Market Implications of Default Prediction

Hong Miao[*], Sanjay Ramchander[†], Patricia Ryan[‡], Tianyang Wang[§]

May 18, 2017

Abstract

This paper proposes a variant application of the Merton distance-to-default model by employing implied volatility and implied cost of capital to forecast defaults. The proposed model's results are compared with predictions obtained from three popular models in different setups. We find that our "best" model, which contains both forward-looking proxies does outperform other models with a default prediction accuracy rate of 89%. Additional analysis using a discrete-time hazard model indicates that the psuedo-R^2 values from regression models that include the two implied measures are as high as 51%. Overall, our results establish the informational relevance of implied cost of capital and implied volatility in predicting defaults.

Keywords: Default Prediction, Distance to Default, Implied Cost of Capital, Implied Volatility

JEL Code: G130

[*]Associate Professor of Finance at Colorado State University, Fort Collins, CO, USA
[†]Professor of Finance at Colorado State University, Fort Collins, CO, USA
[‡]Associate Professor of Finance at Colorado State University, Fort Collins, CO, USA
[§]Assistant Professor of Finance at Colorado State University, Fort Collins, CO, USA

Market Implications of Default Prediction

May 18, 2017

Abstract

This paper proposes a variant application of the Merton distance-to-default model by employing implied volatility and implied cost of capital to forecast defaults. The proposed model's results are compared with predictions obtained from three popular models in different setups. We find that our "best" model, which contains both forward-looking proxies does outperform other models with a default prediction accuracy rate of 89%. Additional analysis using a discrete-time hazard model indicates that the psuedo-R^2 values from regression models that include the two implied measures are as high as 51%. Overall, our results establish the informational relevance of implied cost of capital and implied volatility in predicting defaults.

Keywords: Default Prediction, Distance to Default, Implied Cost of Capital, Implied Volatility

JEL Code: G130

Market Implications of Default Prediction

1 Introduction

Effective estimation and prediction of corporate defaults are critical for asset pricing, credit risk assessment of loan portfolios, and the valuation of other financial products that are exposed to corporate default. This work contributes to a growing literature on predicting corporate bankruptcy. Much of the earlier literature focused on estimating default probabilities using financial accounting data (see, for example, Altman 1968; Beaver 1968; Ohlson 1980). Accounting-based models are often criticized because they are inherently backward-looking thus limiting its value in predicting the probability of bankruptcy. Furthermore, financial statements are thought to be constrained in their predictive ability as they are constructed under the going-concern principle (i.e., under the assumption that the firm will not go bankrupt), are subject to managerial discretion and manipulation, and fail to incorporate important pricing information such as underlying asset volatility (see Hillegeist et al. 2004).

In contrast, more recently, motivated by Merton's (1974) structural model of default risk, there has been renewed interest in the application of market-based bankruptcy prediction models that are based on the contingent claims valuation approach. Market-based measures, at least in theory, are believed to overcome some of the limitations associated with accounting-based models since they are forward-looking by design and incorporate all information relevant to the pricing of securities. In this context, option pricing models provide the necessary structural framework to estimate default probabilities implied by the Merton distance-to-default (DD) model, which provides a measure of the distance between the firm's current value and its bankruptcy threshold. Although extracting DD from the Black-Scholes option pricing model is computationally intensive and requires a number of simplifying assumptions that may not necessarily hold in practice; these challenges not withstanding there is growing empirical evidence that the structural form of the Merton DD model provides significantly superior information than purely accounting-based measures of default probabilities (see for example, Hillegeist et al., 2004; Bharath and Shumway, 2008).

In implementing Merton's (1974) structural model for default prediction Vassalou and Xing

(2004) propose an iterative procedure to estimate the parameters in the model by using historical accounting information and market equity data, and by making some simplifying assumptions. Bharath and Shumway (2008), while maintaining Merton's functional form, propose an alternative in which arbitrary values are used to circumvent the complicated iterative procedure. Results indicate that their *naïve* probability measure outperforms the Merton DD probability estimated by Vassalou and Xing (2004). Adopting a different tack, Campbell et al. (2008) use accounting and equity market variables to develop a reduced form, econometric specification to predict corporate bankruptcies. They show that their model provide meaningful empirical advantages over the bankruptcy risk scores proposed by Altman (1968) and Ohlson (1980), and is furthermore robust to alternative estimates of distance-to-default. The study finds that corporate default rates are strongly associated with lower profitability, higher leverage, lower and more volatile past returns, and smaller cash holdings.

The presumed advantage of the market-based framework vis-à-vis accounting-based models stems from its forward-looking structure; however, it is interesting to note that several important inputs used to estimate the market-based Merton DD themselves are not fully forward-looking. For instance, equity volatility used to infer firm volatility in the option-pricing framework is based on historical stock returns data. In addition, past stock returns are often used as a proxy for expected returns. Even reduced form specifications, such as the one by Campbell et al. (2008), rely on historical excess stock returns and volatility to estimate default probability. We posit that these backward-looking based inputs are not consistent with the forward-looking characteristics of the option-pricing model of the Merton DD measure and may result in sacrificing the model's predictive ability.

There is some related evidence that cast doubt on the historical measures used in bankruptcy studies. For instance, Elton (1999), Pastor et al. (2008), Chava and Purnanandam (2010) document that realized return is at best a noisy proxy of expected returns. Pastor et al. (2008) show that under reasonable assumptions about the dividend growth and expected return processes the implied cost of capital (ICC) is an excellent proxy for expected returns. ICC is the market implied internal rate of return that equates current market price to the discounted future dividend payout based on the analysts' consensus forecasts. Chava and Purnanandam (2010) use ICC as a proxy for expected returns and document a theoretically consistent positive cross-sectional relationship between default risk and stock returns, resolving the anomalous and often puzzling empirical finding among prior studies of a negative relationship between the two variables. Huang et al. (2012) demonstrates that extreme downside risk requires higher expected stock returns, Dungey and Martin (2007) and Yang and Zhou (2013) discuss the market contagion and spillover effect during crisis. Therefore, we hypothesize that the bankruptcy risk should be reflected in the implied cost of capital.

There is also a rich stream of literature that examines the ability of implied volatility (IV) to forecast future volatility in various markets. Recognizing the contribution of Hull, Nelken and White (2004), who present an alternative to the implied volatility in Merton (1974), we extend this work by including and implied cost of capital in the modeling. Studies such as by Latene and Rendleman (1976), Chiras and Manaster (1978), and Beckers (1981) illustrate the superiority of implied volatility as a proxy of future volatility. Jorion (1995) and Covrig and Low (2003) show that implied volatility outperforms past realized volatility as a forecast of future realized volatility in currency markets. Christensen and Prabhala (1998) find that implied option volatility is an unbiased and efficient forecast of future realized volatility in the S&P 100 index. The incremental information in implied volatility relative to past realized volatility in the stock market has found broad support among several studies including Day and Lewis (1992), Canina and Figlewski (1993), Lamoureux and Lastrapes (1993), Fleming (1998), Poteshman (2000), and Blair et al. (2001). Altman and Branch (2015) examine recidivism in Chapter 11 bankruptcy cases, specifically those that filed Chapter 11 reorganization only to revisit the same scenario in what has become known as a Chapter 22. The ability to postulate a firm's distance to default using forward looking measures may help the incidence of recidivism. The recent work of Danielova, Sarkar, and Hong (2013) focuses on corporate risk shifting with implications for default valuation.

The contribution of our paper lies in proposing a new and alternative implementation of the Merton (1974) model to measure default probabilities. Specifically, unlike previous studies, our approach incorporates forward-looking or implied measures of asset returns and volatility into the default prediction model. We hypothesize that the information contained in implied volatility and implied cost of capital would improve the forecasting ability of the default prediction model. Using the relative information content and prediction accuracy testing approaches we compare the performance of our model with three major default probability measurement approaches, as proposed by Vassalou and Xing (2004), Bharath and Shumway (2008) and Campbell et al. (2008), respectively.

Our regression results from a discrete-time hazard model, which evaluates the importance of each default measure in explaining the actual probability of bankruptcy, indicate that the psuedo-R^2 values for models that include implied measures are as high as 51% in comparison to 31.2% for Campbell et al.'s (2008) model, 27% for Shumway's (2001) model, and 15.9% when using Merton distance-to-default. The improvement is even more evident when comparing our results with Hillegeist et al. (2004) who obtain a psuedo-R^2 of only 12% for market-based measure based on the Black-Scholes option pricing model. The log-likelihood statistics from estimating non-nested models show that forecasting approaches that rely on implied cost of capital and implied volatility are the most preferred. We also assess bankruptcy prediction performance by estimating the accuracy ratio (i.e., ratio of Type I and Type II error frequency) for each alternative model specification. Results indicate that the most accurate model is one which incorporates both implied measures, carrying a default prediction accuracy rate of about 89%. Overall, our results support the informational relevance of forward-looking measures of returns and volatility in forecasting firm failures. The results from this paper carry implications for a wide range of studies examining credit risk, credit ratings, and portfolio allocation decisions.

