© The Journal of Risk and Insurance, 1996, Vol. 63, No. 1, 121–130.

Analysis and Prediction of Insolvency in the Property-Liability Insurance Industry: A Comparison of Logit and Hazard Models

Suk Hun Lee
Jorge L. Urrutia

Abstract

This article compares the performance of the logit and hazard models in predicting insolvency and detecting variables that have a statistically significant impact on the solvency of property-liability insurers. The empirical results indicate that the hazard model identifies more significant variables than the logit model and that both models have comparable forecasting accuracy. We conclude that the combined use of both models provides a more complete analysis of the insurance insolvency problem.

Introduction

Insolvency has been a serious problem in the U.S. property-liability insurance industry since the early 1960s (BarNiv and McDonald, 1992). Furthermore, the frequency and severity of insurance failures have increased since the mid-1980s (Harrington, 1992). Among the possible causes of insurers' insolvency are the difficulties in estimating incurred losses and the operation of the insurance firm at undesirable levels of risk (Harrington and Nelson, 1986, and Harrington, 1992). The increase in failures has also led to criticism of the insurance regulatory authorities for insufficient monitoring and inefficient regulatory forbearance (Cummins, Harrington, and Klein, 1995).

Academicians have long been concerned with predicting insurers' insolvency. Numerous prediction methods have been used, including multiple dis-

---

Suk Hun Lee and Jorge L. Urrutia are Associate Professors of Finance at Loyola University of Chicago.

The authors thank two anonymous reviewers and the Editor for their valuable comments that have greatly improved the article. Our special thanks to the late Robert C. Witt for his thoughtful insights and encouragement. Earlier versions of the article have been presented at meetings of the American Risk and Insurance Association, Eastern Finance Association, Southern Finance Association, and Financial Management Association. We are grateful to numerous encouraging comments offered during these presentations. We thank our colleagues George Kaufman and Nicholas Lash for valuable comments. We also thank Scott Harrington, Yehuda Kahane, and Samuel Cox for their useful suggestions and constructive comments. Any remaining errors are the sole responsibility of the authors.

criminant analysis (Trieschmann and Pinches, 1973, and Ambrose and Seward, 1988), regression analysis (Harrington and Nelson, 1986), mean-variance approach (BarNiv and Smith, 1987), portfolio theory (Kahane, Tapiero, and Jacque, 1986), logit and probit models (BarNiv and Hershbarger, 1990, and BarNiv, 1990), and expert systems (Duett and Hershbarger, 1990, and Brockett et al., 1994). BarNiv and Hershbarger (1990) and BarNiv and McDonald (1992) provide excellent surveys of these techniques and their performances. Although most of these models perform relatively well, the general consensus is that more work is needed to improve their reliability and prediction power.

This article compares the performance of the logit model and the hazard model in predicting property-liability insurer insolvencies. Insurance researchers already have employed the logit model (BarNiv and Hershbarger, 1990). The hazard model has been applied to the study of unemployment (Lancaster, 1979) and bank failures (Lane, Looney, and Wansley, 1986). For a previous application to insurer insolvencies, see Kim et al. (1995).

METHODOLOGY

The logit and the hazard model techniques can screen between significant and nonsignificant variables selected by researchers, regulators, or both.

The Logit Model

The logit model is an econometric model that is appropriate when the dependent variable is not a continuous variable but a binary variable that takes a value of one when the insurer becomes insolvent and zero otherwise. The logit model is based on the logit distribution whose cumulative probability distribution is given by the following expression:

$$P_i = F(a + \sum_{j=1}^{m} b_j X_{ij}) = \frac{1}{1 + \exp[-(a + \sum_{j=1}^{m} b_j x_{ij})]}, \tag{1}$$

where P_i is the probability that an event will occur given a set of explanatory variables X_j, and b is a vector of coefficients. The empirical tests require rearranging equation (1) into the following form:

$$z_i = a + \sum_{j=1}^{m} b_j X_{ij} + e_i, \tag{2}$$

where
$$z_i = \ln\left(\frac{P_i}{1 - P_i}\right).$$

The Hazard Model

The survival function in the hazard model, S(t), is defined as the probability that a firm survives longer than t periods:

$$S(t) = P(T > t) = 1 - F(t), \tag{3}$$

where F(t) is the cumulative distribution function of the time to failure. The conditional failure rate is given by the hazard function, h(t):

$$h(t) = \lim_{dt \to 0} \frac{P(t < T < t + dt | T > t)}{dt} = -\frac{S'(t)}{S(t)}. \tag{4}$$

In other words, the hazard function is the probability of failure at a given time period conditional on the fact that the firm has survived to time t.

