The Dynamic Prediction of Company Failure:

The Influence of Time, the Economy and Non-Linearity

MARIA H. **K**IM and **G**RAHAM **P**ARTINGTON[1]

[1]*School of Business, Faculty of Economics and Business, The University of Sydney*

Abstract. A Cox regression model with time-varying variables is used to estimate the survival probabilities of a large sample of Australian firms. The research overcomes the problem of making forecasts from a Cox model when the model contains time-varying variables. In out of sample forecast the model has predictive power. Controlling for the state of the economy does not improve the predictive power of the model but allowing for non-linearity between the predictor variables and financial distress risk does improve predictive power.

JEL Classification: C41, G14, G32, G33

Keywords: Bankruptcy prediction; Time-varying Cox regression model; Baseline hazard; Survival analysis; Macroeconomic risk factors; Non-linearity

---

[1] Maria H. Kim is a PhD student, Graham Partington (g.partington@econ.usyd.edu.au) an Associate Professor in the Finance Discipline of The University of Sydney. We acknowledge the financial support of the Capital Markets CRC and the data service and support of SIRCA. We would like to thank Professor Tony Chen from National Taiwan University and Dr. Li-Sheng Chen from National Yang-Ming University for their SAS Macro programs, test data, and advice. We are grateful to Dr. Philip Gharghori from Monash University for his valuable suggestions and comments, and also for providing us some data. Special thanks are due to Dr. Giang Truong from Lonergan Edward and Associates for her helpful comments.

The Dynamic Prediction of Company Failure:

The Influence of Time, the Economy and Non-Linearity

Abstract. A Cox model with time varying variables is used to estimate the survival probabilities of a large sample of Australian firms. The research overcomes the problem of making forecasts from a Cox model when the model contains time-varying variables. In out of sample forecast the model has predictive power. Controlling for the state of the economy does not improve the predictive power of the model but allowing for non-linearity between the predictor variables and financial distress risk does improve predictive power.

JEL Classification: C41, G14, G32, G33

Keywords: Bankruptcy prediction; Time-varying Cox regression model; Baseline hazard; Survival analysis; Macroeconomic risk factors; Non-linearity

1. Introduction

Much of the previous work in financial distress prediction focuses on static predictions and uses static variables in estimating the predictive model. In this study our goal is to make dynamic predictions, and to use dynamic variables in estimating the model. With dynamic predictions we allow the probability of financial distress to vary over the forecast period. With dynamic variables the model estimation allows for changes in the financial characteristics of a firm over time.

The motivation for the paper is fourfold. First dynamic forecasts of the probability vector for failure f_t to f_{t+n} (where f_t is the probability of failure at time t) have been much less explored than the static forecast of a single failure probability f. Second relatively little use has been made of dynamic forecasting variables. In most applications including a data vector of say the last five years profitability in forming a forecast requires five separate profitability variables in the model and this is not commonly done.[2] In the approach that we use the vector of data is represented by a single profitability variable. Third, one of the most popular techniques for survival analysis is Cox regression, Cox (1972). Unfortunately, for reasons we discuss later, forming forecasts is problematic when a Cox regression contains dynamic variables. We implement a procedure that overcomes this problem. Lastly, it has long been a concern to us that the relation between predictor variables and the risk of financial distress is non-linear. We adopt Loffler and Posch (2007) and develop a simple non-parametric method which allows for non-linearity between the predictor variables and financial distress risk.

[2] A more common approach, as exemplified by Altman (1968,) is to estimate five separate models using data one year before the failure, two years before the failure and so on back to year five.

The work on estimating models which allows for time varying probabilities of financial distress began in the mid 1980s (see for example Crapp and Stevenson, 1987). These models use the techniques of survival analysis, and have attracted increasing attention following the dynamic model of Shumway (2001). Despite the growing use of survival analysis in modeling financial distress, relatively little attention has been given to the use of dynamic (time-varying) variables in estimating these models. Initially this was because of computational difficulties in estimating models with dynamic variables, and even when this problem was overcome a problem remained in making forecasts when using the Cox regression model with time-varying variables.

A key element in forecasts using the Cox regression model is the baseline hazard. In making a forecast the baseline hazard is scaled up, or down, according to the firm's risk factors and this scaled hazard is used to compute the probability of financial distress. When time-varying variables are introduced into the Cox model, forming estimates of the baseline hazard has been problematic. Consequently making forecasts have also been problematic with time-varying variables.

Cox's model has had considerable use in medical studies. Chen, Yen, Wu, Liao, Liou, Kuo and Chen (2005) apply the Cox model with time-varying variables to find the effect of biochemical covariates on death attributed to liver cancer. They implement a method for estimating the baseline hazard and hence are able to make survival forecasts with time-varying variables. Chen et al. (2005) also published the code for implementation of these estimation procedures in SAS. Following the approach of Chen et al., we first construct a time-varying Cox's regression model for the prediction of financial distress using micro level firm specific information.

Then the base model is extended to capture macroeconomic changes through time. In contrast to prior research our model has both time varying firm specific covariates and time varying macroeconomic covariates which track changes in the firm and economy through time.

Last but not least, this study goes on to suggest a non-parametric data transformation of predictor variables based on default rates and incorporate the non-linear relation between predictor variables and the failure risk of a firm into the Cox regression model. In prior bankruptcy studies little attention has been paid to investigating the non-linear pattern of firm characteristics on the default spread. Chan, Faff, Gharghori and Lajbcygier (2008) use a new non-linear technique, called generalized additive models (GAMs). Using the technique suggested by Loffler and Posch (2007) our approach is much simpler, more transparent and less computationally demanding.

Using firm specific data on Australian Securities Exchange (ASX) listed firms from 1995 to 2006; a time-varying Cox's regression model is developed with seven predictor variables measuring profitability, leverage (book and market), liquidity, cash flow, size, and growth opportunities. Each variable captures the impact eight years of data for firms that are in the estimation sample for the full eight years (1995 – 2002). Book leverage, cash flow generating ability, firm size and market leverage are found to be significant predictors across all models. Receiver operating characteristic (ROC) curves and the Brier Score show that the model has modest predictive power and unlike most bankruptcy models the performance of the model improves as the forecast period lengthens.

The remainder of this paper is set out as follows. Section 2 reviews the literature in the area of bankruptcy prediction, introduces survival analysis, and discusses Chen et al. (2005). Section 3 describes the data, introduces variables to be used and explains how predictive accuracy is evaluated. Section 4 presents the methodology to construct a Cox's regression model with time-varying variables using firm specific variables. Section 5 extends the model to incorporate macroeconomic variables. The approach to investigation of non-linear relation between predictor variables and the risk of financial distress is discussed in Section 6. Section 7 presents the results of parameter estimates followed by assessment of predictive accuracy of the models. Section 8 concludes the paper and offers some possible future research directions.

2. Bankruptcy Prediction Literature

2.1 EARLY BANKRUPTCY PREDICTION STUDIES

Research on bankruptcy prediction has been of substantial interest to accounting and finance academics and practitioners for the last four decades. A number of empirical approaches have been applied to the bankruptcy prediction problem since the pioneering work of financial predictive modeling by Beaver (1966), Altman (1968) and Ohlson (1980). The initial approach to predicting corporate bankruptcy has been to apply a statistical classification technique to a set of samples containing both bankrupt and non-bankrupt firms. The principal tools for the early studies have been multivariate discriminant analysis (Altman, 1968) and logit analysis (Ohlson, 1980). The task of predicting bankruptcy of a firm can be posed as a classification problem: given a set of classes (for example, bankrupt and non-bankrupt) and a set of input data vectors, the task is to assign each input data vector to one of the classes. Since the 1980s, the literature has progressed to non-parametric approaches such as recursive partitioning algorithms (Frydman et al., 1985), neural networks techniques (Odom and Sharda, 1990; Coats and Fant, 1992; Tam and Kiang, 1992; Wilson and Sharda, 1994) and survival analysis (Lane, Looney and Wansley, 1986; Crapp and Stevenson, 1987; Chen and Lee, 1993; Bandopadhyaya, 1994).

2.2 SURVIVAL ANALYSIS

In recent studies of financial distress (bankruptcy) prediction, the need to take the time dimension into account is increasingly being recognized. LeClere (2000) points out that qualitative response models such as logistic regression or probit models employ data from the time period directly preceding the occurrence of the event of bankruptcy. Hence, the model is static in that it ignores the entire time period preceding the event. Furthermore, the information provided by the estimated model is limiting as the data used to estimate the probability of financial bankruptcy may only be available immediately prior to the event. Shumway (2001) also points out the discordance between single-period bankruptcy prediction models and multiple-period bankruptcy data. He argues that the single-period classification models that have been commonly used for predicting bankruptcy yield biased and inconsistent estimates because they ignore the fact that the characteristics of firms change through time. Liu (2004) also observes that failure rates change with changes in the time-series of economic data.

Survival analysis is ideally suited to introducing a time dimension into financial distress prediction since the objective is to estimate $S(t) = P(T > t)$, the probability that financial distress will occur at a time T which lies beyond the time horizon t, for a range of values of t. Thus, a time dimension is embedded in the dependent variable of the model. It is also possible to introduce a time dimension into the independent variables by making them time-varying. Thus, for example, a vector of ratios giving the return on assets for a firm over a ten year period would be treated as a single variable, but the value of that variable would be updated as we follow the firm through time in estimating the survival model.

The problem with time-varying variables in the past has been in estimating the baseline hazard and consequently in forming forecasts. Previous studies (e.g., Wheelock and Wilson, 1995; Kim et al., 1995) have not reported the baseline hazard estimates since estimates of the baseline hazard are difficult to obtain when covariates in the model are time-varying.

In the standard proportional hazards model the hazard for each case is a fixed proportion of the hazard of any other case at any point in time. Thus, the ratio of the hazards for any two cases with independent covariates are constant over time such that

$$\frac{h_i(t)}{h_j(t)} = \frac{h_0(t) \cdot \exp\left\{\sum_{m=1}^{p} \beta_m z_m^i\right\}}{h_0(t) \cdot \exp\left\{\sum_{m=1}^{p} \beta_m z_m^j\right\}} = \exp\left\{\sum_{m=1}^{p} \beta_m (z_m^i - z_m^j)\right\} = k. \quad (1)$$

The result is that the plots of the hazard function for all cases are parallel. It is this property that is exploited in estimating the baseline hazard. On the other hand, with the time-varying variables the proportionality no longer applies. Since

$$\frac{h_i(t \mid z(t))}{h_j(t \mid z(t))} = \frac{h_0(t) \cdot \exp\left\{\sum_{m=1}^{p} \beta_m z_m^i(t)\right\}}{h_0(t) \cdot \exp\left\{\sum_{m=1}^{p} \beta_m z_m^j(t)\right\}} = \exp\left\{\sum_{m=1}^{p} \beta_m (z_m^i(t) - z_m^j(t))\right\} \neq k. \quad (2)$$

The result is substantial difficulty in estimating the baseline hazard. Recent advances, however, have made this somewhat less difficult.

