An Empirical Examination of Sample Selection Methods in the Context of Life Insurer Financial Distress

James M. Carson and Robert E. Hoyt*

Abstract: This study empirically examines properties of matched-pair versus non-matched-pair sampling methods in the context of financial distress for the U.S. life insurance industry. While the majority of prior insurer insolvency studies employed matched-pair sampling techniques to identify important variables and to classify and predict firms likely to become financially distressed, we provide empirical evidence that three solvency-related items are sample dependent: variables identified as important measures of insolvency, coefficients, and classification rates. Thus, empirical studies employing relatively small matched-pair samples are likely to yield sample-specific results that are not fully generalizable to the relevant population of firms. Results apply directly to financial distress models and also extend to other research employing choice-based sampling methods that involve binary state models with skewed distribution of the two states of interest. [Keywords: financial distress; choice-based sample; classification; bias]

INTRODUCTION

Empirical models to identify potentially distressed insurers and important solvency variables generally have been developed using matched-pair samples. That is, one firm identified as "financially impaired" is "matched" with one firm that is not distressed, often on the basis of size (surplus or assets) and state of domicile. Results of matched-pair studies

*James M. Carson (jcarson@cob.fsu.edu) is Midyette Eminent Scholar of Risk and Insurance in the College of Business at Florida State University. Robert E. Hoyt (rhoyt@terry.uga.edu) is Head of the Insurance, Legal Studies, and Real Estate Department in the Terry College of Business at The University of Georgia.

Journal of Insurance Issues, 2003, **26**, 2, pp. 114–128.

indicate that classification of solvent and insolvent firms can be performed with a relatively high (over 80 percent accuracy) degree of success. In addition to classification results, these studies identify sets of variables significantly related to a firm's probability of failure. Implicit in using state-based samples is that the studies' results have direct relevance to the populations of companies to which the subsets of firms belong.[1]

The appropriateness of state-based (matched-pair) samples has been examined by Palepu (1986), who suggests that error rate inferences based on state-based samples are not directly generalizable to the population. The non-random nature of state-based sampling results in biased estimates of the expected prediction error rate in the population. The magnitude of this bias is proportional to the difference in the ratios of population and sample shares of solvent and insolvent insurers (see Palepu, 1986, p. 11). Although prior insolvency studies generally avoid stating explicitly that the empirical models be used for prediction, this is clearly the most likely practical application of the results. Regulators and others are interested in predicting financially troubled insurers among the population of insurers. Palepu observes that "since actual use of the model involves the entire population, it is desirable to make the prediction test sample resemble the population as closely as possible."

While the use of state-based samples in empirical insolvency research is the norm, direct empirical comparison of results based on matched-pair versus non-matched-pair samples has been limited. Ohlson (1980) reports that error rates in his unmatched sample are higher than the reported error rates in most of the previous matched-pair samples, but he observes that the differences between his results and those of prior studies "are most difficult to reconcile." No prior study has empirically evaluated this concern in the insurance literature. Thus, the issues of whether significant variables and classification rates obtained from matched-pair samples are representative of the significant variables and classification rates obtained when examining the population of firms still remain.

The goal of this study is to provide direct empirical evidence from solvency models using matched-pair versus non-matched-pair samples. If variables that are identified as important measures of insolvency are sample-dependent, then researchers and others relying on results of studies based on matched-pair samples may be led to focus on a suboptimal set of variables in their search to identify financially distressed firms. This study also provides a direct comparison of the classification results obtained from matched-pair samples (equal numbers of solvent and insolvent firms based on matching criteria) and a non-matched-pair sample (close to the full population of solvent and insolvent firms) of insurers. The data used in this analysis are for 1,685 life insurers, and are obtained from the NAIC database. While this study is done within the context of insolvency in the U.S. life insurance industry, results also provide implications for nonsolvency research that utilizes matched-pair sampling methods.

