國 立 交 通 大 學

財 務 金 融 研 究 所

碩 士 論 文

台灣股票市場違約風險與 Fama-French因子之關係

The Relationship Between Default Risk And Fama-French Factors in Taiwan

研 究 生：王文娟

指導教授: 王克陸 博士

中 華 民 國 九 十 七 年 六 月

台灣股票市場違約風險與Fama-French因子之關係

The Relationship Between Default Risk And Fama-French Factors in Taiwan

研 究 生 ： 王文娟　　Student : Wen-Chuan Wang

指導教授 ： 王克陸 博士　　Advisor : Dr. Kehluh Wang

國 立 交 通 大 學

財 務 金 融 研 究 所 碩 士 班

碩 士 論 文

A Thesis
Submitted to Graduate Institute of Finance
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master of Science
in
Finance

June 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年六月

台灣股票市場違約風險與Fama-French因子之關係

研究生：王文娟　　　　　　　指導教授：王克陸 博 士

國立交通大學財務金融研究所碩士班

2008年 六月

摘要

本研究以台灣股票市場研究違約機率與股票報酬的關係。我們選擇三個預測財務危機的模型，Logit 模型，Probit 模型和離散時間危險模型。利用 ROC 和錯誤分類表比較此三種預測財務危機的模型，實證發現相對於 Logit 和 Probit 模型，離散時間危險模型擁有較準確的預測財務危機的能力。我們以離散時間危險模型計算的違約機率和四因子模型分析違約因子和 Fama-French 因子的關係。實證發現，在考慮投資組合或是考慮個股股票之下，SMB 和 HML 因子係數的比重均與違約機率有相關。因此，實證結果證實 SMB 和 HML 是財務危機相關因子。

關鍵字：違約風險、Logit模型、Probit模型、離散時間危險模型、股票報酬

The Relationship Between Default Risk And
Fama-French Factors in Taiwan

Student: Wen-Chuan Wang Advisor: Dr. Kehluh Wang

Graduate Institute of Finance

National Chiao Tung University

June 2008

ABSTRACT

We study how the default probability relates to the stock returns in Taiwan stock market. Three prediction models are chosen, Logit Model, Probit Model, and Discrete-Time Hazard Model. We find that Discrete-Time Hazard model outperforms the other two models in predicting financial distress. Using the default probabilities predicted by Discrete-Time Hazard Model, we analyze the relationship between default risk and the Fama and French factors, SMB and HML, by running the four-factor model. The empirical results show that both portfolio and individual stock factor loadings are related to the estimated default probabilities. This result supports the interpretation on SMB and HML as distress related factors.

Keywords: default risk; Logit model; Probit model; Discrete-Time Hazard Model; stock return

Acknowledgment

在財金所的兩年裡，老師、學長姐和同學給予我許多珍貴的回憶。感謝所上老師在課業上的教導，讓我獲得寶貴的基礎知識。最要感謝的是指導教授，王克陸老師。在研究過程和論文格式撰寫上，老師給我許多的指導及建議，並且在未來人生規劃方面給予我許多的啟發。此外，也感謝口試委員—彭雅惠老師、許明峰老師及黃星華老師，對於此篇論文給予諸多寶貴的建議和匡正。

還要感謝財金所的學長、學姐和同學們。感謝學長學姐給予我論文的指導及建議。在寫論文的過程中，以文總是給予我許多的幫助，並且和渝薇、詩政總是陪伴著我一起努力。共同奮鬥的佳彣，我們在艱困中一直互相鼓勵，給予信心。也感謝阿 sam 和振綱，和我一起用功讀信用風險的書。

最後還要感謝我的父母，一直陪伴在我身邊，關心鼓勵我。謝謝堂姊們對我的關心、期許與鼓勵。衷心感謝所有幫助我的人。謹以此文獻給關心我的家人、師長與朋友。

王文娟　　謹誌於

交通大學財務金融研究所

中華民國九十七年六月

Contents

List of Tables

List of Figures

1. Introduction

Sharpe (1964) and Lintner (1965) propose the capital assets pricing model (CAPM) to find whether the stock has the excess return, and suggest that there is a linear relation between the expected returns on stocks and market βs and that the market β is the only factor that can explain the expected return on stocks. Since 1980s, some studies find that the rule of firm characteristics, like that firm size (Banz, 1981), book-to-market ratios (Rosenberg, Reid and Lanstein, 1985), and earnings-price ratios (Basu, 1983), leverage (Bhandari, 1988) etc., can explain the cross-section return, those firm characteristics are not concluded in CAPM.

However, in the empirical asset pricing literature, it has long been argued that the cross-section of stock returns is related to risk factors associated with systematic financial distress. The intertemporal capital asset pricing model (ICAPM) proposed by Merton (1973) and the arbitrage pricing theory (APT) proposed by Ross (1976), consider that there are many factors to affect the asset return, not only market risk. Chan, Chen, and Hsieh (1985) and Chen, Roll, and Ross (1986) find that a default factor, the spread between high- and low-grade bonds, has a significant contribution in explaining the cross-section of stock returns. Asset pricing theory claims stocks that underperform when the economy is in a high-distress state should reward the investors who hold them with higher expected returns as a compensation for bearing this non-diversifiable risk. Hence, these high returns would be partially unexplainable by a model that does not account for a distress factor, and therefore would be considered apparent mispricngs.

The existence of “pricing anomalies” such as the size and book-to-market effects has in fact been widely documented. The essence of the pricing anomalies lies in the fact that they cannot be justified by the return’s covariance with the market factor. The difference in market βs cannot explain the return differential between small and large firms, and between stocks with high and with low book-to-market values. Since market risk alone dose not price these stock, some other factors can explain the unexplained part. In their seminal papers, Fama and

French (1993) identify the two stock market factors related to size and book-to-market, i.e., SMB and HML, that in conjunction with the market factor form an impressive pricing model. In a later paper, Fama and French (1995) find that firms with high book-to-market tend to be relatively distressed, coming from a persistent period of negative earnings, and conversely, low book-to-market firms is associated with sustained strong profitability. They suggest that size and in particular book-to-market capture a firm's level of financial distress.

A number of studies have tried to link default risk and stock returns with mixed and contradictory results. Altman (1993) finds that for most distressed firms subsequent average returns are lower. Dichev (1998) finds that bankruptcy risk is not rewarded by higher returns and concludes that a distress factor cannot be at the origin of the size and book-to-market effect. In particular, he finds that portfolios formed on the basis of a distress measure, whether Altman's Z-score or Ohlson's O-score, have returns inversely related to bankruptcy risk, a high probability of default is associated with low average returns. After examining the cross-sectional relation between stock returns and bankruptcy measures, as well as size and book-to-market, he concludes that the fact that firms with low bankruptcy risk outperform firms with high bankruptcy risk can only be explained by a mispricing argument. Griffin and Lemmon (2002) measure bankruptcy risk by using the Ohlson's O-score, and find that the low return of high default-risk firms is driven by low book-to-market stocks with extremely low returns. They attribute these very low returns to mispricing due to a high degree of information asymmetry proxied by low analyst coverage.

By contrary to above studies, Lang and Stulz (1992), and Denis and Denis (1995) find that bankruptcy risk is related to aggregate factors, which implies that bankruptcy risk may be systematic. Fama and French (1996) suggest that small value stocks tend to be firms in distress (with high financial leverage and earnings uncertainty), with higher returns due to a distress premium. Vassalou and Xing (2004) use the distance to default implied by the Merton (1974) model to conclude that the size and book-to-market effects exist only in the quintiles defined by high default risk stocks. They also provide evidence that distress risk is priced in the cross-section and that the Fama and French (FF) factors capture some of the default-related information.

Our study investigates the relationship between SMB, HML, and financial distress risk in Taiwan stock market (exclude banking, security and insurance industries). We focus on testing the hypothesis that the Fama and French (1993) factors are related to a default risk measure. The measurement of the expected default probability is critical in understanding how default risk related to stock return. We compare the performance of Logit model (Ohlson (1980)), Probit model (Zmijewski (1984)), and Discrete-Time Hazard model (Shumway (2001)) to predict default risk. Forecasting precision is imperative because we use the predicted probability of default to examine the relation between distress risk and stock returns. A more informative measure will give us a more complete understanding of what the FF factors represent. In our study, we establish a connection between the probability of default and factor loading. We anticipate that the return and default risk are related. A high probability of default should cause a stock to have high loading on the SMB and HML factors, thus delivering high returns which compensate the investor for holding default risk.