The dependent variable in the hazard model is the duration of solvency years. By assuming that the duration of solvent years follows the Weibull distribution with parameters p and λ, where p and λ are the shape and scale parameters, respectively, the survival function becomes

$$S(t) = \exp[-(\lambda t)^p], \tag{5}$$

and the hazard function is

$$h(t) = \lambda p t^{p-1}. \tag{6}$$

In order to investigate the relationship between the duration of solvent years of property-liability insurers and a set of explanatory variables, we specify the ith insurer's hazard function as

$$h^i(t) = pt^{p-1} \exp\,(B'X_i), \tag{7}$$

where λ is set equal to one, B is the vector of unknown parameters, and X_i is the vector of the ith insurer's characteristics. For a more detailed description of the hazard model, see, for example, Lancaster (1979) and Kiefer (1988).

DEFINITION OF EXPLANATORY VARIABLES

We postulate that the solvency of a property-liability insurer is a function of several variables. Most of these variables have been examined in previous studies; refer to BarNiv and McDonald (1992) for a detailed justification of their use and the hypothesized direction of their impact on insurers' solvency. A brief description of the variables follows:

Leverage Ratio

NPW/S = ratio of net premiums written to surplus.

Profitability Ratios

OM = operating margin, defined as the ratio of net operating income to premiums earned.

RPS = return to policyholders' surplus.

Liquidity Premium

CL = current liquidity ratio.

Product Mix Variables

W_A = ratio of all auto lines net premiums written to total net premiums written.

W_W = ratio of net premiums written for workers' compensation, medical malpractice, and other liabilities to total net premiums written.

LP = ratio of total loss reserve to total net premiums written.

Asset Mix Variable

I_B = ratio of market value of invested bonds to total admitted assets.

Growth Ratios

GRS = rate of growth of statutory surplus.

GRP = rate of growth of net premiums written.

Other Insurers' Characteristics

MUT = a dummy variable equal to one if the insurer is a mutual company and zero otherwise.

DWR = a dummy variable equal to one if the insurer is a direct writer and zero otherwise.

DATA DESCRIPTION

The sample consists of property-liability insurers that became insolvent during the period from January 1980 through June 1991 for which necessary data were available. Insolvent insurers are those declared insolvent by their respective state insurance commissioners and reported in the section on Insurers Retirements and Name Changes of *Best's Reports: Property-Casualty Edition* (A. M. Best).[1] Complete data were available for 82 insolvent insurers, of which 72 were stock companies and 10 were mutual firms. Data for 82 matching solvent insurers were also collected for the same time period. In order to qualify as a solvent firm, an insurer must have operated normally and not been declared or reported insolvent during the entire January 1980 through June 1991 time period. Matching solvent insurers were selected based on state domicile and magnitude of total admitted assets.[2]

---

[1] Our definition of insolvency includes the following: insolvent firms, firms placed in rehabilitation, receivership, under supervision, conservatorship, and dissolved firms. The date of insolvency corresponds to the "declaration of insolvency" reported in *Best's Reports*.

[2] We matched the total admitted assets of solvent and insolvent insurers in the year previous to the last year for which data for the insolvent firm were available in *Best's Reports*. For example, if firm A was declared insolvent in year 1989 and the last year of data for A in *Best's Reports* was 1988, we matched firm A with a solvent firm domiciled in the same state and having admitted assets equal to firm A in the year 1987.

The sampling procedure employed in this study is a choice-based sample and not a random sample. It is known that a choice-based sample leads to oversampling the insolvent firms (see Palepu, 1986, and BarNiv and McDonald, 1992). Oversampling of the insolvent firms without appropriate modification leads to inconsistent and biased estimates of the parameters.[3] We have corrected the oversampling problem by making the appropriate adjustments in the estimation of the parameters following the procedures described in Palepu (1986) and BarNiv (1990).

EMPIRICAL RESULTS

The samples of solvent and insolvent insurers are statistically different in terms of leverage, profitability, liquidity, product mix, asset mix, rate of growth of surplus, and type of distribution system (see Table 1).