PRIOR RESEARCH AND FOCUS OF STUDY

Empirical Insolvency Research

Bankruptcy and insurer insolvency research is briefly reviewed below. Corporate bankruptcy studies utilizing the matched-pair sampling approach (sample size in brackets) include Beaver (1966) [158]; Altman (1968) [66]; Meyer and Pifer (1970) [78]; Dambolena and Khoury (1980) [46]; Zavgren (1985) [90]; and Altman, Eom, and Kim (1995) [68]. The matched-pair approach also is used in areas other than for financial distress classification, such as in Khanna and Poulsen (1995), which used the matched-pair approach to examine managerial decisions of 128 financially distressed and 118 nondistressed firms.

Beginning with Trieschmann and Pinches (1973), insurer insolvency studies used matched-pair samples of relatively small size, applied multiple discriminant analysis (MDA) and/or logit, produced relatively high classification (low error) rates, often used a validation method, and grouped insurer insolvencies over a period of from three to 20 years. Table 1 summarizes several insurer insolvency studies (property-liability and life-health) by sample composition, methodology, classification rates, validation method, time period of insolvencies, and type of insurers examined. Except for Ambrose and Carroll (1994), results of empirical studies based on matched-pair samples generally suggest that classification of insurer failures can be performed with a relatively high degree of accuracy. Other studies considered the bias associated with oversampling the insolvent insurers with the matched-pair approach (see BarNiv and Raveh, 1989; BarNiv, 1990; and BarNiv and McDonald, 1992), but did not empirically evaluate the impact of this bias when the models are applied to the population.

Other recent insolvency research that does not directly address the issue examined here includes Lamm-Tennant, Starks, and Stokes (1996), Grace, Harrington, and Klein (1998), and Pottier (1998). Browne, Carson, and Hoyt (1999) identified economic and market predictors of insolvencies, as opposed to company-specific factors or classification rates, for life insurers. Baranoff, Sager, and Shively (2000) introduce a nonlinear semi-parametric spline model for identifying distressed and sound insurers. Solvency regulation has been strengthened over time, and many observers

Table 1. Summary of Selected Previous Insurance Research: Sample Selection, Methodology, Classification Rates, Validation Method, Time Period of Failures, and Type of Insurers

	Trieschmann & Pinches (1973)	Harrington & Nelson (1986)	Ambrose & Seward (1988)	BarNiv & Hershbarger (1990)	BarNiv & McDonald (1992)	Ambrose & Carroll (1994)	Cummins, et al. (1995)	Carson & Hoyt (1995)
SAMPLE								
Matched sample	No[a]	No	Yes	Yes	Yes	Yes	No[b]	No
# Distressed firms	26	12	29	28	65	26	163	80
# Nondistressed firms	26	69	29	28	70	26	1,596	1,605
PRIMARY METHODOLOGY	MDA	Regression	MDA	Logit	Logit	Logit	Logit	Logit
CLASSIFICATION RATES								
Distressed (Type-I)	.92	.83	.90	.93	.80	.76	.71	.80
Nondistressed (Type-II)	.96	.93	.76	.89	.87	.71	.70	.72
Overall	.94	.91	.83	.91	.84	.74	.70	.72
VALIDATION METHOD	S[c]	None	Jackknife[d]	A[e]	Holdout[f]	Holdout[g]	Jackknife[d]	Holdout[h]
TIME PERIOD OF FAILURES	1966–1971	1976–1981	1969–1983	1975–1985	1974–1988	1969–1991	1990–1993	1989–1991
TYPE OF INSURERS	P-L	P-L	P-L	L-H	P-L	L-H	P-L	L-H

[a]randomly matched
[b]based on 1990 sample
[c]simulated sample
[d]Lachenbruch method
[e]additional solvent sample
[f]later time period holdout sample
[g]same time period holdout sample
[h]25 percent holdout of full sample

regard the U.S. solvency oversight system as the best of any financial regulator in the world (Vaughan, 2002).

