# Forecasting Bankruptcy More Accurately:
# A Simple Hazard Model

Tyler Shumway[*]

University of Michigan Business School

701 Tappan Street, Ann Arbor, Michigan 48109-1234

Phone: (313) 763-4129, E-Mail: *shumway@umich.edu*

Web: *http://www-personal.umich.edu/˜shumway*

July 16, 1999

## Abstract

I argue that hazard models are more appropriate for forecasting bankruptcy than the single-period models used previously. Single-period bankruptcy models give biased and inconsistent probability estimates while hazard models produce consistent estimates. I describe a simple technique for estimating a discrete-time hazard model with a logit model estimation program.

Applying my technique, I find that about half of the accounting ratios that have been used in previous models are not statistically significant bankruptcy predictors. Moreover, several market-driven variables are strongly related to bankruptcy probability, including market size, past stock returns, and the idiosyncratic standard deviation of stock returns. I propose a model that uses a combination of accounting ratios and market-driven variables to produce more accurate out-of-sample forecasts than alternative models.

# 1 Introduction

Economists and accountants have been forecasting bankruptcy for decades.[1] Most researchers have estimated single-period classification models, which I refer to as *static* models, with multiple-period bankruptcy data. By ignoring the fact that firms change through time, static models produce bankruptcy probabilities that are biased and inconsistent estimates of the probabilities that they approximate. Test statistics that are based on static models give incorrect inferences. I propose a hazard model that is simple to estimate, consistent, and accurate.

Static models are inappropriate for forecasting bankruptcy because of the nature of bankruptcy data. Since bankruptcy occurs infrequently, forecasters use samples that span several years to estimate their models.[2] The characteristics of most firms change from year to year. However, static models can only consider one set of explanatory variables for each firm. Researchers that apply static models to bankruptcy have to select when to observe each firm's characteristics. Most forecasters choose to observe each bankrupt firm's data in the year before bankruptcy. They ignore data on healthy firms that eventually go bankrupt. By choosing when to observe each firm's characteristics arbitrarily, forecasters that use static models introduce an unnecessary selection bias into their estimates.

I develop a simple hazard model that uses all available information to determine each firm's bankruptcy risk at each point in time.[3] While static models produce biased and inconsistent bankruptcy probability estimates, the hazard model proposed here is consistent in general and unbiased in some cases. Estimating hazard models with the accounting variables used previously by Altman (1968) and Zmijewski (1984) reveals that half of these variables are statistically unrelated to bankruptcy probability. I develop a new bankruptcy model that uses three market-driven variables to identify failing firms. My new model outperforms alternative models in out-of-sample forecasts.

---

[1] See Altman (1993) for a survey of forecasting models.

[2] For example, Altman's (1968) original bankruptcy sample spans twenty years. The sample used in this paper includes bankruptcies observed over 31 years.

[3] Hazard models are described in Kiefer (1988) and Lancaster (1990).

## 1.1 Advantages of Hazard Models

Hazard models resolve the problems of static models by explicitly accounting for time. The dependent variable in a hazard model is the time spent by a firm in the healthy group. When firms leave the healthy group for some reason other than bankruptcy (e.g., merger), they are considered censored, or no longer observed. Static models simply consider such firms healthy. In a hazard model, a firm's risk for bankruptcy changes through time and its health is a function of its latest financial data and its age. The bankruptcy probability that a static model assigns to a firm does not vary with time.

In econometric terms, there are three reasons to prefer hazard models for forecasting bankruptcy. The first reason is that static models fail to control for each firm's period at risk. When sampling periods are long, it is important to control for the fact that some firms file for bankruptcy after many years of being at risk while other firms fail in their first year. Static models do not adjust for period at risk, but hazard models adjust for it automatically. The selection bias inherent in static bankruptcy models is a result of their failure to correct for period at risk.

The second reason to prefer hazard models is that they incorporate time-varying covariates, or explanatory variables that change with time. If a firm deteriorates before bankruptcy, then allowing its financial data to reveal its changing health is important. Hazard models exploit each firm's time-series data by including annual observations as time-varying covariates. Unlike static models, they can incorporate macroeconomic variables that are the same for all firms at a given point of time. Hazard models can also account for potential *duration dependence*, or the possibility that firm age might be an important explanatory variable.

The third reason that hazard models are preferable is that they may produce more efficient out-of-sample forecasts by utilizing much more data. The hazard model can be thought of as a binary logit model that includes each firm-year as a separate observation. Since firms in the sample have an average of ten years of financial data, approximately ten times more data is available to estimate the hazard model than is available to estimate

2

corresponding static models. This results in more precise parameter estimates and superior forecasts.

## 1.2 Empirical Issues

Hazard models are preferable to static models both theoretically and empirically. Comparing the out-of-sample forecasting ability of hazard models to that of Altman (1968) and Zmijewski (1984), I find that hazard models perform as well as or better than alternatives. Furthermore, hazard models often produce dramatically different statistical inferences than static models. For example, estimating hazard models reveals that about half of the accounting ratios that have been used to forecast bankruptcy are not statistically related to failure. Since previous models use independent variables with little or no explanatory power, I search for a new set of independent variables to develop a more accurate model.

The most accurate out-of-sample forecasts that I can generate are calculated with a hazard model that uses both market-driven and accounting variables to identify bankrupt firms. The market variables include market size, past stock returns, and the idiosyncratic standard deviation of stock returns. I combine these market variables with the ratio of net income to total assets and the ratio of total liabilities to total assets to estimate a model that classifies 75 percent of failing firms in the top decile of firms ranked annually by bankruptcy probability.

## 1.3 Related Research

Precise bankruptcy forecasts are of great interest to academics, practitioners, and regulators. Regulators use forecasting models to monitor the financial health of banks, pension funds, and other institutions. Practitioners use default forecasts in conjunction with models like that of Duffie and Singleton (1997) to price corporate debt. Academics use bankruptcy forecasts to test various conjectures like the hypothesis that bankruptcy risk is priced in stock returns (e.g. Dichev, 1997). Given the broad interest in accurate forecasts, a superior forecasting technology is valuable.

Most previous bankruptcy forecasting models are subject to the criticism of this paper. The models of Altman (1968), Altman, Haldeman, and Narayanan (1977), Ohlson (1980), Zmijewski (1984), Lau (1987) and those of several other authors are misspecified. Some authors have addressed the deficiencies of existing bankruptcy models. Queen and Roll (1987) and Theodossiou (1993) develop dynamic forecasting models. This paper builds on the work of these researchers by explicitly addressing the bias in static models and developing a consistent model.

Bankruptcy forecasters are not the only researchers that can benefit from the results of this paper. Forecasters of corporate mergers have also applied static models to multiple-period data sets. In particular, the merger model of Palepu (1986) is biased and inconsistent in the same way as the bankruptcy studies listed above. Other authors, such as Pagano, Panetta, and Zingales (1998) and Denis, Denis, and Sarin (1997) estimate multiple-period logit models that can be interpreted as hazard models. This paper concentrates on the bankruptcy forecasting literature because it includes some of the most obvious misapplications of single-period models, but the results reported here are relevant for other areas of empirical finance as well.