2 Bankruptcy Prediction Models

The information content of our proposed measures are evaluated by comparing them to three important available models in the literature - the Vassalou and Xing (2004) implementation of the Merton DD model, the Bharath and Shumway (2008) Naïve DD model, and the Campbell et al. (2008) reduced-form specification model. We first describe the different models in this section.

2.1 The Merton (1974) Model

The Merton (1974) model proposes an approach to value corporate debt when a firm has only one outstanding zero coupon bond. With some modification and relaxation of the assumptions this model has been used to predict default. The model assumes the market value of a firm's asset follows a Geometric Brownian motion:

$$dV = \mu V dt + \sigma_V V dW, \tag{1}$$

where, V is the market value of the firm's assets with an instantaneous drift μ, and an instantaneous volatility σ_V, and W is a standard Wiener process.

Assume the firm has outstanding debt with a face value of F and maturity of T. Since shareholders are residual claimants of the firm's assets, the market value of equity, E, can be valued as a call option on V with risk free interest rate r, time to maturity T, and strike price F. If the value of the firm's assets is greater than the face value of the firm's debt, equity holders will exercise the option and pay off the bondholders. On the other hand, if the value of the firm's assets drops below the value of its liabilities, equity holders will let the call option expire and transfer ownership of the firm to debtholders.

The Black-Scholes equation for valuing the firm's equity under these conditions is given by:

$$E = V\mathcal{N}(d_1) - Fe^{-rT}\mathcal{N}(d_2), \tag{2}$$

where

$$d_1 = \frac{\ln(V/F) + (r + 0.5\sigma_V^2)T}{\sigma_V\sqrt{T}}, \tag{3}$$

and $d_2 = d_1 - \sigma_V\sqrt{T}$.

Based on option pricing theory, the risk-neutral probability of default of the firm is the probability that the call option is out-of-the-money, i.e., $\mathcal{N}(-d_2)$. Similarly, the physical probability of default is $\mathcal{N}(-DD)$, where, DD, the distance to default is defined as

$$DD = \frac{\ln(V/F) + (\mu - 0.5\sigma_V^2)T}{\sigma_V\sqrt{T}}. \tag{4}$$

DD can be interpreted as the number of standard deviations of annual asset growth by which the asset level (or expected asset level at a given time horizon) exceeds the face value of the firm's liabilities. The probability of bankruptcy in the Black-Scholes model depends upon the actual distribution of *future* asset values, which is a function of μ. Note V, μ, σ_V are not directly observable and must therefore be estimated. The other three variables, F, T and r can be approximated.

In equation (2), only the market value of equity, E, is observable. The risk free rate, r, is usually represented by the LIBOR rate or the one-year Treasury-bill rate. The other variables, the strike price F, the instantaneous drift μ, the instantaneous volatility σ_V, and the time to maturity T are unobservable since bankruptcy is not known until it is filed. Vassalou and Xing (2004) and Bharath and Shumway (2008) propose alternative implementation approaches of the Merton model. In this paper we use both these approaches as benchmarks to compare our results.

Vassalou and Xing (2004) Implementation Vassalou and Xing (2004) propose an interesting approach to estimate distance-to-default. They use the one-year Treasury rate as a proxy of the risk free rate r and set $T = 1$. To approximate the strike price, F, the model uses the sum of the current debt and 50% of long-term debt. The remaining three variables, namely, the market value of the firm's assets V, the instantaneous drift μ and instantaneous volatility σ_V are estimated by following an iterative procedure using daily equity data from the previous 12 months. First, using daily data from the past 12 months σ_E, the standard deviation of equity returns is estimated and used as an *initial value* for the estimation of σ_V.[1] Then, using equation (2) V is computed for each trading day using the corresponding market value of equity for that day, E. The standard deviation of the daily logarithmic returns of those V's are then used as the value of σ_V for the next iteration. This procedure is repeated until the values of σ_V from two consecutive iterations converge. The estimated σ_V is then used to back out the asset value V for each day. The drift, μ, is the mean of the change in lnV, the logarithm daily "return" of assets. The estimated drift μ is plugged into equation (4) to compute the probability of default.

Bharath and Shumway (2008) Alternative Bharath and Shumway (2008) propose a *naïve* estimation of DD as follows:

$$\text{naïve } DD = \frac{\ln((E+F)/F) + (r_{t-1} - 0.5 \text{ naïve } \sigma_V^2)T}{\text{naïve } \sigma_V \sqrt{T}}, \tag{5}$$

where

$$\text{naïve } \sigma_V = \frac{E}{E+F}\sigma_E + \frac{E}{E+F}\left(0.05 + 0.25\sigma_E\right). \tag{6}$$

The naïve model has the same functional form as Merton's DD probability, and essentially captures the information used by Merton's DD. V is approximated by $E+F$, the one-year Treasury bill rate is used as the proxy for the risk-free rate, and the sum of the current debt and 50% of

[1] σ_E is just used as the initial value for σ_V. Using other reasonable values will converge to the same estimation.

long term debt is used to approximate the strike price. The asset drift, μ, is approximated by the previous year cumulative return (calculated by cumulating monthly returns). Bharath and Shumway (2008) document that their proposed approach is at least as good as the Vassalou and Xing (2004) model.

An Alternative Implementation Since equity is an option on firm value, the volatility of equity, denoted σ_E, is also a function of V and σ. Denoting this function by g, then:

$$\sigma_E = g(V, \sigma_V). \tag{7}$$

If we write $E = f(V, \sigma_V)$, using Ito's lemma,

$$\sigma_E = g\left(V, \sigma_V\right) = \sigma_V V \frac{\frac{\partial f}{\partial V}}{f} = \sigma_V \frac{V}{E} N(d_1). \tag{8}$$

we can then simultaneously solve the system of nonlinear equations (2) and (8) to obtain V and σ_V if we have a good approximation of σ_E. Hillegeist et al. (2004) use this approach to estimate default probability.

2.2 Campbell, Hilscher and Szilagyi (2008) Model

Campbell et al. (2008) predict defaults using a discrete-time hazard model. The hazard rate model uses historical data of realized default to estimate default probability by using a reduced form empirical specification where the marginal probability of bankruptcy or failure over the next period follows a logistical distribution, as follows:

$$P_{t-1}(Y_{it} = 1) = \frac{1}{1 + \exp(-\alpha - \beta x_{i,t-1})}, \tag{9}$$

where Y_{it} is a dummy variable that equals one if the firm goes bankrupt in time t conditional on not failing earlier, and $x_{i,t-1}$ is a vector of explanatory variables known at the end of the previous period, and $(\alpha + \beta x_{i,t-1})$, a linear combination of explanatory variables indicate the default possibility of a firm i at time t. The higher the value of $(\alpha + \beta x_{i,t-1})$, the greater possibility that a firm will default in the forecasting period. The Campbell et al. (2008) model is estimated at the end of each month.

The explanatory variables used in their "best" model include geometrical four quarters moving average of Net Income to Market-valued Total Assets (NIMTAAVG), geometrical 12-month moving average monthly log excess returns relative to the S&P 500 index (EXRETAVG), Total Liability to Market-valued Total Assets – the sum of firm market equity and total liability – (TLMTA), the ratio of Cash and Short-term Assets to the Market-valued Total Assets (CACSHMTA), the standard deviation of each firm's daily stock return over the past 3 months (SIGMA), log price per share, truncated above \$15 (PRICE), the relative size of each firm measured as the log ratio of its market capitalization to that of the S&P 500 index (RSIZE), and market-to-book ratio of equity (MB). The book value of assets are adjusted by adding 10% of the differences between market and book equity to the book value of total assets. The book value of equity is also adjusted in a similar manner. Book value of equity is truncated at \$1 to avoid negative values.[2]

2.3 Proposed New Measures

All the above discussed approaches share a similar drawback in that they rely on historical data to estimate volatility and expected returns. However, by design, the structural-based option pricing equations (2)-(4) are forward looking. Since we are interested in estimating the default probability of a firm in the future, it would seem more informative to use forward-looking estimates of σ_V and μ instead of backward-looking estimates. Therefore, we propose to estimate default probability by replacing historical estimates of volatility and returns with market implied measures of volatility and cost of capital. Specifically, we estimate three alternative specifications of the Merton-based DD model:

1. IV-DD Model (Model 1): We use the OptionMetrics "standard" fixed maturity at-the-money call option implied volatility as the proxy of σ_E, and simultaneously solve the system of nonlinear equations (2) and (8) to obtain V and σ_V. This can be solved each day for the previous 12 months. In order to isolate the marginal information content of implied volatility, this model uses the past one year logarithm returns of the estimated asset value as a proxy for future returns. The default probability measure produced by this model is labeled as IV-DD.

