A Probit Analysis of Best Ratings

Raja Bouzouita and Arthur J. Young*

Abstract: In 1991, the A.M. Best Company changed the schedules and procedures for assigning ratings to property-liability insurers. An explanation of this change is of interest to consumers, regulators, and management. This study uses ordered probit to test for a downshift in the average ratings. The results of the analysis show that there has been a movement toward lower ratings and that the return on surplus is more important with the new ratings. The results are robust with respect to alternative estimation procedures.

INTRODUCTION

Consumers, agents, and brokers are always concerned about the solvency of insurance companies. Policyholders seeking to obtain information on the financial strength of insurance companies are hindered by the complexity of the financial information and the difficulty of predicting insolvency. Published ratings provide an opinion on an insurer's financial strength, including its operating performance and its ability to meet its obligations to its policyholders.[1]

Insurance ratings do not provide any guarantee against default. They do, however, provide policyholders and investors with some information regarding the likelihood of insolvency. Ambrose and Seward (1988) find that the ratings of the A.M. Best Company perform as well as financial ratios in distinguishing between solvent and insolvent insurers; neither one

*Raja Bouzouita is an Assistant Professor of Finance at Central Missouri State University. She received her Ph.D. from the Wharton School, University of Pennsylvania. She has published articles in *The Journal of Risk and Insurance*, the *CPCU Journal*, and the *Journal of Insurance Regulation*. Arthur J. Young is an Assistant Professor of Finance at Central Missouri State University. He received his Ph.D. from Georgia State University. He is director of Central's certificate program in Personal Financial Planning. He has published articles in *Advances in Financial Planning and Forecasting*, the *CPCU Journal*, and the *Journal of Insurance Regulation*.

Journal of Insurance Issues, 1998, **21**, 1, pp. 23–34.

is a perfect predictor of insolvency. The failure of highly rated life insurance companies, such as Executive Life Insurance and Mutual Benefit Life Insurance Companies,[2] generated heavy criticism of the leading rating agencies by the insurance press. These bankruptcies of very prominent insurance companies raised several questions about the motivations and methods of the rating agency. Historically, insurer rating agencies have been criticized for assigning inflated ratings.

The oldest rating agency of insurance companies is the A.M. Best Company. For decades, the A.M. Best Company had a monopoly in rating insurance companies. However, by the middle of the 1980s, several other rating agencies that have a well-established reputation in bond ratings started competing with the A.M. Best Company by providing an opinion on the financial condition of insurance companies. These rating agencies include Standard & Poors, Moody's, and Duff & Phelps.[3] In light of this increased competition, the ratings might be expected to better reflect the underlying characteristics of insurance companies. On the other hand, there are concerns that leading raters have rushed to lower their ratings in order to regain policyholders' confidence (Klein, 1992).

Until 1990 the A.M. Best Company rated the financial strength of insurance underwriters on a scale ranging from A+ to C–. In 1991, the A.M. Best Company expanded the categories from A++ to D, with some additional categories—B++ and C++—in between. The question arises as to whether the A.M. Best Company used the expansion of the rating categories as an opportunity to reduce the inflation in the ratings. If the A.M. Best Company raised the rating standards, then, after controlling for the financial situation of the insurer, the new ratings should be found to be lower, on average, than the previous ratings.

This paper explores the changes in the rating system by examining some key variables related to the likelihood of receiving a certain rating both before and after the change in the rating procedures. The results will help provide some answers to the issues and concerns raised by industry observers and consumers. A less inflated rating system will improve consumers' confidence and will also be of interest to regulators as a monitoring device of the companies on the low end of the rating schedule. On the other hand, if higher standards determine the ratings, managers of insurance companies need to understand these new standards to change practices and be able to explain to policyholders that higher standards may cause lower ratings.

The paper is organized as follows. Section 2 reviews the literature on the Best's ratings as predictors of insolvency. In section 3, data and the methodology are discussed. The empirical results are provided in section

4. The paper concludes with a summary of the major findings, implications, and recommendations for future research.

RELATED LITERATURE

In the insurance literature, two issues arise involving the ratings of insurance companies. The first issue focuses on the statistical models attempting to accurately identify financially distressed companies including Best's ratings in the set of predictor variables. Ambrose and Seward (1988) use a series of dummy variables for Best's letter ratings from A+ to B as one of the predictors of insurer insolvencies. By including the ratings, the percentage of insolvent companies that are correctly classified one year prior to insolvency increases to 90 percent. In a second study on life and health insurers, Ambrose and Carroll (1994) use a dummy variable for firms with A+ or A ratings and firms with lower ratings to predict insolvencies. The authors find the ratings to be one of the significant predictors of insolvency.