Chen et al. (2005) estimate a time-dependent Cox's regression model for patients' deaths from liver cancer. Using a method from Anderson (1992), they estimate the integrated baseline hazard and forecast survival probabilities of each patient. Two SAS Macro programs for time-dependent Cox's regression are introduced in Chen et al. The first program is for parameter estimates on risk factors, deriving the baseline hazard and the prediction of survival on the basis of time-varying covariates. The second program validates the model's predictive accuracy using receiver operating characteristic (ROC) curves. We use the SAS Macro programs in Chen et al., with required modification, to estimate our financial distress models.

3. Data and Methodology

3.1 SAMPLE SELECTION

Our sample includes publicly listed companies on the Australian Securities Exchange (ASX) from 1995 to 2006. Firms which are in the financial sector, as indicated by their GICS code, are excluded from the sample. We obtain annual accounting data from FinAnalysis (Aspect Financial) and annual market capitalization data from SIRCA and Datastream. Macroeconomic data are obtained from two sources; annualized return on All Ordinary Index from SIRCA and other macroeconomic data from the RBA Website. They also cover the entire sample period.

Two data filters are applied to the preliminary dataset. First, a complete set of accounting, market capitalization, and macroeconomic data must be available for every single firm; and secondly, information as to the firm's failure event must be available. In order to classify firms into a group of non-failed and failed firms, we follow the approach of Jones and Hensher (2004) and Chan et al. (2008). Firms are classified as "failed" if (i) they were delisted due to the failure to pay their annual listing fees to the ASX[3], or (ii) there was the appointment of liquidators, insolvency administrators, or receivers[4]. Companies' event of failures in our sample, including the firm's financial statements and the date of its release to the market in Australia, are cross checked against a number of sources: the ASX's Signal G[5] - Company Announcement data obtained from the SIRCA, the deListed Company Database[6], Nothman (1993) and Chan et al. (2008). Resulting from this, 1,703 non-financial firms are identified with available accounting and market capitalization data, which further comprises of 1,570 non-failed firms and 133 failed firms in the final sample.

We note that these are failure events that happen at specific dates, but there may be varying lags between the failure event and the onset of financial distress. We do not have the data to model these lags, but the advantage of dynamic probability forecasts, which give a trajectory to failure, lies in the potential for early warning of problems as the trajectory changes.

---

[3] According to the Australian Securities Exchange Listing Rule 16.5, a firm can be removed from the official list if the firm does not pay an annual listing fee (ASX 2008, Chapter 17).

[4] According to *Australian Corporations Act* (2001), three principal forms of bankruptcy proceeding are available under the legislative provisions; (i) voluntary administration (first introduced in Australia in June 1993 under the *Corporate Law Reform Act* [1992]); (ii) liquidation; and (iii) receivership. Voluntary administration has similarities with Chapter 11 provisions in the U.S., where the company is effectively given a period of time or "breathing space" to reorganize and/or reconstruct. Under Australian voluntary administration laws, once appointed the insolvency administrator has a limited period (28 days) to assess the company and recommend to the creditors whether the company should be wound up or enter into a deed of arrangement (this is a contract that binds the company and creditors and includes such issues as the order by which creditor claims are to be settled from sale of assets). If deed of arrangement stage is not reached, then the legislation provides for an automatic transition to liquidation. With respect to (ii) liquidation, there are essentially two types of the winding-up procedure available: a creditors' voluntary winding up (decided by special resolution of the company) and a court winding up. In the case of (iii) receiverships, the *Corporations Act* (2001) provides that a secured creditor, in the event of a firm's insolvency can appoint a receiver (or a receiver and manager) to recover outstanding claims against the company (extracted from Jones and Hensher (2004, p.1020)).

[5] The company announcements are available via a 'Signal G' service. They detail announcements lodged with the ASX pursuant to the ASX Listing Rules.

[6] deListed is a division of BRG Pacific Pty Limited, holder of Australian Finance Services. deListed provides information on failed companies including companies suspended from ASX, NZX, NSX and BSX, all historical name changes and delistings for these exchanges, and carries administrators/liquidator/receivers declarations for Australian companies (deListed 2006).

We collect annual accounting and market capitalization data for each company. Initially, we collect yearly observations of firm's financial performance from 1989 through 2006. However, application of two filters results in no failed firms between 1989 and 1993 and only one failure observation in 1994[7], and moreover, the sample sizes are very small for 1989 to 1991. It may be that data on firms failing in this period has been deleted from our data sources and, if so, this is likely to raise a survival bias problem. Therefore, we come to a decision that we eliminate those annual observations between 1989 and 1994 from our sample of study.

Table I shows the number of failed and non-failed firms for each year over the sample period of 1995–2006.

[Table I about here]

There are several extreme values among the variables observed. Following the approach of Shumway (2001), all values lower than the first percentile of each variable are set to that value, and analogous treatment is applied to all observations higher than the ninety-ninth percentile of each variable. The data after truncation is described in detail in Section 3.3.

The entire sample period (1995–2006) is divided into two separate samples, an estimation sample (1995–2002) and a holdout sample (2003–2006) for tests of predictive accuracy.

3.2 VARIABLE SET

3.2.1 FIRM SPECIFIC VARIABLES

[7] We expect to have a number of failed observations in these time periods as there was an economy crash in Australia in early 90s. Nevertheless, our data sources rarely show the failure cases in these periods though Chan et al. (2008) can identify numerous defaulted firms through their hand collected dataset.

Key predictors for financial distress with firm specific variables have been identified from previous bankruptcy studies, and we focus on variables used in the recent major studies by Sobehart and Stein (2000), Shumway (2001), and Campbell et al. (2005). As this study is being conducted using Australian data, we also include some of the variables found to be useful in Australian studies by Castagna and Matolscy(1981), Jones and Hensher (2004) and Gharghori, Chan and Faff (2006).

A set of fundamental accounting-based and market-based variables chosen from aforementioned studies is shown in Table II.

[Table II about here]

The accounting-based variables reflect measures of profitability (Net Income/Total Assets (NI/TA)), operating liquidity (Working Capital/Total Assets (WC/TA)), book leverage (Total Liabilities/Total Assets (TL/TA)) and cash flow generating ability (Net Cash Flow from Operations/Total Assets (CF/TA)). As a group, these ratios capture the strength of the firm's financial position.

Shumway (2001)'s and Campbell et al. (2005)'s market-based variables are also used in model estimation. The market-to-book (MB) ratio is commonly used as a proxy for growth opportunities (Rajan and Zingales, 1995; Baker and Wurgler, 2002; Faulkender and Petersen, 2005). Campbell et al. (2005) demonstrate that MB has a positive effect on the risk of failure "when market value is unusually high relative to book value" (p.11). Following Shumway (2001) the size measure we use is the value of the company relative to the value of all companies listed on the ASX. We measure this variable as ln (Firm Market Capitalization $_{i,t}$/Total ASX Market Value $_{t}$), which is denoted as RSIZE. Market Capitalization/Total Liabilities (MC/TL) is used as a measure of market leverage. Bigger values of this variable represent lower levels of leverage and it is expected that this variable will have a negative relationship with the risk of failure.

3.2.2 MACROECONOMIC VARIABLES

For the purpose of capturing economy-wide changes through time, a set of macroeconomic covariates are introduced as control variables in this study (also shown in Table II). Leading indicators found in previous bankruptcy studies at macroeconomic level include yield spread, default spread, return on the market index, consumer confidence index, consumer price index (inflation) and unemployment rate (Partington et al., 2001; Liu, 2004; Bonfim 2009; Rosch and Scheule, 2005; Bellotti and Crook 2009).

Default spread is not included in this study due to a lack of data on corporate bond yields. Unemployment rate is also eliminated due to high level of correlation between yield spread and unemployment rate.

Yield spread, as given by the Australian Government Bond yield less Australian Treasury Notes yield, provides business cycle effects. Yield spread is expected to be widest at the trough of a business cycle when long-term interest rates are high and short-term interest rates are low; thus the high yield spread is expected to increase the risk of default. To construct the yield spread, we obtain monthly yields on 10 year Commonwealth Government Treasury Bonds and monthly yields on 90-day Bank Accepted Bills (BAB) during 1989 and 2008 from the RBA website. We use yields on 90-day BAB as short-term securities [8] because the Commonwealth Government

[8] Ford and Taylor (2005) also use 10-year Government Bond yield and 90-day Bank Accepted Bills to generate the yield spread and Karfakis and Phipps (1996) use 90 day Bank Accepted Bills for the proxy of Australian short term government securities.

temporarily suspended issues of three month Treasury Notes[9] since December 2002, as noted in Brailsford, Handley and Maheswaran (2008).

Return on market index is expected to have negative impact on the failure risk. When the stock market is on a rise, it generally reflects companies' financial conditions are improving, resulting in lower failure rates.

Consumer confidence index (CCI) is another leading indicator to predict economic activity. We expect a negative coefficient on CCI, thus, when CCI goes up, chances of recessions are going down.

Inflation, measured by the change in price from consumer price index (CPI), is usually going strong during the economic expansion. However it should be noted that there are two effects. One is inflation variable standing as an indicator variable for the economy, and the other is this variable as impacts on the individual firm, which is not necessarily picked up in the accounting covariates. That is, on the one hand, strong inflation implies the economy is going well, but in the end inflation is bad for economic activity. If inflation keeps increasing, the real economy starts to deteriorate. On the other hand, if inflation is associated with increase in interest rates, which it is, that puts more pressure on business cash flows and thus, it produces an adverse effect on the firm's propensity to survive. This is supported by Wadhwani's (1986) analysis which shows that higher inflation leads to higher bankruptcy rate in the period 1964-1981. So inflation effect has rather a mixed story.

[9] We first choose three month Treasury Notes as marketable short-term government securities since issue of Treasury Bills ceased in June 1986 in Australia. Present-day Treasury Notes are the successor to Treasury Bills in Australia and are equivalent to the US and UK Treasury Bills which are short-term government obligations. It is noted that Australia's Treasury Notes mature in six months or less whereas Treasury notes in the US are longer-term securities from two to ten years (Carew 1996).

On the whole, an increase in yield spread is an indicator of a declining economy providing conditions for higher risk of default. In contrast, an increase on return on the stock market and on the consumer confidence index is expected to cause reduced risk of financial distress. Inflation remains indecisive at this point.

3.3 SUMMARY STATISTICS

Table III presents descriptive statistics for annual observations of firm specific predictor variables after the data filtering process described in Section 3.1. The minimum and maximum values reported in the table are calculated after truncation. The financial characteristics of non-failed firms are a noticeable contrast to those of failed firms for most variables[10]. For example, failed firms are found to have lower levels of profitability, operating liquidity and cash flow compared to those of non-failed firms. Meanwhile, non-failed firms have a lower level of book leverage and a higher market-to-book ratio. The dispersion of financial ratios among failed firms is also wider than that of non-failed firms, evidenced by higher standard deviations.