2. IV/ICC-DD Model (Model 2): We use the same approach as IV-DD to estimate V and σ_V.

[2]For detailed descriptions of the variables, see Campbell et al. (2008).

However, in this model we employ the implied cost of capital used in Lee et al. (2009), Pastor et al. (2008), and Chava and Purnanandam (2010) as a proxy of expected returns.[3] We expect this approach to perform the best since it uses forward-looking measures for both volatility and expected returns. The default probability produced by this model is labeled as IV/ICC-DD.

3. IV/ICC-Naïve DD Model (Model 3): We plug the option implied volatility into equation (6) as σ_E. Then, we use equation (5) to estimate the naïve distance-to-default by replacing r_{t-1} in Bharath and Shumway's model with ICC_t, the implied cost of capital. The default probability estimate produced by this model is labeled as IV/ICC-Naïve DD.

3 Data and Summary Statistics

This section describes our data sources, sample selection and summary statistics.

3.1 Data Sources

The data for the study covers the time period 1996 to 2013. There is significant variation in the sample period used in prior studies - Vassalou and Xing (2004) cover the period 1971 to 1999, Campbell et al. (2008) study the 1963 to 2003 time period, and Bharath and Shumway's (2008) examination spans 1980 to 2003. The relatively more recent start date of our sample, 1996, is attributed to the availability of historical options data and implied volatility measures from OptionMetrics since this time period. The bankruptcy data for the period 1996 to 2008 is the same data as Chava and Jarrow (2004), Chava and Purnanandam (2010), Chava et al. (2011), and Alanis et. al. (2014), and is provided by Professor Sudheer Chava. Data for the more recent time period, 2009-2013, is collected from Bankruptcy.com and the UCLA-LoPucki bankruptcy research database. Firm accounting information is obtained from COMPUSTAT, and monthly and daily equity markets data are from CRSP. The analysts forecast data used to calculate the implied cost of capital data is from I/B/E/S. To classify firms into industries, we thankfully acknowledge the Fama-French 48 industry classification definition from Professor Kenneth R. French's data library.

[3] Computation of the implied cost of capital is detailed in Appendix A.

3.2 Sample Selection

We start the sample selection process with all common stocks traded on NYSE, AMEX and NASDAQ. We use COMPUSTAT quarterly files to get each firm's "Debt in one Year" and "Long-Term debt" series for all companies. To ensure that all information used to calculate the default measures are publicly available, we do not use book value of debt of the new quarter until 3 months have elapsed from the end of the previous quarter. We require a firm to have at least one year of financial information to be included in the sample. We then combine the accounting information from COMPUSTAT with daily and monthly equity market data from CRSP. Firms without equity market data from CRSP are dropped at this step. Following Bharath and Shumway (2008) financial firms with Fama-French industry code 44 to 47 are excluded. The default measures are computed at the end of each month for the whole sample period, 1996-2013. According to Vassalou and Xing (2004), the default probability for default firms goes up sharply in the 5 years prior to default and many firms are delisted from the exchange about 2 to 3 years prior to default. Therefore, if we only considered firms that defaulted in the coming one month this would result in a significant loss in sample size. Therefore, in all our calculations, we match the last available quarterly accounting data in COMPUSTAT and market equity data in CRSP to the default within the coming five year period. That is, any defaulted firm with both accounting information and market equity data in the five years prior to default is included in the final default sample. A firm is only counted in the default sample only once at the last time when COMPUSTAT and CRSP data are available. It is considered to be solvent at any point of time before then.

Our proposed measures rely on options data to obtain implied volatility. We only consider firms that have at least one year of 182-days implied volatility of "at-the-money" call in the OptionMetrics' standard options database. Following Hillegeist et al. (2004) and others, all default probability measures are winsorized at the 0.001% and 99.999% levels.[4]

Using the sample data we calculate six separate default prediction estimates at the end of each month, with three estimates corresponding to Vassalou and Xing (2004) (the VX-DD approach), Bharath and Shumway (2008) (the BS-Naïve DD approach), Campbell, Hilscher and Szilagyi (2008) (the CHS approach), respectively. Additionally, we calculate three new default prediction estimates

[4]To be consistent with the original work, in the estimation of the Campbell et. al. (2008) logit regression model, we winsorize all the regressors at 5% and 95%.

using the newly proposed measures given by IV-DD (Model 1), IV/ICC-DD (Model 2), and IV/ICC-Naïve DD (Model 3). In this context, it is important to reiterate that the VX-DD and BS-Naïve DD approaches are based on Merton's structural model; whereas, the CHS approach is an empirical measure of default that is estimated by running a dynamic panel model using a logistic specification of various accounting and market equity explanatory variables for solvent and bankrupt firms. We provide a comparative analysis of prediction accuracy and relative information content obtained from the proposed new approaches with the existing default models.

3.3 Summary of Sample

Summary description of our final sample are presented in Tables 1 and 2. Table 1 reports the total number of firms and failure events in our data set on a yearly basis. As with other studies, due to the nature of bankruptcy filings, there is substantial annual variation in the total number of firms and bankruptcy cases. The number of firms vary between 1383 (in the year 2002) and 2020 (in the year 2012). The highest rates of default are witnessed in 2001 when 84 of the 1487 firms in the sample (about 5.65%) defaulted during the year. By contrast, the default rate in 2010 was only about 0.43%. Table 2 breaks the sample firms down into the Fama French 48 industries. We notice that certain types of industries carry greater default risk than others. For instance, 3 of the 11 firms in the Fabricated Products industry (industry code 20) in the sample defaulted. There were no corporate defaults in seven of the 48 industries during the study period. They are Candy and Soda, Beer and Liquor, Tobacco Products, Aircraft, Shipbuilding and Railroad Equipment, Nonmetallic and Industrial Metal Mining, and Shipping Containers.

Table 3 presents summary statistics of the default probabilities for solvent and bankrupt companies in our sample. In general, the six default measures all show significant differences in default probabilities between solvent and bankrupt firms across all statistics. Not surprisingly, the differences in the 1^{st} percentile are relatively small due to winsorization of the data. Among the six measures, Model 1 (IV-DD) has the highest default probability means of 5.29% and 66.39% for the solvent and bankrupt sub-samples, respectively. By contrast, the CHS model provides the lowest default probability means of 0.014% and 0.537% for the two corresponding sample firms. Among the Merton DD-based approaches, Model 2 (IV/ICC-DD) produces the lowest mean and median for bankruptcy firms. It is interesting to note that the 99^{th} percentile for several of the default probability measures of the bankruptcy firms get very close to 100%. For Model 2, the value is 81.41% which is the lowest among the five Merton-DD measures. However, the 99^{th} percentile of CHS default measure for the bankruptcy firms is only 4.56%. We attribute this result to the fact that the CHS measure cannot be technically interpreted as an absolute measure of default probability, but rather a default probability measure that is conditional on survival.[5]

Table 4 reports the correlation matrix across the different default probability measures. The lower triangular and the upper triangular parts of the matrix represent the Pearson product moment correlation and the Spearman rank order correlation respectively. The results show that the correlations across the different measures are all positive ranging from about 0.3 to 0.9. Among the different measures, IV-DD and VX-DD default probabilities share the highest Pearson correlation of 0.86; whereas, IV/ICC-Naïve DD and CHS measures have the lowest Pearson correlation of 0.33. It is also evident that the correlation between IV/ICC-DD and all the other measures are similar, ranging between 0.52 and 0.62.