This study proceeds as follows. Relative researches are reviewed in section 2; the research methodologies are presented in section 3; the data employed are presented in section 4; the empirical results are presented and analyzed in section 5; conclusions are presented in section 6.

2. Literature Review

2.1 Researches Related Financial Distress Prediction Model

Beaver (1966) used the univariate analysis and dichotomous classification test to construct the model. Univariate analysis assumes that a single variable can be used for predictive purposes. He matched the sample by industry and asset size, and chose 79 failed and non-failed firms as a sample from 1954 to 1964. He used individual financial ratios to predict financial failure. The empirical results showed that his model achieved the accuracy rate of 87%, 79%, and 77% in one year, two years, and three years prior to bankruptcy.

Altman (1968) is the first one who proposed the multiple discriminate analysis (MDA) method to predict financial failure. He selected a sample of 33 bankrupt manufacturing that had filed for bankruptcy petition under Chapter 7 of the National Bankruptcy Act between 1946 and 1965, and matched these firms with another 33 non-bankrupt firms selected by both industry and asset size random basis. He chose 22 ratios and divided into five categories by using the stepwise multiple discriminate analysis: liquidity, profitability, leverage, solvency, and activity. Altman's Z-score model contained the five ratios. According to this ratio, if Z score was greater than 2.99, the firms were classified as non-bankrupt. If Z score was below 1.81, the firms were classified as bankrupt. The area between 1.81 and 2.99 was defined as the "gray area" because of the uncertainty. Altman's MDA model proved to be extremely accurate in predicting bankruptcy with an accuracy rate of 94% in one year prior to bankruptcy. However, the accuracy of prediction decreases as the projection period got longer.

Deakin (1972) united Beaver's and Altman's model and formulated the tendency of bankruptcy by a quadratic function. Unlike Beaver, Deakin selected a sample random basis and chose 32 non-bankrupt firms and 32 bankrupt firms. His discriminate model achieved an accuracy rate of 80% three years prior to bankruptcy. Although, the accuracy rate dropped when trying to predict bankruptcy four or more years before it occurred.

Blum (1974) introduced the cash flow concept, considered the trend of ratios, and included the variability, and used three types of financial ratios: liquidity, profitability, and variability as explanatory variables to construct his model. The empirical results showed that his discriminate model achieved an accuracy rate of approximately 94% and 80% in one year and two years prior to bankruptcy, and 70% in three, four and five years prior to bankruptcy.

Altman, Haldeman, and Naraynana (1977) introduced the ZETA model that improved the Z-score model in 1986. They used seven variables: return on asset, stability of earning, debt service, cumulative profitability, liquidity/current ratio, capitalization (five years average of total market value), and size as predictors in the ZETA model. The accuracy rate of the ZETA model was 96% for one year and 70% for five years prior to bankruptcy.

Ohlson (1980) is believed to be the first to used develop a model using Multiple Logistic Regression (Logit) to construct a probabilistic bankruptcy model in prediction bankruptcy. He supported that logit models were preferable over MDA in financial distress prediction because logit regression does not need the assumptions of MDA. He selected the firms traded on OTC or/and stock exchange market between 1970 and 1976, and excluded utilities, transportation companies, and financial services companies. Finally, Ohlson selected 105 bankruptcy firms and matched these firms with 2058 non-bankrupt firms randomly. He used nine variables in the model: SIZE (log(total assets/GNP price-level index)), TLTA (total liabilities/total assets), WCTA (working capital/total assets), CACL(current assets/current liabilities), OENEG (1 if total liabilities exceed total assets, 0 otherwise), NITA (net income/total assets), FUTL(funds provided by operations/total liabilities), INTWO (1 if net income was negative for the last two years, 0 otherwise),and CHIN ($(NI_t - NI_{t-1})/(|NI_t| + |NI_{t-1}|)$). He constructed three logit bankruptcy prediction models which predicted bankruptcy within one, two, and one or two years. The empirical results showed the three models' accuracy rates were 91.12%, 95.55%, and 92.84%, and the model that predicted bankruptcy within one year had better prediction ability.

Zmijewski (1984) used Probit model to construct financial distress prediction model. He selected 76 bankrupt firms and 3880 non-bankrupt firms from 1972 to 1978. His paper examined conceptually and empirically two estimation biases which can result when financial distress models are estimated on nonrandom samples. The first bias results from "oversampling" distressed firms and falls within the topic of choice-based sample biases. The second results from using a "complete data" sample selection criterion and falls within the topic of sample selection biases. His empirical results showed that the bias are clearly exist, but in general they don't appear to affect the statistical inference or overall classification rates.

Shumway (2001) introduced a simple hazard model with a multiple logit model estimation program for forecasting bankruptcy. Shumway referred to models that used multiple period bankruptcy data to estimate single period classification as static models. While static models produce biased and inconsistent bankruptcy probability estimates, his hazard model which used all available information to determine each firm's bankruptcy risk at each point in time was consistent in general and unbiased in some cases. Therefore, the simple hazard model is referred to as discrete-time hazard model. Shumway collected data of firms which began trading from 1962 to 1992 and were in the intersection of the Compustat Industial File and the CRSP Daily Stock Return File for NYSE and AMEX stocks. He incorporated Altman's five independent variables and Zemijewski's three independent variables in his model, and also added three market-driven variables: firms' relative size, firms' abnormal returns, and sigma of firms' stocks. The empirical results showed that firms' trading age wasn't the significant variable and EBIT/TA, ME/TL, NI/TL, and TL/TA were in his model. He found that including accounting ratios and market-driven variables would improve the prediction ability of the discrete-time hazard model.

2.2 Researches Related Stock Return

Sharpe (1964) and Lintner (1965) proposed the Capital Asset Pricing Model (CAPM) and find that there was a positive relation between the expected returns and market βs, and market βs could explain the change of the cross-section returns. In other words, knowing market βs can estimate the asset returns.

Black, Jensen and Scholes (1972) selected monthly stock returns listed on NYSE from January 1926 to March 1966 and estimated βs of stocks. They used βs to sort all stock into ten portfolios. The empirical results showed that there was a linearly positive relation between market risk and stock returns.

Fama and MacBeth (1973) used the monthly returns of common stock listed on NYSE as a sample from January 1926 to June 1968. They demonstrated that market risk can complete explain return and there existed a significantly positive relation between market risk and returns.

Although CAPM has long shaped the way academics and practitioners, since 1980s some researchers found that some phenomena that CAPM cannot explain exist. Market βs cannot complete explain asset pricing, and "anomalies" which affect asset pricing exist, which is demonstrated such as size effect, price earning ratio, book-to-market effect etc.

Banz (1981) tested the relation between total market value of common stocks and returns. Banz used a generalized asset pricing model on a sample which were monthly stock returns listed NYSE from 1936 to 1975, and constructed portfolios by using market value of common stock and market risk, and used generalized least square regression analysis. He found in general small firms had bigger risk-adjusted returns than large firms, and that is, size effect existed.

Reinganum (1981) examined the earning/price and market value effects on stock returns. He collected stock listed NYSE and AMEX between 1963 and 1977 and constructed ten portfolios based on market value at the end of very year. The empirical results showed that the average returns of small firms were bigger 20% than large firm, and the effect existed at least two years.

Chan, Chen and Hsieh (1985) investigated the firm size effect for the period 1958 to 1977 by using a multi-factor pricing model. They found that the risk-adjusted difference in returns between the top five percent and the bottom five percent of the NYSE firms is about one to two percent a year, a drop from about twelve percent per year before risk adjustment. The variable most responsible for the adjustment is the sensitivity of asset returns to the changing risk premium, measured by the return difference between low-grade bonds and long-term government bonds.

Chan and Chen (1991) showed differences in structural characteristics that lead firms of different sizes to react differently to the same economic news. They found that a small firm portfolio contains a large proportion of marginal firms. They found that return indices are important in explaining the time-series return difference between small and large firms.

Fama and French (1992) investigated all NYSE, AMEX, and NASDAQ stocks that met the CRSP-COMPUSTAT data requirements, and divided stocks into ten portfolios sorted by size and book-to-market equity. They found that the two variables, size and book-to-market equity, combined to capture the cross-sectional variation in average stock returns associated with market beta, size, leverage, book-to-market equity, and earnings-price ratios, and when the tests allow for variation in beta that is unrelated to size, the relation between market beta and average return is flat, even when beta is the only explanatory variable.

Fama and French (1993,1995) confirmed that portfolio constructed to mimic risk factors related to size and book-to-market equity add substantially to the variation in stock returns explained by a market portfolio. They showed that a three-factor model that included a market factor and risk factors related to size and book-to-market equity seemed to capture the cross-section of average stock returns.

