보험개발연구 제18권 제1호 2007년 3월

Identification of Insurer Insolvencies Using the Cox Proportional Hazard Model

Cox proportional hazard model을 이용한 보험사 파산 예측

Lee Suk-Ho*

이 석 호

Using another useful statistical model for failure prediction, which is the Cox proportional hazard model, this study attempts to verify and confirm the results of Lee(2005). The results of this study support most of the findings of Lee(2005). In other words, under the frame of the Cox proportional hazard model like the logit model, efficiency measures has tuned out to be important factors in identifying and forecasting insurer insolvencies again. Also, the results of the study support another finding of Lee(2005) that the efficiency variable sets add significant explanatory power to the financial ratio variable sets. Meanwhile, this study also finds that overall, the Cox proportional hazard model has comparable ability to the logit model in identifying and forecasting insurer insolvencies. In the sense that both logit and Cox proportional hazard model convey important information regarding insolvency of property-liability insurers, the combined use of both statistical models in identifying and predicting insolvency of insurers would be desirable. And its performance should be improved with the inclusion of efficiency measures as explanatory factors into the model.

※ Key Words: cox proportional hazard model, efficiency, insurer insolvency, logit model

* 한국금융연구원 연구위원(slee@kif.re.kr)

Ⅰ. Introduction

Insurer insolvency has been one of the major hot issues in insurance industry. Numerous empirical studies have attempted to identify and predict insolvencies of insurers since 1970s. Most of studies used financial ratios such as surplus ratio, expense ratio, loss ratio in identifying insolvencies of insurers. For example, Harrington and Nelson(1986) analyzed the relationship between the premium-to-surplus ratio and insurance company characteristics to identify insolvencies of property-liability insurers. Hershbarger and Miller(1986) examined the effectiveness of the eleven NAIC-IRIS ratios[1)] as a classifying tool for insolvencies of property-liability insurers. Pottier(1998) also used Best's ratings and financial ratio variables (including IRIS ratios) to identify insolvent life insurers. He set up and estimated three different logit models for 42 insolvent life insurers and around 1100 solvent life insurers during the sample period 1990 to 1992.

Meanwhile, few studies used efficiency measures even though almost all causes of insurer insolvencies identified involve some form of operational or managerial inefficiency. Recently, Lee(2005) attempted to employ efficiency measures in identifying and predicting insolvencies of the U.S. property-liability insurers. He used a logit model in identifying insolvencies of the property-liability insurers. He finds that the efficiency measures, especially cost X-efficiency and revenue scale efficiency, are important factors in identifying and forecasting insurer insolvencies in the property-liability

1) The eleven NAIC(National Association of Insurance Commissioners)-IRIS(Insurance Regulatory Information System) ratios are premiums to surplus, change in writings, surplus aid to surplus, two-year adjusted underwriting ratio, investment yield, change in surplus, liabilities to liquid assets, agents balances to surplus, one-year reserve development to surplus, two-year reserve development to surplus, and estimated current reserve deficiency to surplus.

insurance industry. And the study also provided evidence that the efficiency variable sets added significant explanatory or discriminatory power to the financial ratio variable sets.

In order to verify and confirm the findings of Lee(2005), i.e., whether efficiency measures are important factors in identifying and forecasting insurer insolvencies and the efficiency variable sets add significant explanatory or discriminatory power to the financial ratio variable sets, and compare and extend the Lee(2005)'s results, this study employs another useful statistical model for failure prediction, which is the Cox proportional hazard model.

This study also examines the performance of the Cox proportional hazard model itself for insolvency prediction in the insurance industry. While no study has attempted to use the Cox proportional hazard model for failure prediction in the insurance industry, it has been widely used in other industries such as biomedical, engineering, and social sciences. For the cases of insolvency prediction studies in the financial service sectors, there exist two studies that employ this model. Lane, Looney and Wansley(1986) and Whalen(1991) used the Cox proportional hazard model to predict bank failures. The model has been successfully verified as an efficient classification and prediction tool for bank failures. According to Lane et al.(1986), there are two major advantages of using this model. First, since it requires few underlying assumptions, it is less vulnerable to some methodological criticisms. Second, the model is able to incorporate explicitly the time to failure into the modeling effort, which cannot be done in the other traditional classification models such as MDA and the logit model.

Meanwhile, this study, with using more extended data set and adding another set of independent variables(i.e., efficiency variables), also could extend the results of Lee and Urrutia(1996)[2]. They compared the abilities of two statistical models - a logit and hazard model - in identifying and predicting insolvencies of property-liability insurers. The hazard model they used is somewhat different from the Cox proportional hazard model. They employed a form of hazard model with the assumption that the duration of solvency years(the dependent variable) follows the Weibull distribution. The logit model used in their study is similar to that used in previous insolvency studies. The sample used in their study consisted of 82 insolvent and 82 matching solvent insurers from the period 1980 to 1991. They find that the hazard model identifies more significant variables than the logit model and that both models have comparable forecasting accuracy.

This study employs the same independent variables used in Lee(2005), which are four efficiency variables and 39 financial ratios. And it uses the same data set and follows almost the same methodological approaches as Lee(2005). The next section explains descriptions of the methodologies used in this study. Section Ⅲ presents the main results of the study and is followed by the conclusions in section Ⅳ.

Ⅱ. Methodology

The Cox proportional hazard model is used to identify and predict insolvencies of property-liability insurers in this study. For the Cox proportional hazard model, three versions of the model are estimated for

2) Based on previous insurer insolvency studies, they used the following twelve financial variables: net premiums written to surplus, net operating income to premiums earned, return to policyholders' surplus, current liquidity ratio, three product mix variables, market value of invested bonds to total admitted assets, rate of growth of statutory surplus, rate of growth of net premiums written, a mutual dummy variable, and a direct writer dummy variable.

both one and two year-prior data as in Lee(2005). Once the significant predictor variables are identified, classification tests are conducted considering relative misclassification costs. Then, the results of the study are compared with those of Lee(2005) and Lee and Urrutia(1996).

