a reasonable benchmark since both agencies state the primary objective of their rating systems is to provide an opinion about the insurer's ability to meet its contractual obligations to policyholders. We use these probabilities to examine the stringency of the rating system by comparing either the median or mean probability of default for a given rating class.

A shortcoming of analyzing the one-year default probabilities using the hazard model is that the only data available for us to estimate the model is publicly available information. By definition, rating agencies should only exist if their preferential access to private information allows them to better differentiate the quality of insurers. Thus, our second empirical test investigates the determinants of differences in the ratings that are assigned by the incumbent versus the new entrant rating agency while also controlling for the private information the agencies learn through the rating process. The methodology we employ largely follows the work of Cantor and Packer (1997) as we use a Heckman-style sample selection model to investigate differences in the ratings when an insurer chooses to be rated by both agencies. This second methodology allows us to control for private information as we include variables to proxy for the rating assigned the A.M. Best agency in the first stage probit regression. In addition, we add to the literature since, unlike Cantor and Packer , the theory developed in this paper provides strong guidance for the control variables that we should use to explain differences in rating assigned by the two agencies.

5.3 Estimating Default Probabilities

A variety of methods can be used to forecast the likelihood of bankruptcy for an insurance company. U.S. regulatory authorities use three univariate models to forecast bankruptcy. The Insurance Regulatory Information System (IRIS), the oldest system, utilizes a series of twelve audit ratios based upon financial statement data filed with the regulators. The newer Financial Analysis and Surveillance Tracking (FAST) system uses an expanded set of audit rations, approximately thirty, where each ratio is given a corresponding score. Regulators multiply each individual ratio by its corresponding score and then sum over all ratios to produce a FAST score. Insurers with higher FAST scores are more likely to become financially distressed and are subject to greater regulatory scrutiny. Finally the risk-based capital system defines a minimum amount of capital insurers must hold. The individual capital charges depend on the riskiness of the assets and the businesses in which the insurer participates.

In addition to the univariate regulatory models discussed above, economists have developed and implemented a variety of solvency prediction models based upon multivariate statistical techniques. Insolvency forecasting models based upon multiple discriminant analysis (Trieschmann and Pinches (1973)), or logistic regression (Cummins, Grace and Klein (1999)) are common in the literature. In addition, bankruptcy prediction models based upon neural networks (Brockett et al., 1994) and dynamic cash flow simulation models (Cummins, Grace and Phillips 1999) have also been discussed.

The limitation of these methods is that they are based on static models implemented using data that spans only one or just a few years. As a result, these static models are inadequate for the long-term panel data that we assembled for the study[9].

In this study we use the discrete-time hazard model suggested by Shumway (2001) to overcome the biases of the static models and to take advantage of our panel data. The hazard model approach has at least two primary advantages over the more traditional static models. First, hazard models allow for time-varying covariates that explicitly recognize that the financial health of some firms will deteriorate over time even though the firm may not declare bankruptcy for many years. Static models only make comparisons between firms that are classified as healthy or not healthy at just one point in time and they therefore ignore firms that are at risk of bankruptcy even though they have not yet become bankrupt. Shumway (2001) shows that ignoring this information creates a selection bias which leads to inconsistent parameter estimates. Intuitively, hazard models correct this problem by allowing to extract useful information from the times series data on each individual firm. In addition, it can be shown that the parameter estimates from hazard models are unbiased and consistent.

The second reason the hazard model is preferred to static models is because doing so allows us exploit all available information about the condition of the firm rather than just the last year's observations. Thus, the increased amount of data increases the efficiency of the model which yields more reliable parameter estimates and better out-of-sample forecasting results.

Implementing the discrete-time hazard model is rather straightforward. 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 hazard 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

[9] In addition, Theodossiou (1993) suggests that arbitrarily choosing when to observe each firm's characteristics leads to unnecessary selection bias problems and reduced forecasting ability.

data.

5.3.1 Data

The data to estimate the hazard model 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) over the years 1989-2000. Consistent with the literature, we define the year of insolvency as the year that the first formal regulatory action is taken against a troubled insurer. We identify the year of first regulatory against insurers through a variety of sources including the NAIC's *Report on Receiverships* (various years) and the *Status of Single-State and Multi-State Insolvencies* (various years). We also obtain 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 fail between 1990 and 2001.

The explanatory variables we use to estimate the hazard model are the nineteen balance sheet and income statement ratios that make up the NAIC's FAST solvency tracking 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. Thus, the FAST system seems to capture as much predictive power as can be gleaned from financial statement ratios alone. We also include controls for firm size equal to the natural logarithm of the real assets of the firm where the price deflator we use is the Consumer Price Index; and an organization form control variable which is an indicator set equal to 1 if the insurer belongs to a mutual or reciprocal group of insurers.

As discussed above, we estimate the hazard model using all insurers for which we have data to calculate the FAST ratios. Thus, not only do we include insurers rated by A.M. Best and/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 were unable to locate data within 2 years of their first event year. The final data set contains 24,062 solvent firm-year observations and 214 insolvent firm-year observations.

5.3.2 Empirical Results

Summary statistics for the solvent and insolvent company observations are shown in Table 1. Not surprisingly, tests between the means of the solvent and insolvent samples suggest the two groups of insurers differ significantly across a number of dimensions. Insolvent insurers carry significantly higher leverage ratios (the Kenney Ratio and the reserves to policyholder surplus ratio) than do solvent insurers.[10] Insolvent insurers are significantly smaller in terms of asset size than are solvent insurers and are less likely to be members of a mutual. Insolvent insurers pay out significantly more cash relative to premiums collected than do solvent insurers and they much more reliant on reinsurance (see the surplus aid to policyholder surplus ratio).[11]

The results of the discrete-time hazard model are shown in Panel A of Table 2. Overall the explanatory power of the model is reasonable as the pseudo R^2 statistic is 26 percent. The results are consistent with many of the inferences that were discussed after reviewing the summary statistics shown in Table 1. 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.

Panel B of Table 2 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 9.4/4.5 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 here 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).

With the one-year probabilities of default estimated, we can now begin to investigate our hypotheses. However, before we do so we first need to define a mapping between the different rating categories used by the two agencies. Unfortunately a single one-to-one mapping between two systems does not exist and prior research investigating insurance ratings across agencies have used different definitions.[12] For this study we reviewed the verbal descriptions each agency ascribes to their individual rating classes and decided to use the five rating categories shown in

[10] The Kenney Ratio equals the net premiums written by the insurer divided by the insurer's policyholder surplus. Policyholder surplus is the traditional name used in the insurance industry to represent the equity capital of the insurer under statutory (i.e., not GAAP) accounting rules.