Most empirically based insolvency research has employed the logit methodology. Carson and Hoyt (1995) used a large, non-matched-pair sample and obtained relatively low classification rates compared to those studies based on matched-pair samples. In addition, the study reported significant changes in important variables from those identified in previous studies that had examined matched-pair samples and earlier time periods. By directly examining matched-pair and near-population samples from contemporaneous sample periods, the present study controls for variables that may change in importance over time.[2]

Research on Choice-Based Sampling and Insolvency

Choice-based samples, as used in the majority of previous insolvency research, occur when the probability of an observation entering the sample depends on the value (e.g., solvent or insolvent) of the dependent variable. The estimation techniques used in previous insolvency studies, however, generally assume the use of exogenous random sampling in which observations are drawn randomly and the dependent and independent variables are observed ex post. Manski and Lerman (1977) and Palepu (1986) indicate that choice-based sampling violates the random sampling design assumption and causes both parameter and probability estimates to be asymptotically biased. Correlation between the matching variable (such as size) and the independent variables may produce coefficients that are biased and inconsistent.[3]

In defense of smaller samples, Zmijewski (1984) notes that choice-based samples are not necessarily inappropriate, especially in the presence of high search costs, and that estimation techniques exist (weighted exogenous sample maximum likelihood (WESML), conditional maximum likelihood (CML), and full information conditional maximum likelihood (FICML)) that are appropriate for estimating choice-based samples and for obtaining unbiased parameter and probability estimates. Examining six choice-based samples containing decreasing bankrupt firm frequency rates, Zmijewski finds clear evidence of a bias, depending on the degree of oversampling of bankrupt firms, and that the bias is eliminated using the WESML procedure. The author reports that "the bias does not, in general, affect the statistical inferences or the overall classification rates for the financial distress model and the samples tested" (Zmijewski, 1984, p. 77). However, Zmijewski assumes that the non-bankrupt firms are selected randomly. This is not the case in prior insurance studies. Therefore, his study provides no evaluation of the potential problems associated with a methodology based on nonrandom choice-based sampling approaches (matched-pair sampling).

Focus of Study

Despite the theoretical and empirical work done by Manski and Lerman (1977), Ohlson (1980), Zmijewski (1984), and Palepu (1986), prior insolvency studies have not directly tested for disparity in terms of predictors or classification between matched-pair and non-matched-pair samples. We directly examine explanatory variables and classification rates associated with matched-pair and non-matched-pair sampling methods to address these issues. The analysis provides specific evaluation of the impact of the bias introduced by the fact that the matched-pair approach is a subset of the choice-based sampling approach where the non-bankrupt firms are not randomly selected from the population of insurers. This source of bias (correlation between the matching criterion and the independent variables in the prediction model) is not corrected by the adjustment described in Palepu (1986). The adjustment does correct for the oversampling of the insolvent firms by basically revising the estimated probabilities by the ratio of the solvent firms in the sample to the solvent firms in the population. However, it does not adjust for the correlation bias in the matching criterion. BarNiv and McDonald (1992, p. 569) provide a detailed description of this adjustment. The adjustment is commonly applied in prior insurance insolvency studies (see, for example, Ambrose and Carroll, 1994, p. 318; and BarNiv and McDonald, 1992, p. 569). The importance of this uncorrected bias is empirically evaluated in this study.

SAMPLE DESCRIPTION AND METHODOLOGY

The non-matched-pair sample includes 1,605 solvent life insurers and 80 life insurers that became insolvent over a three-year period. To obtain a sufficient number of insolvencies while concurrently minimizing the length of the insolvency period (thereby controlling for the effects of dynamic exogenous factors such as changing interest rates, stock and bond markets, etc.) the most recent three-year period with the greatest number of life insurer insolvencies is the period from 1989 through 1991. This period contains a relatively large number of insurer insolvencies and therefore is especially well suited for financial distress research, including the empirically based methodological analysis here.[4]

Repeated sampling methods are used to generate three matched-pair samples (without replacement), which are subsets of the non-matched-pair sample. Thus, each matched-pair sample consists of 80 solvent and 80 insolvent insurers. Firms are matched on the basis of total assets and state of domicile.[5] Data are ordered to allow analysis corresponding to one year prior to insolvency, with insolvencies occurring in the period 1989 through 1991.[6]