1. Model

Cox(1972) developed a proportional hazard model that has been used widely in biomedical and engineering studies. The basic underlying assumption under the proportional hazard model is that the hazard rate is proportional to the effect of each explanatory variable. In other words, the effect of each explanatory variable is found by multiplying the baseline hazard by some function of the explanatory variable vector, which does not depend on time t(see Cox and Oakes, 1984; Lane et al., 1986).

The main advantages of the Cox proportional hazard model are that it explicitly incorporates the time to failure into the model and it requires no distributional assumption in estimating either the coefficients of the explanatory variables or baseline hazard function. Moreover, it is easy to estimate using common statistical packages. Meanwhile, one should consider two possible problems in employing the Cox proportional hazard model. First, there might exist many ties among survival times for each observation(especially for the censored survival time for nonfailures), which need to be corrected. Second, an assumption of time-constant explanatory variables has to be dealt with. More details will be explained in the later part of this section.

Hazard models, including the Cox proportional hazard model, have been successfully verified as an efficient classification and prediction tool for failure studies in the financial area(e.g., Kim et al., 1995; Lane et al., 1986; Whalen, 1991). The robustness of the model also has been proved in similar studies of other industries, in labor economics, and the social sciences(e.g., Chen and Lee, 1993; Kiefer, 1988; Ng, Cram and Jenkins, 1991).

The proportional hazard model takes the following general form:

$$h(t|z) = \Phi(z)h_0(t) \tag{1}$$

where $h(t|z)$ is a hazard function at time t for an observation conditional on explanatory variable vector z, $\Phi(z)$ is some function of z such that $\Phi(0) = 1$, and $h_0(t)$ is the baseline hazard with $z=0$.

In the Cox proportional hazard model, $\Phi(z)$ is set equal to the exponential function, $exp(\beta' z)$ which simplifies estimation of the coefficients β(Cox and Oakes, 1984). Thus, the general form of the Cox proportional hazard model is as follows:

$$h(t|z) = exp(\beta' z)h_0(t) \tag{2}$$

The Cox proportional hazard model is considered semiparametric in that $exp(\beta' z)$ is parametric and the baseline hazard $h_0(t)$ is nonparametric. And the survivor function that calculates the probability that a particular observation will survive for t periods to failure, can be expressed in the following form:

$$S(t|X) = S_0(t)^{exp(X'\beta)} \tag{3}$$

where, $S_0(t)$ is the baseline survivor function.

As in Lee(2005), three versions of the Cox proportional hazard model are estimated. The Cox proportional hazard models are estimated for both one and two-year prior data. First, the following Cox proportional hazard model with the efficiency variables only is estimated:

$$h_i(t|X)=h_0(t)exp(\beta_1 EFF_{i1}+\cdots+\beta_k EFF_{ik}) \tag{4}$$

The second Cox proportional hazard model includes the financial ratio variables only:

$$h_i(t|X)=h_0(t)exp(\gamma_1 FIN_{i1}+\cdots+\gamma_j FIN_{ij}) \tag{5}$$

Finally, the two sets of variables are combined:

$$h_i(t|X)=h_0(t)exp(\beta_1 EFF_{i1}+\cdots+\beta_k EFF_{ik}+\gamma_1 FIN_{i1}+\cdots+\gamma_j FIN_{ij}) \tag{6}$$

The dependent variable in a Cox proportional hazard model is time to failure(or survival time) of an observation. As in Lane et al.(1986), the time to failure for an insolvent insurer is defined to be the time(in months) from December 31 of the year that the annual financial data are reported to the NAIC(either one or two years prior to failure for this study) until the date of failure. For a solvent insurer, the censored survival time is defined to be the time(in months) from the above reporting date until December 31 of the year of its year-matched insolvent insurer. For the sample observations whose time to failure is unknown until the end of the study, the survival time is censored. In other words, contrary to a failed insurer, because the time to failure or survival time for a nonfailed insurer is unavailable within the scope of the study period(i.e., the sample period), it has to be measured based on a predetermined date. Meanwhile, the solvent insurers are randomly selected(without replacement).

For the Cox proportional hazard model, the coefficient of each variable is estimated using a partial likelihood method. Then the partial likelihood is maximized with respect to the parameters in β or γ. The remarkable thing about the partial likelihood is that the coefficients can be estimated without having to specify the baseline hazard function $h_0(t)$(see Allison, 1995). In other words, the maximum partial likelihood solution of β or γ does not require knowledge of the baseline hazard function. Let $t_1 < t_2 < \cdots < t_N$ be the ordered survival times and let R_j be the set of insurers that survive up to time t. Then the partial likelihood function is defined as follows:

$$L(\beta) = \prod_{j=1}^{N} \frac{exp(X'_{j}\beta)}{\sum_{l \in R_j} exp(X'_{l}\beta)} \tag{7}$$

The partial likelihood estimates are consistent and asymptotically normal even though they are not "ordinary" likelihood estimates(see Allison, 1995; Lawless, 1982).

Two problems must be considered in using the Cox proportional hazard model. First, tied scores or data for survival time have to be dealt with. The partial likelihood function(see Equation (7)) is valid only for data in which no two events occur at the same time(see Allison, 1995). In other words, there exist ties among the observed failure or survival time for each observation(especially for the censored survival time for nonfailures). As in Lane et al.(1986), when tied event(failure) times are present in the study, they are corrected for using the following approximate partial likelihood method, the so-called Breslow's method(Breslow, 1974):

$$L(\beta) = \prod_{j=1}^{N} \frac{exp[(\sum_{q \in D_j} X'_q)\beta]}{[\sum_{l \in R_j} exp(X'_l \beta)]^{d_j}} \quad , \tag{8}$$

where j is the N distinct survival times, X_j is a vector of independent variables for an insurer who actually experienced an insolvency at time t_j , d_j is the number of insolvencies occurring at time t_j, and D_j is the set of d_j observations failing at t_j. For the tied data, the regression coefficients, β are estimated by the value of $\tilde{\beta}$ which maximizes the above equation.