## 2  Hazard versus Static Models

It is important to specify exactly what sort of bankruptcy data is available before discussing alternative models. For simplicity, I assume that bankruptcy can only occur at discrete points in time, $t = 1, 2, 3, \dots$. Most bankruptcy samples contain data on $n$ firms that all existed for some time between $t = 1$ and $t = T$. Each firm either fails during the sample period, survives the sample period, or it leaves the sample for some other reason such as a merger or a liquidation. Define a "failure" time, $t_i$, for each firm (indexed by $i$) as the time when the firm leaves the sample for any reason. Let a dummy variable, $y_i$, equal one if firm $i$ failed at $t_i$ and let it equal zero otherwise, and let the probability mass function of failure be given by $f(t, x; \theta)$, where $\theta$ represents the vector of parameters of $f$ and $x$ represents a

vector of explanatory variables used to forecast failure.

## 2.1   Similarities Between Hazard and Static Models

To facilitate comparison between static and hazard models, only maximum likelihood models are discussed in this section. The static models considered here have likelihood functions of the form

$$\mathcal{L} = \prod_{i=1}^{n} F(t_i, x_i; \theta)^{y_i} [1 - F(t_i, x_i; \theta)]^{1-y_i}, \tag{1}$$

where $F$ is the cumulative density function (CDF) that corresponds to $f(t, x; \theta)$. While there are a number of models with likelihood functions of this form, I refer to all models that pertain to this family as logit models for simplicity.

Describing hazard models requires a few more definitions. Following hazard model conventions, the survivor function, $S(t, x; \theta)$, and the hazard function, $\phi(t, x; \theta)$, are defined as

$$S(t, x; \theta) = 1 - \sum_{j<t} f(j, x; \theta), \quad \phi(t, x; \theta) = \frac{f(t, x; \theta)}{S(t, x; \theta)}. \tag{2}$$

The survivor function gives the probability of surviving up to time $t$ and the hazard function gives the probability of failure at $t$ conditional on surviving to $t$. The hazard model's likelihood function is:

$$\mathcal{L} = \prod_{i=1}^{n} \phi(t_i, x_i; \theta)^{y_i} S(t_i, x_i; \theta). \tag{3}$$

A parametric form for the hazard function, $\phi(t_i, x_i; \theta)$, is often assumed. The model can incorporate time-varying covariates by making $x$ depend on time.

Hazard and static models are closely related. To make the relation between the models clear, I define a multiperiod logit model as a logit model that is estimated with data on each firm in each year of its existence as if each firm-year were an independent observation. The dependent variable in a multiperiod logit model is set equal to one only in the year in which a bankruptcy filing occurred. The following proposition illustrates the link between hazard and multiperiod logit models.

**Proposition 1.** *A multiperiod logit model is equivalent to a discrete-time hazard model with hazard function $F(t, x; \theta)$.*

**Proof:**

Since a multiperiod logit model is estimated with the data from each firm-year as if it were a separate observation, its likelihood function is

$$\mathcal{L} = \prod_{i=1}^{n} \left( F(t_i, x_i; \theta)^{y_i} \prod_{j < t_i} [1 - F(j, x_i; \theta)] \right). \tag{4}$$

As a CDF, $F$ is strictly positive and bounded by one. Since $F$ depends on $t$ and it is positive and bounded, it can be interpreted as a hazard function. Replacing $F$ with the hazard function $\phi$,

$$\mathcal{L} = \prod_{i=1}^{n} \left( \phi(t_i, x_i; \theta)^{y_i} \prod_{j < t_i} [1 - \phi(j, x_i; \theta)] \right). \tag{5}$$

Finally, Cox and Oakes (1984) show that the survivor function for a discrete-time hazard model satisfies

$$S(t, x; \theta) = \prod_{j < t} [1 - \phi(j, x; \theta)]. \tag{6}$$

Substituting equation (6) into (5) demonstrates that the likelihood function of a multiperid logit model is equalivant to the likelihood function of a discrete-time hazard model, (3), with hazard rate $\phi(t, x; \theta) = F(t, x; \theta)$. $\square$

## 2.2   Econometric Properties of Hazard and Static Models

Given the relationship between hazard and static models explained above, it is possible to see both the source and the effect of the selection bias in previous bankruptcy forecasting models. This section illustrates the bias with a simple example. It also presents a fairly general argument for the inconsistency of static models and the consistency of hazard models. Finally, it discusses problems of statistical inference and efficiency inherent in static models.

### 2.2.1 Consistency: A Simple Example

Suppose that there are two periods in which bankruptcy is possible. A dummy variable, $y_{it}$, is set to one if firm $i$ goes bankrupt in period $t$. In each period, each firm has a nonstochastic covariate, $x_{it}$, which only takes on values of zero or one. The covariate is related to the firm's bankruptcy probability by

$$\text{Prob}(y_{it} = 1) = \theta x_{it}. \tag{7}$$

There are $N$ firms for which both $y_{it}$ and $x_{it}$ are observable in period one. In period two, only firms that did not go bankrupt in period one are observable. Each firm's observation is assumed to be independently and identically distributed (i.i.d.). The problem is to estimate $\theta$ given the available data.

Consider first the hazard model estimator for $\theta$. The model of bankruptcy assumed above stipulates that a firm's risk is independent of its age. The discrete-time hazard model described by Proposition 1 has a hazard rate equal to the CDF of $y$. Thus, the hazard function for this problem is equal to the probability of bankruptcy ($\phi = F = \theta x_{it}$), and the (log) likelihood function for the model is

$$\mathcal{L}_H = \ln \left\{ \prod_{i=1}^{N} (\theta_H x_{i1})^{y_{i1}} \left[ (1 - \theta_H x_{i1})(\theta_H x_{i2})^{y_{i2}} (1 - \theta_H x_{i2})^{(1-y_{i2})} \right]^{(1-y_{i1})} \right\}. \tag{8}$$

The terms involving values in period two are raised to the power $(1 - y_{i1})$ because they are only observed when the firm does not go bankrupt in period one.