The multivariate analyses presented below compare the forecasting accuracy of the logit and hazard models and their ability to identify relevant explanatory variables.

Identification of Relevant Explanatory Variables

In order to identify the explanatory variables that have a statistically significant impact on the solvency of property-liability insurers, we use the entire data set, January 1980 through June 1991. In the logit model, four variables are statistically significant at the 5 percent level or better: ratio of net premiums written to surplus (NPW/S), return to policyholders' surplus (RPS), proportion of premiums written in long-tailed lines (W_W), and market value of invested bonds as a proportion of total admitted assets (I_B). The hazard model yields eight variables that are statistically significant at the 5 percent level or better: the same four as the logit model plus operating margin (OM), current liquidity ratio (CL), growth rate of statutory surplus (GRS), and growth rate of the net premiums written (GRP).

Both models are useful tools in detecting variables that have a statistically significant impact on the solvency of property-liability insurers. However, the hazard model detects more relevant variables than the logit model.[4] The hazard model also provides an estimate of the hazard rate. A hazard rate of 2.71 suggests that the probability of an insurer's insolvency increases with the length of the survival period. According to the hazard model, a hazard rate larger than one implies an increasing hazard function.

Forecasting Accuracy of the Logit and Hazard Models

The analysis of forecasting accuracy uses only the statistically significant explanatory variables: four variables for the logit model and eight variables for the hazard

---

[3] We thank an anonymous referee for bringing this important statistical problem to our attention.

[4] We select the set of explanatory variables listed in Table 2 based on results reported in previous studies. The results reported here hold for this specific set of independent variables. Results may differ for a different set of explanatory variables.

model. We divide the sample into an estimation period and an out-of-sample forecasting period. In the estimation period, we estimate the parameters of the logit and hazard models. We use data for the 44 insurers that were declared insolvent during the period 1985 through 1988 and 44 corresponding solvent firms. The estimation results are also presented in Table 2.

The out-of-sample forecast classifies the firms as solvent or insolvent by using data for the 32 firms that were declared insolvent during the period 1989

Table 1
Summary Statistics: Means, Standard Deviations, and the Nonparametric Wilcoxon Ranked Sign Test for Solvent and Insolvent Insurers, January 1980 through June 1991

Variable	Solvent Insurers Mean	Solvent Insurers Standard Deviation	Insolvent Insurers Mean	Insolvent Insurers Standard Deviation	Wilcoxon Ranked z-score
Leverage					
NPW/S	1.40	0.767	2.92	8.675	-5.129**
Profitability					
OM	7.04	13.611	-0.56	66.898	5.830**
RPS	11.74	20.049	-11.62	50.734	6.895**
Liquidity					
CL	161.14	82.715	132.11	159.253	11.464**
Product Mix					
W_A	42.78	32.778	32.41	34.812	3.497**
W_W	22.01	28.856	32.81	34.568	-3.992**
LP	137.65	97.756	117.49	320.315	1.354
Asset Mix					
I_B	63.12	19.311	47.71	22.026	8.126**
Rate of Growth					
GRS	12.21	25.585	-13.93	92.316	3.606**
GRP	8.91	33.754	19.86	126.903	-1.115
Other Insurers' Characteristics					
MUT	0.12	0.706	0.08	0.274	0.168
DWR	0.17	0.379	0.07	0.262	3.520**
Mean of Admitted Assets	$81.92m	110.12	$76.12	115.19	0.330

Note: NPW/S = ratio of net premiums written to surplus. OM = operating margin. RPS = return to policyholders' surplus. CL = current liquidity. W_A = ratio of all auto lines net premiums written to total net premiums written. W_W = ratio of net premiums written for workers' compensation, medical malpractice, and other liabilities to total net premiums written. LP = ratio of total loss reserve to total premiums written. I_B = ratio of market value of invested bonds to total admitted assets. GRS = rate of growth of statutory surplus. GRP = rate of growth of net premiums written. MUT = a dummy variable equal to one if the insurer is a mutual company and zero otherwise. DWR = a dummy variable equal to one if the insurer is a direct writer and zero otherwise.