[Table III about here]

The first panel shows descriptive statistics of firm specific variables for all firm-year observations of the entire sample and the other two panels report descriptive statistics for the estimation and holdout sample. For the whole sample we have 1,703 non-financial firms' information where 11,573 firm-year observations are obtained with 133 failure events. The second panel shows summary statistics for all firm-year observations of the

[10] Wilcoxon-Mann-Whitney test is carried out for each variable to test the significance of differences of firm characteristics between failed and non-failed groups. The test shows the differences are statistically significant at the 1% level for all variables except for the WC/TA for the entire sample and the estimation sample, and at the 5% level for all variables, with the exception of WC/TA, for the holdout sample.

estimation sample. There are a total of 1,267 firms and 6,934 firm-year observations in the estimation sample, of which 87 are failure events. For the holdout sample as shown in the third panel, we have 1,455 firms' information with 4,639 firm-year observations, where there are 46 failure events.

Table III shows that on average profitability (NI/TA) was negative for the full sample, the estimation sample, and the holdout sample. This is not the result of poor profits in a specific period. Panel D in Table III shows that profitability, on average, has been negative across all years in the sample. This result is surprising, but it is driven by small firms. The value weighted mean for NI/TA (not reported here) is positive. As shown in Panel E of Table III, if we restrict our sample to the top quartile of firms by size, the mean and median profitability are positive. If we limit the sample to the top half of firms by size, the mean is negative but the median is positive. Summary statistics of macroeconomic variables for all firm-year observations of the entire sample are presented in Panel F.

It is noted that not all public firms have complete accounting and market information available for estimating the parameters of the model. In this study, any firm-year observations with incomplete data were eliminated from the final sample. Table III, therefore only contains statistics for variables where all values are non-missing. The elimination of missing value cases was done for two reasons. First, handling missing values causes substantial computational problems and second including missing value cases is likely to lead to informative censoring as we discuss below.

In relation to defaulting firms Sobehart and Stein (2000) state, "…financial and market information are less likely to be complete or reliable in the time period leading up to default" (p.12). Thus missing data may be an indicator of failure.

We compare cases with missing and non-missing values using the Mann-Whitney test. The result shows that the missing data is associated with firms that have more negative profits, higher leverage, and more negative cash flow. It appears that the cases with missing data are financially weaker than firms with complete data and therefore are more likely to fail.[11] If this is true, and these firms were included in the study at the times when there was data and then treated as censored when data was not available, this would give rise to informative censoring. That is, the censoring substitutes for the failure event, and this violates the assumptions underlying the analysis.

Correlation matrices of the seven covariates in the model are constructed for the entire sample, the estimation sample and the holdout sample, respectively. The Pearson Product-Moment correlations are examined. All of the correlations are statistically significant at the 1% level, but the correlations are not so large as to cause concerns about collinearity. The highest correlation at about 0.65 is between Net Income/Total Assets and Net Operating Cash Flow/Total Assets (unreported).

3.4 MEASURING MODEL ACCURACY

3.4.1 DISCRIMINATION and CALIBRATION

Model's predictive ability can be commonly assessed in two dimensions: discrimination and calibration. Discrimination refers to the model's ability to distinguish between those companies surviving and those failing at a given point in time. Calibration measures how well the estimated probability of a failure event matches true observation of the event.

[11] We note, however, that we did not find any information that these firms were liquidated, went into receivership, or were delisted for failure to pay fees.

3.4.2 RECEIVER OPERATING CHARACTERISTICS (ROC)

To assess the discriminatory power of the models, we use the survival probabilities to classify each firm as failing or surviving, and then compare the classification with the actual outcome. When the probability prediction is converted to a state prediction, picking the optimal cut-off value becomes an issue. Using ROC curves is one way to sidestep the problem of determining an optimal cut-off point, since it examines the predictive power of the model across the entire spectrum of possible cut-off points. The ROC curve for a particular model is determined by the hit rate (correctly predicting failures) and the false alarm rate (incorrectly predicting non-failures to failures). The ROC curves plot the combinations of the false alarm rate (X-axis) and hit rate (Y-axis) as the cut-off point is varied across all possible values.

Time-varying risk score can be used to identify the optimal cut-off points at different points in time using ROC curve. We form ROC curves at one-year intervals through time for the holdout sample. Table VI shows the area under the ROC (AUROC) curve of the out-of-sample survival functions. The AUROC measures the predictive accuracy of the model and the higher the AUROC, the better the model's prediction. Predictions made at random have an AUROC of 0.5 and models that do not beat this benchmark have no predictive power.

3.4.3 SCORING RULE (BRIER SCORE)

We also use a proper scoring rule, known as the Brier Score to assess the model's performance. While the ROC curves can be used to measure the model's ability to discriminate between those companies surviving and those failing at a given point in time, the Brier Score measures model's calibration, that is, the prediction accuracy at the level of individual company. By calculating the deviation between the predicted probability of a failure event and the actual outcome of the event, the Brier Score shows the relationship between model's prediction and the actual observation of company's status. The Brier Score is calculated as follows:

$$B = \frac{\sum_{n=1}^{N} (p_n - a_n)^2}{N} \tag{3}$$

Where N is the number of predictions, p_n is the predicted probability that a failure event will occur and a_n is the actual observation of the event. When a firm fails, then a_n equals 1, and otherwise it is 0. A Brier score of 1 indicates that the model has no predictive power and a score of 0 shows perfect predictive ability. Thus, the lower the Brier Score, the better the model's predictive power.

4. Base Model Construction

4.1 MODEL 1: TIME-VARYING COX'S REGRESSION MODEL WITH FIRM-SPECIFIC VARIABLES ONLY

Cox's hazards model with time-varying covariates can be expressed as:

$$h_i(t \mid z(t)) = h_0(t) \cdot \exp\left\{ \sum_{j=1}^{p} \beta_j z_j^i(t) \right\}. \tag{4}$$

$h_i(t \mid z(t))$ is the time-dependent hazard function for firm i at time t. $z_j^i(t)$ denotes the value of the jth covariate at time t for the ith firm, β_j is the corresponding coefficient for z_j^i, while $h_0(t)$ is the baseline hazard representing the effect of duration on the hazard in the absence of covariates. Thus, the hazard at time t depends on the value of predictor variables at time t.

In most of the previous bankruptcy literature, each annual observation of firms has been treated as an independent observation and so researchers could not take advantage of all the available multiple-year financial information. Using Equation (4), we are able to "exploit each firm's time-series data by including annual observations as time-varying covariates" (Shumway 2001: p.102). That is, if a firm has been observed for twelve years in the set of firms potentially at risk of financial distress, the values of each covariate, $z_j^i(t)$, for that firm are to be updated twelve times from year to year (t). Consequently, we are able to retain multiple-year financial information for each firm according to its life time (or duration) and make use of all the time-series data within the periods to estimate the regression coefficients.

In a time-dependent Cox's regression model, the value of covariates $z_j^i(t)$, changes with time, and therefore, the hazard ratio (*HR*) also varies with time and is defined as follows.[12]

$$HR(t \mid z(t)) = \frac{h_i(t \mid z(t))}{h_0(t)} = \exp\left\{\sum_{j=1}^{p} \beta_j z_j^i(t)\right\}. \tag{5}$$

It is helpful to introduce some concepts in survival analysis in order to understand the estimation of the model. First, the risk set *R(t)*, is defined as the set of firms (individuals) which are observed at risk of event at time *t*. Firms are said to enter the risk set when they become at risk of experiencing the event and leave the risk set either when they are censored or when the event occurs to them (fail or become financially distressed). Being censored means that a firm leaves the risk set for some other reason than experiencing the

[12] When there are no time-varying variables the ratio of hazards for any two firms is constant over time and so traditionally the model has been known as Cox's proportional hazards model.

event, for example the firm may be taken over, or may still survive at the termination of the study.

Second, it is important to distinguish between calendar time and event time. A graphical demonstration of the difference between arranging the data in terms of "calendar time" and "event time" is presented in Figure 1. It is noted that observations can enter the study at different times in calendar time, however; every observation enters the study at event time 0 in event time. An event time approach looks to the duration (time spent in the risk set) of a firm and sorts observations according to their duration on study. The event time approach is used in our study as is commonly the case in other survival analysis studies.

[Figure 1 about here]

Figure 2 illustrates the arrangement of the data in a time-varying model where yearly observations of firm's financial performance are arranged in event time.

[Figure 2 about here]

We begin with building a base model (Model 1), which contains seven firm-specific covariates. These covariates reflect firm's time series annual data for the entire period as long as the firm is in the risk set; hence values of those covariates vary over time and the dynamic changes of the covariates are captured and exploited in our time-varying Cox regression model .

4.2 INTEGRATED BASELINE HAZARD

To generate survival probabilities at each time t, we need to estimate the baseline hazard function, $h_0(t)$. We follow the approach of Chen et al. (2005), which estimates the integrated baseline hazard function following the equation of Cox's proportional hazards model with time-dependent covariates from Andersen (1992). The integrated baseline hazard function $\hat{H}_0(t)$ can be estimated as follows.

$$\hat{H}_0(t) = \sum_{\tilde{T}_i \le t} \frac{D_i}{\sum_{j \in R(\tilde{T}_i)} \exp\left(\hat{\beta}' \cdot z_j\left(\tilde{T}_i\right)\right)} . \tag{6}$$

D_i is the indicator for whether the firm i experiences the failure, $\tilde{T}_i$ is the failure time for the ith firm, $\hat{\beta}$ is the vector of estimated coefficients, and $z_j\left(\tilde{T}_i\right)$ is the value of the jth covariate at the failure time of the ith firm.

Using the estimated baseline hazard rate, $\hat{h}_0(t)$, the estimated hazard rate of firm i with covariates $z_i(t)$ at time t is derived as:

$$\hat{h}_i(t) = \hat{h}_0(t) \times \exp\left(\hat{\beta} \cdot z_i(t)\right) . \tag{7}$$

It is noted that a time-dependent 'risk score' is defined as $\left(\hat{\beta} \cdot z_i(t)\right)$. The dynamic changes of the risk score over the time horizon studied and the corresponding survival probabilities will be presented in Section 7.

5. Model Extension with Control for Macroeconomic Risk Factors

5.1 MOTIVATION

To this end the base model (Model 1) relies on the event time and does not account for changes in economic climate, so a possible extension is to introduce such variables that we control for the information effects of broad economic conditions. The reason for doing this is two-fold. First, we suspect that the information in the macroeconomic variables, to some extent, would probably be built onto the baseline hazard in the base model (Model 1) in an ad-hoc manner. This has concerned us because it may cause to raise the possibility of model misspecification[13], as being pointed out by Hamerle et al. (2003), Rosch and Scheule (2004) and Rosch and Scheule (2005) that unobservable risk factors are likely to induce uncertainties into the forecasts of default distributions.