Dichev (1998) finds that bankruptcy risk is not rewarded by higher returns and concludes that a distress factor cannot be at the origin of the size and book-to-market effect. In particular, he finds that portfolios formed on the basis of a distress measure, whether Altman's Z-score or

Ohlson's O-score, have returns inversely related to bankruptcy risk, a high probability of default is associated with low average returns. After examining the cross-sectional relation between stock returns and bankruptcy measures, as well as size and book-to-market, he concludes that the fact that firms with low bankruptcy risk outperform firms with high bankruptcy risk can only be explained by a mispricing argument.

Griffin and Lemmon (2002) measured bankruptcy risk by using the Ohlson's O-score, and found that the low return of high default-risk firms was driven by low book-to-market stocks with extremely low returns. They attributed these very low returns to mispricing due to a high degree of information asymmetry proxied by low analyst coverage.

Vassalou and Xing (2004) used a sample from January 1971 to December 1999 and the distance to default implied by the Merton (1974) model to conclude that the size and book-to-market effects exist only in the quintiles defined by high default risk stocks. They also provided evidence that distress risk is priced in the cross-section and that the Fama and French (FF) factors capture some of the default-related information.

3. Empirical Methodology

In this study, we will compare the performance of three models, Logit model, Probit model, and Discrete-time Hazard model. Then we are going to use the default measure (the predicted probability of default), estimated by the model with the best performance in the three models, to examine the relation between distress and the FF factors, SMB and HML. In this section we start with introducing the methodologies used in this study.

The outcomes of the financial distress are between two discrete alternatives, failed or non-failed, so the binary choice model is an appropriate method to apply to. The dependent variable Y_i takes the value of 1 when the company suffers financial distress and takes the value of 0 when otherwise. We start from the concept of the regression model. Let X_i denote a vector of predictors for the i th observation, and β be a vector of unknown parameter and ε_i be error term with zero mean. The dichotomous dependent variable regression model is given by

$$Y_i = X_i'\beta + \varepsilon_i \tag{1}$$

where $Y_i = \begin{cases} 1\ , & \text{if bankruptcy} \\ 0\ , & \text{otherwise} \end{cases}$

Take expectation of both sides in (1), and by given that $X_i = x_i$, we can get the linear probability model which is expressed by

$$\begin{aligned} E(Y_i \mid X_i = x_i) &= 1 \cdot P(Y_i = 1 \mid X_i = x_i) + 0 \cdot P(Y_i = 0 \mid X_i = x_i) \\ &= P(Y_i = 1 \mid X_i = x_i) = x_i'\beta \end{aligned} \tag{2}$$

Although $Y_i \in \{1,0\}$, β often makes $x_i'\beta$ lie out of the range $(0,1)$. Hence, to make the range of the probability $(0,1)$, we can use the cumulative probability function.

$$P(Y_i = 1 \mid X_i = x_i) = F(x_i'\beta) \tag{3}$$

There are many kinds of cumulative probability functions. This study will introduce the Logit Model and Probit Model.

3.1 Logit Model

Financial distress problem can be viewed as a binary situation. Logit Model assumes that the bankruptcy probability has a Logistic distribution. Let p_i denote the bankruptcy probability (p_i is between 0 and 1). Given x_i, the basic Logit specification of p_i is stated by

$$p_i = P(Y_i = 1 \mid X_i = x_i) = F(x_i'\beta) = \frac{1}{1+e^{-\pi_i}} = \frac{1}{1+e^{-x_i'\beta}} = \frac{e^{x_i'\beta}}{1+e^{x_i'\beta}} \tag{4}$$

or written in the form of the logit function of bankruptcy probability

$$\operatorname{logit}\{P(Y_i = 1 \mid X_i = x_i)\} = \ln\left\{\frac{P(Y_i = 1 \mid X_i = x_i)}{1 - P(Y_i = 1 \mid X_i = x_i)}\right\} = x_i'\beta \tag{5}$$

We know that the higher the value of $\pi_i = x_i'\beta$ is, the higher bankruptcy probability it stands for. In equation (4), the range of the probability should be between (0,1).

In the linear regression models, the OLS (ordinary least squares) is always used to estimate the parameters. However, we cannot use the OLS to estimate the coefficients due to bias. Thus, we use the MLE (maximum likelihood estimator). The likelihood function is:

$$L = \prod_{i=1}^{n}\left[F(x_i'\beta)^{y_i}\left(1 - F(x_i'\beta)\right)^{1-y_i}\right] = \prod_{i=1}^{n}\left(\frac{e^{x_i'\beta}}{1+e^{x_i'\beta}}\right)^{y_i}\left(\frac{1}{1+e^{x_i'\beta}}\right)^{1-y_i} \tag{6}$$

By equation (6), we can get the log-likelihood function as

$$\begin{aligned}\ln L &= \sum_{i=1}^{n} y_i \log F(x_i'\beta) + \sum_{i=1}^{n}(1-y_i)\log\left(1 - F(x_i'\beta)\right) \\ &= \sum_{i=1}^{n}\left\{y_i x_i'\beta - \log\left(1+e^{x_i'\beta}\right)\right\}\end{aligned} \tag{7}$$

where y_i equals to one if the firm goes bankruptcy and equals to zero otherwise.

We take differentiating with respect to β for maximizing this equation (7), and set to zero. Then, we can get the normal equations:

$$\frac{\partial}{\partial\beta}\ln L = \sum_{i=1}^{n}\left(y_i x_i - \frac{e^{x_i'\beta}}{1+e^{x_i'\beta}} x_i\right) = 0 \tag{8}$$

By solving this equation (8) for β, we can get the MLE of β.

3.2 Probit Model

Probit model assumes that the bankruptcy probability has a standard normal distribution. Let p_i be the bankruptcy probability (p_i is between 1 and0). Given x_i, we can state the bankruptcy probability by cumulative probability function of the standard normal distribution:

$$p_i = P(Y_i = 1 \mid X_i = x_i) = F(x_i'\beta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\pi_i} e^{-\frac{u^2}{2}} du = \Phi(\pi_i) = \Phi(x_i'\beta) \tag{9}$$

where $\pi_i = x_i'\beta$, and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

As we mentioned above, the difference between the Probit model and Logit model is the cumulative probability function. Therefore, we also use the MLE to estimate the unknown parameters. The likelihood function is expressed as follow:

$$L = \prod_{i=1}^{n} \left[F(x_i'\beta)^{y_i} (1 - F(x_i'\beta))^{1-y_i} \right] = \prod_{i=1}^{n} \left[\Phi(x_i'\beta) \right]^{y_i} \left[1 - \Phi(x_i'\beta) \right]^{1-y_i} \tag{10}$$

According to equation (10), we take this natural logarithm and get the log-likelihood function:

$$\begin{aligned} \ln L &= \sum_{i=1}^{n} y_i \log F(x_i'\beta) + \sum_{i=1}^{n} (1 - y_i) \log(1 - F(x_i'\beta)) \\ &= \sum_{i=1}^{n} \left\{ y_i \log(\Phi(x_i'\beta)) + (1 - y_i) \log(1 - \Phi(x_i'\beta)) \right\} \end{aligned} \tag{11}$$

where y_i equals to one if the firm goes bankruptcy and equals to zero otherwise.

We take differentiating with respect to β for maximizing this equation (11), and set to zero. Then, we can get the normal equations:

$$\frac{\partial}{\partial \beta} \ln L = \sum_{i=1}^{n} \left\{ y_i \frac{\Phi'(x_i'\beta)}{\Phi(x_i'\beta)} x_i - (1 - y_i) \frac{\Phi'(x_i'\beta)}{1 - \Phi(x_i'\beta)} x_i \right\} = 0 \tag{12}$$

By solving this equation (12) for β, we can get the MLE of β.

3.3 Discrete-Time Hazard Model

Different from other models that we discussed above, the Discrete-Time Hazard Model, correcting for period at risk and allowing for the time-varying covariates, is a dynamic model. Shumway's (2001) Discrete-Time Hazard Model uses all available information to produce bankruptcy probability estimates for all firms at each point in time. By using all available data, this model avoids the selection biases inherent in static models. Now, we are going to introduce this Discrete-Time Hazard Model. Let T be a discrete-time random variable, where $T \in \{1, 2, ..., t_i\}$ which represent the time when the firm goes bankrupt, and t_i be the age of the i th firm. The probability mass function of bankruptcy is $f(t_i, x_i; \theta)$, where x_i is the independent variable vector of the i th firm and θ is the unknown parameter vector.