Table 5 compares the out-of-sample default predictability performance of the six default measures. Firms are sorted into deciles at the end of each year based on a particular forecasting variable. Then the number of defaults that occur in each of the decile groups are summarized. We report the results for each of the top five deciles separately and combine the lowest five deciles into one row in the table. The IV/ICC-DD approach, which includes implied estimates of volatility and cost of capital in the Merton DD framework, exhibits the best out-of-sample predictive ability. This model is able to classify about 79% of the default firms into the 1^{st} decile, and classify about 96% of all default firms into the top five deciles. The DD model with implied volatility (IV-DD) classifies about 71% of the default firms into the 1^{st} decile and a total of 93% into the highest five deciles. Examining the top two deciles combined, the rank-order performance of the out-of-sample default predictability are as follows: IV/ICC-DD (87.9%), IV-DD (86.2%), CHS (77.7%), BS-DD (76.7%), VX-DD (75.9%), and IV/ICC-Naïve DD (73.3%), respectively. These results broadly conform with Bharath and Shumway (2008) who also find that the Merton-based VX and naïve approaches classify about 80% of the defaults into the highest quintile. The slightly better out-of-

[5] Theoretically, the CHS measure is not directly comparable with the Merton-DD measures. Default probability estimates based on Merton's model can be calculated using a firm's one period accounting and equity data. In contrast, estimation of the CHS measure requires a robust sample of historical default data, as this model is calibrated using all available information. Arguably, the sample dependent nature of the CHS measure, as well as the Z-score and O-score, is viewed as a shortcoming of this type of model.

sample performance of the CHS approach is also consistent with the findings of Campbell et al. (2008). Most notably, preliminary indications from this particular exercise indicates that the DD model that includes our proposed two new measures - implied volatility and implied cost of capital - stand out as the "best" model among the alternative forecasting approaches. The statistical significance and relative performance of the different default probability measures are more formally investigated in the following section.

4 Model Evaluation Approaches

The bankruptcy prediction performance of the different implementation approaches are evaluated using two different tests: the receiver operating characteristic (ROC) curve to assess predictive accuracy and the relative information content tests using a discrete-time hazard model.

4.1 The ROC Curve

The ROC curve, which has its origins in the field of signal detection, has been recently applied to assess credit rating models and compare different default risk measures (Agrawal and Taffler 2008; Vassslou and Xing 2004). ROC is a graphical plot that illustrates the performance of a binary classifier system with varying discrimination thresholds. The ROC curve is constructed by plotting the fraction of true positives out of the total actual positives versus the fraction of false positives out of the total actual negatives, at various threshold settings. In our case, true positives are defined as the number of correct forecast of defaults and the actual positives are defined as the actual total number of defaults. In a symmetric fashion, false positives are the number of non-default firms forecasted as defaults and the actual negatives reflect the total number of non-defaults.

The prediction accuracy for each model is evaluated by calculating the area under the ROC curve (AUC). In order to calculate AUC, the ROC is smoothed by the binomial model and then the area is calculated using a closed-form expression. The accuracy ratio (AR), which can vary from 50% (random model) to 100% (perfect model), is a linear transformation of the area under the ROC curve:

$$AR = 2 \times (AUC - 0.5). \tag{10}$$

We construct the ROC curve at the end of calendar year where all firms are ranked based on their default from highest risk to lowest risk. For every integer between 0 and 100, we first examine how many firms actually failed within the first $x\%$ of firms with highest bankruptcy risk. This number is then divided by the total number of failures in the sample and plotted against x.

4.2 Relative Information Content Test

The relative information content test is employed to assess the marginal contribution of default related information contained in each model, and the results are then used to rank order the predictive performance of the various models. The discrete hazard model is particularly well-suited for bankruptcy data since they contain binary, cross-sectional and time-series observations, and offer many econometric advantages over static single period models of bankruptcy. The Vuong (1989) test is used to compare the log-likelihood statistics obtained from the different non-nested discrete hazard models, and determine whether the performance differences are statistically significant.

Since we use multiple firm-year observations for each firm we adopt the same form of discrete-time hazard model as in Campbell et al. (2008). The discrete-time hazard model has the following form:

$$p_{i,t}^{j} = \frac{e^{\alpha(t)+\beta X_{i,t}}}{1+e^{\alpha(t)+\beta X_{i,t}}} = \frac{1}{1+e^{-\left(\alpha(t)+\beta_1 X_{i,t}+\beta_2^j S_{i,t}^j\right)}}, \tag{11}$$

where, $\alpha(t)$ is a time-varying, system-wide variable that captures the baseline hazard rate, X is a vector of independent variables for the i^{th} firm at time t, and β is the parameters to be estimated. The inclusion of a time dependent baseline hazard rate in the model allows for default rates to vary each year. However, fluctuations in the baseline hazard rate will cause cross-sectional observations to be correlated across time. Therefore, similar to Hillegeist et al. (2004) we proxy the baseline hazard using the previous year's actual default rate. Our results indicate that in most of the regressions the prior year's default rate is not statistically significant. Therefore, we drop the real default rate by making $\alpha(t)$ constant and discuss only this set of results in the paper. Since default probability is bounded between 0 and 1, it is not suitable as an independent variable for a logistic model. Therefore, we transform all the default probability into a "score" using the inverse logistic function to remove the range restriction. The score of a default probability measure P has the following form:

$$Score(P) = \ln\left(\frac{P}{1-P}\right). \tag{12}$$

Notice as the default probability P approaches one (zero), the score gets to negative (positive) infinity. After the transformation, each of the DD-related default measure becomes a score and the score for CHS default measure is represented by the $(\alpha + \beta x_{i,t-1})$ term. To avoid the impact of extreme values in the regression, we winsorize all the probability measures at the minimum and maximum values of 0.001% and 99.999%. The winsorization results in minimum and maximum values of the scores to be between -11.5 and +11.5. The default scores are labeled as S^{VX}, S^{BS}, S^{CHS}, S^{M1}, S^{M2} and S^{M3} for the scores corresponding with VX-DD, BS-Naïve DD , CHS, IV-DD, IV/ICC-DD and IV/ICC-Naïve DD estimation approaches.

The Vuong (1989) closeness test and the Clarke (2001, 2003) sign test are used to statistically evaluate for the differences in the non-nested model. The Vuong (1989) test is a likelihood-ratio based test with the null hypothesis that the two models are equally close to the actual model, against the alternative that one model is closer. The Clarke (2001, 2003) distribution free test applies a modified paired sign test to evaluate the differences in the individual log-likelihood from two non-nested models. The test determines whether or not the median log-likelihood ratio is statistically different from zero. If the models are equally close to the true model, half the individual log-likelihood ratio should be greater than zero and the other half should be less than zero. The reported test is the difference between the number of positive and negative log-likelihood ratios.

5 Empirical Results

5.1 ROC Results

Figure 1 plots the ROC curves for the six models along the 45° line which represents the ROC benchmark of random guesses. It is easily evident from the figure that the prediction accuracy for all the models are significantly better than the random model. With the exception of the IV/ICC-Naïve DD and BS-Naïve DD models, the ROCs of the remaining models are found to be tightly grouped together. A closer inspection of the area reported under the curve (AUC), at the bottom panel beneath the figure, reveals very little difference in default predictability between the

IV/ICC-DD and the CHS approaches, with both their areas being close to 0.94. Similarly, there is very little distinction between IV-DD and VX-DD, although they each perform better than the BS-Naïve DD model.

Table 6 presents the AUC and AR statistics, as well as the χ^2 test results from comparing each ROC against the benchmark IV/ICC-DD model's ROC. The results show that the default accuracy of IV/ICC-DD is significantly larger (statistically significant at the 1% level) than VX-DD, BS-DD, IV-DD and IV/ICC-Naïve DD. The difference in the prediction accuracy between IV/ICC-DD and CHS, however, is not found to be statistically significant.

5.2 Relative Information Content Test Results

We compare the relative information content of the six models by estimating six separate discrete hazard models with a constant baseline hazard rate as follows:[6]

$$p_{i,t}^{j} = \frac{1}{1 + e^{-\left(\alpha^{j} + \beta^{j} S_{i,t}^{j}\right)}}, \quad j = 1 \text{ to } 6, \tag{13}$$

where, $S_{i,t}^{j}$ is the j^{th} score of firm i at time t.

The results from each model are reported in Table 7. Table 7 presents the estimated coefficients, the log-likelihood values, the McFadden's pseudo-R^2 values (calculated as $1 - L_1/L_0$, where L_1 is the log-likelihood of the estimated model and L_0 is the log-likelihood of the null model that includes only a constant term).

Several interesting results are evident. First, the estimated parameters for all the six scores are positive and statistically significant at the 1% level. Second, the log-likelihood values range from -1300 for the model with S^{M2} to -1848 for Model with S^{M3}. Between these two extremes, the performance rank-order of the different default scores are: S^{CHS}, S^{VX}, S^{BS} and S^{M1}. This rank-order is preserved even in the context of evaluating the pseudo-R^2 values, with S^{M2} generating the highest pseudo-R^2 of 44.8%. The lowest pseudo-R^2, 20%, is associated with the IV/ICC-Naïve DD model. Notably, the explanatory power for all the different models are significantly larger than

[6]We also estimate discrete hazard model using previous year real default rate and the previous five year real default rate as proxy of the baseline rate. In most of the models, the coefficients of the baseline rate are not significant. Thus, we decide to report the results of a constant baseline rate. The results with time-varying baseline rate are available upon request.

those found by Hillegeist et al. (2004) who use a similar discrete hazard model and find a pseudo-R^2 of only 12% for the Merton-based default probability. Overall, results from relative information content corroborate the conclusions derived from the prediction accuracy analysis.