All insurers with complete data from the NAIC data tapes are included in the study, except as described here. An insurer experiencing financial difficulty (but not yet insolvent) may merge with a stronger insurer in order to avoid insolvency. Insurers involved in a merger during this study's sample period are not included in the sample, since the reason for a particular merger may or may not be related to solvency. Excluding merged insurers is consistent with previous studies—e.g., Ambrose and Carroll (1994), Browne and Hoyt (1995), and Browne, Carson, and Hoyt (1999). Voluntary retirements of life insurers also are not included, since these retirements may be due to reasons other than financial impairment. As in A.M. Best Companies' *Best's Insolvency Study* (1992), Blue Cross and Blue Shield organizations, managed care companies, third-party administrators, self-insurers, and fraternal and burial associations also are not included.

The dependent variable in the logistic regression model is specified as 0 for solvent insurers and as 1 for insolvent insurers. The forward stepwise logit is particularly appropriate for our purposes here, in that it singles out and identifies which variables, among those variables listed in Table 2, are most important across each sample. Table 2 includes important variables from previous research as well as other variables.[7]

EMPIRICAL RESULTS

Important Variables

Table 3 presents a comparison of variables chosen by the stepwise logit procedure based on the non-matched-pair versus the matched-pair samples.[8] From the initial list of 23 variables shown in Table 2, the significant variables, while similar to the significant variables found in previous research, differ substantially according to the sample used. Results of the non-matched-pair sample suggest that capital and surplus, reserves to surplus, separate account assets, real estate to assets, change in asset mix, and premium to surplus are important variables in identifying insurer solvency.

However, results of the models developed from the matched-pair samples suggest that different sets of variables are important. The logit model developed from matched-pair Sample 1 indicates that while some of the important variables based on the non-matched-pair sample are

Table 2. List of Variables Used in the Stepwise Logit Model

Variable[a]	Abbreviation (if variable was used)
Log of Capital and Surplus	CAPSUR
Log of Capital and Surplus (mean to variance)	CAPSURXV
Log of Capital and Surplus to Liabilities	CAPSURL
Insurance Leverage	
Reserves to Surplus	RSRVSURP
Financial Leverage	
Premiums to Surplus	PREMSURP
Separate Account Assets to Assets	SEPASSET
Real Estate to Assets	REASSET
Real Estate to Assets (mean to variance)	REASSTXV
Change in Product Mix[b]	CHGPRMIX
Change in Assset Mix[c]	CHGAMIX
Mutual or Stock Insurer (0,1)	FORM
Liabilities to Current Assets	LIABCURR
Change in Capital and Surplus	CHGCASUR
Net Gain from Operations to Capital and Surplus	NGFOCS
Commissions and Expenses to Premium	CMEXPR
Commissions and Expenses to Premium (mean to variance)	CMEXPRXV
Age of Insurer	
Minimum State Capital Requirement	
Business Line Concentration[d]	
Business State Concentration[e]	
Reinsurance to Assets	
Log of Assets	
Net Gain from Operations	

[a]The analysis includes the mean and standard deviation of certain variables, where appropriate. Descriptive statistics are available from the authors.
[b]CHGPMIX is the average of the absolute value of percentage changes in premium for all lines of insurance (same as IRIS ratio).
[c]CHGAMIX is the average of the absolute value of the percentage changes in assets for line items number one through ten (same as IRIS ratio).
[d]This variable is a Herfindahl measure of the squared percentages of premium for each line of business.
[e]This variable is a Herfindahl measure of the squared percentages of premium written in each state.

Table 3. Logit Predictor Coefficients and Classification Rates: Non-Matched-Pair versus Matched-Pair Samples[a]