The first order condition for the maximization of this likelihood function is

$$\frac{\partial \mathcal{L}_H}{\partial \hat{\theta}_H} = \sum_{i=1}^{N} \left\{ \frac{y_{i1}}{\hat{\theta}_H} + (1 - y_{i1}) \left[ \frac{-x_{i1}}{(1 - \hat{\theta}_H x_{i1})} + \frac{y_{i2}}{\hat{\theta}_H} - \frac{(1 - y_{i2})x_{i2}}{(1 - \hat{\theta}_H x_{i2})} \right] \right\} = 0. \tag{9}$$

Using the fact that both $x_{it}$ and $y_{it}$ can only take values of zero or one, this expression can

be simplified to

$$\sum_{i=1}^{N} \frac{y_{i1} + (1 - y_{i1})y_{i2}}{\hat{\theta}_H} = \sum_{i=1}^{N} \frac{(1 - y_{i1})x_{i1} + (1 - y_{i1})(1 - y_{i2})x_{i2}}{(1 - \hat{\theta}_H)}, \qquad (10)$$

which leads to the maximum likelihood estimator,

$$\hat{\theta}_H = \frac{\sum_{i=1}^{N} (y_{i1} + (1 - y_{i1})y_{i2})}{\sum_{i=1}^{N} (y_{i1} + (1 - y_{i1})y_{i2} + (1 - y_{i1})x_{i1} + (1 - y_{i1})(1 - y_{i2})x_{i2})}. \qquad (11)$$

Since firms with $x_{it} = 0$ have no probability of failure and firms with $y_{i1} = 1$ are not observed in period two, this can be simplified to

$$\hat{\theta}_H = \frac{\sum_{i=1}^{N}(y_{i1} + y_{i2})}{\sum_{i=1}^{N}(x_{i1} + x_{i2})}. \qquad (12)$$

Notice that this is a natural estimate of bankruptcy probability. The numerator is equal to the total number of failures observed while the denominator is the total number of firms at risk of failure in both periods. Furthermore, since $E(y_{it}) = \theta x_{it}$, $\hat{\theta}_H$ is unbiased for $\theta$. Under the i.i.d. assumption made above, $\hat{\theta}_H$ is also consistent for $\theta$ by the law of large numbers.

Now consider the static estimator in the same problem. This estimator takes only one input from each firm. Firms that go bankrupt in period one are recorded at bankruptcy, and all other firms are recorded in period two. The (misspecified) likelihood function for this estimator is

$$\mathcal{L}_S = \ln \left\{ \prod_{i=1}^{N} (\theta_S x_{i1})^{y_{i1}} \left[ (\theta_S x_{i2})^{y_{i2}} (1 - \theta_S x_{i2})^{(1 - y_{i2})} \right]^{(1 - y_{i1})} \right\}. \qquad (13)$$

with the first order condition,

$$\frac{\partial \mathcal{L}_S}{\partial \hat{\theta}_S} = \sum_{i=1}^{N} \left( \frac{y_{i1}}{\hat{\theta}_S} + (1 - y_{i1}) \left[ \frac{y_{i2}}{\hat{\theta}_S} - \frac{(1 - y_{i2})x_{i2}}{(1 - \hat{\theta}_S x_{i2})} \right] \right) = 0. \qquad (14)$$

Comparing (14) to (9) reveals that the static estimator's first order condition is missing the

8

$\frac{y_{i2}}{\theta}$ term that is in equation (9). Otherwise, the conditions are identical. Using arguments similar to those above, the static model's maximum likelihood condition can be restated as,

$$\sum_{i=1}^{N} \frac{y_{i1} + (1 - y_{i1})y_{i2}}{\hat{\theta}_S} = \sum_{i=1}^{N} \frac{(1 - y_{i1})(1 - y_{i2})x_{i2}}{1 - \hat{\theta}_S}, \tag{15}$$

which produces,

$$\hat{\theta}_S = \frac{\sum_{i=1}^{N}(y_{i1} + y_{i2})}{\sum_{i=1}^{N}(y_{i1} + x_{i2})}. \tag{16}$$

This static estimator equals the total number of failures divided by the number of failures in period one plus the number of firms at risk of failure in period two. It neglects to consider firms at risk of bankruptcy in period one. Thus, it produces biased and inconsistent estimates. The bias in this estimator can be written as

$$E[\hat{\theta}_S] - \theta = E\left\{ \frac{\sum_{i=1}^{N}(y_{i1} + y_{i2}) \sum_{i=1}^{N}(x_{i1} - y_{i1})}{\sum_{i=1}^{N}(x_{i1} + x_{i2}) \sum_{i=1}^{N}(y_{i1} + x_{i2})} \right\}. \tag{17}$$

Since the denominator in (17) is always positive and the expected value of the numerator is positive, the bias in the static model's estimator is positive. This is consistent with what intuition suggests. The static model's estimates of $\theta$ are too large because they neglect to consider firms that don't go bankrupt even though they are at risk.

This simple example ignores many common complications. It assumes a simple structure and just two periods. In the next subsection, the consistency of more general static estimators is explored.

### 2.2.2 Consistency: The More General Case

The simple example developed above is easily generalized. Before presenting the general argument, three important assumptions must be explained.

**Assumption 1.** *The static model is correctly specified for one period.*

9

In particular, the period-$\tau$ likelihood function,

$$\mathcal{L}_\tau(\theta) = \sum_{i=1}^N \left\{ y_{i\tau} \ln[\mathrm{P}(y_{i\tau} = 1 | x_{i\tau}, \theta)] + (1 - y_{i\tau}) \ln[1 - \mathrm{P}(y_{i\tau} = 1 | x_{i\tau}, \theta)] \right\}, \qquad (18)$$

satisfies all the assumptions that are usually made in order to prove that $\hat{\theta}$ is consistent for $\theta$ (Amemiya, 1985). One of the consistency assumptions made about (18) is that $\mathcal{L}_\tau(\theta)/N$ converges in probability uniformly (as $N \to \infty$) to a nonstochastic function, $Q_\tau(\theta)$, which attains a unique global maximum at the true value of $\theta$.

**Assumption 2.** $Q_\tau(\theta)$ can be represented as the sum $Q_{\tau 1}(\theta) + Q_{\tau 2}(\theta)$, where $Q_{\tau 1}(\theta)$ is the limit of $\frac{1}{N} \sum_{i=1}^N y_{i\tau} \ln[\mathrm{P}(y_{i\tau} = 1 | x_{i\tau}, \theta)]$ and $Q_{\tau 2}(\theta)$ is the limit of $\frac{1}{N} \sum_{i=1}^N (1 - y_{i\tau}) \ln[1 - \mathrm{P}(y_{i\tau} = 1 | x_{i\tau}, \theta)]$.

This assumption is fairly innocuous, but it makes the argument for consistency simple. With Assumptions 1 and 2, the true value of $\theta$ maximizes the function $Q_\tau(\theta) = Q_{\tau 1}(\theta) + Q_{\tau 2}(\theta)$ for any single time period, $\tau$.

**Assumption 3.** *The log likelihood function for each period is sufficiently well specified to satisfy the independence property,*

$$\mathcal{L}(\theta | y_{i\tau}, y_{i\tau+k}, x_{i\tau}, x_{i\tau+k}) = \mathcal{L}_\tau(\theta | y_{i\tau}, x_{i\tau}) + \mathcal{L}_{\tau+k}(\theta | y_{i\tau+k}, x_{i\tau+k}),$$

*for any $k$.*

Assumption 3 is a conditional independence condition that is analogous to the common econometric assumption that the model is sufficiently well specified to guarantee that the error terms of different observations are independent of each other. This assumption will be violated when some unobserved heterogeneity among firms is correlated with failure. Econometricians have developed a number of models that correct this problem (Lancaster, 1990). Rather than complicate the current model with assumptions about unobserved heterogeneity, I assume that all heterogeneity among firms is captured by the variables used to forecast failure, $x_{it}$.