The second issue involves using financial ratios and firm characteristics to explain the ratings. This line of research has appeared primarily in the accounting and finance literature; see Altman et al. (1981) and Edderington (1985) for a survey. In the insurance literature, an earlier study by Brotman (1989) uses ordinary least squares to analyze the factors that determine Best's ratings for a cross section. A somewhat surprising result is that the only significant factors that explain the ratings are the size of the surplus and the percentage change of assets.

The main focus of this paper is to test for a change in the rating standards, not to predict solvency or ratings. This study treats the rating as a categorical variable rather than a binary or a continuous variable. Second, this paper compares the two ratings regimes with an emphasis on finding a concise model that represents both regimes and tests for a shift between the two regimes.

DATA AND METHODOLOGY

A.M. Best states that the performance evaluation assesses insurers' ability to meet obligations to policyholders. The evaluation process involves quantitative and qualitative analysis that is based to some extent on confidential non-public information. Best's financial analysis of insurance companies is based on financial statements from the most recent five-year period.[4] Similarly, this study uses five-year averages of financial ratios. The financial data source is A.M. Best tapes from 1985 to 1992. The sample

consists of stock and mutual property-liability insurers that received a letter rating from the A.M. Best Company in 1989 through 1992 as reported in Best's Key Rating Guide. The firms in the sample were required to have full information for the four years covered by the study.[5]

The focus of this paper is to examine whether, at the time of the expansion of the rating categories, there has been a significant shift toward more conservative ratings. The move toward lower ratings is anticipated as a response to the failure of highly rated life insurance companies and accusations that the ratings are generally inflated. An indicator variable for the change in the rating regime and its interaction with other control variables are used to test for the shift toward moderation in the ratings.

The control variables used in this paper are selected on the basis of the number and type of variables commonly used in the insurance insolvency literature.[6] Ambrose and Seward (1988) find that the ratings and financial ratios are equally successful in predicting insolvency. The financial ratios are used as a control for changes in the firm's characteristics related to insolvency, so that the change in Best's rating standards can be appropriately measured. The model is assumed to have the following form:[7]

$$\begin{aligned} Rating = {} & \beta_0 + \beta_1 GS + \beta_2 ROS + \beta_3 LM + \beta_4 NPW/S + \beta_5 GP \\ & + \beta_6 LIQ + \beta_7 SIZ + \beta_8 DA + \beta_9 STK + \beta_{10} TT + \beta_{11} RD + \beta_{12} GSRD \\ & + \beta_{13} ROSRD + \beta_{14} LMRD + \beta_{15} NPW/SRD + \beta_{16} GPRD \\ & + \beta_{17} LIQRD + \beta_{18} SIZRD + \beta_{19} DARD + \varepsilon \end{aligned} \tag{1}$$

The explanatory variables used in equation (1) capture both the financial and operating performance of the insurer. Profitability (ROS), measured by the ratio of net income and unrealized capital gains to the statutory policyholders' surplus, assesses the insurer's ability to efficiently invest the surplus. It is expected that more profitable firms will have a higher rating. Profitability reflects the ability of management to maintain strong operation and adequate pricing. Policyholders' surplus is a safety cushion against unfavorable fluctuations in the underwriting experience of the insurer and against adverse economic conditions. Growth in surplus (GS) is used to measure this safety net. In case of failure, policyholders will have a claim against the surplus. Companies that experience a sustained increase in the level of their surplus are more likely to receive a better rating. A second measure of growth (GNPW), growth in net premiums written, is included. Insurance companies generate revenues from the sales of policies. A rapid growth in the volume of premiums written, without an adequate increase in surplus, may increase the probability of default.