Predictor Variable	Non-M-P[b] Sample	M-P Sample 1	M-P Sample 2	M-P Sample 3
INTERCEPT	–2.38	–1.96*	–.74	–1.80*
CAPSUR	–.11**			
CAPSURXV	–.01***	–.01*	–.02***	–.02***
RSRVSURP	.04***	.05*	.13*	.10***
SEPASSET	2.05**			
REASSET	13.30***	38.37**	10.45**	9.59
CHGAMIX	17.04***	30.87***	35.97*	32.09***
PREMSURP	.05*			
FORM	.69	1.55**		
LIABCURR		.54		1.05**
NGFOCS		–1.19**	.29***	–1.43***
REASSTXV		–3.53*		
CHGCAPSUR			–1.56***	–.88*
CHGPRMIX			19.06**	3.52
CMEXPR			–.03**	
CMEXPRXV			0.18**	.24**
CAPSURL				.06*
Pseudo-R^2	25.02%	33.28%	41.72%	45.72%
Classification[c]				
Solvent	.75	.81[d](.80)[e]	.76(.68)	.85(.73)
Insolvent	.76	.68(.69)	.86(.86)	.79(.80)
Overall	.75	.74(.79)	.81(.69)	.82(.73)

[a]Models were developed using the Lachenbruch validation method.
[b]M-P = Matched-Pair
[c]Classification based on optimal cutoff points and relative cost of type-I and type-II errors at 20/1. A joint test (chi-square) indicates that overall classification rates are significantly different from each other at the .01 level (test statistic is 50.93).
[d]Classification results from applying the respective M-P models to each of the respective matched-pair samples.
[e]Percentages in parentheses indicate classification rates for the matched-pair sample models applied to the full (non-matched-pair) sample.
NOTE: ***, **, and * indicate significance at the .01, .05 and .10 level, respectively.
NOTE: Variables ending in "XV" are the ratio of mean to variance.

important, other variables also are important, such as liabilities to current assets, net gain from operations to capital and surplus, and real estate to assets (mean to variance). In addition, results from matched-pair Sample 1 indicate that stock insurers may exhibit higher risk than mutual insurers (consistent with Garven, 1992; Lamm-Tennant and Starks, 1993; and Cummins et al., 1995). Matched-pair Sample 2 suggests that four more variables are important in solvency models. Matched-pair Sample 3 indicates yet another significant variable.

Of the four models, three variables are significant and common to all: capital and surplus (mean to variance), reserves to surplus, and change in asset mix. The real estate to assets variable is significant in three of the four models. The set of important variables differs according to the type of sampling method, and even according to the particular sample within the matched-pair sampling method, suggesting that models based on matched-pair samples are not necessarily directly generalizable to other samples or to the population of insurers. The point here is not which variables are important, but rather, that the set of important variables is inconsistent across samples.

Coefficients

Applying the set of variables derived from M-P Sample 1 to the full sample yielded significantly different coefficients (.01 level) from those coefficients derived from the matched-pair sample. That is, evidence was found that coefficients, as well as variables, are sample-dependent. Similar results were found for the models based on M-P Samples 2 and 3.[9]

Classification Rates

Table 3 also presents a comparison of classification results.[10] The full sample model produced an overall classification rate of 75 percent, while the three matched-pair sample models, when applied to their respective estimation samples, produced overall classification rates of 74 percent, 81 percent, and 82 percent.[11] More importantly, however, applying each of the three matched-pair models (same variables and coefficients) to the non-matched-pair sample of insurers resulted in inconsistent overall classification rates, ranging from 69 percent to 79 percent. For example, the M-P 2 model suggests that classification of insurers can be achieved with 81 percent overall accuracy; however, applying the M-P 2 model to the population of insurers indicates an overall success rate of 69 percent. Comparing the overall classification results for M-P 1, M-P 2, and M-P 3 illustrates the variation in results when applying different models to the population of firms.

Results from Table 3 suggest that classification rates from the matched-pair sampling method are inconsistent across samples and models. A joint chi-square test of the classification results in the four models applied to the population of insurers rejects the hypothesis of equal classification rates (for the M-P models, these are the classification rates that are provided in parentheses at the bottom of Table 3).[12] This test is significant at .01 for solvent and overall classification rates, and at .05 for insolvent classification rates. Pairwise chi-square tests were used to compare these classification rates for each of the M-P models to the classification rates for the non-M-P model. For M-P 1 and M-P 2, the overall and solvent classification rates are different from those in the non-M-P model at the .01 level, and the insolvent classification rates are different at the .10 level. For M-P 3, only the solvent classification rate was statistically different in relation to the classification rates in the non-M-P model (.10 level). Pairwise tests for differences in the classification rates among the M-P models all were significant at .01 for the overall and solvent classification rates, and below .10 for the insolvent classification rates.