With these three assumptions, it is easy to show that hazard models are consistent.

**Proposition 2.** *Under Assumptions 1-3, a discrete-time hazard model estimator is consistent for $\theta$, but a simple static model estimator is generally inconsistent.*

**Proof:**

Consider the joint log likelihood function for $y_{i1}, y_{i2}, \ldots, y_{iT}$,

$$\mathcal{L}_{1,T}(\theta) = \mathcal{L}_1 + \mathcal{L}_2 + \ldots + \mathcal{L}_T \tag{19}$$

This is exactly the likelihood function that the hazard model maximizes. Under Assumptions 1 through 3, maximizing this joint likelihood function produces a consistent estimator for $\theta$.

By contrast, consider the objective function of the static estimator in this general framework. In both periods one and two,

$$\begin{aligned} \theta &= \text{arg max}_\theta \ Q_{11}(\theta) + Q_{12}(\theta) \\ \theta &= \text{arg max}_\theta \ Q_{21}(\theta) + Q_{22}(\theta). \end{aligned} \tag{20}$$

Adding periods one and two together, it must be true that

$$\theta = \text{arg } \max_\theta \ Q_{11}(\theta) + Q_{12}(\theta) + Q_{21}(\theta) + Q_{22}(\theta). \tag{21}$$

However, as long as $Q_{12}(\theta)$ is not equal to a constant, the true value of $\theta$ will not maximize the function,

$$\theta \neq \text{arg } \max_\theta \ Q_S(\theta) = Q_{11}(\theta) + Q_{21}(\theta) + Q_{22}(\theta), \tag{22}$$

but this is exactly what the likelihood function of the static model converges to for this two-period problem. Thus, for this problem, the static model's estimate of $\theta$ is generally inconsistent for the true value of $\theta$.

A more general representation of the objective function of the static estimator is

$$G_{S;1,T} = \sum_{i=1}^{N} y_{i1} \ln[\mathrm{P}(y_{i1} = 1|x_{i1}, \theta)] + y_{i2} \ln[\mathrm{P}(y_{i2} = 1|x_{i2}, \theta)] + \dots$$
$$+ y_{iT} \ln[\mathrm{P}(y_{iT} = 1|x_{iT}, \theta)] + (1 - y_{iT}) \ln[1 - \mathrm{P}(y_{iT} = 1|x_{iT}, \theta)].$$

(23)

Under Assumptions 1 through 3, the function $G_{S;1,T}/N$ converges in probability uniformly to a form similar to (22). Since the true value of $\theta$ does not maximize the limitting function of (23), static estimators are not consistent in general. □

## 2.3   Inference and Efficiency

Since the parameter estimates produced by static models are biased and inconsistent, tests of statistical significance performed with static models are invalid. Thus, it is not clear that the variables associated with bankruptcy by static models are significant predictors. This issue is explored in detail in the empirical work below.

The connection between the hazard and logit models implies that even if static models were consistent, hazard models should be more accurate. While each firm has a time-series of annual observations, static models are estimated only with each firm's last observation. Hazard models take advantage of much more data. They are equivalent to logit models in which no firm-year data points have been excluded. Unlike static models, hazard models exploit all of the data available. Thus, for both consistency and efficiency, hazard models are preferable to static models.

# 3   Estimating the Hazard Function

The previous section shows that hazard models are superior to static models for forecasting bankruptcy. In practice, however, many hazard models are difficult to estimate because of their nonlinear likelihood functions and time-varying covariates. Proposition 1 implies that it is possible to estimate discrete-time hazard models with a computer program that estimates logit models. To estimate a hazard model with a logit program, each year in

which the firm survives is included in the logit program's "sample" as a firm that did not fail. Each bankrupt firm contributes only one failure observation ($y_{it} = 1$) to the logit model. Time-varying covariates are incorporated simply by using each firm's annual data for its firm-year logit observations. Estimating hazard models with a logit program is so simple and intuitive that it has been done by academics and regulators without a hazard model justification.[4]

Making statistical inferences in a hazard model estimated with a logit program is simple. Since the logit and hazard models have the same likelihood function, they have the same asymptotic variance-covariance matrix (Amemiya, 1985). However, the test statistics produced by a logit program are incorrect for the hazard model because they assume that the number of independent observations used to estimate the model is the number of firm-years in the data. Calculating correct test statistics requires adjusting the sample size assumed by the logit program to account for the lack of independence between firm-year observations.[5] For the hazard model, each firm's entire life span is one observation. Thus, the correct value of $n$ for test statistics is the number of firms in the data, not the number of firm-years. The $\chi^2$ test statistics produced by logit programs are of the form

$$\frac{1}{n}(\hat{\mu}_k - \mu_0)^{'}\Sigma^{-1}(\hat{\mu}_k - \mu_0) \sim \chi^2(k), \tag{24}$$

where there are $k$ estimated moments being tested against $k$ null hypotheses, $\mu_0$. Dividing these test statistics by the average number of firm-years per firm makes the logit program's statistics correct for the hazard model. Unreported estimates of a proportional hazard model confirm that standard hazard models produce coefficient estimates and test statistics that are similar to those produced by the discrete-time hazard model described here.

Logit models in which several observations exist for each individual usually account

---

[4]Pagano, Panetta, and Zingales (1995) and Denis, Denis and Sarin (1997) use models like the hazard model described here to forecast initial public offerings and executive turnover. The Pension Benefits Guarantee Corporation forecasts bankruptcies by estimating a logit model by firm-year.

[5]The firm-year observations of a particular firm cannot be independent, since a firm cannot fail in period $t$ if it failed in period $t - 1$.

for the lack of independence between observations that is characteristic of panel data (Amemiya, 1985). The logit model used here is already penalized for the lack of independence between firm-year observations by the sample size adjustment described above. Since it does not assume that firm-year observations are i.i.d., no more adjustment for dependence should be necessary.

Interpreting the logit model as a hazard model can clarify the meaning of the model's coefficients. Partitioning $\theta$ into $\theta_1$ and $\theta_2$, the hazard function for the discrete-time hazard model can be written as

$$\phi(t, x; \theta_1, \theta_2) = \frac{1}{1 + \text{Exp}(g(t)'\theta_1 + x'\theta_2)}. \tag{25}$$

If the function of firm age selected, $g(t)$, is the natural logarithm of age then the hazard model is an accelerated failure-time model (see Lancaster, 1990). Coefficients can be interpreted with the regression equation,

$$E[\ln(t)|x] = -\frac{x'\theta_2}{\theta_1}. \tag{26}$$

Alternatively, omitting firm age variables from the model is analogous to estimating an exponential hazard model in which a firm's probability of failure does not depend on its age. In general, any function of age can be included in the model. This makes the discrete-time hazard model more flexible than many common parametric models.