The Vuong and Clarke tests in Panel B suggest that the model with S^{M2} is statistically closer to the "optimal" model than the models with S^{VX}, S^{BS}, $S^{M1}$and S^{M3}. Although the model with S^{M2} has a slightly higher pseudo-R^2 and log-likelihood value than the model with S^{CHS} the difference is not found to be statistically significant. Among the three models with implied measures we find that the model with S^{M2} which incorporates both implied volatility and implied cost of capital outperforms the other two models - S^{M1} and S^{M3}. The relatively weak performance of IV-DD may be attributed to the fact that using historical estimated return of assets to compute distance-to-default significantly reduces the power of the implied volatility measure in the model. Although IV/ICC-Naïve DD incorporates both implied volatility and implied cost of capital, it is a measure which simply (or *naïvely)* replaces the historical volatility and historical equity returns with implied measures into the BS-Naïve DD model. It appears that the simple setup of the Bharath and Shumway model is not able to fully utilize the power of the two forward-looking measures. It may also be the case that the BS-Naïve DD model is optimized with historical measures, and as such it performs poorly compared with alternative model setups that include forward-looking measures. In contrast, the proposed IV/ICC-DD model seems to fully benefit from the power of implied volatility as a proxy of future volatility and ICC as a proxy of expected returns in estimating default probabilities.

We find that the performance of the IV/ICC-DD model to be on par with the CHS model in predicting defaults. However, it must be noted that the estimation results of the CHS model is sample dependent (i.e., specific to the characteristics of the sample used in the regression), which may make it more difficult to establish the model's robustness under alternative samples. By contrast, the DD-based structural model can be computed for any publicly traded firm - therefore, our approach using implied volatility and cost of capital further enhances the attractiveness of using a structural framework to estimate default probabilities.

In Table 8 we conduct a side-by-side comparison of our "best" predictive model (IV/ICC-DD) with each of the remaining estimation models. In other words, in each model, S^{M2} and one of the other five measures are treated as explanatory variables in the discrete-time hazard model.

These results are compared with Tables 7 to confirm the information superiority of the IV/ICC-DD model in predicting defaults. For example, the model with only S^{VX} in Table 7 shows that the estimated coefficient of S^{VX} is 36.6%, with a log-likelihood value of -1466.1 and pseudo-R^2 of 36.5%. In the corresponding regression model that is enhanced with forward-looking measures in the second column of Table 8, the estimated coefficients of S^{VX} and S^{M2} are both significant, with a substantial increase in pseudo-R^2 (36.5% to 46%). In other words, adding forward-looking measures of volatility and expected returns result in a significant improvement in the predictive ability of the model. On the other hand, comparing the result with just S^{M2} as a single factor, we notice that adding S^{VX} only results in a small increase in the pseudo-R^2 from 44.8% to 46%. Similar results hold for all the other models, with one important exception - the inclusion of S^{CHS} to S^{M2} does not greatly improve the model performance when compared with the model containing just S^{CHS} or S^{M2}. This suggests that the forward-looking measures of the Merton DD model is not a superior alternative to the CHS model.

5.3 Additional Relative Information Content Test Results

Platt and Platt (1991) suggest that it is possible to extract additional information by adjusting the default probability measures for industry effects. Thus, similar to Hillegeist et al. (2004), we investigate whether adjusting the default probability measures for industry effects would change the relative performance of the different measures. Using the Fama-French industry classification, the default probability measures are decomposed into the prior year's industry mean and its deviation from its previous year's industry mean. The discrete-time hazard model has the following form:

$$p^j_{i,t} = \frac{1}{1 + e^{-\left(\alpha^j + \beta^j_m S^j_{im,t-1} + \beta^j_d S^j_{id,t}\right)}}, \quad j = 1 \text{ to } 6, \tag{14}$$

where, $S^j_{im,t-1}$ refers to the previous year's mean of the j^{th} "score" for the industry where the i^{th} firm belongs, and $S^j_{id,t} = S^j_{i,t} - S^j_{im,t-1}$ is the deviation of the j^{th} score of firm i at time t.

The results for the industry-adjusted models are presented in Table 9. All the estimated parameters are statistically significant at the 1% level. The results show that there is significant industry variation in default rates; however, incorporating this information into the hazard model results only in a slight improvement of the overall information content of the different models. Importantly, there is no evidence of change in the relative rankings even after adjusting for industry effects. The Vuong and Clarke tests again confirm the superiority of the model containing the IV/ICC-DD measure with respect to the other Merton DD-based implementations. On the other hand, the Vuong and Clarke tests provide conflicting evidence on the relative superiority of IV/ICC-DD versus CHS models.

In an attempt to obtain additional default information we decompose the default measures into lagged levels and changes. The discrete-time hazard model has the following form:

$$p_{i,t}^{j} = \frac{1}{1 + e^{-\left(\alpha^{j} + \beta_{L}^{j} S_{i,t-1}^{j} + \beta_{C}^{j} \Delta S_{i,t}^{j}\right)}}, \quad j = 1 \text{ to } 6, \tag{15}$$

where, $\Delta S_{i,t}^{j} = S_{i,t}^{j} - S_{i,t-1}^{j}$ is the change of the j^{th} score of firm i at time t from time $t-1$. The estimated results from the discrete hazard model regressions are reported in Table 10. The results again confirm the relative superiority of IV/ICC-DD. The Vuong's z-statistics of the difference between the model with S^{M2} and the model with S^{CHS} is 2.397, which is statistically significant at the 5% level.

5.4 Direct Comparison to the Campbell et al. (2008) Hazard Model

All of the five Merton DD-default based specifications provide direct measures of bankruptcy; whereas, the CHS measure is the predicted result from a discrete hazard model. In this section, we first run the Campbell et al. (2008) hazard model that contains the eight factors from their "best" model at the end of each year. We then sequentially add one of the six "scores" and estimate the results from this new model. Notably, in this setup the Campbell et al. (2008) hazard model is nested in the other specifications. To make the comparison consistent we also run a model with the CHS measure. Since the CHS score is from the Campbell et al. (2008) hazard model hazard model itself, we would not expect to find that the CHS score to change the results of the hazard model.

The estimated results of this set of models are in Table 11. The results show that although the CHS model itself should not be considered as an absolute default probability measure, it predicts default quite well. The model has a log-likelihood of -1274.5 and a pseudo-R^2 of 45.9%, which is slightly higher than -1300.1 and 44.8% obtained from the model with S_{M2} as a single factor (as seen in Table 7). These results can be contrasted with those found by Campbell et al. (2008), who find that the pseudo-R^2 of the DD model is only 15.9%, whereas their "best" model has a pseudo-R^2 of 31.2%. In Campbell et al. (2008), when DD is added as an additional factor into their best model, in the 0 horizon (forecasting window of 1-month) model the parameter changed direction, and in the longer horizon the parameters decrease dramatically. For the 12-month horizon, the magnitude of the parameter values decreased by about 74% (from -0.345 to -0.091). The results show that although in longer horizon adding the DD factor contain some additional information, the predictive power of the DD factor is significantly reduced by the other factors.

In our case, adding the eight Campbell et al. variables into the hazard model only improves the performance slightly and reduces the coefficients of all but the CHS parameter. The decrease of the coefficient of S^{M2} is the least, from 0.632 in a single factor setting to 0.446. The observed change is much smaller than what is documented by Campbell et al. (2008). The pseudo-R^2 for the model with S^{M2} and the eight Campbell et al. factors is 51.2% (log likelihood value of -1155.1) compared to 44.8% (log likelihood value of -1300.1) with S^{M2} as a single factor in the discrete hazard model. The results suggest that the DD related models may carry more information for long forecast horizons than short horizons, rendering the parameters to be more resilient to the addition of more factors. In summary, adding the eight explanatory variables into the one-factor hazard model does not significantly strengthen models default prediction performance.

Panel B shows the likelihood ratio test of model differences. The base model is the Campbell et al. (2008)'s best model. The null hypothesis is that alternative models with additional factors (the six other scores) is not different in default prediction from the base model. Since the Campbell et al. (2008) model is nested in the other models, the likelihood ratio test is used. The results in Panel B shows that the models with default scores are better than the model without these scores. In other words, adding the scores strengthen the information content of the Campbell et al. (2008) model.