Second, the Cox proportional hazard model assumes that the effects of the different variables on hazard or survival are constant over time. In other words, it assumes that the hazard for one firm is a fixed proportion of the hazard for any other firm in the study. However, according to Allison (1995), the violation of the assumption creates no real problem for the partial likelihood estimation method. In other words, even if the assumption is not valid for some data, the Cox proportional hazard model still can provide efficient and useful results(see Allison, 1995; Lane et al., 1986; Whalen, 1991).

2. Independent Variables

As in Lee(2005), two sets of the independent variables are used in identifying and predicting insolvencies of property-liability insurers: (1) Financial ratio variables and (2) Efficiency variables. This study considers and examines almost all financial ratio variables that appeared to be statistically significant in previous insolvency prediction studies. Based on previous studies, initially 39 financial ratio variables were considered for explanatory variables[3]. However, they were highly correlated with one another, resulting in a severe multicollinearity problem. To eliminate this problem, factor analysis was conducted to summarize the explanatory content of the financial ratio variables in fewer, uncorrelated variables. Cummins, Grace and Phillips(1999) also conducted a factor analysis for the FAST ratios with the same rationale. By applying a factor analysis, all 39 variables proven to be significant in previous studies are considered and consolidated into one study, rather than just using the best significant ones among them as a final set of variables. Financial ratio factors with eigenvalues greater than one are used as the independent variables[4]. Six financial ratio factors are used in the study.

The rationale for including measures of efficiency is that more efficient firms should be less likely to fail. For the efficiency variables, four types of efficiencies are estimated using SFA(Stochastic Frontier Analysis), which is an econometric approach. In SFA, the log of the cost function can be expressed as the following general form:

$$ln(C_i) = ln(Y_i,\ W_j) + \varepsilon_i,\ where\ \varepsilon_i = v_i + \nu_i. \tag{9}$$

3) 39 financial predictor variables are categorized into the following nine groups: (1) Asset management; (2) Dishonest or fraudulent management; (3) Riskiness by lines-of-business; (4) Leverage, liquidity, and capital adequacy; (5) Loss reserve development and adequacy; (6) Profitability, underwriting results, and operational efficiency; (7) Reinsurance activities; (8) Underwriting and sales practices; and (9) Others.

4) The criteria of retaining factors with eigenvalues greater than one is most popularly used. The rationale is that each of those factors obtained using the criteria accounts for at least as much variances as one of the original variables (see Cummins, Grace and Phillips, 1999 for more details).

C_i is cost of firm i, Y_i is a vector of output quantities, and W_j is a vector of input prices. The error term ε_i is decomposed into v_i and ν_i, where v_i reflects random errors or shocks and ν_i is deemed to be X-inefficiency. The random error term, v_i is assumed to follow a symmetric or normal distribution; the inefficiency term, ν_i is assumed to follow an asymmetric or half-normal distribution(see Bauer, Berger, Ferrier and Humphrey, 1998)[5].

Cost efficiency ranges from zero to one: if firm i is fully cost efficient, its efficiency score is equal to one. The measure of cost efficiency is obtained by estimating the ratio of minimum attainable cost, given output level, input prices, and existing technology, to actual observed cost.

The following translog functional form is used and estimated econometrically to obtain cost X-efficiencies:

$$ln(C_i) = \beta_0 + \sum_{i=1}^{n} \alpha_i ln(y_i) + \sum_{j=1}^{m} \beta_j ln(w_j) + \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{n} \alpha_{ik} ln(y_i) ln(y_k) \qquad (10)$$

$$+ \frac{1}{2} \sum_{i=1}^{m} \sum_{h=1}^{m} \beta_{ih} ln(w_i) ln(w_h) + \sum_{i=1}^{n} \sum_{j=1}^{m} \gamma_{ij} ln(y_i) ln(w_j) + \varepsilon_i$$

Each coefficient parameter is estimated using MLE(Maximum Likelihood Estimation). The coefficient parameters obtained from the above function are used to estimate cost scale efficiency. Scale economies measure the relative change in a firm's cost for a given proportional change in output level. So, cost scale efficiency is estimated at its respective output level as

5) Different distributional assumptions regarding the composed error-term(for example, normal-gamma distribution rather than normal-half normal distribution) under stochastic frontier analysis(SFA) may lead to different results.

follows:

$$Cost\,Scale\,Efficiency_i = \sum_{i=1}^{n} \frac{\partial ln(C)}{\partial ln(y_i)} \tag{11}$$

$$= \sum_{i=1}^{n} \left[\alpha_i + \frac{1}{2} \sum_{k=1}^{n} \alpha_{ik} ln(y_k) + \sum_{j=1}^{m} \gamma_{ij} ln(W_j) \right]$$

If firm i's cost scale efficiency is equal to one, it operates at constant returns to scale, and it is cost scale efficient. When cost scale efficiency is less than one, increasing returns to scale exists, and decreasing returns to scale exists if it is greater than one. For revenue efficiency, similar approach and method in obtaining cost efficiency above are employed[6].

Input quantities and prices, and output quantities and prices are required to estimate efficiency. This study adopts similar input and output estimation as the existing literature; therefore, the discussion is brief(see Berger et al., 1997; Cummins, 1999; Cummins and Weiss, 2000; Cummins, Weiss and Zi, 1999 for more details). Meanwhile, costs are the sum of total expenses(net of loss adjustment expenses) and the cost of equity capital. Revenues are equal to premiums earned minus the present value of losses plus investment income on premiums and equity capital.

Among the three prevailing approaches to measure outputs of financial institutions, this study employs a (modified)value-added approach[7]. Under the value-added approach, any category having significant value added to insurance operations is considered an important output(see Berger et al., 1997; Cummins and Weiss, 2000; Cummins, Weiss and Zi, 1999). The

6) For more details, see Lee(2005).

7) The other two approaches, suggested by Berger and Humphrey(1992), are the asset or intermediation approach and the user-cost approach.

modified value-added approach, first suggested by Berger et al.(1997), classifies insurer's output into three categories: (1) Risk pooling and risk bearing; (2) Intermediation; and (3) Real financial services relating to insured losses[8].