Insurance leverage (NPW/S) measures the ability of the insurer to write new business without jeopardizing the relative financial strength of

the company. A high level of leverage magnifies adverse variations in underwriting and / or economic conditions. Therefore, high levels of leverage would adversely affect the ratings. A second measure of leverage used is the ratio of total debts to assets (DA). Insurance companies' liabilities are largely contingent obligations to their policyholders. Therefore, a high level of financial leverage increases the probability of default to policyholders. It is widely recognized that high leverage results in high return on capital. However, increases in leverage may increase the risk of insolvency.

The composition of the book of business referred as line mix (LM) affects the profitability and the riskiness of the company. The mix of the product lines, measured by the ratio of net premiums written in the long-tail lines to total net premiums written,[8] has significant implications on the overall degree of risk of the loss distribution. Insurers that write a significant proportion of their portfolio in the long-tail lines have a total loss distribution with a relatively high variance. The larger the proportion of premiums written in the long-tail lines, the riskier the insurance portfolio and therefore the ratings will probably be lower.

Insurance companies should be able to meet their financial obligations as they come due. Liquidity (LIQ) is measured by the quick liquidity ratio, which is the amount of cash and short-term investments that can be readily converted into cash. A high degree of liquidity enables the insurer to meet unexpected needs for cash without having to sell assets at a discounted value. However, a large proportion of assets held in the form of liquid assets reduces the overall rate of return on invested assets.

The size (SIZ) of a given company is measured by the log of admitted assets, which captures many aspects or specific characteristics, qualitative as well as quantitative, of the firm. Larger companies have access to skilled and experienced management, which adds value to the company. Also, large companies are better able to sustain unfavorable changes in the underwriting and economic conditions in general. Thus, larger firms are more likely to receive a better rating.

To proxy the qualitative features of management and operating practices, a distinction is made between stock and mutual companies. Mayers and Smith (1997) document that organizational form in the insurance industry is closely related to corporate policy choices and decisions such as risk-taking, distribution systems, and product line. A binary variable (STK) that takes a value of one for stock and zero for mutual is included in all models.

The dependent variable in equation (1) is a categorical variable coded from best to worst by numbers from zero to four (i.e., zero is superior, one is excellent, two is very good, three is good, and four is used for the combination of the lower ratings). The true quality of the insurer is not

observable. Each insurer is given a letter rating as perceived by the A.M. Best Company. Thus, there is a latent continuous variable, rating, for which categorical data exist. Furthermore, there is information contained in the categorization of these responses, which fall into *ordered* or ranked groups. Therefore, this study uses ordered probit, which gives the maximum likelihood estimates of the probabilities of receiving a certain ordered rating. The ordered probit model should be used when there is an underlying relationship between the different categorical responses.[9] In addition to ordered probit, the model is estimated using ordinary least squares to verify the robustness of the results with respect to alternative estimation procedures.

Three different models are estimated in this paper. All three models use a time trend variable to control for the possibility of a steady change over time in the ratings, rather than a shift related to change in the rating process and schedule. Since the objective of this study is to test for changes at the time of the rating change, the second and third models use an indicator variable that measures a shift in the relationship between the explanatory and dependent variables. This indicator variable is equal to 0 prior to the rating change and equal to one thereafter, and is referred to as the rating regime dummy variable (RRD). The third model explores the change in the rating regime by including interaction variables between the rating regime dummy and the explanatory variables, which measure a change in the explanatory variables' coefficients showing a change in the relationship with the dependent variable. A joint test of the rating regime dummy and all interaction terms will show the rating regime's impact on the relationship between financial ratios and Best's ratings.

EMPIRICAL RESULTS

The ratings of the insurance companies in the sample are shown in Table 1. Prior to 1991 there were nine levels of rankings, designated by letters from A+ to C–. From 1991 and later there are 13 levels of rankings, designated by letters from A++ to D. Table 1 shows the distribution of the level of the ratings before and after the rating change. The distribution of ratings is skewed toward the better rating categories. The skewness in the ratings can be due to a number of reasons. The ratings are voluntary, meaning that the rated insurer has the right to appeal or reject the assigned rating before it is made public. There are several categories that are not rated for various reasons, such as not meeting the minimum size requirement, insufficient operating experience, and major change in line of busi-

Table 1. The Distribution of Best's Ratings of Property-Liability Insurers in the Prior and Subsequent Two-year Periods Surrounding the Change of the A. M. Best's Ratings