[11] Surplus aid is a statutory accounting item which equals the increase in the amount the book value equity capital of the insurer due to the purchase of reinsurance.

[12] For example, Pottier and Sommer (1999) use a four category system to map the individual ratings assigned by each agency. Both the GAO (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 categorization system we adopted with the four level system suggested by Pottier and Sommer. The primary conclusions are similar regardless of which categorization is used.

Table 3. Numerical values, also shown in the table, were assigned to each rating category to facilitate comparisons across agencies and over time.

Table 4 shows the extent of the coverage each agency provided of the property-liability insurance industry over the time of this study. The total number of insurance companies in the NAIC data base ranged from a low of 1897 firms in year 1990 to a high of 2100 firms in year 1996. The total assets of the industry grew from $534 billion in 1989 to almost $940 billion by the end of 2000. Of these companies, A.M. Best assigned ratings to approximately 70 – 80 percent of the firms where these firms held approximately 93 percent of the assets of the industry. Obviously during the period of the 1990's, A.M. Best was providing almost complete coverage of the property-liability insurance industry. By comparison, S&P provided ratings on only 18 percent of the firms in the industry in 1992 - 360 insurers - and the number grew to 590 insurers by the end of 2000. Based on assets, S&P does provide greater coverage as they were rating firms that represented almost 70 percent of the assets of the industry by the end of the 2000 up from a low of 24 percent in 1993.

In addition to the coverage statistics, Table 4 also displays the average rating each agency assigned to the firms it oversaw. The difference across the two firms is dramatic. The average rating assigned to insurers by A.M. Best declined slightly over the time period and ranged from a high of 2.8 in 1989 and fell to 2.4 by the end of the time period. S&P stands in stark contrast in two ways. First, unlike Best's, there was a monotonic increase in the average rating assigned by S&P over the time period 1992 – 2000. In 1992, the average rating assigned by S&P was only 0.6 and it more than tripled by 2000 to be 2.1. Second, S&P appears dramatically more pessimistic about the overall financial health of the property-liability insurance industry over this time period than did A.M. Best – especially during the early part of the 1990's.

One possible explanation for the difference of opinion regarding the average health of the industry across the agencies could be because the firms tracked by A.M. Best were, on average, of higher financial quality than the firms tracked by S&P. An empirical result that would be inconsistent with the hypotheses we develop based upon the model presented in this paper. To consider this possibility, we calculate the average and median probability of default using the results from the hazard model for insurers tracked by A.M. Best and by S&P over the time period of this study. The results shown in Table 5 suggest the average and median probability of default statistics are always lower for S&P than they are for A.M. Best suggesting the firms tracked by S&P were typically of higher financial quality firms – not lower. The non-parametric Wilcoxon-Mann-Whitney test rejects the null hypothesis of equal medians for all nine years and the parametric t-test rejects the null hypothesis of equal means in seven out of nine years. Thus, it appears that, on average, S&P provided rating opinions on insurers of higher average quality and requiring higher standards in order to obtain any particular rating. However - before we can make that conclusion we need to consider the manner by which S&P entered the market for insurance ratings.

Prior to 1991, S&P provided coverage to only a small number of property-liability insurers (approximately 100). However, in 1991, S&P dramatically expanded their coverage by introducing a service they called "Insurance Solvency Review." The primary enhancement in the new service was that S&P increased the number of firms it covered by offering "qualified ratings" in addition to their traditional ratings. The methodology S&P used to determine a qualified rating for an insurer differed in at least three important ways from the traditional manner. First, qualified ratings were solely based upon publicly available data. Thus, unlike the traditional method, S&P analysts did not interview or speak to the management of an insurer prior to issuing the qualified rating. Second, individual insurers were not required to request the rating nor pay a fee to receive the qualified rating. Finally, when the system of qualified ratings was introduced, S&P maintained a policy which said no insurer could receive above a BBB rating – regardless of the characteristics of the company. S&P ultimately relaxed this position following significant criticism from the industry and began to issue qualified ratings above BBB in 1994.

Table 6 shows summary statistics regarding the types of ratings, qualified versus unqualified, given by S&P over this time period. In 1992, S&P issued 360 ratings of which 337, or 94 percent, were determined using the qualified rating system. Only 23 firms received a full rating in 1992. Over time more firms agreed to obtain a full rating and by 2000 over 300 property-liability insurers paid to receive an unqualified rating. Similar to Best's, the average full rating declined slightly over time from a high of 3.2 in 1992 to 2.8 by the end of the time period. The average qualified rating increased over time from a low of 0.5 in 1992 to 1.2 by year 2000. However, 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 using the traditional methodology.

We know from Table 6 the average ratings issued by S&P differ significantly across the two rating methodologies (qualified vs. unqualified). But does financial quality of the insurers in each category differ significantly? In Table 7 we report summary statistics regarding the default probabilities of firms rated by A.M. Best's and those rated by S&P's on a qualified and unqualified basis. Table 7 clearly display a natural ordering within each rating technology: firms that received higher ratings had, on average, lower probabilities of default. For example, the average probability of default for firms rated by A.M. Best increases monotonically by rating category from a low of 0.25 percent for firms rated "Extremely Strong" to a high of 3.11 percent for firms that received the lowest rating "Marginal." A similar pattern can be seen for S&P firms that received either a full or qualified rating. The results suggest that at least, on average, each of the three technologies required firms to be less likely to default in order to receive a higher rating.

Now consider the standards necessary to achieve a rating across each technology.

Figure 4.1 HERE

Figure 4.1 graphically displays the average probability of default of the firms over the time period of this study by rating category across each of the three rating technologies (the data can be seen in Table 7). It is easy to see the stringency employed by A.M. Best and S&P is similar when S&P issued a full unqualified rating although in three of the five categories the standards to achieve a rating are slightly more strict. The average probability of default for firms was slightly lower (and statistically significant) in the Extremely Strong, Adequate and Marginal categories. In the Good and Strong categories, the average probability of default for S&P rated companies was slightly higher than the average A.M. Best company (note - these differences are also statistically significant). These results stand in stark contrast, however, to the case where S&P issued an unqualified rating. In this case we find the average probability of default was substantially lower in each rating category relative to Best's and even relative to the standard S&P's employed on its own full ratings. For example, firms that received an Adequate rating (BBB) from S&P on a qualified basis had an average probability of default equal to 0.22 percent. A firm with a default probability of 0.22 percent likely would have received either an Extremely Strong (AAA) or a Strong (AA) rating if S&P was using their full rating standards.