Since insurer outputs are mostly intangible, and it's not possible to directly measure them, proxy variables are used. Losses incurred should be closely correlated with categories (1) and (3). In that sense, losses incurred are used as output proxies for both risk pooling and risk bearing and real financial services relating to insured losses. Because risks or services vary by line of business, losses incurred are distinguished by four lines: (1) short-tail personal lines; (2) short-tail commercial lines; (3) long-tail personal lines; and (4) long-tail commercial lines. Because losses incurred are reported at undiscounted values, the losses incurred are discounted using estimated industry-wide payout patterns. For intermediation, invested assets are used as an output proxy.

As another output, net reinsurance brokerage is used. Some sample insurers ceded large amounts of reinsurance relative to assumed premiums. Thus these insurers reported negative commissions(i.e., a negative of an expense item). In these cases, reinsurance commissions are treated as an output. This treatment is appropriate under the value-added approach for measuring inputs and outputs, because these expenses resemble an output rather than an input. All output measures are deflated by the Consumer Price Index(CPI) to make results comparable across years. The methods to estimate output prices are same as the existing literature.

8) Under the modified value-added approach, those activities having significant value-added, as judged using operating cost allocations, are counted as important outputs(see Berger et al., 1997; Cummins and Weiss, 2000; Cummins, Weiss and Zi, 1999 for more details).

Insurer inputs can be categorized into four groups: agent services, (non-agent)labor services, materials, and financial equity capital(see Cummins and Weiss, 2000; Cummins, Weiss and Zi, 1999). For both agent and non-agent labor services, the quantity of labor is estimated with current dollar labor costs divided by a salary deflator. Then, the price of labor can be measured as the current wage rate divided by the Consumer Price Index(CPI). In a similar way, the quantity and price of materials(using a business service deflator) can be computed.

The quantity of equity capital can be defined as statutory policyholders' surplus, excess of statutory over statement reserves, and unauthorized reinsurance, deflated using the CPI. The cost of equity capital is measured as the ratio of the expected net income to the average value of equity capital, i.e., the value of expected return on equity(ROE).

Concerning classification of insurer insolvency, as in Lee(2005), based on the estimates from the estimation sample, holdout sample insurers are classified using a preset cutoff point[9]. Classification accuracy across the different model specifications is compared. Meanwhile, misclassification costs, i.e., the ECM(Expected Costs of Misclassification) should be considered and compared[10].

3. Sample

The number of insolvent insurers is 83 for the one-year prior data and 93

9) The cutoff point 0.5 is used for this study. The value is known to be appropriate for qualitative models such as logit model as well as it was used in many previous insolvency studies(see BarNiv and McDonald, 1992; Brockett, et al., 1994; Eck, 1982; Lee, 2005).

10) For details of the ECM, see Lee(2005).

for the two-year prior data[11]. Certain insurers that have unusual data or missing values for some key variables(e.g., insurers with negative or zero amounts for assets or premiums) are eliminated[12]. To avoid the oversampling problem, this study adopts the random sampling approach. All property-liability insurers reporting to the NAIC from 1991 to 2002 are used in the study. Some insurers were deleted due to missing or unusual values for some variables(e.g., negative premiums). First, similar to that used in BarNiv and Hathorn(1997), a "proportionate" random sample is constructed by randomly selecting solvent insurers, without replacement so that their yearly proportion matches that of the insolvent insurers. As another way of randomly selecting solvent insurers, the same number of solvent insurers(300) for each year is randomly selected, yielding the "non-proportionate" random sample[13].

Ⅲ. Results

In order to verify and confirm the findings of Lee(2005), i.e., whether efficiency measures are important factors in identifying and forecasting insurer insolvencies and the efficiency variable sets add significant explanatory or discriminatory power to the financial ratio variable sets,

11) Insolvent insurers are from A. M. Best Company(1991, 1996, 2001, 2002 and 2003)

12) The definition of insolvency for this study is similar as the previous studies (e.g., see Cummins, Grace and Phillips, 1999; Lee and Urrutia, 1996; Lee 2005). In this study, "insolvent" insurers include insurers placed in rehabilitation, receivership, conservatorship, and insurers under supervision, and (involuntarily)dissolved insurers.

13) Cummins, Grace and Phillips(1999) also use the same 300(randomly selected) solvent insurers for each base year.

regression analysis using another useful statistical model for failure prediction, which is Cox proportional hazard model, is conducted. And the performances of two statistical models – a logit and Cox proportional hazard model – in identifying and predicting insolvencies of property-liability insurers are compared.

First, the proportionate random sample set(the sample period from 1991 to 2002) is analyzed. 〈Table 1〉 shows the regression results based on one-year prior data for the sample set. Most of the results are very similar with those of Lee(2005) as expected. For the Cox proportional hazard model like the logit model in Lee(2005), the coefficients for the cost X-efficiency and revenue scale efficiency variable are negative and significant at the 1 percent level when the model includes the efficiency variables only. And the coefficients for the cost X-efficiency and revenue scale efficiency variable remain negative and statistically significant for the Cox proportional hazard model when the efficiency variables are combined with the financial ratio factors. Moreover, while revenue X-efficiency is insignificant in Lee(2005), i.e., logit model, it comes out to be significant variable in the Cox proportional hazard model when the efficiency variables are combined with the financial ratio factors.

These results indicate that using an another useful prediction model(the Cox proportional hazard model), efficiency measures are verified and confirmed to be not only significant predictors themselves but also add significant explanatory power or information to the financial ratio factors in predicting insolvencies of property-liability insurers.

Meanwhile, concerning financial ratio variables, the Cox proportional hazard model identifies more significant variables than the logit model in Lee(2005)'s; the Cox proportional hazard model identifies all six financial ratio variables while the logit model identifies four financial ratio variables.

This result is consistent with the findings of Lee and Urrutia(1996).

〈Table 2〉 presents the results based on two-year prior data for the proportionate random sample set. As with the one-year prior data, for the Cox proportional hazard model, the coefficients for both cost X-efficiency and revenue scale efficiency are still negative and significant at the 1 percent level when the model includes the efficiency variables only. Also, when the efficiency variables are combined with the financial ratio factors, those efficiency variables remain negative and highly significant for the Cox proportional hazard model.