** Significant at the 1 percent level.

Table 2
Identification of Relevant Explanatory Variables for Predicting Insurer Insolvency

	Logit Model				Hazard Model			
Variable	1980-1991		1985-1988		1980-1991		1985-1988	
Constant	-1.951	(1.051)	-2.176	(0.797)	1.747	(0.046)	1.549	(0.045)
NPW/S	0.507	(0.190)**	0.456	(0.245)	-0.008	(0.003)*	-0.00001	(0.003)
OM	0.0002	(0.022)			0.002	(0.0003)**	0.00003	(0.0004)
RPS	-0.022	(0.011)*	-0.025	(0.01)**	0.003	(0.0002)**	0.002	(0.0003)**
CL	-0.002	(0.005)			0.001	(0.0001)**	0.0004	(0.0002)*
W_A	0.002	(0.006)			0.0008	(0.0005)		
W_W	0.023	(0.009)**	0.017	(0.008)*	-0.003	(0.0005)**	-0.002	(0.0005)**
LP	0.0000	(0.0002)			-0.00005	(0.00006)		
I_B	-0.036	(0.007)**	-0.033	(0.01)**	0.006	(0.0007)**	0.006	(0.0009)**
GRS	-0.003	(0.007)			0.0006	(0.0001)**	0.002	(0.0003)**
GRP	-0.002	(0.006)			0.0004	(0.0002)*	0.00004	(0.0002)
MUT	-0.112	(0.109)			0.167	(0.093)		
DWR	-0.123	(0.456)			-0.157	(0.095)		
Hazard Rate					2.713	(0.105)**	3.156	(0.172)**
Chi-squared	95.046		40.651					
Significance Level	0.1E-06		0.31E-07					

Note: The dependent variable in the logit model is a binary-valued variable which takes a value of one if an insurer was declared insolvent and zero otherwise. The dependent variable for the hazard model is the duration of solvency years. Figures in parentheses represent standard errors. For variable definitions, see Table 1.

* Significant at the 5 percent level. ** Significant at the 1 percent level.

through 1991 and 32 corresponding solvent firms. The classification results for the hazard and logit models are shown in Table 3.[5] Panel A reports classification results for cutoff values ranging from 5 to 90 percent. Taking into account both the one-year- and two-year-ahead forecast errors, the minimum error rates are obtained at the 30 percent cutoff value for the hazard model and at the 15 through 25 percent cutoff values for the logit model. At a cutoff value of 50 percent, the hazard model misclassifies 36 percent and 27 percent of the insurers in the one-year- and two-year-ahead forecasts, respectively, while the logit model misclassifies 28 percent and 27 percent of the firms in the one-year- and two-year-ahead forecasts, respectively. Thus, the logit model performs slightly better than the hazard model in the one-year-ahead forecast.

Panel B reports the expected cost of misclassification for several values of misclassification costs for Type I error, C_1, holding the misclassification cost for Type II error, C_2, equal to one. There is no substantial difference in expected cost of misclassification for the logit and hazard models for low values of C_1 (1 to 25). However, for high values of C_1, the misclassification costs are higher for the logit

[5] BarNiv (1990) and BarNiv and McDonald (1992) provide a detailed discussion on the classification and prediction based on different criteria, such as minimizing the number of misclassifications and the expected cost of misclassification.

model. Overall, the results indicate that the two models have similar forecasting accuracy.

Table 3
Comparison of Insolvency Forecasting Accuracy of the Hazard and Logit Models: Out-of-Sample Data, 1989 Through 1991

Panel A: Misclassifications for Different Cutoff Values

	Hazard Model				Logit Model			
	One-Year Forecast		Two-Year Forecast		One-Year Forecast		Two-Year Forecast	
Cutoff Value (Percent)	N	%	N	%	N	%	N	%
5	13	20	28	44	16	25	15	23
15	15	23	24	38	14	22	15	23
20	16	25	20	31	14	22	15	23
25	17	27	19	30	14	22	15	23
30	18	28	17	27	16	25	16	25
35	19	30	17	27	16	25	16	25
40	21	33	15	23	16	25	16	25
50	23	36	17	27	18	28	17	27
60	23	36	19	30	18	28	17	27
70	25	39	19	30	18	28	17	27
90	26	41	25	39	17	27	18	28

Panel B: Expected Cost of Misclassification for Various Misclassification Costs of Type I Error, C_1, Holding the Misclassification Cost for Type II Error, C_2, Equal to One

	Hazard Model		Logit Model	
C_1	One-Year Forecast	Two-Year Forecast	One-Year Forecast	Two-Year Forecast
1	0.033	0.033	0.025	0.030
5	0.164	0.162	0.125	0.147
25	0.406	0.549	0.480	0.558
50	0.484	0.829	0.870	1.027
75	0.563	0.918	1.261	1.495
100	0.642	0.957	1.652	1.964

CONCLUSION

This article compares the performance of the logit and hazard models in predicting insolvency and identifying variables that have a significant impact on the solvency of property-liability insurers.