5.4 Selection Bias Model to Explain Rating Differences

The results so far suggest insurers that opted to receive a fully unqualified rating from S&P were, on average, of higher financial quality. In addition, we have also shown the average probability of default of insurers in each rating class was slightly lower for S&P in three of the five rating categories consistent with the hypothesis the new entrant agency would maintain higher standards. Unfortunately the analysis thus far has two shortcomings. First, we've 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 under or over stated because of 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 agency to learn private information that is shared with the agency during the rating process. Thus, it would be advantageous to design an empirical methodology allows us to capture some of this private information. In this section we present a methodology that attempts to control for both shortcomings.

To begin the discussion, assume the following model is used by the incumbent rating agency to determine the rating for a particular firm:

$$r_{if} = \alpha_i + \boldsymbol{\beta}_i' \mathbf{X}_f + \varepsilon_{if} \tag{7}$$

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

$$r_{ef} - r_{if} = (\alpha_e - \alpha_i) + (\beta_e - \beta_i)X_f + (\varepsilon_{ef} - \varepsilon_{if}), \qquad (8)$$

where all variables subscribed with an e represent the rating and/or the technology of the new entrant firm. As discussed by Cantor and Packer (1997), estimating equation (8) directly using OLS will lead to biased results if the decision to seek a second rating from the new agency is correlated with the ratings assigned by that agency. Failure to correct for this selection bias will make it impossible to understand if the differences we see between the two rating systems are due to actual differences between the two systems 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 unqualified rating from S&P will be those that are of higher than average financial quality 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 additional regressor, known as an inverse Mill's ratio, that, when included in the ratings difference model, will control for the selection bias. Thus, in the second stage we estimate via OLS,

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

where the constant term measures the mean difference in ratings standards across the two agencies and the inverse mills term (IMR) captures the sample selection effect.[13] We hypothesize

[13] 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, we only include an intercept term and a control for sample selection bias in the second stage rating difference regressions.

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 have several hypotheses as to why an insurer would seek a second rating from S&P that we test using the Probit model. First, consistent with our theory, higher quality insurers within each rating category assigned by A. M. Best, have a stronger demand to receive a second rating from S&P. To test this hypothesis, we measure the difference between the median probability of default for all insurers within the Best's rating category assigned to each insurer, and the estimated default probability for that insurer. H1 implies that there will be greater demand for a second rating, the higher the value of this variable. We also look to see whether demand for a second rating is simply higher for insurers with a higher credit quality. Thus, we include indicator variables for each rating category assigned by A. M. Best (we omit the indicator for the Marginal class to avoid singularity). Including the Best's rating variables also allows us to control for the private information that Best's learns during the rating process.

As discussed earlier, our one-year default probabilities are calculated using publicly available information only. Thus, in an attempt to control for the private information that the rating agencies learn by going through the rating process, we include indicator variables for each rating category assigned by A.M. Best (we omit the indicator for the Marginal rating class to avoid singularity). We hypothesize that controlling for the financial quality of the insurer using the one-year default probabilities estimated using publicly available information, insurers with higher A.M. Best ratings will have greater demand to seek a second rating than insurers that received a lower rating.

We include three variables designed to test the Millon and Thakor (1985) hypothesis that more opaque or complex firms have a stronger demand for ratings. Specifically, we hypothesize larger firms and insurers with more diverse operations will be more likely to seek a second rating. Our proxy for firm size is the natural logarithm of the firm's assets. Firm complexity is proxied by geographical concentration of the insurer's business. Namely, we use 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.

We control for the organizational form of the insurer by including an indicator set equal to one if the insurer is a mutual or reciprocal insurer and zero if it is a stock insurance company. We have two competing hypotheses this variable. First, the managerial discretion literature predicts mutual insurers should underwrite less risky lines of insurance and have more transparent business models due to the reduced ability of a diffuse set of owner/policyholders to monitor management (Mayers and Smith 1987). Consistent with that literature, we hypothesize a negative relationship between the insurer being organized as a mutual and demand for a second rating. An alternative to the managerial discretion hypothesis suggests a positive relationship between mutual ownership form and the desire to seek a second rating because the only information available for potential customers to judge the financial quality of a mutual consists of regulatory accounting data. Under this rationale, stock insurers would have reduced incentives to seek an additional rating since presumably the amount of information about the insurer is readily available given the additional information available to policyholders due to analysts issuing reports that follow stock companies and because of information conveyed to the market through movements in the insurer's share price.

The final control variable we include tests the hypothesis that the risk aversion of the insurer's primary customer base provides incentives for the monopoly agency to disclosure more or less information about the true financial quality of the insurer. Recall our theory predicts that rating agencies have incentive to reveal more information to the market when the participants are more sensitive to differences in financial quality. To test this hypothesis we include a variable equal to the percentage of the insurer's premiums in retail lines of insurance.[14] We expect a positive sign on this variable consistent with the hypothesis that state guaranty funds provide greater protection to the retail policyholders of insurers that become bankrupt. Thus, retail policyholders should place less value on information and therefore the monopolist agency's optimal strategy should be to design a system that reveals little information to market participants. Based upon our theory, a new entrant has more value to add to the process due to the incumbent agency's incentive to withhold information.

5.4.1 Data and Empirical Results

The data for the rating difference tests includes any insurer that received a rating from A.M. Best over the time period 1994 - 2000. We eliminate the years 1992 and 1993 from the analysis due to S&P's policy of not assigning any firm a rating above BBB on an unqualified basis. We estimate both the first stage probit regression and the second stage OLS regressions separately for insurers that receive qualified vs. unqualified ratings from S&P since our earlier work suggests the standards across the two methodologies differs quite substantially. There are 6587 insurer-year observations in the Best's sample, 1925 observations in the S&P qualified rating sample, and 1439 observations in the S&P full rating sample.

Table 8 displays summary statistics for all the variables used in the ratings differences tests. The first-stage probit regression results are shown in Table 9. The most important results concern the test for Hypothesis 1. In Model 1 for the full rating sample (and indeed for the qualified rating sample), the variable measuring the difference between the medium insurers default in the Best's rate category and the estimated default for the insurer, is both positive and significant as predicted. While the significance remains with the inclusion of control variables for the S&P qualified rating sample, this is no longer the case for the S&P full rating sample.

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

However, with the full, as with the qualified sample, the results for the separate indicator variables for Best's rating classes reveal, as expected that the likelihood of requesting a second rating from S&P increases as the quality of the insurer (as rated by Best's) also increases. This second result also is suggested by Hypothesis 1. For example, focussing on the marginal effects on the model shown in Column 4 of Table 9, a firm rating Extremely Strong 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. We also note the likelihood that an insurer requests a full rating from S&P is monotonically increasing in A.M. Best rating category. Also consistent with our theory, large insurers, insurers in more retail lines of business, 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.