The results for the two-year prior sample also suggest that the efficiency variables under the Cox proportional hazard model like the logit model, especially cost X-efficiency and revenue scale efficiency add explanatory power to the financial ratio factors as well as being significant predictors themselves. On the other hand, for two-year prior data, both Cox proportional hazard model and logit model identify all six financial ratio variables.

Second, 〈Table 3 and 4〉 show the results for the non-proportionate random sample set. For one-year prior data, the results are almost same as those in the one-year data for the proportionate random sample. In other words, for the Cox proportional hazard model like the logit model in Lee(2005), the coefficients for both the cost X-efficiency and revenue scale-efficiency variables are negative and highly significant both when the model includes the efficiency variables only and when the efficiency variables are combined with the financial ratio factors. And, as in the proportionate random sample set, revenue X-efficiency is significant variable under the Cox proportional hazard model again when the efficiency variables are combined with the financial ratio factors while it is insignificant in Lee(2005).

The results for the two-year prior data for the non-proportionate random sample are almost identical to those of the one-year prior data. Therefore, the results for the non-proportionate random sample using the Cox proportional hazard model, also confirm the findings of Lee(2005). Meanwhile, both Cox proportional hazard model and logit model identify almost same financial ratio variables for the non-proportional random sample.

〈Table 5〉 presents prediction results. In overall, the prediction accuracy rates are lower for the Cox proportional hazard model than for the logit model in Lee(2005), which is similar with the findings of Lee and Urrutia(1996). For the Cox proportional hazard model, the prediction accuracy rates are unexpectedly lower when the efficiency variables are combined with the financial ratio factors than when the model includes the financial ratio variables only(except for the non-proportionate random sample based on two-year prior data) although the differences are small. Meanwhile, for the Cox proportional hazard model, mostly the ECM(Expected Costs of Misclassifications) are significantly lower when the financial ratio factors are combined with the efficiency variables as expected, than when the model includes the financial ratio factors alone(see 〈Table 6 and 7〉)[14].

14) In terms of the ECM, the lower is better.

〈Table 1〉 Regression Results for the Proportionate Random Sample from 1991 to 2002(One-year Prior Data)

Classification	Cox Model			Lee(2005)		
	EFF	FIN	EFFFIN	EFF	FIN	EFFFIN
Intercept				5.1673 ** (4.0940)	-4.2237 *** (261.5385)	5.0008 (2.3870)
CXE	-6.6009 *** (64.4309)		-3.1635 *** (8.2780)	-7.1186 *** (56.2047)		-3.1455 ** (5.1579)
RXE	-0.7144 (0.3362)		-3.5055 ** (6.3048)	-0.7842 (0.3134)		-0.7248 (0.1333)
CSE	1.3716 (0.6512)		-0.7344 (0.1758)	1.6177 (0.7900)		-3.1352 (1.7997)
RSE	-5.1706 *** (12.3497)		-3.3957 ** (5.1036)	-5.6201 *** (12.1336)		-5.2845 *** (7.0935)
Factor 1		0.5896 ** (4.4071)	0.5496 (2.5010)		0.5254 (1.4324)	0.3842 (0.5651)
Factor 2		0.6007 *** (39.8677)	0.5905 *** (29.9782)		0.8358 *** (31.7927)	0.8120 *** (27.7435)
Factor 3		-9.5447 * (3.7598)	-7.6713 (2.4116)		-10.1815 * (3.1760)	-10.0054 * (2.9243)
Factor 4		-0.3762 *** (96.0027)	-0.3194 *** (47.7261)		-1.0998 *** (65.9329)	-1.0497 *** (50.0774)
Factor 5		-2.4916 *** (7.1097)	-1.0921 (0.8914)		-1.6955 (0.7717)	-1.3881 (0.5000)
Factor 6		-0.4762 *** (11.5065)	-0.2135 (1.7999)		-0.6302 *** (11.1454)	-0.5118 *** (7.1059)
Pseudo R^2	0.07	0.16	0.18	0.13	0.39	0.41

Note:

1) For the Cox model, the dependent variable is the duration of solvency years.
2) EFF refers to models that include the efficiency variables only, FIN pertains to the models with the financial ratio factors only, and EFFFIN refers to models in which the efficiency variables are combined with the financial ratio factors.
3) CXE = cost X-efficiency, RXE = revenue X-efficiency, CSE = cost scale efficiency, and RSE = revenue scale efficiency.
4) Figures in parentheses represent Chi-square values.
5) *** significant at the 1 % level,
   ** significant at the 5 % level,
   * significant at the 10 % level.
6) Factor1: premium written related variables,
   Factor2: loss related variables,
   Factor3: net income related variables,
   Factor4: surplus related variable,
   Factor5: invested assets related variables,
   Factor6: receivables or recoverable related variables.

〈Table 2〉 Regression Results for the Proportionate Random Sample from 1991 to 2002(Two-year Prior Data)

Classification	Cox Model			Lee(2005)		
	EFF	FIN	EFFFIN	EFF	FIN	EFFFIN
Intercept				2.4751 (0.9187)	-3.7583 *** (288.2816)	0.7288 (0.0595)
CXE	-5.7196 *** (52.5219)		-4.0694 *** (12.4669)	-6.5212 *** (43.2498)		-4.2949 *** (10.4090)
RXE	2.0952 (1.5466)		0.3019 (0.0377)	2.3324 (1.8245)		3.3310 (2.6944)
CSE	1.7706 (1.1847)		0.9394 (0.2672)	1.8692 (1.1804)		-1.1245 (0.2897)
RSE	-5.5325 *** (14.8399)		-4.9327 *** (10.1909)	-5.7820 *** (14.1268)		-4.3265 ** (5.7247)
Factor 1		0.6492 *** (17.4531)	0.6213 *** (14.7299)		1.345 *** (14.1821)	0.8681 *** (10.4257)
Factor 2		0.7650 *** (61.7947)	0.7313 *** (31.1446)		0.8743 *** (30.0680)	0.7720 *** (20.4394)
Factor 3		-21.2957 *** (13.8543)	-26.5108 *** (15.9280)		-19.8855 *** (7.9348)	-22.8678 *** (7.2473)
Factor 4		-0.3911 *** (48.3420)	-0.3826 *** (36.4178)		-0.8771 *** (44.2808)	-0.8408 *** (35.7283)
Factor 5		-2.1467 *** (10.5269)	0.2105 (0.0183)		-1.9786 ** (4.7143)	-1.9232 * (2.7451)
Factor 6		-0.6396 *** (33.0240)	-0.5545 *** (21.1540)		-0.5296 *** (11.0349)	-0.4721 *** (7.9854)
Pseudo R^2	0.06	0.12	0.15	0.11	0.25	0.29