Thus, results indicate that empirical models based on matched-pair samples produce results that are inconsistent in relation to the percentage of insurers that are likely to be correctly classified when these models are applied to the population of insurers. The disparity in classification rates likely is a manifestation of the variation in significant variables produced from heterogeneity across choice-based samples.

CONCLUSIONS AND IMPLICATIONS

The goal of this study was to provide empirical evidence on the bias in variable selection, coefficients, and classification rates introduced by the matched-pair approach as a subset of the choice-based sampling approach. The study examined results from solvency models derived from a non-matched-pair sample and subsets of matched-pair samples of life insurers. A limitation of the study was that one matched-pair sample necessarily had a closer match than the other matched-pair samples. Models developed from matched-pair samples were applied to the full sample and resulted in inconsistent classification rates, suggesting that functions based on small samples do not map correspondingly to larger samples. If varying classification rates were the only concern with using matched-pair samples, expectations of matched-pair model performance on larger samples could simply be adjusted. However, important measures of insolvency differed across samples, indicating that researchers, regulators, and others relying on results of studies based on matched-pair samples may be led to focus on a suboptimal set of variables in their search to identify financially distressed firms. The study illustrated the importance of utilizing the largest sample possible when developing solvency models applicable to the population of U.S. insurers. In this sense, the development of large, electronic-format databases (e.g., NAIC database) is particularly important. Results here are not inconsistent with previous research (e.g., Zmijewski, 1984; Palepu, 1986), since those studies examined the case of random sampling.

The study's findings are relevant to solvency research and to other types of analyses utilizing matched-pair sampling methods. The results of our study highlight the potential bias associated with the use of matched-pair sampling methods. Since these methods have been applied to other issues in insurance (e.g., mergers and acquisitions), our results suggest that researchers should proceed with caution when relying on matched-pair approaches.

ENDNOTES

[1] The terms choice-based sampling, state-based sampling, and matched-pair sampling have been used interchangeably in prior classification studies. Importantly, matched-pair sampling has been the term most commonly used in prior insurance insolvency studies. These prior studies used a procedure by which nonbankrupt insurers were "matched" with bankrupt insurers through reference to a matching criterion. The terms choice-based sampling and stated-based sampling do not necessarily imply the use of a matching criterion, but instead refer to the oversampling of one segment of the population (e.g., bankrupt firms).

[2] A comparison of results from BarNiv and Hershbarger (1990), Ambrose and Carroll (1994), and Carson and Hoyt (1995) indicates that of 14 different significant variables, only the Change in Assets variable is statistically significant in all studies.

[3] We find evidence that this is an important concern: the financial matching criterion for the matched-pair samples in our study, log of assets, was significantly correlated (.05 level) with 14 of the 23 variables listed in Table 2 and with 9 of the 16 variables in Table 3. Similar correlation exists between an alternative size measure, surplus, and the independent variables.

[4] Note that the purpose of the research is not to indicate the best set of important predictor variables. Rather, the focus here is methodological and thus, in a sense, less dependent on the particular time period analyzed (given a sufficient number of insolvencies).

[5] Although results varied somewhat, constructing matched-pair samples with replacement yielded similar results. In addition, using surplus, as opposed to assets, as the matching variable for the matched-pair samples yielded similar findings to the results reported in Table 3. An inherent limitation of the study is that, since matching was done, one matched-pair sample necessarily had a closer match than the other two matched-pair samples. Using variables other than total assets and state of domicile to produce the matched pairs likely would lead to the same general findings, albeit different significant variables, coefficients, and classification rates, thus further reinforcing the article's conclusions. Including the FORM variable in the matching process would be a logical matching variable to employ as well.

[6] The A.M. Best definition of "financially impaired company" (FIC) is the basis for determining failed and nonfailed insurers. The date of FIC is that of the first official action taken by the department of insurance in the insurer's state of domicile. When the focus is on developing predictive models, variables that are available for more than one year prior to insolvency are

desirable. The focus of this study is on the effects of alternative sampling methods. Therefore, variables for one year prior to insolvency are considered sufficient.