The logit model detects four variables that have statistically significant impacts on the probability of insolvency of a property-liability insurer: ratio of net premiums written to surplus, return to policyholders' surplus, proportion of premiums written in long-tailed lines, and market value of invested bonds as a proportion of total admitted assets. On the other hand, the hazard model detects eight statistically significant variables: the four identified by the logit model plus operating margin, current liquidity ratio, rate of growth of surplus, and rate of growth of premiums written.

The tests of forecasting accuracy indicate that the logit model has a somewhat lower misclassification rate than the hazard model, but the logit model also has somewhat higher misclassification costs for high costs of Type I errors, C_1. Thus, the logit and hazard models seem to have comparable forecasting accuracy.

Since both models provide important information about the insolvency of property-liability insurers, the combined use of the logit and hazard models allows for a more thorough analysis of the incidence of insolvency in the insurance industry.

REFERENCES

A. M. Best Company, various years, *Best's Reports: Property-Casualty Edition* (Oldwick, N.J.: A. M. Best).

A. M. Best Company, various years, *Best's Trend Report: Property-Casualty Edition* (Oldwick, N.J.: A. M. Best).

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

BarNiv, Ran, 1990, Accounting Procedures, Market Data, Cash-flow Figures and Insolvency Classification: The Case of the Insurance Industry, *Accounting Review,* 65: 578–604.

BarNiv, Ran and Robert Hershbarger, 1990, Classifying Financial Distress in the Life Insurance Industry, *Journal of Risk and Insurance,* 57: 110–136.

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,* 59: 543–573.

BarNiv, Ran and Michael L. Smith, 1987, Underwriting Investment and Solvency, *Journal of Insurance Regulation,* 5: 409–428.

Brockett, Patrick L., William W. Cooper, Linda L. Golden, and Utai Pitaktong, 1994, A Neural Network Method for Obtaining an Early Warning of Insurer Insolvency, *Journal of Risk and Insurance,* 61: 402–424.

Browne, Mark J. and Robert E. Hoyt, 1995, Economic and Market Predictors of Insolvencies in the Property-Liability Insurance Industry, *Journal of Risk and Insurance,* 62: 309–327.

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

Duett, Edwin H. and Robert A. Hershbarger, 1990, Identifying Financial Distress in the Property-Casualty Industry, *Journal of the Society of Insurance Research,* 21: 33–45.

Harrington, Scott E., 1992, The Solvency of the Insurance Industry, Paper presented to the Conference on Bank Structure and Competition, Chicago, May.

Harrington, Scott E. and Jack M. Nelson, 1986, A Regression-Based Methodology for Solvency Surveillance in the Property-Liability Insurance Industry, *Journal of Risk and Insurance,* 53: 583–605.

Kahane, Yehuda, Charles S. Tapiero, and Laurent Jacque, 1986, Concepts and Trends in the Study of Insurer's Solvency, Paper presented to the International Conference on Insurance Solvency, Wharton School, Philadelphia.

Kiefer, Nicholas M., 1988, Economic Duration Data and Hazard Functions, *Journal of Economic Literature,* 26: 646–679.

Kim, Yong Duck, Dan R. Anderson, Terry L. Amburgey, and James C. Hickman, 1995, The Use of Event History Analysis to Examine Insurer Insolvencies, *Journal of Risk and Insurance,* 62: 94–110.

Lancaster, Tony, 1979, Econometric Methods for the Duration of Unemployment, *Econometrica,* 47(4): 939–956.

Lane, Williams R., Stephen W. Looney, and James W. Wansley, 1986, An Application of the Cox Proportional Hazards Model to Bank Failure, *Journal of Banking and Finance,* 10: 511–531.

Palepu, Krishna G., 1986, Predicting Takeover Targets: A Methodological and Empirical Analysis, *Journal of Accounting and Economics,* 8: 3–35.

Trieschmann, James J. and George E. Pinches, 1973, A Multivariate Model for Predicting Financially Distressed P-L Insurers, *Journal of Risk and Insurance,* 40: 327–338.