Note:

1) For the Cox model, the dependent variable is the duration of solvency years.
2) EFF refers to models that include the efficiency variables only, FIN pertains to the models with the financial ratio factors only, and EFFFIN refers to models in which the efficiency variables are combined with the financial ratio factors.
3) CXE = cost X-efficiency,
   RXE = revenue X-efficiency,
   CSE = cost scale efficiency,
   RSE = revenue scale efficiency.
4) Figures in parentheses represent Chi-square values.
5) *** significant at the 1 % level,
   ** significant at the 5 % level,
   * significant at the 10 % level.

〈Table 3〉 Regression Results for the Non-Proportionate Random Sample from 1991 to 2002(One-year Prior Data)

Classification	Cox Model			Lee(2005)		
	EFF	FIN	EFFFIN	EFF	FIN	EFFFIN
Intercept				5.1890 ** (4.5340)	-5.0775 *** (378.0617)	4.3300 (1.6603)
CXE	-6.6563 *** (66.6855)		-3.2605 *** (11.3058)	-6.9688 *** (60.4528)		-3.0920 *** (5.8637)
RXE	-1.2239 (0.8421)		-4.5264 *** (10.3966)	-1.2506 (0.8218)		-1.1089 (0.3067)
CSE	1.0180 (0.3850)		0.5816 (0.1062)	1.1991 (0.4969)		-2.7371 (1.4561)
RSE	-5.5691 *** (14.6025)		-2.8795 ** (3.7534)	-5.8505 *** (14.6833)		-5.5885 *** (7.7473)
Factor 1		1.6210 *** (21.2924)	1.7475 *** (16.9334)		1.9612 *** (18.5473)	1.8311 *** (13.6574)
Factor 2		0.6763 *** (54.7982)	0.61479 *** (74.4237)		0.9162 *** (65.7066)	0.8735 *** (57.2250)
Factor 3		-3.8955 (1.0767)	-0.02317 (0.0006)		0.0855 (0.0107)	0.0132 (0.0002)
Factor 4		-0.3503 *** (87.8748)	-0.2943 *** (46.1743)		-1.2538 *** (72.4978)	-1.1992 *** (57.5937)
Factor 5		-0.5304 (1.3103)	0.00589 (0.0010)		-0.0261 (0.0018)	0.0187 (0.0023)
Factor 6		-0.4899 *** (15.0673)	-0.21428 (2.4964)		-0.6650 *** (17.7899)	-0.5292 *** (9.8188)
Pseudo R^2	0.07	0.19	0.21	0.11	0.44	0.46

Note:

1) For the Cox model, the dependent variable is the duration of solvency years.
2) EFF refers to models that include the efficiency variables only, FIN pertains to the models with the financial ratio factors only, and EFFFIN refers to models in which the efficiency variables are combined with the financial ratio factors.
3) CXE = cost X-efficiency,
   RXE = revenue X-efficiency,
   CSE = cost scale efficiency,
   RSE = revenue scale efficiency.
4) Figures in parentheses represent Chi-square values.
5) *** significant at the 1 % level,
   ** significant at the 5 % level,
   * significant at the 10 % level.

〈Table 4〉 Regression Results for the Non-Proportionate Random Sample from 1991 to 2002(Two-Year Prior Data)

Classification	Cox Model			Lee(2005)		
	EFF	FIN	EFFFIN	EFF	FIN	EFFFIN
Intercept				1.3226 (0.2616)	-4.4416 *** (426.5727)	-0.8271 (0.0892)
CXE	-5.5222 *** (42.3848)		-2.3674 *** (6.5841)	-5.8563 *** (39.3868)		-2.6556 ** (5.1667)
RXE	4.3509 ** (5.9226)		2.9989 ** (4.2914)	4.4893 ** (5.9922)		5.2592 *** (8.3684)
CSE	0.6905 (0.1977)		0.0540 (0.0010)	0.7138 (0.1944)		-2.0143 (1.0425)
RSE	-6.5063 *** (21.0235)		-5.6206 *** (13.9962)	-6.6551 *** (19.9562)		-5.5162 *** (10.6809)
Factor 1		0.3244 *** (10.9288)	0.2740 *** (8.6459)		0.3828 *** (11.9313)	0.2615 *** (5.3924)
Factor 2		0.8474 *** (101.5075)	0.7716 *** (75.3792)		0.9339 *** (53.3907)	0.8905 *** (42.2791)
Factor 3		-26.7790 *** (24.0797)	-27.1529 *** (26.9246)		-24.8664 *** (12.9121)	-24.7279 *** (12.3311)
Factor 4		-0.4831 *** (90.1540)	-0.4446 *** (60.8622)		-0.7695 *** (57.8661)	-0.7239 *** (42.1342)
Factor 5		-2.2256 *** (22.8034)	-2.2505 *** (20.0075)		-2.0184 *** (11.4526)	-2.1060 *** (9.6466)
Factor 6		-0.8223 *** (43.8096)	-0.7615 *** (37.7454)		-0.7945 *** (26.6180)	-0.7220 *** (20.4074)
Pseudo R^2	0.06	0.12	0.15	0.10	0.22	0.26

Note:
1) For the Cox model, the dependent variable is the duration of solvency years.
2) EFF refers to models that include the efficiency variables only, FIN pertains to the models with the financial ratio factors only, and EFFFIN refers to models in which the efficiency variables are combined with the financial ratio factors.
3) CXE = cost X-efficiency,
   RXE = revenue X-efficiency,
   CSE = cost scale efficiency,
   RSE = revenue scale efficiency.
4) Figures in parentheses represent Chi-square values.
5) *** significant at the 1 % level,
   ** significant at the 5 % level,
   * significant at the 10 % level.