Accordingly, our second goal is to obtain better control of systemic risk factors by taking out of the baseline the latent effects of macroeconomic risk factors and placing them directly into the model[14]. In the previous research, Rosch and Scheule (2005) demonstrates in their multi-factor model that default rates fluctuate cyclically and part of the cyclical variances in default rates can be attributed to systematic risk indicators. Thereafter, we expect identifying these risk factors and incorporating them into the base model may reduce part of uncertainty from unobservable factors, thus reduce the chance of model misspecification and even enhance the power of model performance.

However, we should note that it is not the major focus of this study to have macroeconomic variables as predictors; rather are they treated as control variables. The main reason we include these variables is we take such controls for the state of the economy so that firm observations set to be arranged in event time are able to be assessed based on the adequate benchmarks according to their corresponding states of the economy.

5.2 MODEL 2: TIME-VARYING COX REGRESSION MODEL WITH MACROECONOMIC CONTROL VARIABLES

[13] Model misspecification refers to a model that yields biased and more inconsistent coefficients and standard errors in this context.

[14] Hillegeist et al. (2004) also suggest that the time-varying hazard rate should be measured by including the system-level variables, such as macro-economic factors and incorporated into the bankruptcy model so as to reduce the likelihood of biased and inconsistent coefficients and standard errors.

Adding macroeconomic variables into the model as control variables for the general economic conditions gives rise to a question as to whether their impact is contemporaneous or lagged. And if there is a lag, a further question arises as to how long the lag would be. When we judge the condition of Australian economy over time, we do not just focus on the current figure; rather we interpret the current figure in the context of the trend. One sensible way trying to capture the whole effect of the status of the economy over a period of time is to use an exponentially weighted moving average (Holt, 2004), so that we reflect the current condition of the economy into the model while still not discard the information from the previous observations entirely.

Nevertheless, as noted at the outset, the motivation behind the present study is to tackle the issue of handling the firm specific covariates and the macroeconomic covariates at the consistent level; so to narrow the focus, we take a alternative to incorporate macroeconomic factors as time-varying variables[15].

With this end, we add a set of macroeconomic risk factors as time-varying covariates to the base model. Four macroeconomic variables described in Section 3.2.2, return on ASX All Ordinary Index, yield spread, return on consumer confidence index, and return on consumer price index (inflation effect) are considered in this study. Given that every firm year observation in the existing sample is complemented by a set of macroeconomic data measured as at the date of firm's financial year end, each value of macroeconomic covariates is being updated with the function of time-varying variables. With this end, the values of percentage changes (or rate of changes) of macroeconomic variables are

[15] In Rosch and Scheule (2005), time-varying variables have been used to observe lagged systematic risk factors.

converted to those in the form of decimal fractions to be measured on the same scale that the firm-specific data are assessed on.

In short, we now have eleven time-varying covariates, including seven idiosyncratic firm-specific covariates and four systematic macroeconomic variables in Model 2, which basically runs on the same platform (the time-varying Cox's regression analysis) as the base model.

6. An Inquiry into the non-linear relation between predictor Variables and the Firm's Failure Event

6.1 WHY A NON-LINEAR APPROACH TO DEFAULT RISK MODELING?

It has long been a concern to us that the relation between predictor variables and the risk of financial distress may be non-linear. For certain accounting ratios, it is certainly reasonable to expect a non-linear relation with the default risk. For example, intuition suggests that there would be a negative relationship between liquidity and failure risk. A company may fail, for it suffers from insufficient liquidity to pay its bills. But a company may also fail if it has too many current assets because it has to finance these assets (say inventory) and these are not generating cash flow. So both too much liquidity and too little liquidity can cause a problem.

Loffler and Posch (2007) also point out it is quite likely that there is non-linear relation between explanatory variables and the default probabilities and demonstrate that the model's fit has greatly improved after the variables are remodeled with logit transformations. Further, in Chan et al. (2008), non-linearities between accounting variables and default risk are scrutinized through new non-linear models, known as logistic generalized additive models, and the superiority of these non-linear models predictive power has been validated over the linear counterparts. This prospect of non-linear relationship between predictor variables and the financial distress instigates us to move a further step.

6.2 MODEL 3: TIME-VARYING COX REGRESSION MODEL WITH NON-PARAMETRIC DATA TRANSFORMATION BASED ON DEFAULT RATES

We adopt Loffler and Posch's (2007) approach to develop a simple non-parametric alternative. Admitting that the function of covariates in the aforementioned base model (Model 1), whose outcome is the time-dependent risk score ($\hat{\beta} z_i(t)$), is still a function of linear combination of a set of variables, we now aim at absorbing non-parametric data transformation into our base model.

As macroeconomic variables are included as control variables, we restrict non-parametric transformation to firm-specific variables only. First, we divide the original data range up into percentile classes for each of firm-specific variables, in our study we choose to divide it to twenty classes, and note the range of observations in each class. Then we count the incidents of defaults and the number of observations in that class and work out the default rate for each class. Lastly, instead of using the original values of the variables, we recode them to the default rates that match the range that the original values are assigned to. Thus, basically we are transforming the original variable to the default rate. In this way, we effectively embed the non-linear relation between predictor variables and the failure risk of a firm into the Cox regression model.

However, this approach may give rise to an issue of data mining problem, the overfitting, which is an inherent problem of the non-parametric model approach. As we use the data-driven transformation in building the model, the model relies heavily on available sample data only. Consequently, it can cause lack of generalization in the holdout sample, which can be evidenced by a decline in the out-of-sample prediction accuracy of the model (Richardson et al., 1998). This issue will be verified with the result of out-of-sample prediction in the following section.

7. Empirical Analysis

7.1 MODEL ESTIMATION

The time-varying Cox regression model with firm-specific variables (Model 1), the model with macroeconomic control variables (Model 2) and the model with non-parametric transformation based on default rates (Model 3) have been estimated using the estimation sample (1995–2002) of 1,267 listed Australian firms, with 87 failure observations.

The estimated parameters and the goodness of fit measure of the models are presented in Table IV. Panel A in Table IV shows the total number of firms used in estimating the model parameters, and each number of failed and censored firms. The resulting coefficient estimates of three models are shown in Panel B in Table IV, with their expected signs and their respective p-values[16] (based on a chi-squared statistic for the significance of each coefficient). Panel C provides an in-sample goodness of fit measure.

From the third column of Panel B in Table IV, we see that all of the variables in Model 1 have coefficients of the sign expected and five of the variables are statistically

[16] The p-values are based on a Wald test-statistics as given by the squared ratio of the estimated coefficient to its estimated standard error.

significant in explaining failure risk. Less operating liquidity, higher book leverage, less cash flow generating ability, smaller size and less market value relative to debt increase the probability of failure as expected. RSIZE, which was an important predictor in previous studies (Shumway, 2001; Campbell et al., 2005) turns out to be significant in this model as well.

Moving to the next column (Model 2), the parameter estimates of all firm specific variables remain unchanged from Model 1 in terms of sign, and almost consistent in term of magnitude and statistical significance. The macroeconomic variables, return on market index, inflation and consumer confidence index are insignificant.

Notably, yield spread turns out be a significant macroeconomic covariate in large magnitude with a negative coefficient, contrary to expectations, which explains higher yield spread would result in a decrease in the risk of financial distress[17].

Reflecting upon the effect of the yield spread on the failure risk, we suspect that the problem arises from the potential opposite signs of the current and future effects. For instance if we take the current case in Australia, the yield spread has widened as the government has cut short-term rates in an attempt to stimulate the economy, which is projected to recover in a year or two. Therefore, the wider yield spread reflects poor current economic conditions and improving future economic conditions. This suggests that the wider yield spread is associated with an increased failure risk now but a reduced risk in the future.

---

[17] The negative coefficient on yield spread is supported by a standard economic explanation for the relationship between the yield spread and future economic growth. "When the economy is strong, there will be an expectation of higher average short-term interest rates in the future. Expectations of higher average short-term rates will lead to bond yields being higher than present short-term rates and thus to a higher yield spread. Conversely, when the economy is weak there will be an expectation of lower average short-term interest rates in the future, leading to a lower (and possibly negative) yield spread." (Ford and Taylor, 2005: p.114)

That being the case, the negative sign on yield spread appears to be driven by the future effect of yield spread on the likelihood of failure risk. We speculate that part of the reasons for this result can be explained by the possibility of covariates' lagged effects on the actual default event. If this is the case, it is a significant finding because it supports the argument that there would be lagged effects of some macroeconomic factors on the risk level of corporate failures, which has been questioned earlier in section 5.2. We will discuss more on the timing issue in the following section 7.3.

The last column (Model 3), on the whole, produces similar findings in terms of statistical significance of variables to the base model (Model 1), but with one exception. As was the case in the Model 1, book leverage, cash flow generating ability, size and market value relative to debt remain to be significant. But in this modeling operating liquidity is driven to insignificance, instead profitability is found to be a significant factor. The sign and magnitude of the estimated parameters cannot be intuitively expected nor interpreted because the covariates used in this modeling have been non-parametrically transformed from their raw values to the corresponding default rates. By the nature of non-parametric methods, the result is quite data-driven.

To affirm the possible anomaly among some predictor variables, we examine univariate relationship between default rates and predictor variables. And the result ascertains the non-linearities as we observe there is marked U-shape relation between WC/TA, MB and the default rates.

Panel C shows the results of log likelihood statistics. $-2 \log L$ measures the goodness of fit of the presented model, and smaller values indicate a more desirable model. The likelihood ratio is calculated as $LR = 2(\ln L - \ln L_0)$, where ln L is the log likelihood of the estimated model and ln L_0 is the log likelihood of the restricted model only with a constant. The more likelihood is lost by imposing the restriction, the larger the Likelihood Ratio statistic will be. The model fit results set out in Panel C suggests that a model that uses transformed data based on default rates (Model 3) can explain corporate failure better than the same failure prediction model with raw accounting and market information and also better than the model with macroeconomic risk factors controlled.

Though the non-parametric modeling shows better measurement of the estimation function over the in-sample dataset, it might result in an overfit. This will be verified by the result in the out-of-sample prediction accuracy of the model in the following section.

[Table IV about here]

Table V exemplifies the dynamic changes of risk scores and corresponding survival probabilities over the lifetime of a company, resulting from the time-varying Cox regression model with firm specific variables (Model 1). The time-varying risk scores can be calculated for each firm as $\hat{\beta} z_i(t)$. Following the approach of Chen et al. (2005), $\hat{\beta}$ is a vector of estimated coefficients shown in Table IV and $z_i(t)$ is a vector of values of firm specific covariates for firm *i* at time *t*. For example, the risk score of Firm 1 at time 2 is estimated using the estimated coefficients from Table IV and the values of seven firm specific covariates for firm 1 at the second year of the firm's life time.

Panel A in Table V presents the resulting risk scores and survival probabilities for ten randomly selected firms in the non-failed group and Panel B shows those for ten firms in the failed group.[18] Comparing these survival probabilities for the failed firms with those for the surviving firms at the same time horizons (life time), the failing firms have lower

[18] Risk scores and survival probabilities are presented for only twenty firms in our study sample due to limited space.

probabilities. However, in most cases the differences are not great and the survival probabilities for the failed firms are generally high, with several above 0.9.