We denote the survival function as follow:

$$S(t_i, x_i; \theta) = 1 - \sum_{j<t_i} f(j, x_i; \theta) = P(T \geq t_i \mid x_i; \theta)$$

$$= \int_{t_i}^{\infty} f(u, x_i; \theta) du \tag{13}$$

where $S(t_i, x_i; \theta)$ can be interpreted that before the firm year t_i, the firm does not go bankrupt.

And define the hazard function (the instantaneous rate of default per unit of time) as follow:

$$h(t_i, x_i; \theta) = \frac{f(t_i, x_i; \theta)}{S(t_i, x_i; \theta)} = P(T = t_i \mid T \geq t_i, x_i; \theta) \tag{14}$$

It can be interpreted that the firm goes bankrupt at age t_i.

And $F(t_i, x_i; \theta)$ is the cumulative probability density function of $f(t_i, x_i; \theta)$. Then, according to equation (13),

$$F(t_i, x_i; \theta) = \int_0^{t_i} f(u, x_i; \theta) du = 1 - \int_{t_i}^{\infty} f(u, x_i; \theta) du = 1 - S(t_i, x_i; \theta) \tag{15}$$

We can take differentiating and get

$$\frac{d}{dt} F(t_i, x_i; \theta) = \frac{d}{dt} \left[1 - S(t_i, x_i; \theta)\right]$$

$$\Rightarrow\ F'(t_i,x_i;\theta) = -S'(t_i,x_i;\theta) \tag{16}$$

Then, we can use equation (14) to translate the survival function as follow:

$$h(t_i,x_i;\theta) = \frac{f(t_i,x_i;\theta)}{S(t_i,x_i;\theta)} = \frac{F'(t_i,x_i;\theta)}{S(t_i,x_i;\theta)} = \frac{-S'(t_i,x_i;\theta)}{S(t_i,x_i;\theta)} = -d\ln S(t_i,x_i;\theta)$$

$$\Rightarrow\ -\int_0^{t_i} h(u,x_i;\theta)\,du = \int_0^{t_i} d\ln S(u,x_i;\theta)$$

$$\Rightarrow\ -\int_0^{t_i} h(u,x_i;\theta)\,du = \ln S(t_i,x_i;\theta) - \ln S(0)$$

Because the survival probability of a firm at time $t=0$ should be one, $S(0)=1$ $\Rightarrow\ \ln S(0)=0$. Therefore, we can get

$$\Rightarrow\ -\int_0^{t_i} h(u,x_i;\theta)\,du = \ln S(t_i,x_i;\theta)$$

$$\Rightarrow\ S(t_i,x_i;\theta) = e^{-\int_0^{t_i} h(u,x_i;\theta)du} = e^{-H(t_i,x_i;\theta)} \tag{17}$$

where $H(\cdot)$ is the cumulative probability function of $h(t_i,x_i;\theta)$.

According to Cox and Oakes (1984), we can show that the hazard function for a discrete-time hazard model as follow:

$$\because\ \ h(t_i,x_i;\theta)\to 0\ \ \Rightarrow\ h(t_i,x_i;\theta) \approx -\ln\left(1-h(t_i,x_i;\theta)\right)$$

$$\therefore\ \ H(t_i,x_i;\theta) = -\sum_{j=1}^{t_i}\ln\left(1-h(j,x_i;\theta)\right) = -\ln\prod_{j=1}^{t_i}\left(1-h(j,x_i;\theta)\right)$$

Substituting above equation into equation (17), we can show the survival function as follow:

$$S(t_i,x_i;\theta) = e^{-H(t_i,x_i;\theta)} = e^{\ln\prod_{j=1}^{t_i}\left(1-h(j,x_i;\theta)\right)} = \prod_{j=1}^{t_i}\left(1-h(j,x_i;\theta)\right) \tag{18}$$

Shumway (2001) proved that interpreting the logit model as a hazard model can clarify the meaning of the model’s coefficients, and a discrete-time hazard model is equivalent to a multi-period logit model. So, the hazard function and likelihood function of a discrete-time hazard model can be shown as follows:

$$h(t_i, x_i; \theta) = \frac{e^{x_i'\theta}}{1+e^{x_i'\theta}} \tag{19}$$

$$L = \prod_{i=1}^{n} \left\{ \left(h(t_i, x_{t_i}; \theta) \right)^{y_i} \left[1 - h(t_i, x_{t_i}; \theta) \right]^{1-y_i} \prod_{j=1}^{t_i - 1} \left[1 - h(j, x_j; \theta) \right] \right\}$$

$$= \prod_{i=1}^{n} \left\{ \left[\frac{h(t_i, x_i; \theta)}{1 - h(t_i, x_i; \theta)} \right]^{y_i} \prod_{j=1}^{t_i - 1} \left[1 - h(j, x_j; \theta) \right] \right\}$$

$$= \prod_{i=1}^{n} \left\{ \left[\frac{h(t_i, x_i; \theta)}{1 - h(t_i, x_i; \theta)} \right]^{y_i} S(t_i, x_i; \theta) \right\} \tag{20}$$

where y_i equals to one if the firm goes bankruptcy and equals to zero otherwise.

According to equation (20), we take this natural logarithm and get the log-likelihood function:

$$\ln L = \sum_{i=1}^{n} \left\{ y_i \ln h(t_i, x_i; \theta) + (1 - y_i) \ln\left(1 - h(t_i, x_i; \theta)\right) + \sum_{j=1}^{t_i - 1} \ln\left(1 - h(j, x_j; \theta)\right) \right\}$$

$$= \sum_{i=1}^{n} \left[y_i \ln\left(\frac{h(t_i, x_i; \theta)}{1 - h(t_i, x_i; \theta)} \right) \right] + \sum_{i=1}^{n} \sum_{j=1}^{t_i} \left[\ln\left(1 - h(j, x_j; \theta)\right) \right] \tag{21}$$

Substituting equation (19) into equation (21), we can get

$$\ln L = \sum_{i=1}^{n} (y_i x_i' \theta) + \sum_{i=1}^{n} \sum_{j=1}^{t_i} \left[\ln \frac{1}{1+e^{x_j'\theta}} \right] \tag{22}$$

We take differentiating with respect to θ for maximizing this equation (22), and get the first-order condition as follow:

$$\frac{\partial}{\partial \theta} \ln L = \sum_{i=1}^{n} (y_i x_i') + \sum_{i=1}^{n} \sum_{j=1}^{t_i} \left(\frac{-e^{x_j'\theta} x_j'}{1+e^{x_j'\theta}} \right) = 0 \tag{23}$$

By solving this equation (24) for θ, we can get the MLE of θ.

3.4 Error Classification

Logit model, Probit model, and Discrete-Time Hazard model can classify every firm as

Default group or Non-default group by using a cut-off point. According to the error classification, we can easily observe the predicted ability of a model. In order to analyze the performance of models, we first find a value for the bankruptcy probability that make the total error minimum. The value is called optimal cut-off point value. If the predicted bankruptcy probability of a firm is higher than the cut-off point value, we will classify the firm as the default group; if the predicted bankruptcy probability of a firm is lower than the cut-off point value, we will classify the firm as the non-default group. Then, we define two type errors: type I error and type II error. Type I error is that firms with a predicted default probability higher than the cut-off point value are classified as the non-default group; type II error is that firms with a predicted default probability lower than the cut-off point value are classified as the default group. In addition, α is the probability of type I error, β is the probability of type II error.

To search the optimal cut-off point value, we must find a cut-off point value that makes sum of α and β minimum. Using the error classification table, we can observe the sum of α and β for all cut-off point values, and then find the optimal cut-off point value.

3.5 ROC (Receiver Operating Characteristic) Curve

ROC analysis is widely used in medicine, radiology, psychology and other areas for many decades. In the social sciences, ROC analysis is often called the ROC Accuracy Ratio, a common technique for judging the accuracy of default probability models.

Assume someone has to use the rating score for bankrupt and non-bankrupt firms to predict which firm will go bankrupt in the next period. One possible way for the decision-maker would be propose a cut-off point value C. If the rating score is higher than the cut-off point value, the firm might go bankrupt; while if the rating score is lower than the cut-off point value, the firm is considered non-bankruptcy. The four outcomes can be formulated in a 2×2 contingency table in Table 1.

We define the hit ratio $HR(C)$ as

$$HR(C) = P(S_D \leq C) \tag{25}$$

where S_D means the random variable of the distribution of bankruptcy firms.

And define the false alarm rate $FAR(C)$ as

$$FAR(C) = P(S_{ND} \leq C) \tag{26}$$

where S_{ND} means the random variable of the distribution of non-bankruptcy firms.