〈Table 5〉 Prediction Results for Holdout Sample(Correct %)

One-year Prior

Sample Set	Cox Model			Lee(2005)		
	EFF	FIN	EFFFIN	EFF	FIN	EFFFIN
Proportionate Random Sample	50.0	61.5	54.4	54.3	74.0	75.7
Non-Proportionate Random Sample	50.0	66.5	56.5	52.5	73.9	75.8

Two-year Prior

Sample Set	Cox Model			Lee(2005)		
	EFF	FIN	EFFFIN	EFF	FIN	EFFFIN
Proportionate Random Sample	50.0	58.5	51.6	51.6	61.6	61.5
Non-Proportionate Random Sample	50.0	58.0	58.2	50.0	54.9	58.2

Note:
1) Cutoff point is 0.5.
2) EFF refers to models that include the efficiency variables only, FIN pertains to the models with the financial ratio factors only, and EFFFIN refers to models in which the efficiency variables are combined with the financial ratio factors.

〈Table 6〉 The Results of the Expected Costs of Misclassification (One-year Prior Data)

Sample Set	C1	Cox Model			Lee(2005)		
		EFF	FIN	EFFFIN	EFF	FIN	EFFFIN
Proportionate Random Sample	1	0.0100	0.7615	0.0091	0.0138	0.0052	0.0119
	5	0.0500	0.7615	0.0456	0.0504	0.0260	0.0310
	25	0.2500	0.7615	0.2282	0.2330	0.1304	0.1266
	50	0.5000	0.7615	0.4565	0.4612	0.2608	0.2462
	75	0.7500	0.7615	0.6847	0.6895	0.3912	0.3657
	100	0.7094	0.7615	0.9130	0.9177	0.5217	0.4853
Non-proportionate Random Sample	1	0.0100	0.6630	0.0080	0.0100	0.0070	0.0090
	5	0.0500	0.6630	0.0430	0.0480	0.0280	0.0290
	25	0.2500	0.6630	0.2170	0.2390	0.1320	0.1240
	50	0.5000	0.6630	0.4340	0.4780	0.2620	0.2430
	75	0.7500	0.6630	0.6510	0.7170	0.3920	0.3630
	100	1.0000	0.6630	0.8690	0.9560	0.5220	0.4820

Note:
1) EFF refers to models that include the efficiency variables only, FIN pertains to the models with the financial ratio factors only, and EFFFIN refers to models in which the efficiency variables are combined with the financial ratio factors.
2) Various misclassification costs of type I error(C1), which ranges from 1 to 100, holding the misclassification cost of type II error(C2) constant, are used.
3) Type Ⅰ error is defined to be the misclassification of an insolvent insurer as a solvent insurer while type Ⅱ error is the misclassification of a solvent insurer as an insolvent insurer.

〈Table 7〉 The Results of the Expected Costs of Misclassification (Two-year Prior Data)

Sample Set	C1	Cox Model			Lee(2005)		
		EFF	FIN	EFFFIN	EFF	FIN	EFFFIN
Proportionate Random Sample	1	0.0100	0.7894	0.0116	0.0136	0.0096	0.0116
	5	0.0500	0.7907	0.0503	0.0522	0.0403	0.0422
	25	0.2500	0.7974	0.2436	0.2456	0.1936	0.1956
	50	0.5000	0.8057	0.4852	0.4872	0.3852	0.3872
	75	0.7500	0.8141	0.7269	0.7289	0.5769	0.5789
	100	1.0000	0.8224	0.9685	0.9705	0.7685	0.7705
Non-proportionate Random Sample	1	0.0100	0.7992	0.0122	0.0100	0.0099	0.0112
	5	0.0500	0.8004	0.0454	0.0500	0.0459	0.0444
	25	0.2500	0.8064	0.2114	0.2500	0.2259	0.2104
	50	0.5000	0.8139	0.4189	0.5000	0.4509	0.4179
	75	0.7500	0.8214	0.6264	0.7500	0.6759	0.6254
	100	1.0000	0.8289	0.8339	1.0000	0.9009	0.8329

Note:
1) EFF refers to models that include the efficiency variables only, FIN pertains to the models with the financial ratio factors only, and EFFFIN refers to models in which the efficiency variables are combined with the financial ratio factors.
2) Various misclassification costs of type I error(C1), which ranges from 1 to 100, holding the misclassification cost of type II error(C2) constant, are used.
3) Type Ⅰ error is defined to be the misclassification of an insolvent insurer as a solvent insurer while type Ⅱ error is the misclassification of a solvent insurer as an insolvent insurer.

Ⅳ. Conclusions

The results of this study confirm the findings of Lee(2005). In other words, under the frame of the Cox proportional hazard model like the logit model, efficiency measures has tuned out to be important factors in identifying and forecasting insurer insolvencies again. Also, the results of the study supports another finding of Lee(2005) that the efficiency variable sets add significant explanatory or discriminatory power to the financial ratio variable sets. Meanwhile, this study also finds that overall, the Cox proportional hazard model has comparable ability to the logit model in identifying and forecasting insurer insolvencies albeit the prediction accuracy rate is a little better for the logit model.

Lastly, in the sense that both logit and Cox proportional hazard model convey important information regarding insolvency of property-liability insurers, the combined use of both statistical models in identifying and predicting insolvency of insurers would be desirable. And its performance should be improved with the inclusion of efficiency measures as explanatory factors into the model.

References

Allison, P.D., *Survival Analysis Using the SAS System: A Practical Guide*, Cary: SAS Institute Inc., 1995.

A. M. Best Company, *Best's Insolvency Study Property/Casualty Insurers 1969-1990*, Oldwick: A. M. Best Company, 1991.

A. M. Best Company, *Best's Key Rating Guide*, Oldwick: A. M. Best Company(1996, 2001, 2002 and 2003).

BarNiv, R. and Hathorn, J., "The Merger or Insolvency Alternative in the Insurance Industry", *Journal of Risk and Insurance*, Vol.64, 1997, pp.89~113.