[7] The focus of this study is on alternative sampling methods and their effect on estimation results. Solvency-related hypotheses for several of the variables are discussed in BarNiv and Hershbarger (1990) and Ambrose and Carroll (1994).

[8] Results are validated using the Lachenbruch (jackknife) method, as in Ambrose and Seward (1988) and other studies. The procedure uses a one-step approximation to the coefficient vector that would be obtained if the observation was excluded from the sample (see Cummins et al., 1995). Judge et al. (1988) discusses the jackknife and bootstrap procedures as forms of resampling methods. The bootstrap method, while computationally more cumbersome than the jackknife, provides parameters of the empirical model based on random resampling to form N matched-pair samples. The ratio of type-I errors (misclassifying an insolvent insurer as solvent) and type-II errors (misclassifying a solvent insurer as insolvent) was varied from 1/1 to 40/1, and results presented are for a ratio of 20/1. The results in Table 3 are based on optimal cutoff points, as described in BarNiv and McDonald (1992), for classifying a solvent versus an insolvent insurer.

[9] Welch's paired t-tests were performed to check for significant differences in the coefficients across samples. Welch's t-test does not assume that the two population variances are equal. See Ott and Hildebrand (1983) for reference to Welch's test, p. 299. Paired t-tests indicated that all coefficients except one were significantly different at the one percent level; the other coefficient was significantly different at the five percent level.

[10] To examine the robustness of the results across alternative methodologies, a nonparametric method, recursive partitioning (see Frydman et al., 1985), also was examined. Results were similar to the results found for the logit model: variables chosen by the method varied by sample, and models based on matched-pair samples produced inconsistent overall classification rates when applied to the non-matched-pair sample. Consistent with previous research, classification rates were lower for the recursive partitioning model compared to the logit model.

[11] Another inference that could be drawn from the results of this study, as compared with the higher classification rates from earlier studies, is that life insurer insolvency prediction has become less tractable over time. Use of various instruments and "window dressing" techniques may have increased over time, and may mask an insurer's soundness—e.g., financial reinsurance, loss reserve adjusting, and the use of surplus notes that increase an insurer's assets without a similar increase in liabilities.

[12] The joint and pairwise chi-square tests for differences in proportions used here are described in Conover (1971, pp. 141–154), and were used previously in an insolvency study by BarNiv (1990). The test statistic is compared to a chi-square distribution with one degree of freedom for the pairwise tests and a chi-square distribution with three degrees of freedom for the joint tests.

REFERENCES

Altman, E. I. (1968) "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *Journal of Finance*, 23, pp. 589–609.

Altman, E. I., Y. H. Eom, and D. W. Kim (1995) "Failure Prediction: Evidence from Korea," *Journal of International Financial Management and Accounting*, 6, pp. 230–249.

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

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

A.M. Best Company. (1992) *Best's Insolvency Study, Life-Health*, Special Report (June).

Baranoff, E. G., T. W. Sager, and T. S. Shively (2000) "A Semiparametric Stochastic Spline Model as a Managerial Tool for Potential Insolvency," *Journal of Risk and Insurance*, 67, pp. 369–396.

BarNiv, R. (1990) "Accounting Procedures, Market Data, Cash-Flow Figures, and Insolvency Classification: The Case of the Insurance Industry," *The Accounting Review*, 65, pp. 578–604.

BarNiv, R. and R. A. Hershbarger (1990) "Classifying Financial Distress in the Life Insurance Industry," *Journal of Risk and Insurance*, 57, pp. 110–136.

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

BarNiv, R. and A. Raveh (1989) "Identifying Financial Distress: A New Nonparametric Approach," *Journal of Business Finance and Accounting*, 16, pp. 361–384.

Beaver, W. H. (1966) "Financial Ratios as Predictors of Failure," *Journal of Accounting Research*, Supplement, pp. 71–111.

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, pp. 309–327.