PRIOR				SUBSEQUENT			
		1989	1990			1991	1992
				A++	Superior	12	14
A+	Superior	119	117	A+	Superior	94	91
A	Excellent	97	92	A	Excellent	97	95
A–	Excellent	58	69	A–	Excellent	68	65
				B++	Very Good	13	20
B+	Very Good	44	36	B+	Very Good	27	26
B	Good	11	11	B	Good	15	14
B–	Good	5	10	B–	Good	6	5
				C++	Fair	0	0
C+	Fairly Good	2	1	C+	Fair	4	2
C	Fair	0	0	C	Marginal	0	2
C–	Fair	0	0	C–	Marginal	0	1
				D	Below Minimum Standards	0	1

Source: A.M. Best *Key Rating Guide,* 1993.

ness or ownership. After the change in the rating schedule, the ratings are less skewed. There are fewer firms in the superior category and more firms in the categories below the good category.

The results from ordered probit are reported in Table 2. In the first model, all of the explanatory variables are significant, except for growth in net premiums written, and have the expected effect on the ratings. The parameters from the ordered probit do not lend themselves to the same interpretation as the estimates from linear models. Instead, the marginal impacts of the explanatory variables are calculated using a procedure provided in Greene (1990).[10] The marginal effects for the three models show that if the estimated coefficient is positive, then the marginal impact is negative on the highest category and positive on the remaining four categories and vice versa. The coefficient of the organizational dummy shows that stock companies tend to have a lower rating than mutual

Table 2. Ordered Probit Estimates of the Impact of Financial Ratios on Best's Ratings, 1989–1992

Independent Variables	Model 1	Model 2	Model 3
Intercept	4.777	4.823	4.142
	(10.9)***	(11.0)***	(7.41)***
Growth in Surplus	–1.278	–1.276	–1.145
	(–6.53)***	(–6.51)***	(–4.67)***
Return on Surplus	–1.598	–1.599	–1.054
	(–10.18)***	(–10.2)***	(–5.08)***
Line Mix	0.706	0.704	0.670
	(6.31)***	(6.29)***	(4.20)***
NPW to Surplus	0.510	0.510	0.522
	(8.80)***	(8.79)***	(6.59)***
Growth in Net Premuims	0.003	0.003	–0.022
	(0.04)	(0.04)	(–0.19)
Quick Liquidity	0.327	0.328	0.383
	(6.25)***	(6.24)***	(6.02)***
Admitted Assets	–0.481	–0.481	–0.461
	(–18.5)***	(–18.5)***	(–13.3)***
Debts to Assets	3.802	3.805	3.984
	(11.1)***	(11.1)***	(8.49)***
Stock	0.302	0.302	0.321
	(4.69)***	(4.69)***	(4.96)***
Time Trend	0.086	0.048	0.046
	(3.07)***	(0.80)	(0.77)***
Rating Regime Dummy (RRD)		0.095	1.561
		(0.70)	(2.03)**
Growth in Surplus*RRD			0.227
			(0.45)
Return on Surplus*RRD			–1.911
			(–4.14)***
Line Mix*RRD			0.103
			(0.46)
NPW to Surplus*RRD			–0.011
			(–0.09)
Growth in Premiums Written*RRD			0.033
			(0.23)
Liquidity*RRD			–0.129
			(–1.23)
Admitted Assets*RRD			–0.038
			(–0.86)
Debts to Assets*RRD			–0.504
			(–0.74)
MU(1)	1.865	1.865	1.882
	(30.1)***	(30.1)***	(29.9)***
MU(2)	2.750	2.750	2.774
	(33.1)***	(33.1)***	(33.2)***
MU(3)	4.010	4.011	4.043
	(29.3)***	(29.3)***	(28.8)***
Pseudo R^2	0.3688	0.3689	0.3731

*** significant at the 1% level
** significant at the 5% level

Table 3. OLS Estimates of the Impact of Financial Ratios on Best's Ratings, 1989–1992