BarNiv, R. and McDonald, J.B., "Identifying Financial Distress in the Insurance Industry: A Synthesis of Methodological and Empirical Issues", *Journal of Risk and Insurance*, Vol.59, 1992, pp.543~574.

Bauer, P.W., Berger, A.N., Ferrier, G.D. and Humphrey, D.B. "Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Efficiency Methods", *Journal of Economics and Business*, Vol.50, 1998, pp.85~114.

Berger, A.N. and Humphrey, D.B., "Measurement and Efficiency Issues in Commercial Banking", *Output Measurement in the Service Sectors, National Bureau of Economics Research, Studies in Income and Wealth*, In Z. Griliches(Eds.), Chicago: University Chicago Press, 1992, pp.245~279.

Berger, A.N., Cummins, J.D. and Weiss, M.A., "The Coexistence of Multiple Distribution Systems for Financial Services: The Case of Property-Liability Insurance", *Journal of Business*, Vol.70, 1997, pp.515~546.

Breslow, N.E., Covariance Analysis of Censored Survival Data, *Biometrics*, Vol. 30, 1974, pp.89~99.

Brockett, P.L., Cooper, W.W., Golden, L.L. and Pitaktong, U., "A neural Network Method for Obtaining an Early Warning of Insurer Insolvency", *Journal of Risk and Insurance*, Vol.61, 1994, pp.402~424.

Chen, K.C.W., and Lee, C.W.J., "Financial Ratios and Corporate Endurance: A Case of the Oil and Gas Industry", *Contemporary Accounting Research*, Vol.9, 1993, pp.667~694.

Cox, D.R. "Regression Models and Life Tables", *Journal of Royal Statistical Society*, B34, 1972, pp.187~220.

Cox, D.R., and Oakes, D., *Analysis of Survival Data*, London: Chapman and Hall, 1984.

Cummins, J.D., and Weiss, M.A., "Analyzing Firm Performance in the Insurance Industry Using Frontier Efficiency and Productivity Methods", *Handbook of Insurance Economics*, In G. Dionne(Eds.), Boston: Kluwer Academic Publishers, 2000, pp.767~829.

Cummins, J.D., Grace, M.F. and Phillips, R.D., "Regulatory Solvency Prediction in Property-Liability Insurance: Risk-Based Capital, Audit Ratios, and Cash Flow Simulation", *Journal of Risk and Insurance*, Vol.66, 1999, pp.417~458.

Cummins, J.D., Weiss, M.A., and Zi, H., "Organizational Form and Efficiency: The Coexistence of Stock and Mutual Property-Liability Insurers", *Management Science*, Vol.45, 1999, pp.1254~1269.

Eck, J.R., "Detecting Financially Troubled Property-Liability Insurers", *Journal of Business Research*, Vol.10, 1982, pp.445~458.

Harrington, S.E., and Nelson, J.M., "A Regression-Based Methodology for Solvency Surveillance in the Property-Liability Insurance Industry", *Journal of Risk and Insurance*, Vol.53, 1986, pp.583~605.

Hershbarger, R.A., and Miller, R.K., "The NAIC Information System and the Use of Economic Indicators in Predicting Insolvencies", *Journal of Insurance Issues and Practices*, Vol.9, 1986, pp.21~43.

Kiefer, N.M., "Economic Duration Data and Hazard Functions", *Journal of Economic Literature*, Vol.26, 1988, pp.646~679.

Kim, Y., Anderson, D.R., Amburgey, T.L. and Hickman, J. C., "The Use of Event History Analysis to Examine Insurer Insolvencies", *Journal of Risk and Insurance*, Vol.62, 1995, pp.94~110.

Lane, W.R., Looney, S.W., and Wansley, J.W., "An Application of the Cox Proportional Hazards Model to Bank Failure", *Journal of Banking and Finance*, Vol.10, 1986, pp.511~531.

Lawless, J. F., *Survival Models and Methods for Lifetime Data*, New York: John Wiley & Sons, Inc., 1982.

Lee, S.H., "Efficiency Measures and Insolvencies of Property-Liability Insurers", *The Journal of Risk management*, Vol.16, No.2, 2005, pp.3~37.

Lee, S.H. and Urrutia, J.L., "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, 1996, pp.121~130.

Ng, S.H., Cram, F., and Jenkins, L., "A Proportional Hazards Regression Analysis of Employee Turnover Among Nurses in New Zealand", *Human Relations*, Vol.44, 1991, pp.1313~1329.

Pottier, S.W., "Life Insurer Financial Distress, Best's Ratings and Financial Ratios", *Journal of Risk and Insurance*, Vol.65, 1998, pp.275~288.

Whalen, G., "A Proportional Hazards Model of Bank Failure: An Examination of Its Usefulness as an Early Warning Tool", *Federal Reserve Bank of Cleveland*, 1991, pp.21~31.

요 약

본 연구의 목적은 logit model을 사용한 기존 Lee(2005)의 연구 결과를, 즉 효율성 변수가 손보사 파산의 중요한 예측 변수인지, 그리고 이들 효율성 변수들이 보험사 파산 예측 모델에서 기존에 주로 사용되어졌던 재무변수들의 설명력과 분별력을 증진시키는지의 여부를 다른 파산예측 모델인 Cox proportional hazard model을 이용하여 검증 · 확인하는 데에 있다. 본 연구의 결과, 효율성 변수가 보험사 파산의 중요한 예측 변수이고, 재무변수들의 설명력을 증진시키는 것으로 확증되고 있다. 본 연구의 또 하나의 의의로서, 보험사 파산예측에서는 본 연구가 처음 시도하는 것이지만, 사회과학, 생물의학 등의 분야 및 은행 파산예측에서 이미 효과적인 모델로 판명된 Cox proportional hazard model이 보험사 파산예측에서도 유용한 모델이 되는 것으로 나타나고 있다. 향후 보험사 파산예측에서 logit model과 Cox proportional hazard model, 그리고 기존 재무변수들에 효율성 변수가 결합된다면 보다 의미 있는 결과가 도출될 것으로 기대되어 진다.

※ 국문 색인어: 로짓 모델, 보험사 파산, 칵스비례위험모델, 효율성