The explanation for the foregoing seems to lie in the interaction between the risk score and the baseline hazard. Since the incidence of failure in the estimation sample is small, the risk of failure for an average firm is small. Consequently, although the baseline hazard rises through time, it remains small. Thus, to obtain a small probability of survival requires a substantial scaling up of the baseline hazard by the risk score. It appears in this analysis that the risk scores for failed firms are not often large enough to achieve the required scaling up.

[Table V about here]

7.2 MODEL VALIDATION

In order to test the predictive power of the dynamic model proposed in this study, we show comparisons of the performance of our dynamic model with that of the conventional static model. To do this, we have chosen the logit model, a widely used static technique. We use three versions of the logit model designed to match the time-varying Cox's regression models (Model 1, Model 2 and Model 3). This section presents the out-of-sample prediction results of each model and compares the performance of a model against one another.

Panel A and Panel B in Table VI presents each model's predictive accuracy out of sample based on ROC curves and the Brier Score, respectively. Predictive probabilities of failures out of sample are measured in the same mechanism as the model is constructed with the estimation sample. That is, based on multiple years of financial data available before the event of failure in the holdout sample, a set of firms' annual observations as at their financial year ends is injected into the model where each value of covariates is being updated with the function of time-varying variables, and the model generates the survival probabilities as far as the data allows. Based on the changing survival probabilities through time, we measure the area under ROC curve and the Brier Score at each predictive point in time since the beginning of out of sample period. For example a firm survives two years and fails in year three. The first year predictive survival probability based on the data observed at the end of year one in the holdout period will be evaluated at time horizon $t = 1$, and the second year forecast based on the data at the end of year two is assessed at $t = 2$. In year three, the third year forecast at $t = 3$ is based on the data at the end of year three, then the firm drops out of the holdout sample, because there is no data beyond year three.

In general terms, our dynamic model has stronger discriminatory power over the static logit model (See Panel A). There is a tendency for the dynamic model's ROC estimates to become even better at longer horizons, but not for Model 2. This result is in direct contrast to most bankruptcy studies, where predictive accuracy deteriorates sharply as the time horizon lengthens. Table VI shows that the static logit model exhibits diminishing accuracy as the time horizon increases.

Indeed, this result informs us that the underlying hazard plays a significant role in explaining the predictive power. The area under the ROC (AUROC) is increasing with time in the out-of-sample forecast because the pattern of baseline hazard in the estimation sample gets into line well with the pattern of baseline in the holdout sample. That is, both in the estimation sample and in the holdout sample, the number of failures is increasing over the period from year 1 to 3 and dropping at year 4 (refer to Table I), and this is expected to happen to the baseline hazard.

The result from Brier Score (Panel B in Table VI) also verifies that the dynamic models perform better than the static logit model; although, unlike the result from ROC, it does not provide the evidence that models become better calibrated as the time horizon gets longer.

[Table VI about here]

7.3 IMPLICATIONS OF MODELING WITH MACROECONOMIC RISK FACTORS AND WITH A NON-LINEAR APPROACH

Allowing for non-linearity between predictor variables and the risk of financial distress significantly enhances the model's discriminatory power shown in Panel A of Table VI. The superior predictive accuracy of Model 3 (approximately on average 80%) against that of Model 1 (approx. on average 70%) also provides the evidence to reject the argument that there might be overfitting effect leading to poor generalization in the holdout sample. This suggests that researchers should consider a non-linear approach to developing a more powerful bankruptcy prediction model.

As for the effect of macroeconomic covariates, contrary to expectations, a model that controls for the state of the economy (Model 2) does not improve predictive power over the same failure prediction model with no economy wide indicators (Model 1), as shown in Table VI. It is possible that the firm specific variables already embed the impact of the macroeconomic variables. If so, there would be very little need to control for the macroeconomic variables because they are already captured in the firm specific time-varying covariates. At the outset, we suggest that the information of economy wide conditions might be confounded in the baseline hazard, but the result provides no evidence to support this conjecture.

8. Summary and conclusions

Problems of time-varying predictor variables and baseline hazard estimates have been major obstacles to the application of survival analysis into multiple-period bankruptcy data. This study has taken a step towards solving these problems by applying the time-varying Cox's regression model to Australian financial distress prediction. We use seven firm-specific and four macroeconomic covariates, whose values are updated on a yearly basis from 1995 to 2006. The attractive feature of time-varying survival modeling is that it allows for dynamic changes of firm's risk levels and its corresponding survival probabilities through time.

We start by exploring the interaction between the financial distress risk and firm-specific accounting-based and market-based ratios. The result obtained suggests that firms with less operating liquidity, higher book leverage, less cash flow generating ability, smaller size and less market value relative to debt are more likely to fail, which is partly in line with the results found in Shumway (2001). In particular, the time-varying Cox model outperforms the static logit model in both short and long time horizons.

The inclusion of macroeconomic covariates does not make dramatic difference to the result of the base model, which affirms that we have a proper dynamic model and we are not imposing any restriction on the underlying hazard.

Finally when non-linearity between predictor variables and the risk of failure is taken into account, the result of the model improves considerably. The result obtained enables us to conclude that a non-linear approach makes an important additional contribution in predicting which firms would fail or not.

While the model has some predictive power, there is scope for improvement. It is still possible that there is a problem with untimely or less than reliable financial statement information, especially for those firms approaching financial distress. It may, therefore, be worthwhile to include more market-driven variables such as past stock returns, and stock returns volatility which can be observed more frequently than accounting data. Also, as suggested earlier, the exponentially weighted moving average should be used for each data point of the macroeconomic covariates to take the whole effect of the state of the economy, thus we properly interpret the current figure in the context of the trend.

References

Allison, P. D. (1995) *Survival analysis Using SAS: A Practical Guide*, SAS Institute Inc., Cary, North Carollina.

Altman, E. I. (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance* **23**(4), 589-609.

Altman, E. I., Haldeman, R. G. and Narayanan, P. (1977) Zeta™ analysis: A new model to identify bankruptcy risk of corporations, *Journal of Banking and Finance* **1**, 29-54.

Andersen, P. K. (1992) Repeated assessment of risk factors in survival analysis, *Statistical Methods in Medical Research* **1**, 297-315.

Australian Corporations Act 2001. Available at: http://www.austlii.edu.au/au/legis/cth/consol_act/ca2001172.

Australian Securities Exchange (2008) *Listing Rules*, Sydney, Australia, Available at: http://www.asx.com.au/supervision/rules_guidance/changes/pre_20080331/listing_rules1.htm

Baker, M. and Wurgler, J. (2002) Market timing and capital structure, *Journal of Finance* **57**, 1-32.

Bandopadhyaya, A. (1994) An estimation of the hazard rate of firms under Chapter 11 protection, *The Review of Economics and Statistics* **76**, 346-350.

Beaver, W. H. (1966) Financial Ratios as Predictors of Failure, *Journal of Accounting Research* **4**, 71-111.

Bellotti, T., and Crook, J. (2009) Credit Scoring with Macroeconomic Variables using Survival Analysis, *Journal of the Operational Research Society* to be updated.

Bonfim, D. (2009) Credit risk drivers: Evaluating the contribution of firm level information and of macroeconomic dynamics, *Journal of Banking & Finance* 33, 281-299.

Brailsford, T. J., Handley, J. C. and Maheswaran, K., 2008, Re-examination of the Historical Equity Risk Premium in Australia, *Accounting and Finance* 48, 73-97.

Campbell, J. Y., Hilscher, J. and Szilagyi J. (2005) In search of distress risk, unpublished working paper, Harvard University.

Carew, E. (1996) *The Language of Money 3,* St. Leonards, N.S.W, Allen & Unwin.

Castagna, A. D. and Matolcsy, Z. P. (1981) The Prediction of Corporate Failure: Testing the Australian Experience, *Australian Journal of Management* **6**, 23-50.

Chan, H., Faff, R., Gharghori, P. and Lajbcygier, P. (2008) Default Risk Modeling: A New Non-Linear Approach, Unpublished Working paper, University of Melbourne and Monash University

Chen, K. C. and Lee, C. J. (1993) Financial ratios and corporate endurance: A case of the oil and gas industry, *Contemporary Accounting Research* **9**(2), 667-694.

Chen, L., Yen, M., Wu, H., Liao, C., Liou, D., Kuo, H. and Chen, T. H. (2005) Predictive survival model with time-dependent prognostic factors: development of computer-aided SAS Macro program, *Journal of Evaluation in Clinical Practice* **11**(2), 181-193.

Coats, P.K and Fant, L.F. (1992) A neural network approach to forecasting financial distress, *The Journal of Business Forecasting*, 9-12.

Cox, D. R. (1972) Regression and life tables, *Journal of the Royal Statistical Society* **34**, 187-220.

Crapp, H. and Stevenson, M. (1987) Development of a method to assess the relevant variables and the probability of financial distress, *Australian Journal of Management* **12**(2), 221-236.

deListed (2006) *Delisted Companies*, NSW, Australia, Available at: http://www.delisted.com.au/delisted.aspx

Faulkender, M. and Petersen, M. A. (2005) Does the source of capital affect capital structure?, *Review of Financial Studies* **19**, 45-79.

Ford, B., and Taylor, K. (2005) Recent developments in Australian bond yields, in The Treasury, ed., *Economic Roundup*, Spring, 111-120.

Frydman, H., Altman E. I. and Kao, D. (1985) Introducing recursive partitioning for financial classification: The case of financial distress, *Journal of Finance* **40**(1), 269-291.

Gharghori, P., Chan, H. and Faff, R. (2006) Investigating the Performance of Alternative Default-Risk Models: Option-Based Versus Accounting-Based Approaches, *Australian Journal of Management* **31**, 207-234.

Hanley A. and McNeil B. (1982) The meaning and use of the area under a receiver operating characteristics (ROC) curve, *Radiology* **143**(1), 29-36.

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

Jones, S. and Hensher, D. A. (2004) Predicting firm financial distress: A mixed logit model, *The Accounting Review* **79**(4), 1011-1038.

Karfakis, C. and Phipps, A. (1996) Treasury note and bank bill rates, the risk premium and Australian monetary policy, *Australian Economic Paper* **35**, 321-333.

Koh, H. C. (1992) The sensitivity of optimal cutoff points to misclassification costs of Type I and Type II Errors in the going-concern prediction context, *Journal of Business Finance and Accounting* **19**, 187-197.

Lane, W. R., Looney S. W. and Wansley, J. W. (1986) An application of the Cox proportional hazards model to bank failure, *Journal of Banking and Finance* **10**(4), 511-531.

LeClere, M. J. (2000) The occurrence and timing of events: Survival analysis applied to the study of financial distress, *Journal of Accounting Literature* **19**, 158-189.

Liu, J. (2004) Macroeconomic determinants of corporate failures: evidence from the UK, *Applied Economics* **36**, 939-945.