Browne, M. J., J. M. Carson, and R. E. Hoyt (1999) "Economic and Market Predictors of Insolvencies in the Life-Health Insurance Industry," *Journal of Risk and Insurance*, 66, pp. 643–659.

Carson, J. M. and R. E. Hoyt (1995) "Life Insurer Financial Distress: Classification Models and Empirical Evidence," *Journal of Risk and Insurance*, 62, pp. 764–775.

Conover, W. J. (1971) *Practical Nonparametric Statistics.* New York, NY: Wiley and Sons.

Cummins, J. D., S. E. Harrington, and R. Klein (1995) "Insolvency Experience, Risk-based Capital, and Prompt Corrective Action in Property-Liability Insurance," *Journal of Banking and Finance*, 19, pp. 511–527.

Dambolena, I. G., and S. J. Khoury (1980) "Ratio Stability and Corporate Failure," *Journal of Finance*, 35, pp. 1017–1026.

Frydman, H., E. Altman, and D. Hao (1985) "Introducing Recursive Partitioning Analysis for Financial Classification: The Case of Financial Distress Classification," *Journal of Finance*. 40, pp. 269–291.

Garven, J. R. (1992) "An Exposition of the Implications of Limited Liability and Asymmetric Taxes for Property-Liability Insurance," *Journal of Risk and Insurance*, 59, pp. 34–56.

Grace, M. F., S. E. Harrington, and R. W. Klein (1998) "Risk-Based and Solvency Screening in Property-Liability Insurance: Hypotheses and Empirical Tests," *Journal of Risk and Insurance*, 65, pp. 213–243.

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

Judge, G. G., R. C. Hill, W. E. Griffiths, H. Lutkepohl, and T. Lee (1988) Introduction to the Theory and Practice of Econometrics. New York, NY: Wiley and Sons.

Khanna, N. and A. B. Poulsen (1995) "Managers of Financially Distressed Firms: Villains or Scapegoats?," *Journal of Finance*, 50, pp. 919–940.
Lamm-Tennant, J. and L. T. Starks (1993) "Stock versus Mutual Ownership Structures: The Risk Implications," *Journal of Business*, 66, pp. 29–46.
Lamm-Tennant, J., L. Starks, and L. Stokes (1996) "Considerations of Cost Tradeoffs in Insurance Solvency Surveillance Policy," *Journal of Banking and Finance*, 20, pp. 835–852.
Manski, C. F., and S. R. Lerman (1977) "The Estimation of Choice Probabilities From Choice Based Samples," *Econometrica*, 45 (Nov), pp. 1977–1988.
Meyer, P.A., and H. W. Pifer (1970) "Prediction of Bank Failures," *Journal of Finance*, 25, pp. 853–868.
Ohlson, J. A. (1980) "Financial Ratios and the Probabilistic Prediction of Bankruptcy," *Journal of Accounting Research*, 18, pp. 109–131.
Ott, L. and D. K. Hildebrand (1983) *Statistical Thinking for Managers*. Boston, Mass: Duxbury Press.
Palepu, K. G. (1986) "Predicting Takeover Targets: A Methodological and Empirical Analysis," *Journal of Accounting and Economics*, 8, pp. 3–35.
Pottier, S. W. (1998) "Life Insurer Financial Distress, Best's Ratings, and Financial Ratios," *Journal of Risk and Insurance*, 65, pp. 275–288.
Trieschmann, J. S. and G. E. Pinches (1973) "A Multivariate Model for Predicting Financially Distressed P-L Insurers," *Journal of Risk and Insurance*, 40, pp. 327–347.
Vaughan, Terri (2002) "The NAIC's 2002 Agenda: Toward a More Efficient System of Insurance Regulation," *Journal of Insurance Regulation*, 20(3), pp. 251–270.
Zavgren, C. V. (1985) "Assessing the Vulnerability to Failure of American Industrial Firms: A Logistic Analysis," *Journal of Business Finance and Accounting*, 12, pp. 19–45.
Zmijewski, M. (1984) "Methodological Issues Related to the Estimation of Financial Distress Prediction Models," *Journal of Accounting Research*, Supplement (22), pp. 59–82.