6 Conclusions

Default risk is a critical risk factor for both shareholders and debtholders. Further, in the literature Jeon and Nishihara (2016) develop a structural model that examines capital structure and credit risk. In this paper, we develop a forward looking default model to forecast defaults. Thus, any variable that contains (possible) future information should be preferred in default risk models. Among the different types of default models, the Merton (1974) distance-to-default based model is one of the most important and widely used model. However, in empirical applications, most implementations of the Merton distance-to-default model use historical return and volatility as proxies of expected return and volatility. We propose an important and variant application of the Merton distance-to-default model that uses implied cost of capital and option implied volatility as forward-looking proxies to improve the prediction of defaults.

Studies show that option implied volatility is a better predictor of future volatility than the historical volatility. Hammoudeh, Bhar, and Liu (2013) examine credit default swaps as a measure of risk. In the likelihood of default, the expectation of the market participants about the future underlying asset price movements is uncertain. Similarly, the implied cost of capital is measure implied by the current equity price and the market consensus represented by a group of financial analysts who follow the firm closely. Thus, it also contains market expectations about future equity returns. We hypothesize that the combination of both measures into the Merton distance-to-default model should provide a more accurate prediction of default risk.

Empirical data for the period 1996 to 2012 are used to evaluate the results of our proposed approach with existing implementations of the Merton DD-based model by Vassalou and Xing (2004) and Bharath and Shumway (2008), as well as the dynamic logistic empirical specification of Campbell et al. (2008) in different setups. Results from receiver operating characteristic and relative information content tests with discrete-time hazard model indicate that our proposed measures perform well in all settings. The model's out-of-sample performance compared with other Merton-based models is superior and statistically significant at the 1% level. The results also show that our proposed implementation is at least as good as, and in some cases better than, the best model specification in Campbell et al. (2008).

We conclude that implied cost of capital and implied volatility contain important information and should be considered as proxies of future expected returns and volatility in default prediction models.

References

[1] Agarwal, V., Taffler, R., 2008. Comparing the Performance of Market-based and Accounting-based Bankruptcy Predication Models. Journal of Banking & Finance 32, 1541-1551.

[2] Alanis, E., Chava, S., Kumar, P., 2014. Shareholder bargaining power and debt vverhang. Working Paper, Georgia Institute of Technology.

[3] Altman, E.I., Haldeman, R.G., Narayanan, P., 1977. A new model to identify bankruptcy risk of corporations. Journal of Banking and Finance 1, 29-54.

[4] Beaver, W.H., 1968. The information content of annual earnings announcements. Journal of Accounting Research 6, 67-92.

[5] Beckers, S., 1981. Standard deviation implied in option prices as predictors of future stork price variability. Journal of Banking and Finance 5, 363-381.

[6] Berndt, A., Douglas, R., Duffie, D., Ferguson, M., Schranz, D., 2008. Measuring default risk premia from default swap rates and EDFs. Carnegie Mellon University Working Paper.

[7] Bharath, S.T., Shumway, T., 2008. Forecasting default with the Merton distance to default model. Review of Financial Studies 21, 1339-1369.

[8] Blair, B.J., Poon, S., Taylor, S.J., 2001. Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high-frequency index returns? Journal of Econometrics 105, 5-26.

[9] Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. Journal of Finance 63, 2899-2939.

[10] Canina, L., Figlewski, S., 1993. The informational content of implied volatility? Review of Financial Studies 6, 659-681.

[11] Chava, S., Purnanandam, A. 2010. Is default risk negatively related to stock returns? Review of Financial Studies 23, 2523-2559.

[12] Chava, S., Jarrow, R.A., 2004. Bankruptcy prediction with industry effects. Review of Finance 8, 537-569.

[13] Chava, S., Stefanescu, C., Turnbull, S., 2011. Modeling the loss distribution. Management Science 57, 1267-1287.

[14] Chiras, D.P., Manaster, S., 1978. The information content of option prices and a test of market efficiency. Journal of Financial Economics 6, 213-234.

[15] Christensen B.J., Prabhala, N.R., 1998. The relation between implied and realized volatility, Journal of Financial Economics 50, 125-50.

[16] Clarke, K.A., 2001. Testing nonnested models of international relations: Reevaluating realism. American Journal of Political Science, 45, 724-744.

[17] Clarke, K.A., 2003. Nonparametric model discrimination in international relations. Journal of Conflict Resolution 47, 72–93.

[18] Covrig, V., Low, B. S., 2003. The quality of volatility traded on the over-the-counter market: A multiple horizons study? Journal of Futures Markets 23, 261-285.

[19] Day, T.E., Lewis, C.M., 1992. Stock market volatility and the information content of stock index options? Journal of Econometrics 52, 267-287.

[20] Danielova, A., Sarkar, S., Hong, G., 2013. Empirical evidence on corporate risk-shifting. The Financial Review 28, 443-460.

[21] Dungey, M. and Martin, V.L., 2007. Unravelling financial market linkages during crises. Journal of Applied Econometrics, 22(1), pp.89-119.

[22] Elton, E.J., 1999. Expected return, realized return and asset pricing tests. Journal of Finance 54, 1199-1220.

[23] Fleming, J.. 1998. The quality of market volatility forecasts implied by S&P 100 index option prices? Journal of Empirical Finance 5, 317-345.

[24] Friewald, N., Wagner, C., Zechner, J. 2014. The cross-section of credit risk premia and equity returns. Journal of Finance 59, 2419-2468.

[25] Hammoudeh, S., Bhar, R., Liu, T. 2013. Relationships between financial sectors CDS spreads and other gauges of risk: Did the Great Recession change them? The Financial Review 48, 151-178.

[26] Hillegeist, S.A., Keating, E.K., Cram, D.P. and Lundstedt, K.G., 2004. Assessing the probability of bankruptcy. Review of accounting studies, 9(1), 5-34.

[27] Huang, W., Liu, Q., Rhee, S.G. and Wu, F., 2012. Extreme downside risk and expected stock returns. Journal of Banking & Finance, 36(5), 1492-1502.

[28] Hull, J.C., Nelken, I., White, A. 2004/05. Mertons model, credit risk, and volatility skews. Journal of Credit Risk 1:1, 3-27.

[29] Jeon, H., Nishihara, M. 2016. The effects of reversible investment on capital structure and credit risks, The Financial Review 51, 263-293.

[30] Jorion, P., 1995. Predicting volatility in the foreign exchange market? Journal of Finance 50, 507-628.

[31] Lamoureux, C.G., Lastrapes, W.D., 1993. Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities? Review of Financial Studies 6, 293-326.

[32] Latane, H., Rendleman, R., 1976. Standard deviation of stock price ratios implied in option prices. Journal of Finance 31, 369-381.

[33] Lee, C.M., Ng, C.D. , Swaminathan, B., 2009. Testing international asset pricing models using implied cost of capital. Journal of Financial and Quantitative Analysis 44, 307-335.

[34] Merton, R.C. 1974. On the pricing of corporate debt: The risk structure of interest rates. Journal of Finance 29, 449-470.

[35] Hillegeist, S.A., Keating, E.K., Cram, D.P., Lundstedt, K.G., 2004. Assessing the probability of bankruptcy. Review of Accounting Studies 9, 5-34.

[36] Ohlson, J.A. 1980. Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research 18, 109-131.

[37] Pastor, L., Sinha, M., Swaminathan, B., 2008. Estimating the intertemporal risk-return trade-off using the implied cost of capital. Journal of Finance 63, 2859-2897.

[38] Platt, H., Platt, M., 1991. A note on the use of industry-ralative ratios in bankruptcy prediction. Journal of Banking and Finance 15, 1183-1194.

[39] Poteshman, A. M., 2000. Forecasting future volatility from option prices? Working Paper University of Illinois at Urbana-Champaign.

[40] Vassalou, M., Xing, Y., 2004. Default risk in equity returns. Journal of Finance 59, 831-868.

[41] Vuong, Q.H., 1989. Likelihood ratio tests for model selection and non-nested hypothesis. Econometrica 57, 307-333.

[42] Yang, J. and Zhou, Y., 2013. Credit risk spillovers among financial institutions around the global credit crisis: Firm-level evidence. Management Science, 59(10), 2343-2359.

Table 1: Bankruptcy by Year

The table reports the number of bankruptcy cases year by year.

Year	Bankrupt	Firms	Rate
1997	8	1670	0.479
1998	12	1772	0.677
1999	17	1582	1.075
2000	39	1494	2.610
2001	84	1487	5.649
2002	68	1383	4.917
2003	42	1405	2.989
2004	19	1561	1.217
2005	20	1606	1.245
2006	9	1758	0.512
2007	10	1836	0.545
2008	17	1709	0.995
2009	39	1787	2.182
2010	8	1875	0.427
2011	10	1963	0.509
2012	15	2020	0.743

Table 2: Bankruptcy by Industry

The table reports the number of bankruptcy by the Fama French 48 Industry (financial industry excluded).