Independent Variables	Model 1	Model 2	Model 3
Intercept	3.225	3.252	2.775
	(15.5)***	(15.4)***	(9.66)***
Growth in Surplus	–0.804	–0.802	–0.698
	(–5.40)***	(–5.39)***	(–3.96)***
Return on Surplus	–1.138	–1.138	–0.778
	(–7.69)***	(–7.69)***	(–4.38)***
Line Mix	0.389	0.388	0.358
	(5.76)***	(5.74)***	(3.66)***
NPW to Surplus	0.338	0.338	0.345
	(9.68)***	(9.67)***	(7.17)***
Growth in Net Premuims	0.020	0.020	0.009
	(0.99)	(0.94)	(0.34)
Quick Liquidity	0.181	0.181	0.210
	(6.54)***	(6.54)***	(5.51)***
Admitted Assets	–0.262	–0.262	–0.242
	(–24.0)***	(–24.0)***	(–15.9)***
Debts to Assets	1.960	1.962	1.970
	(9.94)***	(9.95)***	(7.19)***
Stock	0.167	0.167	0.180
	(4.46)***	(4.46)***	(4.77)***
Time Trend	0.047	0.024	0.023
	(2.92)***	(0.68)	(0.67)
Rating Regime Dummy (RRD)		0.058	1.031
		(0.75)	(2.49)**
Growth in Surplus*RRD			0.073
			(0.21)
Return on Surplus*RRD			–1.180
			(–3.45)***
Line Mix*RRD			0.077
			(0.57)
NPW to Surplus*RRD			–0.010
			(–0.14)
Growth in Premiums Written*RRD			0.010
			(0.25)
Liquidity*RRD			–0.068
			(–1.24)
Admitted Assets*RRD			–0.035
			(–1.61)
Debts to Assets*RRD			–0.123
			(–0.31)
Adjusted R^2	0.4359	0.4357	0.4399

*** significant at the 1% level
** significant at the 5% level

companies.[11] The time trend is positive and significant only in the first model.

The main observation concerning the second model is the striking consistency with the first model in the signs and the significance of the

explanatory variables except for the time trend, which becomes insignificant with the inclusion of the rating regime dummy variable. On the other hand, model three, which includes interaction terms between the rating regime dummy and the explanatory variables, reveals that in the two-year period subsequent to the change in the rating regime there is shift toward lower ratings and more emphasis on the return on surplus. A Wald test on the rating regime dummy variable and all the interaction terms is significant,[12] which implies that there is a significant change in the relation between the explanatory variables and the Best's ratings at the time Best's introduced the new ratings schedule.

To see how strong the results are with respect to the specification of the model, the model was estimated using OLS. The results from the linear estimates, presented in Table 3, show that the results are qualitatively the same as those found using ordered probit. The joint test for the coefficients of the rating regime dummy and all the interaction terms being equal to zero is statistically significant. Therefore, the results obtained are robust with respect to the estimation procedure used.

SUMMARY AND IMPLICATIONS

In 1991, A.M. Best Company changed the rating procedures and schedules. Controlling for the relations among financial ratios and the ratings, this paper examines the impact of the change in procedures and schedules on the average ratings assigned to insurers. The financial ratios used in this study characterize the insurers' financial and operating performance. In analyzing this change in a ranked categorical variable, this study uses ordered probit. The empirical results show that the A.M. Best Company changed their ratings significantly, increasing the emphasis on profitability and shifting to more stringent standards. A high return on surplus is more likely to be associated with a higher rating under the revised scale and procedures. Consistent results were found using ordinary least squares as an alternative estimation procedure.

The results of this study indicate that the A.M. Best Company enhanced their ratings. This has important implications to consumers, policymakers, and management of insurance companies. A strengthened rating system should increase policyholders' confidence in the rating agency. As the financial world is becoming riskier, insurance consumers will demand a better value for the protection they buy. Reliable ratings will help policyholders distinguish between healthy and unhealthy insurers, reducing information asymmetry between insurance managers and consumers. As the ratings become more reliable, regulators may pay closer

attention to companies that are at the lower end of the rating scale. The results of the study show that after the change in the rating regime, there is more emphasis on profitability and ratings are lower on average. In the future, increased competition among rating agencies may induce additional rating standard changes to more accurately portray insurers' performance.

NOTES

[1]The incentives of consumers to inform themselves about the financial condition of insurance companies may be reduced by the existence of state guaranty funds, which compensate policyholders of an insolvent company. However, the compensation is often delayed and incomplete due to specified state limits.