Focussing on the Probit regression results based upon qualified rating sample, the first conclusion we draw is that the exogenous variables used to explain why a firm received a qualified rating from S&P have significantly less explanatory power as the pseudo R^2 statistics for these regressions are much lower than the corresponding statistics for the full rating sample. Obviously the theory developed in this paper, which takes as a fundamental assumption that firms voluntarily will choose to seek an additional rating, does not adequately explain the manner by which S&P targeted firms to receive a qualified rating. Therefore, although the pattern among many of the estimated beta coefficients is similar in the qualified sample relative to the full rating sample, the overall explanatory power of the model is much weaker. That being said, it is reasonable to expect S&P would target insurers to receive a qualified rating in the hopes they ultimately would have demand to purchase a full rating at some point in the future. For example, insurers among the top two A.M. Best rating categories are still the most likely receive a qualified rating by S&P. In addition, more complex insurers, as proxied by the geographical concentration of their business, were also more likely to receive a qualified rating. However, there are notable differences across the two samples including mutual insurers being more likely to receive a qualified rating than stock insurers and insurers with more commercial business also more likely to receive a qualified rating.

The results from the second stage OLS regressions are shown in Table 10. Our first conclusion is that we find strong evidence of the selection effect described by H1, as the coefficient on the Inverse Mills Ratio is always positive and significantly different from zero. Insurers that seek a second rating to differentiate themselves, do expect on average to receive a more favorable rating from S&P when we control for the amount of publicly available information and for the private information revealed to A. M. Best. For example, based on the results shown in Model 1 for the S&P full rating, insurers strategically choosing a second rating expect to receive, on average, a 0.76 higher rating from S&P than A. M. Best. Qualitatively similar results are obtained for the S&P qualified rating category. Both results support the contention that, better than average insurers within any pre-existing Best's rating category, are seeking 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 on Model 1 for the full rating, we see that the estimated mean difference in rating standards is 1.36 grades lower on the S&P scale than under the A. M. Best. 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 had a rating system that required higher standards in order to maintain the same rating. We see a similar pattern for the unqualified rating methodology although the ratings scale for the new entrant on this basis was much more stringent.

6 Conclusion

The objective of the paper is twofold. First, it analyzes optimal information disclosure of a monopoly rating agency depending on the marginal value of information to buyers. Second, it characterizes the optimal entry strategy of a new rating agency to the market dominated by the incumbent. The qualitative results of the paper are that information disclosure choice 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 is beneficial for companies who voluntarily obtain the second rating. However, it decreases the payoff of the companies that are on the bottom side of each rating class. These results are strongly supported by our empirical analysis of the insurance industry.

An interesting question for further research is the optimal information disclosure in the industry whether the two agencies offer ratings simultaneously. Lizzeri (1999) establishes that a simultaneous one -shot competition between information intermediaries results in full information disclosure and zero fee for rating. However, agencies offer ratings repeatedly, and the nature of the repeated relation may allow companies to sustain positive profits.

References

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

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

[3] Brockett, Patrick L., William W. Cooper, Linda L. Golden, Utai Pitaktong, 1994, "A Neural Network Method for Obtaining an Early Warning of Insurer Insolvency", *Journal of Risk and Insurance*, Vol. 61(3):402-424.

[4] 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.

[5] 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.

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

[7] Green, Jerry R., and Nancy L. Stokey, 2007, "A Two-Person Game of Information Transmission", *Journal of Economic Theory*, 135(1): 90-104.

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

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

[10] Merton, Robert C., 1974, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance* 29: 449-470.

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

[12] 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.

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

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

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

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

[17] 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.

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

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

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

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

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

[23] Theodossiou, Panayiotis, 1993, "Predicting Shifts in the Mean of a Multivariate Time Series Process: An Application in Predicting Business Failures," *Journal of the American Statistical Association*, 88(422): 441-449.

[24] Trieschmann, James S. and George E. Pinches, 1973, "A Multivariate Model for Predicting Financially Distressed P-L Insurers", *Journal of Risk and Insurance*, 40(3): 327-338.

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

Appendix: Proofs

Proof of Proposition 1. To prove uniqueness, consider a set of types V_N who do not go to the intermediary. Then the reservation price of these types 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 this reservation price is equal to $max(0, \frac{1}{2} - \frac{1}{12}a)$. If type $v = 1$ is the only one rated, it is paid a reservation price equal to 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. ■

Proof of Proposition 2. We distinguish between two cases, $u_N > 0$ and $u_N < 0$. $u_N > 0$ is equivalent to

$$\frac{1}{2} - \frac{1}{12}av_F > 0. \tag{9}$$

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

$$(1 - v_F)(\frac{1}{2}v_F + \frac{1}{12}av_F^2)$$
$$\text{subject to (9).}$$

Denote $\lambda \geq 0$ the Lagrangian multiplier of (9). 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 (9) is not binding, $\lambda \geq 0$. Then $v_F = \frac{a-6+\sqrt{a^2+6a+36}}{3a}$, and (9) 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 3. 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

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

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

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

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

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 4. 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

$$\max_{(b_M, v_M)} (1 - v_M)(u_R(v_M, v_M + b_M) - u_N) = (1 - v_M)(\frac{1}{2}v_M + \frac{1}{12}av_M^2 + \frac{1}{2}b_M - \frac{1}{12}ab_M^2)$$

$$\frac{1}{2} - \frac{1}{12}av_M \geq 0, \quad (10)$$

$$1 - b_M - v_M \geq 0. \quad (11)$$

Constraint (10) is equivalent to $u_N > 0$, and (11) 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 (11) 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 (11) 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

$$\begin{aligned} \max_{(b_M, v_M)} (1 - v_M)u_R(v_M, v_M + b_M) &= (1 - v_M)(v_M + \frac{1}{2}b_M - \frac{1}{12}ab_M^2) \\ -\frac{1}{2} + \frac{1}{12}av_M &\geq 0 \text{ and (11).} \end{aligned}$$

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 (11) 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 Corollary 1. The proof follows immediately from comparing the monopoly profit obtained in Propositions 2 and 4. ■

Proof of Proposition 5. 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 4 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

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

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

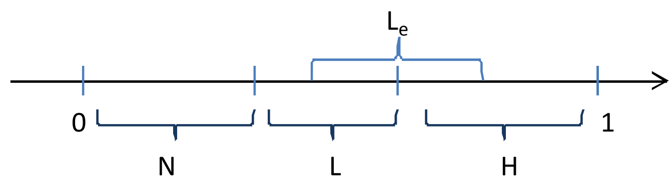

Figure 5: Ratings of the entrant must be nested in ratings of the incumbent

Proof of Proposition 6. In each rating interval k denote $\widehat{v}_k$ the type that is indifferent between full disclosure and pooling,

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

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 7. In each interval N, L, H types at the top of the interval obtain the valuation below their true types (Proposition 6). 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 3).