Table 1 Decision Results Given the Cut-off Point Value C

		Bankruptcy	Non-bankruptcy
Rating Score	$< C$	Correct Prediction (hit)	Wrong Prediction (false alarm)
	$> C$	Wrong Prediction (miss)	Correct Prediction (correct rejection)

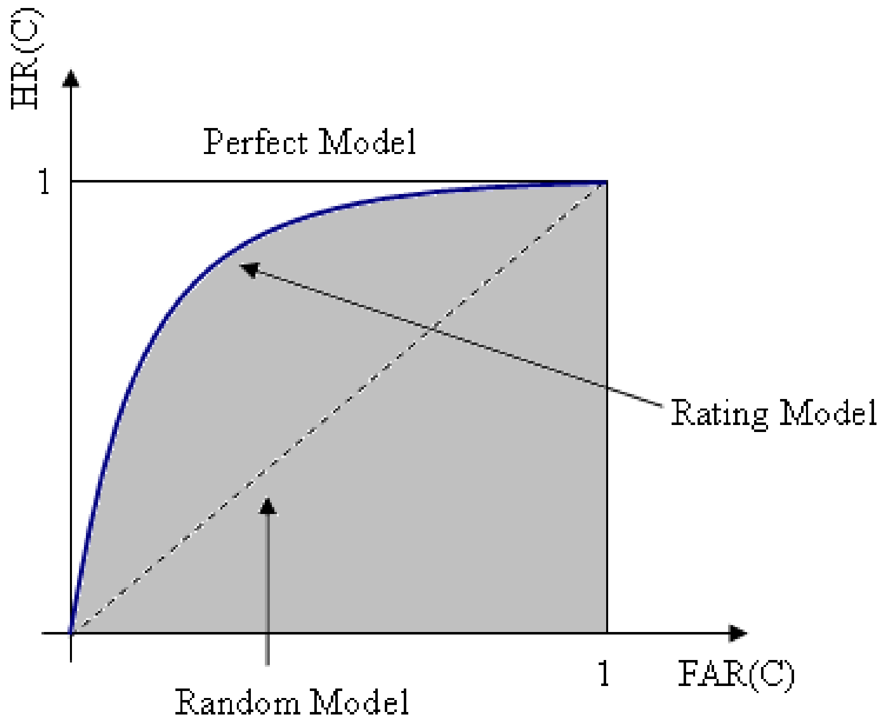

Figure 1 ROC Curve

The construction of ROC is as follows. For all cut-off point values C that are contained in the range of the rating scores the quantities $HR(C)$ and $FAR(C)$ are computed. The ROC curve is a plot of $HR(C)$ versus $FAR(C)$ for all cut-off point values C. The ROC curve is illustrated in Figure 1.

A performance of a rating model is the better the steeper the ROC curve is at the left end and the closer the ROC curve's position is to the point (0,1). Similarly, the model is the better the larger the area under the ROC is. Hence, we denote this by AUC (area under curve) which is the area under the ROC curve. It can be interpreted as the average power of the tests on bankruptcy/non-bankruptcy corresponding to all possible cut-off point values C. The area AUC is 0.5 for a random model without discriminative power and is 1.0 for a perfect model. It's between 0.5 and 1.0 for any reasonable rating model in practice.

3.6 Four-Factor Model

Fama and French (1993) three-factor model is widely used to explain the cross-section stock return. Since it has been reported that there exists a momentum effect in stock prices (Jegadeesh and Titman (1993)) we would likely observe slightly higher returns corresponding to the group of firms that performed well in the past and slightly lower returns in the group of firms that performed poorly in the past. Therefore, a momentum factor is added to the Fama and French (1993) three-factor model because the default measure loads on the past economic performance as well as past stock returns. Accordingly, we consider a four factor model: the Fama and French (1993) three factor model plus a momentum factor. The Fama and French (1993) three-factor model is expressed as follow:

$$R_{it} - R_{ft} = \alpha_i + \beta_{MKT,i}\left(R_{mt} - R_{ft}\right) + \beta_{SMB,i} SMB_t + \beta_{HML,i} HML_t + \varepsilon_{it}$$

Then, the four-factor model can be shown:

$$R_{it}-R_{ft}=\alpha_i+\beta_{MKT,i}\left(R_{mt}-R_{ft}\right)+\beta_{SMB,i}SMB_t+\beta_{HML,i}HML_t+\beta_{MOM,i}MOM_t+\varepsilon_{it}$$

where, R_{it} is the value-weighted monthly return of a portfolio at time t.

R_{ft} is the risk-free interest rate at time t.

R_{mt} is the return on the valued-weighted portfolio of the stocks in the six benchmark portfolios at time t.

SMB_t is the difference between the returns on small- and big-stock benchmark portfolios with about the same weighted-average book-to-market equity at time t.

HML_t is the difference between the simple average of the return on the two high-BE/ME (book-to-market equity) benchmark portfolios and the average of the returns on the two low- BE/ME benchmark portfolios at time t.

MOM_t is the average of the returns on two high prior return benchmark portfolios minus the average of the returns on two low prior return benchmark portfolios at time t.

According to Fama and French (1992), we use the same way to construct six benchmark portfolios formed from sorts of stocks on ME (stock price times number of shares) and BE/ME. In June of each year t, we rank the stock on size (price times shares), and split stocks into two groups, small and big (S and B). And we also break stocks into three book-to-market equity groups based on the breakpoints for the bottom 30% (Low), middle 40% (Medium), and top 30% (High) of the ranked values of BE/ME for stocks. Then, we construct six benchmark portfolios (S/L, S/M, S/H, B/L, B/M, B/H) from the intersections of the two ME and the three BE/ME groups. The six benchmark portfolios is illustrated in Table 2. Monthly value-weighted returns on the six benchmark portfolios are calculated from July of year t to June of year $t+1$.

The portfolio SMB (small minus big) is the difference, each month, between the simple average of the returns on the three small-stock portfolios (S/L, S/M, and S/H) and the simple average of the returns on the three big-stock portfolios (B/L, B/M, and B/H). Thus, SMB is the difference between the returns on small- and big-stock portfolios with about the same weighted-average book-to-market equity.

Table 2 Six Benchmark portfolios

		BE/ME		
		High	Medium	Low
ME	Small	S/H	S/M	S/L
	Big	B/H	B/M	B/L

The portfolio HML (high minus low) is defined the difference, each month, between the simple average of the returns on the two high-BE/ME portfolios (S/H and B/H) and the average of the returns on the two low-BE/ME portfolios (S/L and B/L).

The R_m is the return on the value-weighted portfolio of the stocks in the six size-BE/ME benchmark portfolios.

The MOM factor is defined by the way from Kenneth French's web page. The MOM is the average of the returns on two (big and small) high prior return portfolios minus the average of the returns on two low prior return portfolios. The six value-weight portfolios are formed using independent sorts on size and prior return of stocks. The portfolios are constructed monthly. Big means a firm is above the median market cap at the end of the previous month; small firms are below the median stock market cap. Prior return is measured from month -12 to -2. Firms in the low prior return portfolio are below the 30th percentile. Those in the high portfolio are above the 70th percentile (Details about construction of a momentum factor can be obtained from Kenneth French's web page).

3.7 Fama and MacBeth Procedure

We use the two-stage Fama and MacBeth (1973) procedure for investigating the relationship between default probability and factor loading at the individual stock level.

In the first stage, for each stock in July of each year, we run a time-series regression using 60-month window of monthly data to obtain estimates of the stock loadings on the factors. In the second stage, in July of each year, we run a regression of a logistic transformation of the default probability,

$$\overline{P_{it}} = \log\left(\frac{P_{it}}{1-P_{it}}\right)$$

, on the individual stock factor loading, $\widehat{\beta}_{SMB,it}$ and $\widehat{\beta}_{HML,it}$, and a set of control variables (size and book-to-market). Specifically:

$$\overline{P_{it}} = \lambda_{0,t} + \lambda_{SMB,t}\widehat{\beta}_{SMB,it} + \lambda_{HML,t}\widehat{\beta}_{HML,it} + \lambda_{x,t}x_{it} + \varepsilon_{it}$$

Following Fama and MacBeth (1973), t-statistic is applied to test the null hypothesis that there is no relation between the factor loadings and the default probability against the alternative that there is a positive relation for λ_{SMB} and λ_{HML}:

H_0 : $\lambda = 0$ against H_1 : $\lambda > 0$

4. Data

4.1 Definition of Financial Distress

A company encounters financial difficulties and defaults when it falls to service its debt obligation. Many researchers have studied corporate bankruptcy, different people have come up with different definitions that basically reflect their special interest in the field.