[2]Similarly, several property and liability insurers assigned a high rating failed in 1984 and 1985.

[3]For a complete description of methods and procedures used by those rating agencies, see Klein (1992).

[4]See Best's *Key Rating Guide Property and Casualty,* 1991 edition, page xiii.

[5]Affiliated companies that received the rating of the parent company and those assigned rating based on the consolidated financial statements of the entire group were excluded from the sample. Also, insurers that received the rating of their reinsurer were left out of the analysis.

[6]See, for example, BarNiv and McDonald (1992) for a survey of the insolvency studies in the property-liability insurance industry. For a more recent study, see Lee and Urrita (1996).

[7]In the ordered probit model, the dependent variable is a latent rating, r^*, and is estimated as a categorical variable, r, as follows: $r = 0$ if $r^* < 0$; $r = 1$ if $0 \le r^* < \mu_1$; $r = 2$ if $\mu_1 \le r^* < \mu_2$; $r = 3$ if $\mu_2 \le r^* < \mu_3$; $r = 4$ if $\mu_3 \le r^*$. The μ_i thresholds and the β_i coefficients are estimated jointly.

[8]The lines of insurance that are considered to have a long payout tail are: general liability, auto liability, workers' compensation, general liability, and medical malpractice.

[9]An alternative estimation procedure, the multinomial probit, treats the choice between alternative levels of the dependent variable as if there were no underlying relationship between the response categories.

[10]For the formula for calculating the marginal impact of explanatory variables, see Greene (1990), pp. 703–706.

[11]The marginal effect of the stock dummy on the different rating categories is: Category A++, A+: –0.05186; A, A–: 0.02127; B++, B+: 0.02365; B, B–: 0.00067; and C++, C+, C, C–, D: 0.00021.

[12]Chi squared with eleven degrees of freedom was 22.3.

REFERENCES

A.M. Best Company (1993) *Key Rating Guide Property-Casualty Insurers.* New Jersey: A.M. Best Company, Inc. 1993 Edition.

A.M. Best Company (1991) *Best's Insolvency Study: Property/Casualty Insurers 1969–1990* New Jersey: A.M. Best Company.

Aldrich, H. John, and Forest D. Nelson (1984) *Linear Probability, Logic and Profit Models.* Newbury Park, California: Sage University, Paper 45.

Altman, E., R. Avery, R. Eisenbeis, and J. Sinkey (1981) "The Application of Statistical Classification Methods to Bond Quality Ratings," in *Application of Classification Techniques in Business, Banking and Finance*. Westport, CT: JAI Press.

Ambrose, Jan Mills, and J. Allen Seward (1988) "Best's Ratings, Financial Ratios and Prior Probabilities in Insolvency Prediction," *Journal of Risk and Insurance*, vol. 55, pp. 229–244.

Ambrose, Jan Mills, and Anne M. Carroll (1994) "Using Best's Ratings in Life Insurer Insolvency Prediction," *Journal of Risk and Insurance*, vol. 61, pp. 317–327.

BarNiv, Ran, and James B. McDonald (1992) "Identifying Financial Distress in the Insurance Industry: A Synthesis of Methodological and Empirical Issues," *Journal of Risk and Insurance*, vol. 60, pp. 543–574.

Brotman, Billie Ann (1989) "Reliability of Best's Insurer Ratings Using Financial Information Published in the Annual Report," *Journal of Insurance Issues and Practices*, vol. XII, no. 1, pp. 58–70.

Ederington, Louis H. (1985) "Classification Models and Bond Ratings," *The Financial Review*, vol. 20, pp. 237–262.

Greene, H. William (1990) *Econometric Analysis.* New York: Macmillan Publishing.

Klein, Robert W. (1992) "Insurance Company Rating Agencies: A Description of their Methods and Procedures," National Association of Insurance Commissioners.

Lee, Suk Hun, and George L. Urrita (1996) "Analysis and Prediction of Insolvency in the Property-Liability Insurance Industry: A Comparison of Logit and Hazard Models," *Journal of Risk and Insurance*, vol. 63, pp. 121–130.

Mayers, David and Clifford Smith (1997) "Organizational Forms Within the Insurance Industry: Theory and Evidence." Paper presented at the American Risk and Insurance Association, San Diego, California.