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 (6). 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: 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 214 firm-year observations in the insolvent sample and 24,062 in the solvent sample.

	Solvent Insurers		Insolvent Insurers		Test Statistic
Variable	μ_{sol}	σ_{sol}	μ_{ins}	σ_{ins}	H_0: $\mu_{sol} = \mu_{ins}$
Kenney Ratio: NPW to Policyholder Surplus	1.13	0.85	1.87	1.12	9.591
Reserves to Policyholder Surplus	1.03	0.94	1.64	1.25	7.237
1 Yr. Growth in NPW (%)	11.87	41.62	11.69	61.21	0.042
1 Yr. Growth in GPW (%)	11.94	37.63	11.06	52.93	0.244
Surplus Aid to Policyholder Surplus	2.05	4.34	6.07	7.52	7.816
Investment Yield (%)	5.71	1.38	5.41	1.55	2.778
1 Yr. Growth in Policyholder Surplus (%)	8.82	16.30	-8.50	19.87	12.710
Two-year Reserve Development to Policyholder Surplus (%)	-2.73	10.80	4.00	11.62	8.449
Gross Expenses to GPW	0.58	0.76	0.55	0.64	0.843
1 yr. Change in Gross Expenses (%)	0.05	0.47	0.09	0.58	1.006
1 yr. Change in Liquid Assets (%)	1.17	2.66	0.37	1.79	6.518
Investments in Affiliates to Policyholder Surplus	0.58	1.32	0.94	1.74	3.038
Receiv's. from Affiliates to Policyholder Surplus	0.02	0.04	0.04	0.05	5.243
Misc. Recoverables to Policyholder Surplus	0.03	0.05	0.07	0.08	6.691
Non-investment Grade Bonds to Policyholder Surplus	0.65	2.37	0.68	2.49	0.183
Other Invested Assets to Policyholder Surplus	0.01	0.03	0.02	0.04	3.414
Dummy = 1 if insurer has a large single agent	0.12	0.33	0.22	0.42	3.480
Dummy = 1 if insurer has a large single agent they control	0.08	0.28	0.12	0.32	1.502
Losses, Exp's, Div's and Taxes Paid to Premiums Collected	1.29	0.73	1.59	0.84	5.205
Total Assets (000000's in 2000 $)	433.65	2215.43	100.76	519.92	8.691
Ind. = 1 if insurer is part of a mutual group	0.26	0.44	0.08	0.28	8.965

Table 2: Hazard Model Regression Results

Panel A:

Table displays the results of the discrete-time hazard model regression model. The dependent variable y_{it} = 1 for each insurer that has a formal regulatory action taken against the insurer in either year t+1. Otherwise y_{it} = 0 for all other observations. There are 24,062 healthy firm-year observations and 214 insolvent company observations.

Variable	Coefficient Estimate	Standard Error	χ^2 Statistic	
Intercept	-0.7577	1.1653	0.4228	
Kenney Ratio: NPW to Policyholder Surplus	0.0047	0.0015	10.3304	***
Reserves to Policyholder Surplus	189.3000	121.0000	2.4455	
1 Yr. Growth in NPW (%)	0.0055	0.0026	4.3943	**
1 Yr. Growth in GPW (%)	0.5068	0.2575	3.8733	**
Surplus Aid to Policyholder Surplus	0.0415	0.0128	10.5759	***
Investment Yield (%)	-0.0117	0.0646	0.0328	
1 Yr. Growth in Policyholder Surplus (%)	-0.0390	0.0063	38.6798	***
Two-year Reserve Development to Policyholder Surplus (%)	0.0311	0.0086	13.0963	***
Gross Expenses to GPW	0.2654	0.1796	2.1834	
1 yr. Change in Gross Expenses (%)	-0.1195	0.1961	0.3714	
1 yr. Change in Liquid Assets (%)	-0.0462	0.0517	0.7956	
Investments in Affiliates to Policyholder Surplus	0.0000	0.0000	10.9740	***
Receiv's. from Affiliates to Policyholder Surplus	3.3208	1.6962	3.8327	*
Misc. Recoverables to Policyholder Surplus	2.1060	1.2641	2.7755	*
Non-investment Grade Bonds to Policyholder Surplus	0.0556	0.0317	3.0818	*
Other Invested Assets to Policyholder Surplus	6.7624	2.2026	9.4257	***
Dummy = 1 if insurer has a large single agent	0.6341	0.2205	8.2740	***
Dummy = 1 if insurer has a large single agent they control	-0.3205	0.2870	1.2477	
Losses, Exp's, Div's and Taxes Paid to Premiums Collected	0.6960	0.1589	19.1885	***
Ln(Total Assets in $2000)	-0.4707	0.0665	50.0860	***
Ind. = 1 if insurer is part of a mutual group	-0.8337	0.2709	9.4694	***
Log Likelihood Function Value	-908.617			
Pseudo R^2	25.86%			

*** - 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	24,062	0.81%	0.20%	2.46%	0.01%	11.08%
Insolvent	214	9.35%	4.46%	12.78%	0.09%	66.45%

Table 3
Insurer Rating Categories: A.M Best vs. Standard & Poor's

Number	Description	A.M. Best	S&P
4	Extremely Strong	A++,A+	AAA
3	Strong	A	AA
2	Good	A-	A
1	Adequate	B++,B+	BBB
0	Marginal	B and below	BB and below

Table 4
Number of Companies Rated and Average Rating
A.M. Best vs. Standard & Poor's 1989 - 2000

Table displays the number of companies in the NAIC database, the number of firms rated by A.M. Best and 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 final two columns display the average rating of the companies rated by agency.