# 4    The Data

To compare hazard to static model forecasts, I estimate both hazard and static models and examine their out-of-sample accuracy. Only firms in the intersection of the *Compustat Industrial File* and the *CRSP Daily Stock Return File* for NYSE and AMEX stocks are included in the sample. Firms that began trading before 1962 or after 1992 are excluded. Firms with CRSP SIC codes from 6000 to 6999 (financial firms) are also excluded. Table

I provides summary statistics for all of the independent variables described below.

## 4.1 Bankruptcy Data

I collected bankruptcy data from the *Wall Street Journal Index*, the *Capital Changes Reporter*, and the *Compustat Research File*. I also searched for firms whose stock was delisted from the NYSE or AMEX in the *Directory of Obsolete Securities* (1993) and *Nexis*. All firms that filed for any type of bankruptcy within five years of delisting are considered bankrupt. The final sample contains 300 bankruptcies between 1962 and 1992.

The variable of interest in the hazard model is firm age. In this paper, a firm's age is defined as the number of calendar years it has been traded on the NYSE or AMEX. So, for example, if a firm began trading on the NYSE in 1964 and then merged in 1965, it would contribute two firm-year observations to the logit model. One observation would give the firm's age as one year and the other would indicate that the firm's age was two years. The dependent variable associated with both of these observations would be equal to zero, indicating no bankruptcy occurred. If the firm filed for bankruptcy, only its second firm-year observation would have a dependent variable value of one.

I use the firm's trading age as the variable to be explained because there is no attractive alternative to measure how long the firm has been a viable enterprise. Since a firm must meet a number of requirements to be listed by an exchange, firms are fairly homogeneous when initially listed. However, a firm can be incorporated as a small speculative concern or as a large holding company, making the firm's age since incorporation less economically meaningful than its age since listing. In the hazard models I estimate, firm age is never statistically significant after controlling for other firm characteristics.

## 4.2 Independent Variables

I estimate models with several different sets of independent variables. The forecasting models incorporate Altman's (1968) and Zmijewski's (1984) independent variables, as well as some new market-driven independent variables described in this section.

Altman's variables are described extensively in Altman (1993). They include the ratios of working capital to total assets (WC/TA), retained earnings to TA (RE/TA), earnings before interest and taxes to TA (EBIT/TA), market equity to total liabilities (ME/TL), and sales to TA (S/TA). The COMPUSTAT item numbers that I used to construct Altman's variables appear with the variables' summary statistics in Table I.

In order to make my forecasting exercise realistic, I lag all data to ensure that the data are observable in the beginning of the year in which bankruptcy is observed. To construct Altman's (and Zmijewski's) variables, I lag COMPUSTAT data to ensure that each firm's fiscal year ends at least six months before the beginning of the year of interest. I lag the market-driven variables described below in a similar fashion.

There are a number of extreme values among the observations of Altman's ratios constructed from raw COMPUSTAT data. To ensure that statistical results are not heavily influenced by outliers, I set all observations higher than the 99th percentile of each variable to that value. All values lower than the first percentile of each variable are truncated in the same manner. Zmijewski's variables and the market-driven variables I introduce below are also truncated to avoid outliers. Unreported results with untruncated data are generally similar to the results I report. The minimum and maximum numbers reported in Table I are calculated after truncation.

Zmijewski's variables include the ratio of net income to total assets (NI/TA), the ratio of total liabilities to TA (TL/TA), and the ratio of current assets to current liabilities (CA/CL). As with Altman's variables, the COMPUSTAT item numbers used to construct each of these variables appears in Table I. The data are lagged and truncated as described above.

Because the market equity of firms that are close to bankruptcy is typically discounted by traders, firm size is a very important bankruptcy predicting variable. Each firm's market capitalization is measured at the end of the year before the observation year. To make size stationary, the logarithm of each firm's size relative to the total size of the NYSE and AMEX market is used. These data are all readily available in the CRSP database. The

average of relative size is negative because it is the logarithm of a generally small fraction.

If traders discount the equity of firms that are close to bankruptcy then a firm's past excess returns should predict bankruptcy as well as its market capitalization. I measure each firm's past excess return in year $t$ as the return of the firm in year $t-1$ minus the value-weighted CRSP NYSE/AMEX index return in year $t-1$. Each firm's annual returns are calculated by cumulating monthly returns. When some of a firm's monthly returns are missing, the value-weighted CRSP NYSE/AMEX index return is substituted for the missing returns. The average excess return reported in Table I is a small positive number because equal-weighted returns are typically higher than value-weighted returns.

The last market-driven variable that I use is the idiosyncratic standard deviation of each firm's stock returns, denoted sigma in the tables below. Sigma is strongly related to bankruptcy both statistically and logically. If a firm has more variable cash flows (and hence more variable stock returns) then the firm ought to have a higher probability of bankruptcy. Sigma may also measure something like operating leverage. I calculate each firm's sigma for year $t$ by regressing each stock's monthly returns in year $t-1$ on the value-weighted NYSE/AMEX index return for the same year. Sigma is the standard deviation of the residual of this regression. I drop values calculated with regressions based on less than twelve months of returns. To avoid outliers, relative size, past returns, and sigma are all truncated at the 99th and 1st percentile values in the same manner as all other independent variables.

Since a complete set of explanatory variables is not always observable for each firm-year, I substitute variable values from past years for missing values in some cases. This does not present an econometric problem because, for example, accounting ratios observed in year $t$ are still observable in years $t+1$ and $t+2$. By filling in missing data, the number of firm-years available to estimate Altman's model rises from 27,665 to 28,226. The number of bankruptcies available to identify rises from 201 to 229.

# 5 Forecasting Results

In this section, I report parameter estimates for various forecasting models and I compare the out-of-sample accuracy of all the models considered. Unreported estimates of analogous proportional hazard models are approximately proportional to the estimates reported.

## 5.1 Models with Altman's Variables

Table II reports the results of estimating models with Altman's variables. Panel A displays Altman's original discriminant function,[6] a new set of coefficients calculated by Begley, Ming, and Watts (1996)[7] and two functions calculated with my bankruptcy data. The coefficients in the third line of the table (1962-1983 data) are calculated only with data available in 1983. Those data consist of 1822 firms that have complete data for at least one year between 1962 and 1983, 118 of which filed for bankruptcy by 1983. The discriminant analysis (DA) calculations use either the last data reported (at least six months) before bankruptcy or the last data available in 1983 for each firm in the sample. Forecasts based on this function are compared to hazard model forecasts formed with data available in 1983 in Table III. The fifth line in Panel A (1962-1992 data) reports DA coefficients calculated with data available in 1992. Again, the calculations use only the last available set of data for each firm. Panel B reports hazard model coefficients for the same variables. The hazard model estimates are based on all available data (each firm-year) from 1962 to 1992. In unreported results with untruncated data, ME/TL loses its significance in the hazard model but the results are otherwise quite similar to those reported.

Both the hazard model and the DA coefficients confirm that firms with higher earnings relative to assets are less likely to fail. Larger firms with less liabilities and firms with higher working capital are also relatively safe. The effects of retained earnings and sales vary from model to model. The log of firm age is not statistically significant in the hazard

---

[6]Altman (1993) discusses the interpretation of discriminant analysis coefficients extensively.