Mason, S. J. and Graham, N. E. (1999) Conditional probabilities, relative operating characteristics, and relative operating levels. *Weather and Forecasting* **14(**5), 713-725.

Nothman, P. (1993) *Record of Liquidations and Receiverships*, Financial Analysis Publications, Sydney, NSW.

Odom, M. D. and Sharda, R. (1990) A neural network model for bankruptcy prediction, *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN)* **2**, 163-168.

Ohlson, J. A. (1980) Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* **18**(1), 109-131.

Pacey, J. W. and Pham, T. M. (1990) The predictiveness of bankruptcy models: methodological problems and evidence, *Australian Journal of Management* **15(**2), 315-337.

Partington, G., Russel, P., Stevenson, M. and Torbey, V. (2001) Predicting return outcomes to shareholders from companies entering chapter 11 bankruptcy, *Managerial Finance* **27**, 78-96.

Rajan, R. G. and Zingales L. (1995) What do we know about capital structure? Some evidence from international data, *Journal of Finance* **50**, 1421-1460.

Richardson, F.M., Kane, G.D. and Lobingier, P. (1998) The Impact of Recession on the Prediction of Corporate Failure, *Journal of Business Finance and Accounting* **25**, 167-186.

Rosch, D. and Sheule, H. (2005) A Multi-Factor Approach for Systematic Default and Recovery Risk, *The Journal of Fixed Income* **15**, 63-75.

Sobehart J. R. and Keenan S. (2001) Measuring default accurately, Credit Risk Special Report, *Risk*, 31-33.

Sobehart, J. R. and Stein, R. M. (2000) Moody's public firm risk model: A hybrid approach to modeling short term default risk, Moody's Rating Methodology.

Shumway, T. (2001) Forecasting bankruptcy more accurately: A simple hazard model, *The Journal of Business* **74**, 101-124.

Tam, K. Y. and Kiang, M. Y. (1992) Managerial applications of neural networks: The case of bank failure predictions, *Management Science* **38**, 926-947.

Tan, C. N. and Dihardjo H. (2001) A study on using artificial neural networks to develop an early warning predictor for credit union financial distress with comparison to the probit model, *Managerial Finance* **27**, 56-77.

Wadhwani, S. B. (1986) Inflation bankruptcy, default premia and the stock market, Economic Journal, 96, 120–38.

Wheelock, D. C. and Wilson, P.W. (1995) Explaining bank failures: Deposit insurance, regulation, *Review of Economics & Statistics* **77**, 689-700.

Wilson, R. L. and Sharda, R. (1994) Bankruptcy prediction using neural networks, *Decision Support Systems* **11**, 545-557.

Wong, B., Partington, G., Stevenson, M. and Torbey, V. (2007) Surviving chapter 11 bankruptcies: Duration and payoff?, *ABACUS* **43**(3), 363-387.

Table I
Data sample

This table shows the total number of firm-year observations in our study sample, the number of non-failed firm-year observations and the number of failed firm-year observations and the percentages of failed to non-failed firm-year observations for every year over the sample period of 1995 – 2006. Panel A shows the failure distribution of firm-year observations in calendar years, and Panel B shows that of firm-year observations in the risk set arranged according to the event time. Our study sample includes financially distressed (failed) firm data of publicly traded companies on the Australian Securities Exchange (ASX) between 1995 and 2006. Firms which are in the financial sector, as indicated by their GICS code, are excluded from the sample. In our study sample, we have 13,387 firm-year observations in total where 133 failed observations are found.

Year	No of firm-year observations	No of non-failed firm-year observations	No of failed firm-year observations	Percentage of Failed firms to the total number of firms
1995	674	671	3	0.45%
1996	728	723	5	0.69%
1997	773	762	11	1.42%
1998	806	799	7	0.87%
1999	873	862	11	1.26%
2000	983	968	15	1.53%
2001	1,043	1,025	18	1.73%
2002	1,054	1,037	17	1.61%
2003	1,061	1,051	10	0.94%
2004	1,128	1,115	13	1.15%
2005	1,210	1,194	16	1.32%
2006	1,240	1,233	7	0.56%

Table II
Predictor variables used in previous bankruptcy studies

This table shows predictor variables used in a number of previous bankruptcy studies in the USA and Australia. We take into consideration those found to be useful from the recent major studies by Sobehart and Stein (2000), Shumway (2001), and Campbell et al. (2005). As this study is being conducted using Australian data, we also include some of the variables found to be useful in Australian studies by Castagna and Matolscy (1981), Jones and Hensher (2004) and Gharghori et al. (2006). Nine predictor variables are initially considered measuring profitability, leverage (book and market), liquidity, cash flow generating ability, size and growth opportunity, excess return and market sensitivity. Two market variables, firm's past excess return and price volatility, are excluded in our model as there is insufficient data on failed observations.

Description	Ratio / Variable	Abbreviation	Source
Accounting-based Variables			
Return on Assets (Profitability)	Net Income / Total Assets	NI / TA	Shumway (2001) Sobehart (2000)
Operating Liquidity	Working Capital / Total Assets	WC / TA	Castagna (1981) Gharghori et al.(2006)
Leverage	Total Liabilities / Total Assets	TL / TA	Gharghori et al. (2006) Shumway (2001)
Cash flow generating ability	Net Cash Flow from Operations / Total Assets	CF / TA	Jones & Hensher (2004)
Market-based Variables			
Market-to-Book Ratio	Market Capitalization / Total Equity	MB	Campbell et al. (2005) Gharghori et al. (2006)
Firm Relative Size	ln (Firm Market Capitalization $_{i,t}$ / Total ASX Market Value $_{t}$)	RSIZE	Campbell et al. (2005) Shumway (2001)
Market Leverage	Market Capitalization / Total Liabilities	MC / TL	Altman (1968) Gharghori et al. (2006)
Macroeconomic Variables			
Return on ASX AOI	$(r_{m,t} / r_{m,t-1}) - 1$	AOI	Bonfim (2009) Rosch & Scheule (2005)
Yield Spread	10 Year Government Bond yield – 90 day Bank Accepted Bills	YIELD SPREAD	Partington et al. (2001) Rosch & Scheule (2005)
Rate of Change in Consumer Confidence Index	$(CCI_t / CCI_{t-1}) - 1$	CCI	Bellotti & Crook (2009)
Rate of Change in Consumer Price Index	$(CPI_t / CCI_{t-1}) - 1$	CPI	Liu (2004)

Table III
Descriptive statistics of firm specific variables

This table shows descriptive statistics of firm specific variables for firm-year observations of the ASX listed firms. Each firm has multiple observations according to firm age (duration). The data is reported after truncation of the top and bottom one percent of the distribution for each variable. NI/TA is the firm's net income divided by its total assets; WC/TA is the firm's working capital divided by its total assets; TL/TA is the ratio of firm's total liabilities to its total assets; CF/TA is the ratio of firm's net operating cash flow to its total assets; MB is the market-to-book ratio of the firm's market capitalization to its total equity; RSIZE is the firm's relative size measured as the natural logarithm of the ratio of each firm's market capitalization to that of the ASX All Ordinary Index; MC/TL is the firm's market capitalization divided by its total liabilities. Panel A shows descriptive statistics for all firm-year observations for the entire sample over the period of 1995 – 2006. There are a total of 1,703 non-financial firms and 11,573 firm-year observations in the sample. The description of Panel B is as for Panel A except that it applies to an estimation sample over the period of 1995 – 2002. There are a total of 1,267 non-financial firms and 6,934 firm-year observations in the sample. The description of Panel C is also as for Panel A except that it applies to a holdout sample for the period of 2003 – 2006. There are a total of 1,455 non-financial firms and 4,639 firm-year observations in the sample.

Panel A: Descriptive statistics for the entire sample

Variables	Distress group	N	Mean[1]	Median	Std. Dev.	Minimum	Maximum
NI / TA	Non-failed	11,440	-0.2113530	-0.0214504	0.6333925	-4.3934734	0.3786786
	Failed	133	-1.3751883	-0.2150772	4.7324547	-36.621599	0.3057557
WC / TA	Non-failed	11,440	0.0448330	0.0089766	0.1975598	-0.7373647	0.6930535
	Failed	133	-0.1210392	0.0073113	0.8486489	-4.6596426	0.7887640
TL / TA	Non-failed	11,440	0.3762798	0.3454462	0.3387747	0.0048200	2.1600516
	Failed	133	1.0329740	0.5850310	1.9887918	0.0092766	14.322398
CF / TA	Non-failed	11,440	-0.0733434	-0.0034486	0.3192611	-1.9000658	0.4206350
	Failed	133	-0.4447830	-0.0666129	1.7017677	-10.618299	0.5881350
MB	Non-failed	11,440	2.5028849	1.5393512	3.8137582	-6.4046426	25.955423
	Failed	133	1.3837192	1.0521490	6.1556743	-24.823141	20.988845
RSIZE	Non-failed	11,440	-9.9007978	-10.249197	2.0450529	-13.455038	-4.2356670
	Failed	133	-10.567063	-10.629307	1.7811200	-14.316423	-5.4669017
MC / TL	Non-failed	11,440	22.682022	3.6451045	53.539340	0.1109045	368.24548
	Failed	133	16.426137	1.4309966	72.212903	0.0061208	573.39857

Panel B: Descriptive statistics for estimation sample							
Variables	Distress group	N	Mean[1]	Median	Std. Dev.	Minimum	Maximum
NI / TA	Non-failed	6,847	-0.1983632	-0.0118899	0.5985206	-4.0460359	0.3069298
	Failed	87	-0.7708188	-0.1507157	1.8185842	-10.924480	0.3057557
WC / TA	Non-failed	6,847	0.0579390	0.0177774	0.1954693	-0.6430925	0.7296216
	Failed	87	-0.0978951	0.0476481	0.8724769	-4.6596426	0.8152995
TL / TA	Non-failed	6,847	0.3812875	0.3694325	0.3153611	0.0045697	1.8801400
	Failed	87	0.9143171	0.5890799	1.6906323	0.0092766	14.322398
CF / TA	Non-failed	6,847	-0.0531791	0.0051411	0.2791978	-1.6299930	0.3890125
	Failed	87	-0.2797293	-0.0394291	1.3287582	-10.588315	3.1344902
MB	Non-failed	6,847	2.2376136	1.3472436	3.3228769	-4.5730477	22.863319
	Failed	87	1.3804670	0.8772364	6.0826064	-26.449203	20.988845
RSIZE	Non-failed	6,847	-9.7719011	-10.122221	2.0445546	-13.344636	-4.1501507
	Failed	87	-10.383899	-10.402640	1.7987942	-14.316423	-5.3528381
MC / TL	Non-failed	6,847	20.642607	2.9206647	52.377842	0.0861023	364.40868
	Failed	87	6.6312887	0.9641672	18.015574	0.0020494	118.89981