In Taiwan, definitions of bankruptcy can be found in Operating Rule of the Taiwan Stock Exchange Corporation and Company Law, which are in Operating Rule of the Taiwan Stock Exchange Corporation Article 49, 50, 50-1. And we also find the definitions of financial distress and quasi financial distress in TEJ database, and these definitions conform to Operating Rule of the Taiwan Stock Exchange Corporation Article 49, 50, 50-1. Therefore, we will use the definitions of financial distress and quasi financial distress in TEJ database as default event in our study.

4.2 Sample Data

Our sample firms must be listed on Taiwan Stock Exchange Corporation (TSE) or GreTai Security Market (GTSM or OTC). Because the characteristics of banking, security and insurance industries are different from others, we exclude these industries from our sample firms. Besides, we also exclude the firms of which financial reports are incomplete.

We collect data of the sample firms from TEJ database. The study period is 1988-2006. The non-default firms are firms that remain traded on TSE or GTSM during 1988-2006. If the firms experience the financial distress situations mentioned in section 4.1 during this period, we classify firms as default group. We use the observations between 1988 and 2002 as the estimation sample, and the observations from 2003 to 2006 as the out-sample validation group to examine the models' accuracy. Finally, there are 728 non-default firms and 109 default firms in the in-sample, and 956 non-default firms and 63 default firms in the out-sample. The number of in-sample and out-of-sample firms is shown in Table 3.

Table 3　Numbers of Sample Firms

	Sample period	No. of non-default firms	No. of default firms
In-sample	1988～2002	728	109
Out-sample	2003～2006	956	63

In order to estimate the coefficients of each financial prediction model, we collect firm's financial information and calculate the market-driven variables from TEJ database each year in this sample period. The variables used by Logit Model (Ohlson, 1980), Probit Model (Zmijewski, 1984), and Discrete-Time Hazard Model (Shumway, 2001) are illustrated as follows:

1. Variables of Logit Model (Ohlson, 1980)：

SIZE：log(total assets).

TLTA：total liabilities / total assets.

WCTA：working capital / total assets.

CACL：current assets / current liabilities.

OENEG：1 if total liabilities exceed total assets, 0 otherwise.

NITA： net income/total assets.

FUTL：funds provided by operations / total liabilities.

INTWO：1 if net income was negative for the last two years, 0 otherwise.

CHIN：$(NI_t - NI_{t-1})/(|NI_t| + |NI_{t-1}|)$

2. Variables of Probit Model (Zmijewski, 1984)：

NITA：net income / total assets.

TLTA：total liabilities / total assets.

CACL：current assets / current liabilities.

3. Variables of Discrete-Time Hazard Model (Shumway, 2001)

NITA, TLTA and the market-driven variables which are as follows:

Rel-Size：the logarithm of the ratio of the market capitalization of a firm to the total market capitalization.

Ex-Ret：a firm's excess return in year t is that the return of the firm in year $t-1$ minus the value –weighted TAIEX index return in year $t-1$.

SIGMA：regress each stock's monthly returns in year $t-1$ on the value -weighted TAIEX index return in year $t-1$. Sigma is the standard deviation of the residual of this regression.

Table 4 on next page reports the descriptive statistics of variables.

Table 4 Descriptive Statistics

Variable	Mean	Std. Dev.	Maximum	Minimum
SIZE	9.547195	0.495793	11.7586	8.205242
TLTA	0.40086	0.158352	1.513943	0.015454
WCTA	0.201678	0.184536	0.866156	-1.07126
CACL	2.161351	2.081144	92.96316	0.152729
OENEG	0.000516	0.02271	1.000000	0.000000
NITA	0.042586	0.102783	0.525305	-1.68247
FUTL	0.792146	1.401729	47.42159	-0.84235
INTWO	0.209854	0.407231	1.000000	0.00000
CHIN	-0.00633	0.514159	1.000000	-1.00000
Rel-Size	-3.15507	0.601494	-0.68152	-5.7928
Ex-Ret	0.103095	0.779272	28.60954	-1.29635
SIGMA	0.124292	0.075963	1.124059	0.003719

For the four-factor model, we collect monthly stock returns from TEJ database in the sample period. All variables are described as follows. The risk-free rate is the 30-day deposit interest rate of First Bank. The market equity, ME or size, is price times shares outstanding. The book equity, BE, is the book value of stockholder' equity minus the book value of preferred stock. Then, the book-to-market ratio, BE/ME, used to form portfolios in June of year t is book equity for December of year $t-1$, divided by market equity at December of year $t-1$. As mentioned in section 3.6, we can construct six benchmark portfolios formed form sorts of stocks on ME and BE/ME. Monthly value-weighted returns on the six portfolios are calculated from July of year t to June of $t-1$, and the six benchmark portfolios are reformed in June of each year. Then using the way mentioned in section 3.6, we can get returns of SMB, HML, and MOM every month.

5. Empirical Results

5.1 Predicting Default

In this study, we use MLE to estimate the coefficients in Logit model, Probit model, and Discrete-Time Hazard model. These coefficient estimates of each model are shown in Table 5.

Table 5 Coefficient Estimates of Model

	Logit Model	Probit Model	Discrete-Time Hazard Model
Intercept	-16.9998**	-2.8396**	-8.5026**
	(-6.3880)	(-8.2741)	(-11.4487)
SIZE	1.2284**		
	(4.8348)		
TLTA	7.0393**	3.81057**	6.6777**
	(5.9854)	(7.0585)	(8.5981)
WCTA	-0.7054		
	(-0.3939)		
CACL	-0.0698	-0.11637	
	(-0.2559)	(-1.4642)	
OENEG	19.9539		
	(0.0000)		
NITA	-9.8193**	-3.90448**	-6.7811**
	(-4.7382)	(-5.6686)	(-6.3059)
FUTL	-0.0506		
	(-0.0570)		
INTWO	-0.6132		
	(-1.6680)		
CHIN	0.1898		
	(0.6796)		
Rel-Size			-0.3455
			(-1.8830)
Ex-Ret			-0.4201
			(-1.8719)
SIGMA			2.5800*
			(2.0595)

1. z-statistics are presented in the parenthesis.

2. ** and * indicate statistically significance at 1% and 5% statistical level, respectively.

First, we see the coefficients in Logit model in Table 5. Three variables, NITA, SIZE, and TLTA, are statistically significant. Our expectation is that the coefficients of six variables (NITA, SIZE, WCTA, FUND, CACL, CHIN) are negative and the coefficients of two variables (TLTA, INTWO) are positive. Most of the coefficients are consistent of our expectation except for the signs of SIZE, CHIN, and INTWO. However, CHIN and INTWO are not statistically significant. Logit model indicates that NITA and TLTA are related to corporate defaults. The higher the income is, the lower the probability of default.

In Probit model shown in Table 5, we have two statistically significant variables (TLTA, NITA). We expect that the coefficient of TLTA is positive and the coefficients of two variables (CACL, NITA) are negative. And all the coefficients are consistent with our expectation. Like Logit model, TLTA and NITA are statistically significant in Probit model. It may indicate that the variables, TLTA and NITA, are important for predicting financial distress.

Eventually, in Discrete-Time Hazard model we expect that the coefficients of two variables (TLTA, SIGMA) are positive and the coefficients of three variables (NITA, Ex-Ret, Rel-Size) are negative. All variable are consistent with our expectation. And there are three statistically significant variable (TLTA, NITA, SIGMA). As similar to Logit model and Probit model, TLTA and NITA are also statistically significant in Discrete-Time Hazard model. This shows that TLTA and NITA are related to corporate default and important for predicting financial distress.

With coefficient estimates of each model, we can calculate the optimal cut-off point value of each model by using data in in-sample period and compare the performance of each model. First, given each cut-off point value between 0 and 1, we can calculate the corresponding type I and type II error. Then, following Begley, Ming, and Watts (1996), where is the minimum sum of type I and II error r is corresponding to the optimal cut-off point value. Table 6 shows the optimal cut-off point value of each model and type I and type II error. In table 6, we can see that Discrete-Time Hazard Model has a minimum cut-off point value and minimum type I error, but maximum type II error with in-sample data.

Table 6 Optimal Cut-off Point Value

	Logit Model	Probit Model	Discrete-Time Hazard Model
Optimal cut-off point	0.2044	0.2058	0.0313
Type I error rate	0.2569	0.2752	0.2018
Type II error rate	0.1126	0.1304	0.1703

By given optimal cut-off point values, we can compare the performance of each model for predicting corporate default. Type I and II error in out-sample are calculated by using optimal cut-off point values. The results are shown in Table 7.