[7]The published version of Begley, Ming, and Watts (1996) contains two typographical errors. The coefficients reported in Table II are correct.

model and its coefficient is quite small. There appears to be little duration dependence in bankruptcy probability.

The T-statistics reported in panel A are tests of the differences in the means of bankrupt and healthy firms. They indicate that all of Altman's variables are strong bankruptcy predictors. Performing the same test, Altman (1993) finds that all of his variables except S/TA are statistically significant predictors. Unfortunately, since the samples of healthy and bankrupt firms are not chosen randomly, the T-statistics of Panel A are biased and inconsistent. According to the hazard model, the only statistically significant variables are EBIT/TA and ME/TL. The discrepancy between these two findings is due to the bias inherent in Altman's method. As firms approach bankruptcy, their financial condition deteriorates. Altman's model drops observations on firms that will be bankrupt in two or three years. It neglects, for example, firms that have low values of WC/TA in a particular year, which go bankrupt the following year. Omitting such observations inflates test statistics.

Table III compares the out-of-sample accuracy of the models described above. To construct the table, I sort all firms each year from 1984 to 1992 into deciles based on their fitted probability values. Fitted probabilities (or rankings) are created by combining the coefficients from models estimated with 1983 data with the data available in each subsequent year. The table reports the percentage of bankrupt firms that are classified into each of the five highest probability deciles in the year in which they failed. It also lists the percentage of bankrupt firms classified among the least likely fifty percent of firms to fail. There are 111 bankrupt firms that have the accounting data required to evaluate the discriminant function between 1984 and 1992. Again, results with untruncated data are very similar to those in Table III.

By a reasonable margin, the most accurate model listed in Table III is the hazard model. The hazard model classifies almost seventy percent of all bankruptcies in the highest bankruptcy probability decile. It classifies 96.6 percent of bankrupt firms above the median probability. The discriminant analysis models cannot match this accuracy.

19

## 5.2  Models with Zmijewski's Variables

Table IV reports the results of estimating three models with Zmijewski's (1984) variables. Panel A reports Zmijewski's original estimates as well as estimates of Zmijeski's model calculated with my data. The second line reports the coefficients for Zmijewski's model estimated with data available in 1983 while the fourth line lists coefficients using data from 1962 to 1992. As with Altman's model, all of these estimates are calculated by observing only each firm's last available set of accounting data, and the test statistics reported on lines three and five are biased and inconsistent. The hazard model in Panel B exploits the entire time-series of accounting data available from 1962 to 1992.

The coefficients for Zmijewski's model are remarkably similar across models. Even the hazard model's estimated coefficients are close to the coefficients of the simple logit models in Panel A. As expected, firms with high income and low liabilities are less likely to fail than other firms. The current ratio (CA/CL) is not significantly related to bankruptcy in any of the estimates. Zmijewski also reports that CA/CL is not statistically significant in his model. The log of firm age is insignificant in the hazard model, confirming that there is little or no duration dependence in bankruptcy data.

While the coefficients are quite similar, the test statistics associated with each model are quite different. As in the case of discriminant analysis, Zmijewski's model appears to vastly overstate the statistical significance of the parameters. While according to Zmijewski's model both NI/TA and TL/TA are excellent bankruptcy predictors, according to the hazard model only the coefficient on NI/TA is significantly different from zero at the 99% level. This fact, combined with the fact that TL/TA and NI/TA are strongly correlated ($\rho = 0.40$), suggests that Zmijewski's model is essentially a one-variable model.

Curiously, when the data are not truncated, the sample correlation between NI/TA and TL/TA is -0.98. Neither NI/TA nor TL/TA are significant in the (unreported) hazard model with untruncated data, and the hazard model's forecasting accuracy is poor. Truncating the data to control for outliers is important for Zmijewski's model.

Table V compares the accuracy of logit and hazard estimates of Zmijewski's model.

Like Table III, Table V examines how many bankruptcies between 1984 and 1992 can be identified with the fitted values of a model estimated with data available before 1984. It also examines the accuracy of a simple ranking of firms by their ratio of net income to total assets. Unlike Table III, Table V does not report that the hazard model dominates alternative models. The hazard model does not even perform better than the NI/TA sort. Each of the models appears fairly accurate, assigning between 54% and 56% of bankrupt firms to the highest bankruptcy probability decile. However, none of the three models appears to add much explanatory power to NI/TA. This is not surprising, given that each of these models only includes one strong bankruptcy predictor. Thus, while it is a little disappointing that the hazard model does not outperform the logit model, it is not possible for one (monotonic) model to outperform another model if both are based on only one important bankruptcy predictor.

None of the forecasts made with Zmijewski's model are as successfull as the hazard model that uses Altman's variables in Table III. Still, the variables in these two models measure similar things. Both EBIT/TA and NI/TA measure the profitability of the firm, while both ME/TL and TL/TA measure the firm's leverage. A critical difference between Altman's and Zmijewski's variables is that Altman's ME/TL contains a value determined in equilibrium by market traders rather than by accounting conventions. In an effort to build bankruptcy models with more power, two models that incorporate other market-driven variables are described in the next section.

## 5.3   Models with Market-Driven Variables

Parameter estimates for two hazard models that include market-driven variables appear in Table VI. The model reported in Panel A forecasts bankruptcies with market-driven variables exclusively while the model in Panel B combines market-driven variables with two accounting ratios from Zmijewski's model. Because there is no evidence of duration dependence in bankruptcy probability, neither model contains the log of firm age as an explanatory variable. Both models are estimated with all data (each firm-year) from 1962

to 1992. An important advantage of the model that is based solely on market-driven variables is that firms without COMPUSTAT data can remain in the model's sample. The model in Panel A is estimated with 33,621 firm-years and 291 bankruptcies, while the model in panel B is estimated with only 28,664 firm-years and 239 bankruptcies. Estimates calculated with untruncated data are quite similar to those reported.

All of the coefficients in both models have the expected signs. Larger, less volatile firms with high past returns are safer than small, volatile firms with low past returns. High net income and low liabilities are again associated with low risk. While all three of the market-driven variables are statistically significant in Panel A, both NI/TA and sigma become insignificant when market variables and accounting ratios are combined in Panel B.

The accuracy of these models is examined in Table VII. As in Tables III and V, firms are sorted annually based on their implied bankruptcy probability, formed by combining parameter estimates based on 1983 data with the data available after 1983. The number of bankruptcies occuring bewteen 1984 and 1992 in each probability decile is reported. Combining accounting and market variables results in the most accurate model documented in this paper. This model classifies three-quarters of bankrupt firms in the highest bankruptcy decile, and it only classifies 3.5 percent of bankrupt firms below the bankruptcy probability median. The model based solely on market-driven variables performs quite well also, classifying 69 percent of bankrupt firms in the highest probability decile and 95 percent of bankrupt firms above the probability median. Bankruptcy forecasts can be improved dramatically by conditioning on market-driven variables.