	Number of Companies			Total Assets ($ billions)			Average Rating	
Year	NAIC	A.M. Best	S&P	NAIC	A.M. Best	S&P	A.M. Best	S&P
1989	1903	1110 (58.3%)		534.6	491.6 (91.9%)		2.840	
1990	1897	1175 (61.9%)		566.5	511.6 (90.3%)		2.750	
1991	1968	1261 (64.1%)		615.2	565.1 (91.9%)		2.604	
1992	2012	1352 (67.2%)	360 (17.9%)	659.3	597.5 (90.6%)	218.9 (33.2%)	2.583	0.664
1993	2061	1437 (69.7%)	349 (16.9%)	698.2	635.8 (91.1%)	164.5 (23.6%)	2.555	0.693
1994	2065	1515 (73.4%)	391 (18.9%)	729.3	668.7 (91.7%)	177.9 (24.4%)	2.434	1.192
1995	2084	1551 (74.4%)	393 (18.9%)	783.9	724.0 (92.4%)	200.0 (25.5%)	2.386	1.417
1996	2100	1577 (75.1%)	572 (27.2%)	830.2	774.0 (93.2%)	554.6 (66.8%)	2.382	1.818
1997	2096	1598 (76.2%)	568 (27.1%)	911.0	860.9 (94.5%)	612.9 (67.3%)	2.416	1.871
1998	2096	1620 (77.3%)	587 (28.0%)	949.4	897.5 (94.5%)	644.1 (67.8%)	2.474	2.082
1999	2042	1620 (79.3%)	583 (28.6%)	953.0	903.1 (94.8%)	655.7 (68.8%)	2.500	2.163
2000	1952	1570 (80.4%)	590 (30.2%)	938.5	888.2 (94.6%)	645.4 (68.8%)	2.461	2.110

* - S&P provided ratings on property-liability insurers over the years 1989-1991. We were unable to locate this data in electronic format.

Table 5
Summary Statistics One-Year Probability of Default
A.M Best vs. Standard & Poor's: 1989 - 2000

Table displays the average and median probability of default of the firms that receive ratings by A.M. Best and Standard & Poor's. The T-test column reports the results testing the average proability of default for A.M. Best is different than S&P assuming unequal variances. The column labeled "Non-Par." reports the results of the non-parametric Wilcoxon-Mann-Whitney difference in medians test. The chart below displays the average and median statistics for each agency over time period of this study.

	A.M. Best				Standard & Poor's				Test Statistics $H_0: \mu_{Best} = \mu_{S\&P}$	
Year	Num	Mean	Median	Std. Dev.	Num	Mean	Median	Std. Dev.	T-Test	Non-Par.
1989	1110	0.40%	0.14%	1.39%						
1990	1175	0.60%	0.27%	1.25%						
1991	1261	0.50%	0.14%	1.80%						
1992	1352	0.64%	0.15%	2.55%	360	0.37%	0.11%	1.07%	3.050	4.427
1993	1437	0.47%	0.12%	2.15%	349	0.34%	0.09%	0.82%	1.767	4.027
1994	1515	0.46%	0.14%	1.18%	391	0.41%	0.10%	1.57%	0.601	4.953
1995	1551	0.42%	0.10%	2.14%	393	0.24%	0.06%	0.89%	2.604	5.744
1996	1577	0.88%	0.19%	3.44%	572	0.41%	0.14%	0.88%	4.988	5.028
1997	1598	0.52%	0.15%	1.72%	568	0.32%	0.11%	1.35%	2.771	4.767
1998	1620	0.69%	0.16%	2.91%	587	0.28%	0.12%	0.48%	5.421	5.108
1999	1620	0.80%	0.20%	2.55%	583	0.45%	0.16%	1.49%	3.935	3.898
2000	1570	0.68%	0.20%	2.05%	590	0.36%	0.17%	0.78%	5.129	3.548

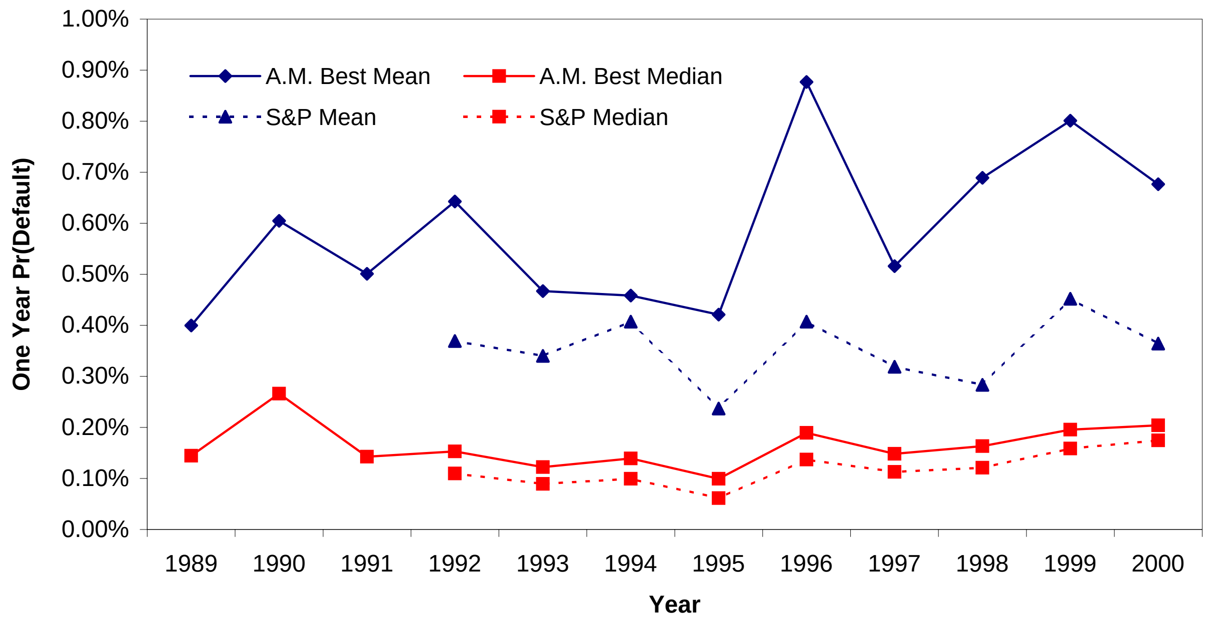

Table 6
Standard & Poor's Qualified vs. Full Ratings: 1992 - 2000

Table displays summary statistics of ratings S&P issued property-liability insurers using the qualified vs. the unqualified rating system over the years 1992 - 2000.