Table 7 Out-sample Type I and II Error

	Logit Model	Probit Model	Discrete-Time Hazard Model
Type I error rate	0.3650	0.1904	0.1269
Type II error rate	0.1056	0.1202	0.1422
Sum of type I and II errors	0.4707	0.3107	0.2692

We can see that Discrete-Time Hazard Model has a minimum type I error and maximum type II error, while Logit Model has a maximum type I error and minimum type II error. However, the minimum sum of type I and II errors is in Discrete-Time Hazard Model. Moreover, for investors and obliges, the cost that a firm with financial distress is misclassified as a non-default firm is bigger than that a non-default firm is misclassified as a default firm. This means that the cost of type I error is more serious than that in type II error. Hence, we consider that Discrete-Time Hazard Model has a more power for predicting corporate default.

Besides, we also compare the performance of these three models by ROC curve. Figure 2 shows the ROC curve with out-sample data and we can see the AUC of each

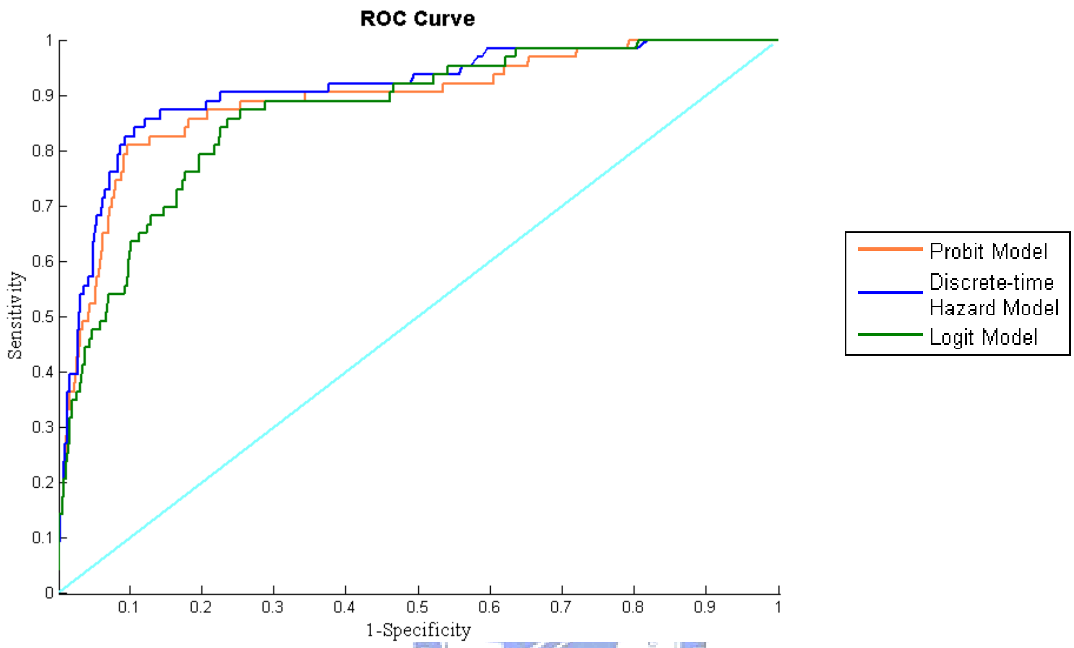

Figure 2 ROC Curve For Each Model

Table 8 Area Under Curve

	Logit Model	Probit Model	Discrete-Time Hazard Model
AUC	0.8660	0.8889	0.9081

model in Table 8. In Figure 2, we can find that the position of ROC curve of Discrete-Time Hazard Model is closer to the point (0,1) than the other two models. And in Table 8, the Discrete-Time Hazard Model with the highest area under the ROC curve (AUC) exhibits the best diagnostic performance. Again, we consider the Discrete-Time Hazard Model the best in predicting the financial distress among three models. Since forecasting precision is imperative, we use Discrete-Time Hazard Model to calculate probability of default for each firm with out-sample data, and examine the relation between distress risk and the FF factor, SMB and HML.

5.2 Stock's Default Probability and FF Factors

In June of every year, we form seven portfolios by sorting according to the predicted default probability obtained from Discrete-Time Hazard model. The portfolio stock composition is kept for one year (from July to the following June) with monthly rebalancing the portfolio weights. These seven portfolios are obtained by sorting the out-sample estimates of the probability of default. These predicted default probabilities are determined from accounting and market information since the end of the previous year.

We are going to test the hypothesis that the FF factors are related to a default risk measure. In particular, we try to establish a connection between returns and the probability of default, and between the probability of default and factor loadings. We anticipate that returns and default risk are related. A high probability of default should cause a stock to have high loading on the SMB and HML factors, thus delivering high returns which compensate the investor for holding default risk.

Table 9a reports the average characteristics of seven portfolios using the probabilities obtained from Discrete-Time Hazard Model. We can find that, generally, the portfolios with higher default probability have relatively higher returns. The average size (stock price times number of shares) is almost monotonically decreasing

Table 9a Characteristics of Default Portfolios

The table shows reports average characteristics of seven portfolios when using the probability obtained from Discrete-Time Hazard Model. The average value-weighted returns are reported in percentage terms. The average size of each portfolio is reported in billions of NT dollars.

	Low	2	3	4	5	6	High
$\overline{ret}$	1.9328	2.4728	3.0523	2.0030	2.1799	3.1359	2.7590
$\overline{size}$	28390.22	10158.65	13915.43	10299.67	4944.04	3932.51	2410.34
$\overline{BM/ME}$	0.5362	0.7692	0.8112	0.9352	1.0223	1.1005	1.1611

Table 9b Four-Factor Model on Default Portfolios

For each portfolios, we run a time-series regression of the value-weighted returns on a four-factor model. The table reports loading estimates along with robust Newey and West t-statistics. ** and * indicate statistically significance at the 1% and 5% statistical level, respectively.

	Low	2	3	4	5	6	High
α	-0.0009	0.0002	0.0058	-0.0039	-0.0088	0.0008	-0.0059
	(-0.7004)	(0.0917)	(1.7588)	(-1.5798)	(-2.0024)	(0.1863)	(-1.7620)
β_{MKT}	0.9281**	1.0678**	1.3352**	1.0473**	1.1456**	1.0487**	1.0005**
	(23.4592)	(12.5938)	(19.2555)	(9.9220)	(16.8274)	(11.0575)	(12.7675)
β_{SMB}	-0.1279**	0.1923*	0.0232	0.0376	0.3737**	0.2779*	0.4352**
	(-3.4090)	(2.1908)	(0.2359)	(0.4672)	(4.2024)	(2.0693)	(4.5351)
β_{HML}	-0.2068**	-0.2241**	-0.1583*	-0.0309	0.1719*	0.3773**	0.5458**
	(-6.8942)	(-3.1370)	(-2.1155)	(-0.3896)	(2.4637)	(3.5815)	(7.5426)
β_{MOM}	0.1646*	-0.0387	-0.1341	-0.4877**	0.1895	-0.1077	0.0682
	(2.5689)	(-0.1829)	(-0.8036)	(-3.1050)	(0.8785)	(-0.3883)	(0.2773)
Adjusted R^2	0.9273	0.8520	0.8689	0.7908	0.8241	0.7563	0.8220

along the default dimension. And the average of book-to-market is increasing in the default measure. Firms with high default probabilities have lower size (2410.34) and higher book-to-market (1.1611) on average, while firms with low default probabilities have relatively high average size (28390.22) and low book-to-market (0.5362).

For each portfolios, we run a time-series regression of the value-weighted returns on a four-factor model (FF factors plus a momentum factor).

$$R_{it} - R_{ft} = \alpha_i + \beta_{MKT,i}\left(R_{mt} - R_{ft}\right) + \beta_{SMB,i}SMB_t + \beta_{HML,i}HML_t + \beta_{MOM,i}MOM_t + \varepsilon_{it}$$

Table 9b represents the estimates of the factor loadings along with robust Newey and West (1987) t-statistics. We can find that all β_{MKT} are statistically significant, but they don't have a trend. If SMB and HML are related to financial distress, we would expect the portfolio loadings to increase along the dimension of default risk. The β_{SMB} s are not strongly increasing along the default measure, but generally they seem increasing. The result is much stronger for HML than SMB. The β_{HML} almost monotonically increase along the default measure (except second portfolio). In general, this result is consistent with our anticipation that a high probability of default should cause a stock to have high loading on the SMB and HML factors, thus delivering high returns which compensate the investor for holding default risk.