# 6   Conclusion

This paper develops a hazard model for forecasting bankruptcy. The hazard model is theoretically preferable to the static models used previously because it corrects for period at risk and allows for time-varying covariates. It uses all available information to produce bankruptcy probability estimates for all firms at each point in time. By using all the

available data, it avoids the selection biases inherent in static models.

The hazard model is simple to estimate and interpret. A logit estimation program can be used to calculate maximum likelihood estimates. Test statistics for the hazard model can be derived from the statistics reported by the logit program. The hazard model can be interpreted either as a logit model done by firm-year or it can be viewed as a discrete accelerated failure-time model.

Estimating the hazard model with a set of bankruptcies observed over thirty-one years, I find that while half of the accounting ratios used previously are poor predictors, several previously neglected market-driven variables are strongly related to bankruptcy probability. A firm's market size, its past stock returns, and the idiosyncratic standard deviation of its stock returns all forecast failure. Combining these market-driven variables with two accounting ratios, I estimate a model that is quite accurate in out-of-sample tests.

# References

Altman, E. I., 1968, Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy, *Journal of Finance* 23, 589-609.

Altman, E. I., 1993. *Corporate Financial Distress and Bankruptcy: A Complete Guide to Predicting and Avoiding Distress and Profitting from Bankruptcy* (John Wiley and Sons, Inc., New York).

Altman, E. I., R. Haldeman, and P. Narayanan, 1977, ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations, *Journal of Banking and Finance* 10, 29-54.

Amemiya, T., 1985. *Advanced Econometrics* (Harvard University Press, Cambridge, MA)

Begley, J., J. Ming, and S. Watts, 1996, Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models, *Review of Accounting Studies* 1, 267-284.

Cox, D. R., and D. Oakes, 1984. *Analysis of Survival Data* (Chapman and Hall, New York).

Financial Stock Guide Service, 1993. *Directory of Obsolete Securities* (Financial Information, Inc., Jersey City, NJ).

Denis, D. J., D. K. Denis, and A. Sarin, 1997, Ownership Structure and Top Executive Turnover, *Journal of Financial Economics* 45, 193-221.

Dichev, I., 1998, Is the Risk of Bankruptcy a Systematic Risk?, *Journal of Finance* 53, 1131-1148.

Duffie, D. and K. Singleton, 1997, Modeling Term Structures of Defaultable Bonds, working paper, Stanford University.

Kiefer, N. M., 1988, Economic duration data and hazard functions, *Journal of Economic Literature* 26, 646-679.

Lancaster, T., 1990. *The Econometric Analysis of Transition Data* (Cambridge University Press, New York).

Lau, A. H. L., 1987, A Five-State Financial Distress Prediction Model, *Journal of Accounting Research* 18, 109-131.

Ohlson, J. S., 1980, Financial Ratios and the Probabilistic Prediction of Bankruptcy, *Journal of Accounting Research* 19 109-31.

Pagano, M., F. Panetta, and L. Zingales, 1998, Why Do Companies Go Public? An Empirical Analysis, *Journal of Finance* 53, 27-64.

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

Queen, M., and R. Roll, 1987, Firm mortality: Using market indicators to predict survival, *Financial Analysts Journal* May-June 1987, 9-26.

Theodossiou, P. T., 1993, Predicting shifts in the mean of a multivariate time series process: An application in predicting business failures, *Journal of the American Statistical Association* 88, 441-449.

Zmijewski, M. E., 1984, Methodological issues related to the estimation of financial distress prediction models, *Journal of Accounting Research* 22, 59-82.

# Table I
## Summmary Statistics

Table I reports summary statistics for all of the variables used to forecast bankruptcy. Each observation represents a particular firm in a particular year. The column labeled COMP lists the Compustat variable numbers used to construct the variables that are taken from Compustat. The column labeled Source indicates whether the variable in question has been used by previous researchers, with Alt indicating that the variable was used by Altman (1968) Zmi indicating that the variable was used by Zmijewski (1984) and Shu indicating that the variable is new to this paper. In the entire sample, there are 300 bankruptcies among 3182 firms and 39745 firm-years. The sample period is from 1962 to 1992. All variables are truncated at the 99th and 1st percentile, so the Minimum and Maximum quantities reported are actually those quantiles.

| Variable | COMP | Source | Mean | Median | Std. | Minimum | Maximum |
|----------|------|--------|------|--------|------|---------|---------|
| WC/TA | 179/6 | Alt | 0.289 | 0.300 | 0.200 | -0.165 | 0.736 |
| RE/TA | 36/6 | Alt | 0.255 | 0.264 | 0.256 | -0.912 | 0.781 |
| EBIT/TA | 178/6 | Alt | 0.105 | 0.104 | 0.096 | -0.225 | 0.386 |
| ME/TL | ME/181 | Alt | 2.799 | 1.223 | 4.717 | 0.034 | 31.893 |
| S/TA | 12/6 | Alt | 1.493 | 1.361 | 0.921 | 0.141 | 5.475 |
| NI/TA | 172/6 | Zmi | 0.048 | 0.054 | 0.079 | -0.329 | 0.231 |
| TL/TA | 181/6 | Zmi | 0.507 | 0.512 | 0.192 | 0.090 | 1.028 |
| CA/CL | 4/5 | Zmi | 2.439 | 2.108 | 1.460 | 0.447 | 9.214 |
| Default Spread | NA | Shu | 2.251 | 2.170 | 0.827 | 0.590 | 3.860 |
| Sigma | NA | Shu | 0.110 | 0.097 | 0.055 | 0.030 | 0.325 |
| $r_{it-1} - r_{mt-1}$ | NA | Shu | 0.057 | -0.029 | 0.511 | -0.817 | 2.181 |
| Relative Size | NA | Shu | -10.076 | -10.146 | 1.638 | -13.616 | -6.115 |
| Ln(Age) | NA | Shu | 1.937 | 2.079 | 0.920 | 0.000 | 3.434 |

# Table II
## Forecasting Bankruptcy with Altman's Variables

Table II presents parameter estimates for models that employ Altman's (1968) forecasting variables. Panel A reports discriminant analysis coefficients and Panel B reports hazard model estimates. While the DA calculations reported in Panel A use only each firm's last available set of accounting ratios, the hazard model estimates in Panel B exploit all of the data available from 1962 to 1992. The 1962-1983 data consist of 1822 firms with complete data in or before 1983. 118 of the 1822 firms filed for bankruptcy by 1983.

### Panel A: Discriminant Analysis Coefficients

| Coefficients | WC/TA | RE/TA | EBIT/TA | ME/TL | S/TA |
|---|---|---|---|---|---|
| Altman (1968) | 1.2 | 1.4 | 3.3 | 0.6 | 1.00 |
| BMW (1996)[a] | 10.4 | 1.0 | 10.6 | 0.3 | -0.17 |
| 1962-1983 data | 0.4 | 2.8 | 11.1 | 0.01 | -0.35 |
| T-Statistics* | (-5.66) | (-11.96) | (-13.92) | (-5.93) | (2.12) |
| 1962-1992 data | 1.2 | 0.6 | 10.0 | 0.05 | -0.47 |
| T-Statistics* | (-7.38) | (-11.59) | (-7.21) | (-3.41) | (10.53) |

*These t-statistics are inconsistent because of selection bias.