Ind Code	Industry	Firm Years	Number of Firms	Defaults	Percentage of Firm-years Bankrupt	Percentage of firms Bankrupt
1	Agriculture	115	22	2	1.739	9.091
2	Food Products	534	73	4	0.749	5.479
3	Candy and Soda	73	11	0	0.000	0.000
4	Beer and Liquor	129	15	0	0.000	0.000
5	Tobacco Products	72	10	0	0.000	0.000
6	Recreation	138	27	5	3.623	18.519
7	Entertainment	457	87	18	3.939	20.690
8	Printing and Publishing	219	33	7	3.196	21.212
9	Consumer Goods	445	58	5	1.124	8.621
10	Apparel	372	57	6	1.613	10.526
11	Healthcare	532	105	15	2.820	14.286
12	Medical Equipment	867	164	7	0.807	4.268
13	Pharmaceutical Products	2058	400	16	0.777	4.000
14	Chemicals	847	116	8	0.945	6.897
15	Rubber and Plastic Products	145	30	2	1.379	6.667
16	Textiles	68	16	4	5.882	25.000
17	Construction Materials	403	64	6	1.489	9.375
18	Construction	333	49	6	1.802	12.245
19	Steel Works Etc	578	93	10	1.730	10.753
20	Fabricated Products	39	11	3	7.692	27.273
21	Machinery	1012	160	10	0.988	6.250
22	Electrical Equipment	399	68	8	2.005	11.765
23	Automobiles and Trucks	464	81	13	2.802	16.049
24	Aircraft	161	21	0	0.000	0.000
25	Shipbuilding and Railroad Equipment	66	13	0	0.000	0.000
26	Defense	61	6	1	1.639	16.667
27	Precious Metals	269	53	1	0.372	1.887
28	Non-Metallic and Metal Mining	223	40	0	0.000	0.000
29	Coal	120	22	1	0.833	4.545
30	Petroleum and Natural Gas	1989	331	16	0.804	4.834
31	Utilities	1311	164	9	0.686	5.488
32	Communication	1599	337	59	3.690	17.507
33	Personal Services	286	56	8	2.797	14.286
34	Business Services	2965	727	55	1.855	7.565
35	Computers	1021	222	19	1.861	8.559
36	Electronic Equipment	2036	367	20	0.982	5.450
37	Measuring and Control Equipment	575	84	3	0.522	3.571
38	Business Supplies	450	61	5	1.111	8.197
39	Shipping Containers	109	14	0	0.000	0.000
40	Transportation	1084	182	18	1.661	9.890
41	Wholesale	954	153	16	1.677	10.458
42	Retail	1763	280	37	2.099	13.214
43	Restaurants, Hotels and Motels	548	88	5	0.912	5.682
48	Almost Nothing	452	88	13	2.876	14.773
Total		28341	5059	441	1.556	8.717

Table 3: Summary Statistics of Default Probability Measures

The table reports the summary statistics of the different default probability measures. Here, VX, BS, CHS, and M1, M2 and M3 refer to the Vassalou and Xing (2004), Bharath and Shumway (2008) Naive measure, the Campbell et al., (2004) best model, the proposed IV-DD, IV/ICC-DD and IV/ICC Naive DD measures, respectively.

Variables	Status	Frequency	Mean (%)	Median (%)	Std. Dev. (%)	P1 (%)	P25 (%)	P75 (%)	P99 (%)
VX	Solvent	27900	4.704	0.001	16.149	0.001	0.001	0.051	91.479
	Bankrupt	441	59.529	72.551	38.309	0.001	18.102	95.896	99.999
BS	Solvent	27900	3.323	0.001	12.471	0.001	0.001	0.006	73.116
	Bankrupt	441	50.159	62.828	37.098	0.001	7.758	83.649	99.998
CHS	Solvent	27900	0.014	0.002	0.093	0.000	0.001	0.005	0.235
	Bankrupt	441	0.537	0.166	0.887	0.001	0.041	0.601	4.560
M1	Solvent	27900	5.293	0.001	18.227	0.001	0.001	0.010	99.158
	Bankrupt	441	66.385	84.958	37.550	0.001	33.669	98.632	99.999
M2	Solvent	27900	0.707	0.001	4.107	0.001	0.001	0.003	18.168
	Bankrupt	441	24.686	17.388	23.447	0.001	4.531	40.650	81.406
M3	Solvent	27900	3.199	0.001	12.239	0.001	0.001	0.041	71.836
	Bankrupt	441	31.627	22.267	32.128	0.001	0.540	55.877	99.994

Table 4: Correlation Matrix of Default Probability Measures

The table reports the correlation matrix of the new default probability measures. The lower triangular and upper triangular parts of the matrix represent the Pearson correlation and the Spearman rank correlation. Here, VX, BS, CHS, and M1, M2 and M3 refer to the Vassalou and Xing (2004), Bharath and Shumway (2008) Naive measure, the Campbell et al., (2004) best model, the proposed IV-DD, IV/ICC-DD and IV/ICC Naive DD measures, respectively.

Variables	VX	BS	CHS	M1	M2	M3
VX		0.775	0.724	0.817	0.763	0.799
BS	0.749		0.646	0.717	0.699	0.680
CHS	0.488	0.503		0.688	0.670	0.711
M1	0.859	0.674	0.417		0.797	0.756
M2	0.611	0.619	0.597	0.591		0.868
M3	0.708	0.561	0.326	0.553	0.522	

Table 5: Out of Sample Forecasts

The table reports the out-of-sample forecasts for the different default probability measures. Here, VX, BS, CHS, and M1, M2 and M3 refer to the Vassalou and Xing (2004), Bharath and Shumway (2008) Naive measure, the Campbell et al., (2004) best model, the proposed IV-DD, IV/ICC-DD and IV/ICC Naive DD measures, respectively.

Decile	VX	BS	CHS	M1	M2	M3
1	57.143	61.050	60.302	70.748	78.458	62.528
2	18.802	15.654	17.407	15.420	9.524	10.835
3	8.756	6.446	8.259	4.308	4.082	4.063
4	5.069	4.512	4.796	1.814	2.268	2.483
5	3.134	2.394	2.931	0.907	1.134	2.032
6-10	7.097	9.945	6.306	6.803	4.535	18.059

Table 6: Receive Operating Characteristic Analysis Results

The table reports the results of the Receiver Operating Characteristic analysis. Here, VX, BS, CHS, and M1, M2, and M3 refer to the Vassalou and Xing (2004), Bharath and Shumway (2008) Naive measure, the Campbell et al., (2004) best model, the proposed IV-DD, IV/ICC-DD and IV/ICC Naive DD measures, respectively. AUC and AR are the area under the curve and accuracy rate. The χ^2 statistics is the statistics test results of comparing each ROC against the benchmark IV/ICC-DD model's ROC. *** indicates significance at the 1% level.

Measures	AUC	StdErr	χ^2	AR
VX	0.922	0.007	33.98***	0.843
BS	0.902	0.009	34.10***	0.803
CHS	0.939	0.007	0.54***	0.879
M1	0.926	0.007	18.97***	0.853
M2	0.943	0.006		0.885
M3	0.831	0.012	113.69***	0.662

Table 7: Logistic Regression with Default Scores as Single Factors

The table reports the results of the Hazard model:

$$p^j_{i,t} = \frac{1}{1 + e^{-\left(\alpha^j + \beta^j S^j_{i,t}\right)}}, \quad j = 1 \text{ to } 6, \tag{16}$$

where, $S^j_{i,t}$ is the j^{th} score of firm i at time t. Here, VX, BS, CHS, and M1, M2, and M3 refer to the Vassalou and Xing (2004), Bharath and Shumway (2008) Naive measure, the Campbell et al., (2004) best model, the proposed IV-DD, IV/ICC-DD and IV/ICC Naive DD measures, respectively. *, **, and *** indicates significance at the 10%, 5% and 1% level.