	Full Ratings						Qualified Ratings						Test Statistic H_0: $\mu_{full} = \mu_{qual}$	
Year	Num	Percent	μ_{full}	σ_{full}	Min	Max	Num	Percent	μ_{qual}	σ_{qual}	Min	Max	T-Stat	
1992	23	(6.4%)	3.22	0.74	1.00	4.00	337	(93.6%)	0.49	0.50	0.00	1.00	17.50	***
1993	25	(7.2%)	2.96	1.14	0.00	4.00	324	(92.8%)	0.52	0.50	0.00	1.00	10.67	***
1994	34	(8.7%)	3.03	1.03	0.00	4.00	357	(91.3%)	1.02	0.87	0.00	3.00	11.03	***
1995	55	(14.0%)	2.89	1.07	1.00	4.00	338	(86.0%)	1.18	0.91	0.00	4.00	11.27	***
1996	232	(40.6%)	2.78	0.91	1.00	4.00	340	(59.4%)	1.16	0.91	0.00	4.00	20.79	***
1997	255	(44.9%)	2.75	0.85	1.00	4.00	313	(55.1%)	1.16	0.85	0.00	4.00	22.09	***
1998	320	(54.5%)	2.78	0.81	1.00	4.00	267	(45.5%)	1.24	0.87	0.00	4.00	22.04	***
1999	343	(58.8%)	2.80	0.73	0.00	4.00	240	(41.2%)	1.25	0.81	0.00	4.00	23.64	***
2000	339	(57.5%)	2.77	0.77	0.00	4.00	251	(42.5%)	1.22	0.82	0.00	4.00	23.37	***

*** - significant at the 1 percent level

Number of Insurers that Received Qualified and Unqualified Ratings by Standard & Poor's 1992 - 2000

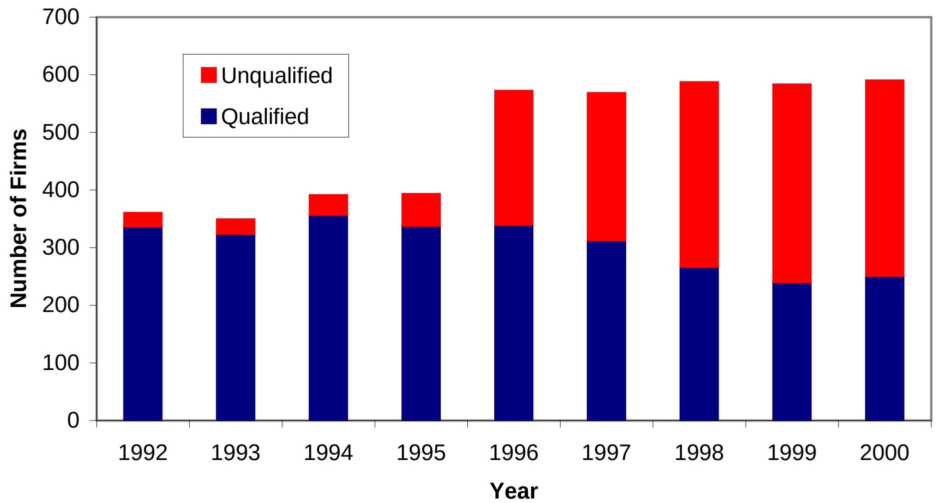

Table 7
Stringency and Accuracy of A.M. Bests vs. Standard and Poor's Ratings

Each panel shows the distribution of ratings issued by a particular rating agency over the time period of this study as well as summary statistics of the probability of default by rating category. Panel A displays statistics for firms that receive a full unqualified rating from S&P during the years 1992 - 2000. Panel B displays statistics for firms that received qualified ratings from S&P during the years 1992 - 2000. Panel C displays statistics for firms that received ratings from A.M. Best during the years 1989 - 2000.

Panel A: Firms that Receive a Full Rating from Standard & Poor's

					Percentiles			
Rating	**Num**		μ	σ	**10^{th}**	**Median**	**90^{th}**	**90^{th} - 10^{th}**
Extremely Strong	339	(20.8%)	0.16%	0.22%	0.01%	0.09%	0.33%	0.32%
Strong	697	(42.9%)	0.36%	1.40%	0.03%	0.13%	0.61%	0.58%
Good	515	(31.7%)	0.51%	0.72%	0.06%	0.29%	1.16%	1.10%
Adequate	69	(4.2%)	0.51%	0.93%	0.04%	0.22%	1.15%	1.11%
Marginal	6	(0.4%)	1.11%	1.94%	0.01%	0.43%	5.03%	5.02%
	1626							

Panel B: Firms that Receive a Qualified Rating from Standard & Poor's

					Percentiles			
Rating	**Num**		μ	σ	**10^{th}**	**Median**	**90^{th}**	**90^{th} - 10^{th}**
Extremely Strong	9	(0.3%)	0.01%	0.02%	0.00%	0.01%	0.06%	0.06%
Strong	92	(3.3%)	0.05%	0.06%	0.01%	0.02%	0.13%	0.12%
Good	673	(24.3%)	0.15%	0.46%	0.02%	0.06%	0.26%	0.24%
Adequate	1131	(40.9%)	0.22%	0.45%	0.02%	0.09%	0.44%	0.42%
Marginal	862	(31.2%)	0.70%	1.84%	0.03%	0.20%	1.62%	1.58%
	2767							

Panel C: A.M. Best Ratings

					Percentiles			
Rating	**Num**		μ	σ	**10^{th}**	**Median**	**90^{th}**	**90^{th} - 10^{th}**
Extremely Strong	4790	(27.6%)	0.25%	0.74%	0.02%	0.10%	0.48%	0.46%
Strong	4593	(26.4%)	0.31%	0.90%	0.03%	0.13%	0.64%	0.61%
Good	4067	(23.4%)	0.43%	1.06%	0.04%	0.18%	0.91%	0.88%
Adequate	2683	(15.4%)	0.77%	2.16%	0.06%	0.28%	1.66%	1.60%
Marginal	1253	(7.2%)	3.11%	6.61%	0.12%	0.89%	7.91%	7.79%
	17,386							

Table 8
Summary Statistics of Insurer's Receiving Full vs. Qualified Ratings from Standard & Poor's: 1994 - 2000

The sample includes insurer-year observations for all firms that receive an A.M. Best rating and any firm that receives a rating from S&P over the years 1994 - 2000. We eliminate all observations from 1992 and 1993 because of S&P's policy of not assigning any firm a rating above BBB when the insurer is being rated on a qualified basis.