### Panel B: Hazard Model Estimates

| Variable | Coefficient | $\chi^2$ | P-Value |
|---|---|---|---|
| Intercept | -3.226 | 11.86 | 0.000 |
| WC/TA | -0.732 | 0.31 | 0.577 |
| RE/TA | -0.818 | 1.02 | 0.312 |
| EBIT/TA | -8.946 | 13.32 | 0.000 |
| ME/TL | -1.712 | 6.25 | 0.012 |
| S/TA | 0.158 | 0.58 | 0.446 |
| Ln(age) | 0.015 | 0.00 | 0.967 |

2496 firms, 28226 firm-years, 229 failures

[a]Note that the published version of Begley, Ming, and Watts (1997) contains two typographical errors. The coefficients reported above are correct.

# Table III
## Forecast Accuracy with Altman's Variables

Table III presents a comparison of the out-of-sample accuracy of various bankruptcy models. All of the models use the independent variables identified by Altman (1968), and all of the models are estimated with data available between 1962 and 1983. Parameter estimates calculated with 1983 data are combined with annual data between 1984 and 1992 to forecast bankruptcies occurring between 1984 and 1992.

Probability Rankings versus Actual Bankruptcies
Percent Classified out of 111 Possible

| Decile | Altman | BMW | New DA | Hazard |
|--------|--------|------|--------|--------|
| 1 | 42.3 | 52.3 | 60.4 | 67.6 |
| 2 | 12.6 | 11.7 | 11.7 | 15.3 |
| 3 | 12.6 | 8.1 | 8.1 | 3.6 |
| 4 | 9.0 | 6.3 | 8.1 | 3.6 |
| 5 | 8.1 | 5.4 | 4.5 | 3.6 |
| 6-10 | 15.4 | 16.2 | 7.2 | 6.3 |

# Table IV
## Forecasting Bankruptcy with Zmijewski's Variables

Table IV presents parameter estimates for models that employ Zmijewski's (1984) forecasting variables. Panel A reports simple logit analysis coefficients and Panel B reports hazard model estimates. While the logit estimates reported in Panel A are calculated with only each firm's last available set of accounting ratios, the hazard model estimates in Panel B exploit all of the data available from 1962 to 1992. The 1962-1983 data consist of 1897 firms with complete data in or before 1983. 130 of the 1897 firms filed for bankruptcy by 1983.

### Panel A: Simple Logit Model Coefficients

| Coefficients | Intercept | NI/TA | TL/TA | CA/CL |
|---|---|---|---|---|
| Zmijewski (1984) | -4.336 | -4.513 | 5.679 | 0.004 |
| 1962-1983 data | -5.112 | -5.222 | 4.579 | -0.166 |
| $\chi^2$ test statistic* | (57.1) | (25.2) | (34.1) | (1.52) |
| 1962-1992 data | -4.201 | -4.701 | 3.106 | -0.119 |
| $\chi^2$ test statistic* | (110.3) | (62.5) | (47.9) | (2.03) |

*These test statistics are inconsistent because of selection bias.

### Panel B: Hazard Model Estimates

| Variable | Coefficient | $\chi^2$ | P-Value |
|---|---|---|---|
| Intercept | -7.811 | 25.08 | 0.000 |
| NI/TA | -6.307 | 10.80 | 0.001 |
| TL/TA | 4.068 | 6.53 | 0.011 |
| CA/CL | -0.158 | 0.28 | 0.599 |
| Ln(age) | 0.307 | 1.11 | 0.292 |

2657 firms, 32524 firm-years, 241 failures

# Table V
**Forecast Accuracy with Zmijewski's Variables**

Table V presents a comparison of the out-of-sample accuracy of various bankruptcy models and the ratio of net income to total assets (NI/TA). All of the models use the independent variables identified by Zmijewski (1984), and all of the models are estimated with data available between 1962 and 1983. Parameter estimates calculated with 1983 data are combined with annual data between 1984 and 1992 to forecast bankruptcies occurring between 1984 and 1992. All data, including net income to total assets, are lagged by at least six months.

Probability Rankings versus Actual Bankruptcies
Percent Classified out of 111 Possible

| Decile | Zmijewski | New Logit | Hazard | NI/TA |
|--------|-----------|-----------|--------|-------|
| 1 | 54.1 | 55.9 | 55.0 | 56.7 |
| 2 | 19.8 | 17.1 | 15.3 | 16.2 |
| 3 | 8.1 | 7.2 | 10.8 | 8.1 |
| 4 | 4.5 | 6.3 | 8.1 | 1.8 |
| 5 | 5.4 | 6.3 | 0.9 | 6.3 |
| 6-10 | 8.1 | 7.2 | 9.9 | 10.9 |

# Table VI
## Forecasting Bankruptcy with Market-Driven Variables

Table VI presents parameter estimates for two hazard models that forecast bankruptcy with a new set of market-driven variables. Each firm-year in the sample, from 1962 to 1992, is included in these models.

### Panel A: Market Variables Only

| Variable | Coefficient | $\chi^2$ | P-Value |
|---|---|---|---|
| Intercept | -12.027 | 39.27 | 0.000 |
| Relative Size | -0.503 | 8.06 | 0.005 |
| $r_{it-1} - r_{mt-1}$ | -2.072 | 11.14 | 0.001 |
| Sigma | 9.834 | 11.03 | 0.001 |

2894 firms, 33621 firm-years, 291 failures

### Panel B: Market and Accounting Variables

| Variable | Coefficient | $\chi^2$ | P-Value |
|---|---|---|---|
| Intercept | -13.303 | 30.79 | 0.000 |
| NI/TA | -1.982 | 0.88 | 0.348 |
| TL/TA | 3.593 | 6.90 | 0.009 |
| Relative Size | -0.467 | 5.24 | 0.022 |
| $r_{it-1} - r_{mt-1}$ | -1.809 | 6.52 | 0.011 |
| Sigma | 5.791 | 2.47 | 0.116 |

2497 firms, 28664 firm-years, 239 failures

# Table VII
## Forecast Accuracy with Market-Driven Variables

Table VII presents a comparison of the out-of-sample accuracy of the bankruptcy models that contain market-driven variables. All of the models are estimated with data available between 1962 and 1983. Parameter estimates calculated with 1983 data are combined with annual data between 1984 and 1992 to forecast bankruptcies occurring between 1984 and 1992.

Probability Rankings versus Actual Bankruptcies

| Decile | Market | Acct. and Market |
|--------|--------|------------------|
| 1 | 69.0 | 75.0 |
| 2 | 10.6 | 12.5 |
| 3 | 7.8 | 6.3 |
| 4 | 5.0 | 1.8 |
| 5 | 2.8 | 0.9 |
| 6-10 | 4.8 | 3.5 |
| Possible | 142 | 112 |