We also investigate how individual stocks (as opposed to portfolios) relate to default risk. If the factors are pricing systematic distress risk, it should be that individual firm loadings on the factors are related to their estimated default probability. Every month we estimate individual stock loading for four-factor model using a 60-month window and requiring that any stock has at least 36 monthly observations in every window. In June of each year, we sort the stocks into seven portfolios based on the default probability. Hence, for each portfolio, we compute the average loading for the stock in that portfolio in July of each year of the out-sample. The result is reported in Table 10. As we can see from the table, the average loadings of SMB and HML are increasing along the default measure. The average firm in the low default probability has a negative loading on HML and that in the high default probability has a positive loading. This result is similar to result reported in Table 9b and conforms our anticipation.

Table 10 Average Individual Stocks Loadings on Factors

	Low	2	3	4	5	6	High
$\overline{\beta_{SMB}}$	0.4543	0.5903	0.6661	0.6354	0.7097	0.7604	1.0232
$\overline{\beta_{HML}}$	-0.0462	0.0487	0.0426	0.0832	0.1561	0.1853	0.2125

A more rigorous test of the relation between default probability and factor loading at the individual stock level involves a two-stage Fama and MacBeth (1973) procedure mentioned in section 3.7.

$$\overline{P_{it}} = \lambda_{0,t} + \lambda_{SMB,t}\widehat{\beta}_{SMB,it} + \lambda_{HML,t}\widehat{\beta}_{HML,it} + \lambda_{x,t}x_{it} + \varepsilon_{it} \qquad \text{where } \overline{P_{it}} = \log\left(\frac{P_{it}}{1-P_{it}}\right)$$

Table 11 Fama-MacBeth Procedure

	Model(1)	Model(2)	Model(3)
λ_0	-1.8288 (-2.1116)	-5.1378 (-48.1914)	-2.6767 (-3.6138)
λ_{SMB}		0.6658** (4.3007)	0.3477** (3.9915)
λ_{HML}		0.8370** (3.8740)	0.3752* (2.1829)
λ_{Size}	-0.3894 (-4.1890)		-0.3167 (-3.6742)
$\lambda_{BM/ME}$	0.3127 (1.4785)		0.3073 (1.6860)

Table 11 reports the results. Model (1) gives the estimates of the model which uses only size and book-to-market as opposed to the factor loadings. The base case regression Model (2), relates the SMB and HML loadings to the default probability and highlights a positive and statistically significant relationship. We re-estimate Model (2) by including size and book-to-market as control variables. If the FF factor loadings are related to the default probability, we expect that the parameter estimates retain statistically significance. In Model (3) of Table 11, we find that both of λ_{SMB} and λ_{HML} maintain their signs and significance, although the significance of λ_{HML} reduces. Therefore, we consider that SMB and HML are related to default risk.

6. Conclusion

Fama and French (1996) suggest that small value stocks tend to be firms in distress (with high financial leverage and earnings uncertainty), with higher returns due to a distress premium. Vassalou and Xing (2004) conclude that the size and book-to-market effects exist only in the quintiles defined by high default risk stocks. They also provide evidence that distress risk is priced in the cross-section and that the Fama and French (FF) factors capture some of the default-related information.

In this study, we investigate the relationship between FF factors and financial distress risk in Taiwan stock market. Our sample data are collected form TEJ database form 1988 to 2006. We compare the performance of Logit model (Ohlson (1980)), Probit model (Zmijewski (1984)), and Discrete-Time Hazard model (Shumway (2001)) to predict default risk. By using error classification and ROC curve to measure the performance of each model, we find that Discrete-Time Hazard model, when compared to other two models, is more accurate in predicting financial distress. Then, using the default probabilities predicted by Discrete-Time Hazard model, we analyze the relationship between default risk and the Fama and French factors. Our results are consistent with our expectation that a high probability of default should cause a stock to have high loading on the SMB and HML factors, thus delivering high returns which compensate the investor for holding default risk. We provide evidence that both portfolio and individual stock factor loading of SMB and HML are related to the default probability, and consider that SMB and HML are related to default risk. Further, the future research can look at the role of the aggregate measure of financial distress in directly pricing the cross-section of average stock returns.

7. Reference

7.1 English Reference

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

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

Banz, R. W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.

Basu, S., 1983, The relationship between earnings yield, market value, and return for NYSE common stocks: Further evidence, *Journal of Financial Economics* 12, 129-156.

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

Bhandari, L. C., 1988, Debt/equity ratio and expected common stock returns: empirical evidence, *Journal of Finance* XLIII, 507-528.

Black, F., Jensen, M. and Scholes, M., 1972, The capital asset pricing model: some empirical tests, in Michael Jensen, editor, Studied in The Theory of Capital Market, *Praeger Publishers Inc.*, New York

Blum, M., 1974, Failing company discriminant analysis, *Journal of Accounting Research* 12, 1-25.

Chan, K. C. and Chen, N., 1991, Structural and return characteristics of small and large firms, *Journal of Finance* 46, 1467-1484.

Chan, K. C., Chen, N. and Hsieh, H. A., 1985, An exploratory investigation of the firm size effect, *Journal of Financial Economics* 14, 451-471.

Chen, N., Roll, R. and Ross, S. A., 1986, Economic forces and the stock market, *Journal of Business* 59, 383-403.

Cox, D. R. and D. Oakes, 1984, Analysis of survival data, *New York: Chanman & Hall*.

Denis, D. J. and Denis, D., 1995, Causes of financial distress following leveraged recapitalization, *Journal of Financial Economics* 27, 411-418.

Deakin, E. B., 1972, A discriminant analysis of predictors of business failure, *Journal of Accounting Research* 10, 167-179.

Dichev, I., 1998, Is the risk of bankrupt a systematic risk?, *Journal of Finance* 53, 1131-1147.

Fama, E. F. and MacBeth, J. D., 1973, Risk return and equilibrium : empirical test, *Journal of Political Economy* 81,607-636.

Fama, E. F. and French K. R., 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Fama, E. F. and French K. R., 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fama, E. F. and French, R. K., 1995, Size and book-to-market factors in earnings and returns, *Journal of Finance* 50, 131-155.

Fama, E. F. and French, R. K., 1996, The CAPM is wanted, dead or alive, *Journal of Finance* 5, 1947-1958.

Griffin, J. M. and Lemmon, M. L., 2002, Book-to-market equity, distress risk, and stock returns, *Journal of Finance* 57, 2317-2336.

Jegadeesh, N. and Titman, S., 1993, Returns to buying winners and selling losers : implications for stock market efficiency, *Journal of Finance* 47, 65-91.

Lang , R. and Stulz, R., 1992, Contagion and competitive intraindustry effects of bankruptcy announcements : an empirical analysis, *Journal of Financial Economics* 32, 45-60.

Lintner, J., 1965, The valuation of risk assets and the selection of risky investment in stock portfolio and capital budgets, *Review of Economics and Statistics* 47, 13-37.

Merton, R. C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867-887.

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

Reinganum, Marc R., 1981, A new empirical perspective on the CAPM, *Journal of Financial and Quantitative Analysis*, 16 439-462.

Rosenberg, B., Reid, K., and Lanstein, R., 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9-11.

Ross, S. A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341-360.

Sharp, W. F., 1964, Capital asset prices: a theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425-442.

Shumway, T., 2001, Forecasting bankruptcy more accurately: a simple hazard rate model, *Journal of Business* 74, 101-124.

Saretto, A. A., 2004, Predicting and pricing the probability of default, Working paper.

Vassalou, M and Xing, Y., 2004, Default risk in equity returns, *Journal of Finance* 59, 831-868.

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

7.2 Chinese Reference

黃瑞卿，魏曉琴，李昭勝，李正福，「使用離散型存活模式預測公司財務危機機率」，交通大學財務金融所研討論文，民國九十三年。

黃瑞卿，蕭兆祥，李昭勝，「產業效應與市場導出變數在離散型財務危機模式之研究」，

管理與系統，第十四卷，第一期，71-94 頁，民國九十六年。

陳孟雅，「公司財務結構與違約機率之分析」，交通大學財務金融研究所，碩士論文，民國九十四年。

陳坉熙，「股票報酬率與違約風險關係之研究—以台灣股票市場為例」，交通大學管理科學研究所，碩士論文，民國九十五年。

詹益宗，「財務危機預警模型之比較」，交通大學財務金融研究所，碩士論文，民國九十五年。

蕭兆祥，「使用離散型比率危險模式建立財務危機預警模式」，東華大學應用數學系研究所，碩士論文，民國九十四年。