Panel C: Descriptive statistics for holdout sample							
Variables	Distress group	N	Mean[1]	Median	Std. Dev.	Minimum	Maximum
NI / TA	Non-failed	4593	-0.2313561	-0.0373489	0.6925602	-4.9642924	0.4921973
	Failed	46	-4.3772405	-0.3663879	18.607822	-122.19686	0.3667659
WC / TA	Non-failed	4593	0.0238634	-0.0009276	0.2046794	-0.9616614	0.5967544
	Failed	46	-1.2710254	-0.0252066	8.1961797	-55.571998	0.7085983
TL / TA	Non-failed	4593	0.3697353	0.3047606	0.3792409	0.0054318	2.5779545
	Failed	46	2.1730091	0.5696045	8.3126787	0.0019821	56.448152
CF / TA	Non-failed	4593	-0.1040425	-0.0221553	0.3753381	-2.2797480	0.4554392
	Failed	46	-1.5850209	-0.1037934	7.6858770	-51.255926	0.5881350
MB	Non-failed	4593	2.9034679	1.8408969	4.5570747	-9.2704149	30.959837
	Failed	46	3.3070721	1.1888777	17.208894	-24.823141	110.80619
RSIZE	Non-failed	4593	-10.092743	-10.447834	2.0304011	-13.553226	-4.3916314
	Failed	46	-10.914813	-11.106567	1.7290229	-14.491692	-7.2078032
MC / TL	Non-failed	4593	25.787648	5.0740957	55.503579	0.1332688	375.51033
	Failed	46	38.319946	1.8335542	135.53783	0.0239425	728.36605

Panel D: Descriptive Statistics of NI/TA grouped by Year						
Year	N	Mean	Median	Std. Dev.	Minimum	Maximum
1995	674	-0.1155786	0.0190306	0.5374183	-4.6944134	0.3698276
1996	728	-0.0774989	0.0132043	0.3641876	-4.6944134	0.3698276
1997	773	-0.1080678	0	0.4014774	-4.9638071	0.3698276
1998	806	-0.1733446	0	0.6198979	-4.6944134	0.3698276
1999	873	-0.1681899	0	0.5646330	-4.6944134	0.3698276
2000	992	-0.1396347	-0.0049617	0.4931552	-4.6944134	0.3698276
2001	1,034	-0.3593912	-0.0401468	0.9277242	-10.9244802	0.3698276
2002	1,054	-0.3696295	-0.0649399	0.9007334	-8.5540541	0.3698276
2003	1,061	-0.3203504	-0.0386937	0.8473606	-7.5795358	0.3698276
2004	1,128	-0.2209747	-0.0294531	0.7211106	-9.2613779	0.3698276
2005	1,210	-0.2237759	-0.0410236	0.6274303	-4.6944134	0.3698276
2006	1,240	-0.2827157	-0.0384873	1.5482209	-36.6215998	0.3698276

Panel E: Descriptive Statistics of NI/TA grouped by Firm Size						
Quartile	N	Mean	Median	Std. Dev.	Minimum	Maximum
1	2893	-0.5247709	-0.1765108	1.1743217	-36.6215998	0.3698276
2	2893	-0.2784907	-0.0794690	0.9163825	-36.6215998	0.3698276
3	2893	-0.0939643	0.0137022	0.3983990	-4.6944134	0.3698276
4	2894	0.0341840	0.0499120	0.1886210	-4.6944134	0.3698276

Panel F: Descriptive Statistics of Macroeconomic variables					
Variables	Mean	Median	Std. Dev.	Minimum	Maximum
AOI	0.0955128	0.1012200	0.0873570	-0.0861800	0.2993700
YIELD SPREAD	0.0093166	0.0107000	0.0063333	-0.0074000	0.0195000
CCI	-0.0006173	0.0142595	0.0990238	-0.1910149	0.3227138
CPI	0.0280150	0.0284006	0.0185175	-0.0033306	0.0607780

[1] Wilcoxon-Mann-Whitney test is carried out for each variable to test the significance of differences of firm characteristics between failed and non-failed groups. The test shows the differences are statistically significant at the 1% level for all variables except for the WC/TA for the entire sample and estimation sample, and at the 5% level for all variables, with the exception of WC/TA, for holdout sample.

Table IV
Parameter Estimates and Model Fit Summary for Model 1, 2 and 3

Panel A shows the total number of firms, the number of failed firms and the number of censored (non-failed) firms and percentages of censored to the total number of firms in our estimation sample over the period of 1995 – 2002. Panel B reports the parameter estimates of Cox hazards model with time-varying covariates, $h_i(t \mid z(t)) = h_0(t) \cdot \exp\left\{\sum_{j=1}^{P} \beta_j z_j^i(t)\right\}$. NI/TA is the firm's net income divided by its total assets; WC/TA is the firm's working capital divided by its total assets; TL/TA is the ratio of firm's total liabilities to its total assets; CF/TA is the ratio of firm's net operating cash flow to its total assets; MB is the market-to-book ratio of the firm's market capitalization to its total equity; RSIZE is the firm's relative size measured as the natural logarithm of the ratio of each firm's market capitalization to that of the ASX All Ordinary Index; MC/TL is the firm's market capitalization divided by its total liabilities. Return on ASX AOI is the annualized return on the ASX All Ordinaries Index; Yield spread is 10 year Commonwealth Government Treasury Bond yields less 90-day Bank Accepted Bills (BAB) yields; Return on CCI is annualized return on Consumer Confidence Index; Return on CPI is annualized return on Consumer Price Index indicating the inflation effect. A positive coefficient on a particular variable implies that the hazard rate is increasing in that variable. Parameter estimates are given first followed by p-values. The chi-square of the likelihood ratio test for the model fit is reported in Panel C. The last seven variables suffixed "_d" are the firm specific variables whose values has been replaced by the default rates.
*** significant at 1%, ** significant at 5%, * significant at 10%

Panel A: Number of failed and censored firms in the estimation sample

Total	Failed	Censored	Percent Censored
1,267	87	1,180	93.13

Panel B: Parameter estimates

Variables	Expected Sign	Model 1	Model 2	Model 3
NI / TA	-	-0.03198	-0.08537	
WC / TA	-	-0.42753*	-0.40896*	
TL / TA	+	0.45184***	0.46205***	
CF / TA	-	-0.52675***	-0.48696***	
MB	+	0.01182	0.01176	
RSIZE	-	-0.10635*	-0.11700**	
MC / TL	-	-0.02247**	-0.02258**	
Return on ASX AOI	-		2.32016	
Yield Spread	+		-71.15026**	
Return on CCI	-		-0.77179	
Return on CPI	(?)		-0.30485	
NI / TA_d	+			25.07251**
WC / Ta_d	+			12.57555
TL / Ta_d	+			23.13604***
CF / Ta_d	+			45.15919**
MB_d	+			-2.41531
RSIZE_d	+			35.31048*
MC / TL_d	+			33.80514***

Panel C: Model Goodness of fit				
	Without Covariates	Model 1	Model 2	Model 3
-2 LOG L[1]	1176.910	1104.337	1093.277	1076.892
Likelihood Ratio[2]		72.5728***	83.6336***	100.0184***
Degrees of Freedom		7	11	7

[1] The term -2 LOG L is the logarithm of the maximum likelihood estimator for the model. It measures the goodness of fit of the presented model, and smaller values indicate a more desirable model.
[2] The likelihood ratio tests the null hypothesis that all coefficients except for the constant are zero. It is calculated as $LR = 2(\ln L - \ln L_0)$, where $\ln L$ is the log likelihood of the estimated model and $\ln L_0$ is the log likelihood of the restricted model only with a constant. The more likelihood is lost by imposing the restriction, the larger the Likelihood Ratio statistic will be.

Table V

The time-dependent risk scores and survival probabilities of twenty firms selected from estimation sample

This table presents dynamic changes of risk scores (Score) and survival probabilities (P(S)) by time horizon using a time-dependent Cox regression model, $h_i(t \mid z(t)) = h_0(t) \cdot \exp\left\{\sum_{j=1}^{P} \beta_j z_j^i(t)\right\}$, for twenty randomly selected firms. The time-dependent risk score can be calculated for each firm as $\hat{\beta} * z_i(t)$, where $\hat{\beta}$ is a vector of estimated coefficients shown in

Table IV and $z_i(t)$ is a vector of values of covariates for firm *i* at time *t*. The survival probabilities are calculated using $\hat{h}_i(t) = \hat{h}_0(t) \times \exp\left(\hat{\beta} \cdot z_i(t)\right)$ and $\hat{S}_i(t) = \exp\left[-\int \hat{h}_i(u)du\right]$ in Section 3.3 and the estimated risk scores. Panel A shows the resulting risk scores and survival probabilities for ten non-failed firms, and Panel B shows those for ten failed firms. Each firm has its life time and the risk score and the survival probabilities are generated for different time horizons.

Panel A: 10 Non-failed firms from Estimation Sample

Firm (*i*)	Firm 1		Firm 2		Firm 3		Firm 4		Firm 5		Firm 6		Firm 7		Firm 8		Firm 9		Firm 10	
Lifetime (yrs)	3		4		5		6		6		6		7		7		7		8	
Status	Non-failed		Non-failed		Non-failed		Non-failed		Non-failed		Non-failed		Non-failed		Non-failed		Non-failed		Non-failed	
Time horizon	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)
0	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
1	0.8028	0.9995	0.4400	0.9995	0.6588	0.9995	0.6066	0.9995	0.0218	0.9995	0.6523	0.9995	1.3563	0.9995	0.7058	0.9995	0.4239	0.9995	0.4048	0.9995
2	0.7545	0.9984	0.4351	0.9987	0.7027	0.9985	0.4902	0.9986	0.6080	0.9990	0.8000	0.9985	0.9727	0.9976	0.6949	0.9985	0.4559	0.9987	0.4308	0.9988
3	1.4228	0.9923	1.2470	0.9958	0.9608	0.9941	0.6850	0.9951	0.5995	0.9959	1.0386	0.9939	0.217	0.9941	0.6907	0.9948	-0.043	0.9959	0.292	0.9961
4			1.2470	0.9933	1.0440	0.9922	0.6362	0.9936	0.3294	0.9945	1.5352	0.9919	0.2149	0.9931	0.5620	0.9934	0.2558	0.9952	0.2837	0.9951
5					1.0440	0.9900	0.4957	0.9922	0.5185	0.9935	0.6586	0.9883	0.0675	0.9922	0.4963	0.9920	0.4934	0.9942	0.347	0.9941
6							0.4957	0.9908	0.5185	0.9921	1.3838	0.9867	-0.665	0.9913	0.5995	0.9906	0.3288	0.9928	0.367	0.9929
7													0.5837	0.9795	0.7558	0.9160	0.7962	0.8962	0.5261	0.9302
8																			0.5217	0.8266