	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	$H_0: \mu_a = \mu_q$	$H_0: \mu_a = \mu_f$	$H_0: \mu_q = \mu_f$
Ind. = 1 for Marginal A.M. Best Rating	0.089	0.285	0.035	0.183	0.001	0.026	10.0 ***	24.7 ***	8.1 ***
Ind. = 1 for Adequate A.M. Best Rating	0.216	0.412	0.115	0.319	0.015	0.122	11.5 ***	33.6 ***	12.6 ***
Ind. = 1 for Good A.M. Best Rating	0.284	0.451	0.193	0.395	0.153	0.360	8.6 ***	12.0 ***	3.1 ***
Ind. = 1 for Strong A.M. Best Rating	0.244	0.430	0.368	0.482	0.278	0.448	10.1 ***	2.7 ***	5.6 ***
Ind. = 1 for Extremely Strong A.M. Best Rating	0.166	0.372	0.289	0.454	0.553	0.497	10.9 ***	28.0 ***	15.9 ***
Median Pr(Default | A.M. Best) - Insurer Pr(Default)	-0.005	0.025	-0.001	0.007	-0.002	0.011	11.6 ***	5.9 ***	4.7 ***
S&P Rating	-		1.215	0.862	2.770	0.806			53.9 ***
A.M. Best Rating	2.182	1.204	2.762	1.095	3.368	0.798	20.0 ***	46.3 ***	18.6 ***
S&P Rating - A.M. Best Rating	-		-1.547	1.016	-0.598	0.648			33.0 ***
Ind. = 1 if insurer is part of a mutual group	0.319	0.466	0.473	0.499	0.175	0.380	12.1 ***	12.4 ***	19.7 ***
Total Assets (000000's in 2000 $)	352.3	1861.7	414.9	885.9	1,860.4	5,963.9	2.0 **	9.6 ***	9.2 ***
% NPW in Retail Lines of Insurance	0.360	0.368	0.356	0.360	0.322	0.328	0.4	3.9 ***	2.8 ***
Ind. = 1 if year = 1994	0.154	0.361	0.166	0.372	0.020	0.140	1.2	23.3 ***	15.8 ***
Ind. = 1 if year = 1995	0.163	0.370	0.162	0.368	0.036	0.185	0.2	19.2 ***	13.0 ***
Ind. = 1 if year = 1996	0.138	0.345	0.163	0.369	0.150	0.357	2.6 ***	1.2	1.0
Ind. = 1 if year = 1997	0.133	0.339	0.146	0.353	0.160	0.367	1.4 *	2.6 ***	1.1
Ind. = 1 if year = 1998	0.136	0.343	0.126	0.332	0.197	0.398	1.2	5.4 ***	5.5 ***
Ind. = 1 if year = 1999	0.137	0.344	0.116	0.321	0.220	0.414	2.5 ***	7.1 ***	7.9 ***
Ind. = 1 if year = 2000	0.138	0.345	0.122	0.327	0.217	0.413	1.8 **	6.9 ***	7.3 ***

*** - 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 only one rating from A.M. Best was 6587. The number of firm-year observations that received both an A.M. Best rating and a qualified/full rating by Standard & Poor's was 1925/1459, respectively.

Table 9
Probit Regression Results Predicting Whether Insurer Received a Full or Qualified Rating from Standard & Poor's: 1994 - 2000

Table displays Probit regression results where the dependent variable for the first two regression models equaled 1 when insure i was assigned a qualified rating by Standard & Poor's in year t and 0 otherwise. In the last two regressions the indicator variable equals 1 when insurer i requested a full unqualified rating from Standard & Poor's in year t. Panel A displays the estimated coefficients on the independent variables. Panel B displays the marginal effects.

Panel A: Regression Results	**Did Insurer Receive Qualified Rating from S&P?**		**Did Insurer Request Full Rating from S&P?**	
Independent Variable	**Model 1**	**Model 2**	**Model 1**	**Model 2**
Intercept	-0.9503 ***	-2.4668 ***	-0.7661 ***	-5.6412 ***
	(0.040)	(0.201)	(0.037)	(0.393)
Median Pr(Default | A.M. Best) - Insurer Pr(Default)	16.8910 ***	8.5616 ***	4.2096 ***	0.1443
	(2.057)	(0.080)	(1.216)	(1.941)
Ind. = 1 for Adequate A.M. Best Rating		0.0795		0.6393 **
		(0.081)		(0.325)
Ind. = 1 for Good A.M. Best Rating		0.0913		1.4601 ***
		(0.078)		(0.315)
Ind. = 1 for Strong A.M. Best Rating		0.4509 ***		1.6510 ***
		(0.077)		(0.314)
Ind. = 1 for Extremely Strong A.M. Best Rating		0.2861 ***		2.1744 ***
		(0.080)		(0.080)
Ln(Total Assets in $2000)		0.0630 ***		0.1915 ***
		(0.010)		(0.012)
Ind. = 1 if insurer is part of a mutual group		0.4996 ***		-0.6318 ***
		(0.033)		(0.047)
% NPW in Retail Lines of Insurance		-0.1408 ***		0.1701 ***
		(0.046)		(0.058)
State of Business of Herfindahl		-0.1037 **		-0.5685 ***
		(0.046)		(0.059)
Log-likehood function value	-4811.9	-4592.2	-3846.8	-2879.6
Pseudo R^2	1.64%	6.13%	7.32%	30.63%
Panel B: Estimated Marginal Effects				
Median Pr(Default | A.M. Best) - Insurer Pr(Default)	4.4935 ***	2.2128 ***	0.8483 ***	0.0158
Ind. = 1 for Adequate A.M. Best Rating		0.0205		0.0702 **
Ind. = 1 for Good A.M. Best Rating		0.0236		0.1603 ***
Ind. = 1 for Strong A.M. Best Rating		0.1165 ***		0.1813 ***
Ind. = 1 for Extremely Strong A.M. Best Rating		0.0739 ***		0.2387 ***
Ln(Total Assets in $2000)		0.0163 ***		0.0210 ***
Ind. = 1 if insurer is part of a mutual group		0.1291 ***		-0.0694 ***
% NPW in Retail Lines of Insurance		-0.0364 ***		0.0187 ***
State of Business of Herfindah		-0.0268 **		-0.0624 ***

*** - 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
OLS Regression Results Explaining Rating Difference Between Standard & Poor's and A.M. Best Ratings: 1994 - 2000

Table displays OLS regression results where the dependent variable for the first two regression models equaled the difference between the qualified rating assigned by Standard & Poor's minus the rating assigned by A.M. Best in year t. The dependent variable for the last two regressions the indicator variable equaled the difference between the full rating assigned by Standard & Poor's minus the rating assigned by A.M. Best in year t.

	Did Insurer Receive Qualified Rating from S&P?		Did Insurer Request Full Rating from S&P?	
Independent Variable	**Model 1**	**Model 2**	**Model 1**	**Model 2**
Intercept	-1.8229 ***	-2.6280 ***	-1.3560 ***	-0.9242 ***
	(0.265)	(0.127)	(0.120)	(0.041)
Inverse Mills Ratio	0.1889	0.0924 ***	0.0808 ***	0.0338 ***
	1.0460	8.7910	6.4290	8.8150
R^2	0.57%	4.40%	2.91%	5.17%
Expected increase in rating due to insurer strategic choice	0.276	1.081	0.758	0.326
Estimated difference in ratings due difference in S&P vs. A.M. Best standards	-1.823	-2.628	-1.356	-0.924
Average Rating Difference S&P Rating - A.M. Best Rating	-1.547	-1.547	-0.598	-0.598