Panel B: 10 Failed firms from Estimation Sample																				
Firm (*i*)	Firm 11		Firm 12		Firm 13		Firm 14		Firm 15		Firm 16		Firm 17		Firm 18		Firm 19		Firm 20	
Lifetime (yrs)	2		3		3		4		5		5		6		7		8		8	
Status	Failed		Failed		Failed		Failed		Failed		Failed		Failed		Failed		Failed		Failed	
Time horizon	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)	Score	P(S)
0	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
1	1.4936	0.9995	1.4474	0.9987	-3.898	0.9995	-1.037	0.9995	0.6032	0.9995	0.4369	0.9995	0.6551	0.9995	1.0157	0.9991	-2.332	0.9999	1.5933	0.9985
2	5.8930	0.7824	1.2389	0.9976	5.0650	0.9951	2.2117	0.9872	1.5925	0.9942	0.9689	0.9946	-0.493	0.9960	1.0093	0.9982	-0.509	0.9997	1.3987	0.9972
3			5.0935	0.9335	5.0650	0.8861	0.8663	0.9806	4.0695	0.9907	0.9926	0.9927	0.2650	0.9955	1.2771	0.9967	1.3474	0.9982	1.3139	0.9956
4							0.8663	0.9788	4.0695	0.9461	1.0542	0.9906	2.0366	0.9945	1.5099	0.9949	1.5407	0.9963	1.3159	0.9941
5									4.0695	0.9013	1.0542	0.9883	2.0732	0.9882	2.0687	0.9917	1.4478	0.9946	1.6356	0.9920
6													1.1564	0.9754	4.5098	0.9470	1.2462	0.9928	2.2953	0.9871
7															4.5099	0.9006	2.7002	0.9847	2.6083	0.9797
8																	6.2260	0.0500	5.0260	0.3993

Table VI
Predictive Accuracy

This table shows the area under the ROC (AUROC) curve and the Brier Score of the out-of-sample survival functions and logit functions. The AUROC measures the discriminatory power of the models and the higher the AUROC, the better the model. Predictions made at random have an AUROC of 0.5 and models that do not beat this benchmark have no predictive power. Panel A examines predictive accuracy using the ROC curves and describes the area under the ROC curves for holdout sample from 2003 to 2006. Holdout sample is reserved for the purpose of out-sample prediction, and whose predictive accuracy is tested against the estimated time-varying dynamic Cox hazards mode and the conventional logit model. Panel B shows the Brier Score, which measures the deviation between the predicted probability of a failure event and the actual outcome of the event at the level of an individual company. The smaller the Brier Score, the better the model. . The naïve forecast is based on the proportion of defaults to the estimation sample.

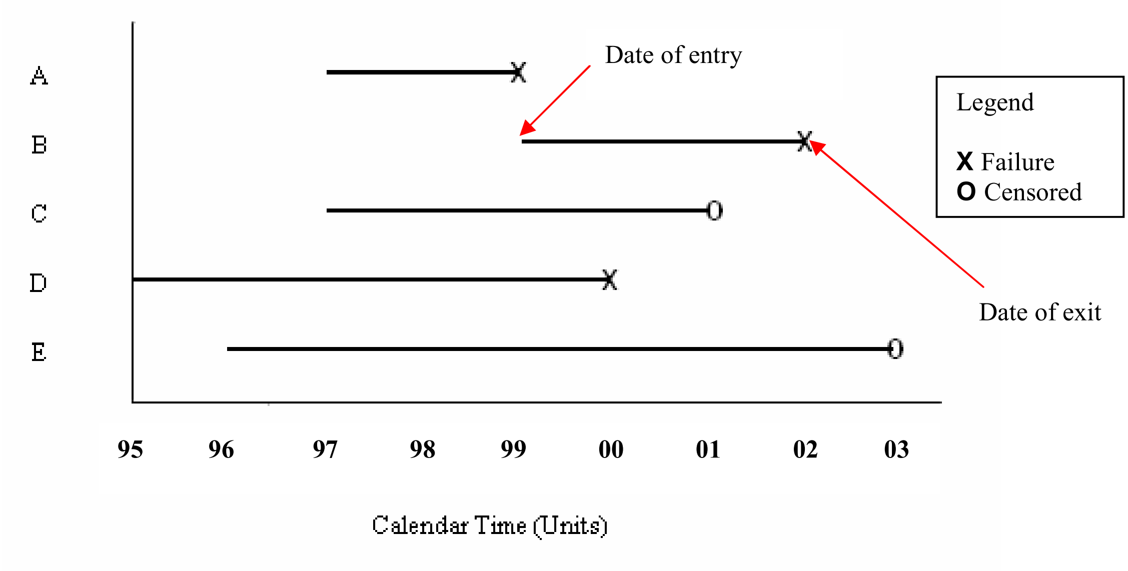
Panel A: Predictive accuracy over holdout sample – Area under the **ROC** curve for out-of-sample prediction

		Model 1		Model 2		Model 3	
Time Horizon	Random Forecast	Dynamic Cox	Logit	Dynamic Cox	Logit	Dynamic Cox	Logit
t = 1	0.5	0.680	0.644	0.683	0.488	0.784	0.684
t = 2	0.5	0.709	0.669	0.704	0.666	0.787	0.696
t = 3	0.5	0.720	0.506	0.678	0.646	0.819	0.771
t = 4	0.5	0.701	0.572	0.584	0.536	0.814	0.848

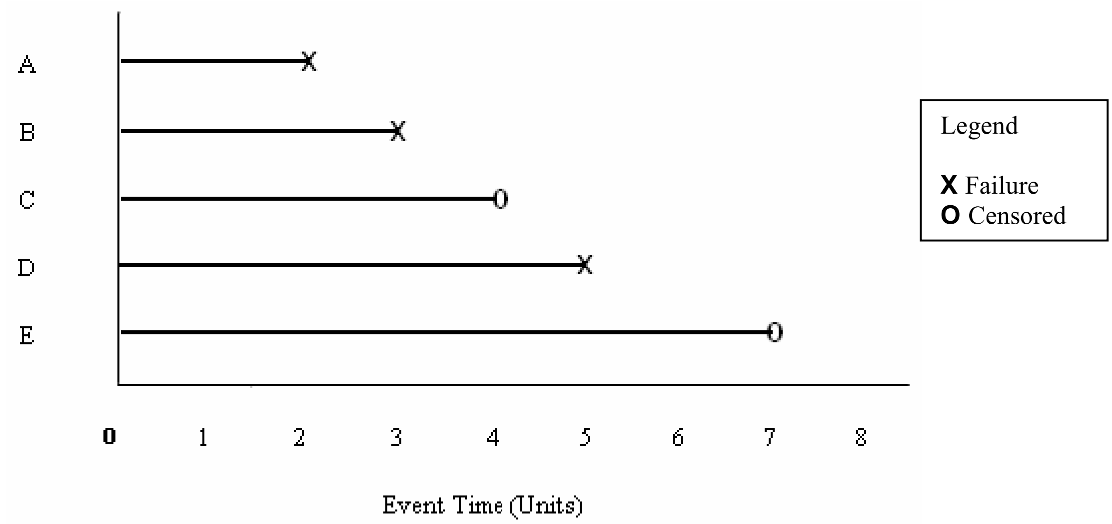
Panel B: Predictive accuracy over holdout sample – **Brier Score** for out-of-sample prediction

		Model 1		Model 2		Model 3	
Time Horizon	Naive Forecast	Dynamic Cox	Logit	Dynamic Cox	Logit	Dynamic Cox	Logit
t = 1	0.069	0.031	0.034	0.031	0.211	0.031	0.033
t = 2	0.070	0.024	0.028	0.024	0.407	0.025	0.026
t = 3	0.064	0.019	0.026	0.018	0.019	0.018	0.022
t = 4	0.053	0.005	0.012	0.019	0.013	0.025	0.007

Figure 1
Calendar time vs. Event time
Panel A: Arrangement of Firms in the Risk set according to Calendar Time

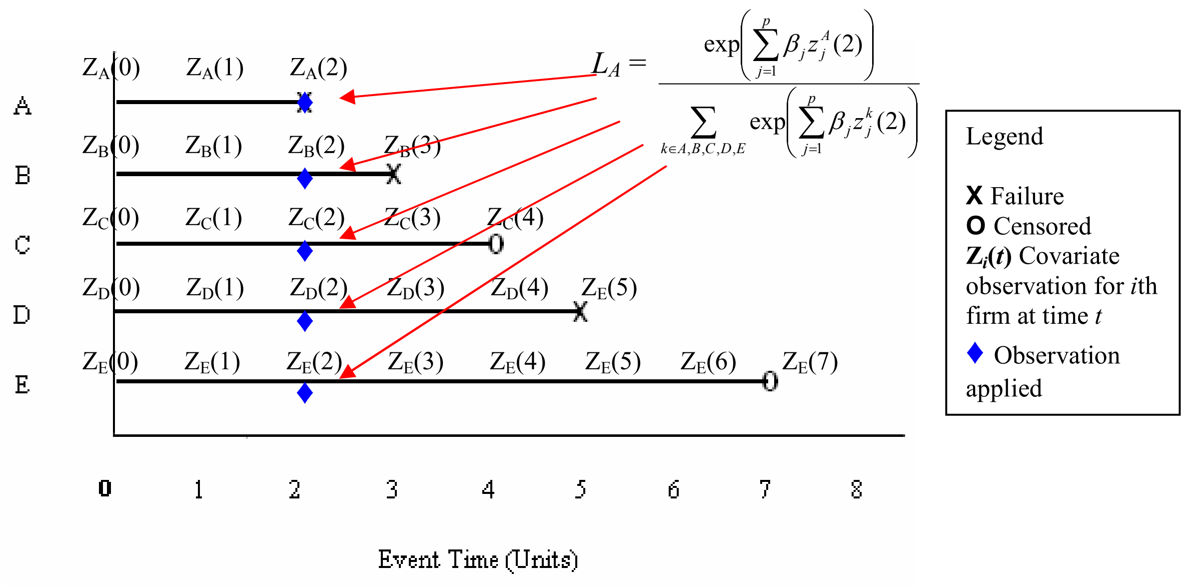


Panel B: Arrangement of Firms in the Risk set according to Event Time



This Figure presents a graphical demonstration of the difference between arranging the data in terms of “calendar time” and “event time”. Panel A illustrates observations of A to E arranged in calendar time while Panel B in event time. It is noted that observations can enter the study at different times in Panel A, however; every observation enters the study at event time 0 in Panel B. The length of the line indicates the life time of the observation. An “X” market at the end of the line denotes the event of failure, whereas an “O” marker indicates that the observation has been censored for reasons other than failure.

Figure 2
Calculation of the Likelihood for the Failure of Firm A in a Time-Varying Model



The figure illustrates the arrangement of the data in forming the likelihood function in a time-varying model. Observations of A to E are arranged in event time. The model inputs are refreshed at every incident of failure to reflect the observation's covariate values in the risk set at that particular time. Every time an event of failure is recorded, the vector of coefficients is re-estimated. The length of the line indicates the life time of the observation. An "X" market at the end of the line denotes the event of failure, whereas an "O" marker indicates that the observation has been censored for reasons other than failure.