Panel A: Logistic Regression Estimation Results

Parm	S^{VX}	S^{BS}	S^{CHS}	S^{M1}	S^{M2}	S^{M3}
Constant	-2.467***	-1.950***	4.900***	-2.852***	-0.298	-2.350***
S^{VX}	0.366***					
S^{BS}		0.408***				
S^{CHS}			1.018***			
S^{M1}				0.263***		
S^{M2}					0.632***	
S^{M3}						0.275***
Log Likelihood	-1466.1	-1503.9	-1332.5	-1575.3	-1300.1	-1848.2
Pseudo-R^2	0.365	0.348	0.433	0.324	0.448	0.200
Observations	28341	28341	28341	28341	28341	28341
Bankruptcy	441	441	441	441	441	441

Panel B: Vuong Test and Clarke Sign Test Results

	Vuong's Z	Clarke
VX versus M2	7.227***	10398.0***
BS versus M2	7.107***	10745.5***
CHS versus M2	1.331	7455.5***
M1 versus M2	9.208***	10704.5***
M3 versus M2	13.183***	11863.5***

Table 8: Logistic Regression with S^{M2} as a Common Factor

The table reports the results of the Hazard model:

$$p_{i,t}^{j} = \frac{1}{1 + e^{-\left(\alpha^{j} + \beta_{M2} S_{i,t}^{M2} + \beta^{j} S_{i,t}^{j}\right)}}, \quad j = 1 \text{ to } 5, \tag{17}$$

where, $S_{i,t}^{M2}$ refers to the score of the IV/ICC DD measure, and $S_{i,t}^{j}$ is the score of the other five measures. Here, VX, BS, CHS, and M1, M2, and M3 refer to the Vassalou and Xing (2004), Bharath and Shumway (2008) Naive measure, the Campbell et al., (2004) best model, the proposed IV-DD, IV/ICC-DD and IV/ICC Naive DD measures, respectively. *, **, and *** indicates significance at the 10%, 5% and 1% level.

Parm	S^{VX}	S^{BS}	S^{CHS}	S^{M1}	S^{M3}
Constant	-0.636***	-0.335	2.518***	-0.743***	-0.323
S^{M2}	0.496***	0.544***	0.381***	0.484***	0.703***
S^{VX}	0.119***				
S^{BS}		0.105***			
S^{CHS}			0.489***		
S^{M1}				0.121***	
S^{M3}					-0.096***
Log Likelihood	-1255.5	-1249.9	-1237.9	-1265.3	-1284.2
Pseudo Rsq	0.460	0.463	0.475	0.463	0.455
Observations	28341	28341	28341	28341	28341

Table 9: Logistic Regression with Industry Means and Deviations

The table reports the results of the Hazard model:

$$p^j_{i,t} = \frac{1}{1+e^{-\left(\alpha^j+\beta^j_m S^j_{im,t-1}+\beta^j_d S^j_{id,t}\right)}}, \quad j = 1 \text{ to } 6, \tag{18}$$

where, $S^j_{im,t-1}$ refers to the previous year's mean of the j^{th} "score" for the industry where the i^{th} firm belongs, and $S^j_{id,t} = S^j_{i,t} - S^j_{im,t-1}$ is the deviation of the j^{th} score of firm i at time t. Here, VX, BS, CHS, and M1, M2, and M3 refer to the Vassalou and Xing (2004), Bharath and Shumway (2008) Naive measure, the Campbell et al., (2004) best model, the proposed IV-DD, IV/ICC-DD and IV/ICC Naive DD measures, respectively. *, **, and *** indicates significance at the 10%, 5% and 1% level.

Panel A: Logistic Regression Estimation Results

Parm	S^{VX}	S^{BS}	S^{CHS}	S^{M1}	S^{M1}	S^{M1}
Constant	-2.723***	-1.711**	4.539**	-2.286***	-0.983	-1.453**
S^{VX}_{id}	0.376***					
S^{VX}_{im}	0.338***					
S^{BS}_{id}		0.401***				
S^{BS}_{im}		0.433***				
S^{CHS}_{id}			1.026***			
S^{CHS}_{im}			0.984***			
S^{M1}_{id}				0.253***		
S^{M1}_{im}				0.332***		
S^{M2}_{id}					0.654***	
S^{M2}_{im}					0.566***	
S^{M3}_{id}						0.254***
S^{M3}_{im}						0.376***
Log Likelihood	-1464.4	-1503.1	-1332.3	-1566.2	-1295.7	-1836
Pseudo-R^2 Rsq	0.366	0.348	0.433	0.329	0.450	0.205
Observations	28341	28341	28341	28341	28341	28341
Bankruptcy	441	441	441	441	441	441

Panel B: Vuong Test and Clarke Sign Test Results

	Vuong's Z	Clarke
S^{VX}_{id} and S^{VX}_{im} versus S^{M2}_{id} and S^{M2}_{im}	7.167***	10412.0***
S^{BS}_{id} and S^{BS}_{im} versus S^{M2}_{id} and S^{M2}_{im}	7.192***	10802.5***
S^{CHS}_{id} and S^{CHS}_{im} versus S^{M2}_{id} and S^{M2}_{im}	1.468	7568.5***
S^{M1}_{id} and S^{M1}_{im} versus S^{M2}_{id} and S^{M2}_{im}	8.959***	10738.5***
S^{M3}_{id} and S^{M3}_{im} versus S^{M2}_{id} and S^{M2}_{im}	13.139***	11869.5***

Table 10: Logistic Regression with Lags and Changes

The table reports the results of the Hazard model:

$$p_{i,t}^{j} = \frac{1}{1+e^{-\left(\alpha^{j}+\beta_{L}^{j}S_{i,t-1}^{j}+\beta_{C}^{j}\Delta S_{i,t}^{j}\right)}}, \quad j = 1 \text{ to } 6, \tag{19}$$

where, $\Delta S_{i,t}^{j} = S_{i,t}^{j} - S_{i,t-1}^{j}$ is the change of the j^{th} score of firm i at time t from time $t-1$. Here, VX, BS, CHS, and M1, M2, and M3 refer to the Vassalou and Xing (2004), Bharath and Shumway (2008) Naive measure, the Campbell et al., (2004) best model, the proposed IV-DD, IV/ICC-DD and IV/ICC Naive DD measures, respectively. *, **, and *** indicates significance at the 10%, 5% and 1% level.

Panel A: Logistic Regression Estimation Results

Parm	S^{VX}	S^{BS}	S^{CHS}	S^{M1}	S^{M2}	S^{M3}
Constant	-2.161***	-1.716***	5.991***	-2.118***	0.177	-2.433***
S_{t-1}^{VX}	0.412***					
ΔS^{VX}	0.346***					
S_{t-1}^{BS}		0.442***				
ΔS^{BS}		0.377***				
S_{t-1}^{CHS}			1.133***			
ΔS^{CHS}			0.954***			
S_{t-1}^{M1}				0.375***		
ΔS^{M1}				0.268***		
S_{t-1}^{M2}					0.726***	
ΔS^{M2}					0.652***	
S_{t-1}^{M3}						0.275***
ΔS^{M3}						0.272***
Log Likelihood	-1203	-1239	-1138.1	-1078.2	-900.25	-1374.6
Pseudo-R^2	0.377	0.357	0.447	0.369	0.478	0.199
Observations	26908	26908	26908	26908	26908	26908
Bankruptcy	417	417	417	417	417	417

Panel B: Vuong Test and Clarke Sign Test Results

	Vuong's Z	Clarke
S_{t-1}^{VX} and ΔS^{VX} versus S_{t-1}^{M2} and ΔS^{M2}	6.489***	8537.0***
S_{t-1}^{BS} and ΔS^{BS} versus S_{t-1}^{M2} and ΔS^{M2}	6.560***	8747.0***
S_{t-1}^{CHS} and ΔS^{CHS} versus S_{t-1}^{M2} and ΔS^{M2}	2.397**	7397.0***
S_{t-1}^{M1} and ΔS^{M1} versus S_{t-1}^{M2} and ΔS^{M2}	7.483***	8714.5***
S_{t-1}^{M3} and ΔS^{M3} versus S_{t-1}^{M2} and ΔS^{M2}	11.713***	9441.5***

Table 11: Logistic Regression with Campbell et al. (2008) Factors

The table reports the results of the Hazard model:

$$p_{i,t}^{j} = \frac{e^{\alpha(t)+\beta X_{i,t}}}{1+e^{\alpha(t)+\beta X_{i,t}}} = \frac{1}{1+e^{-\left(\alpha(t)+\beta_1 X_{i,t}+\beta_2^j S_{i,t}^j\right)}}, \tag{20}$$

where, $X_{i,t}$ are the eight explanatory variables in the Campbell et al. (2008) best model, and $S_{i,t}^{j}$ is the j^{th} score of firm i at time t. Here, VX, BS, CHS, and M1, M2, and M3 refer to the Vassalou and Xing (2004), Bharath and Shumway (2008) Naive measure, the Campbell et al., (2004) best model, the proposed IV-DD, IV/ICC-DD and IV/ICC Naive DD measures, respectively. *, **, and *** indicates significance at the 10%, 5% and 1% level.

Panel A: Logistic Regression Estimation Results

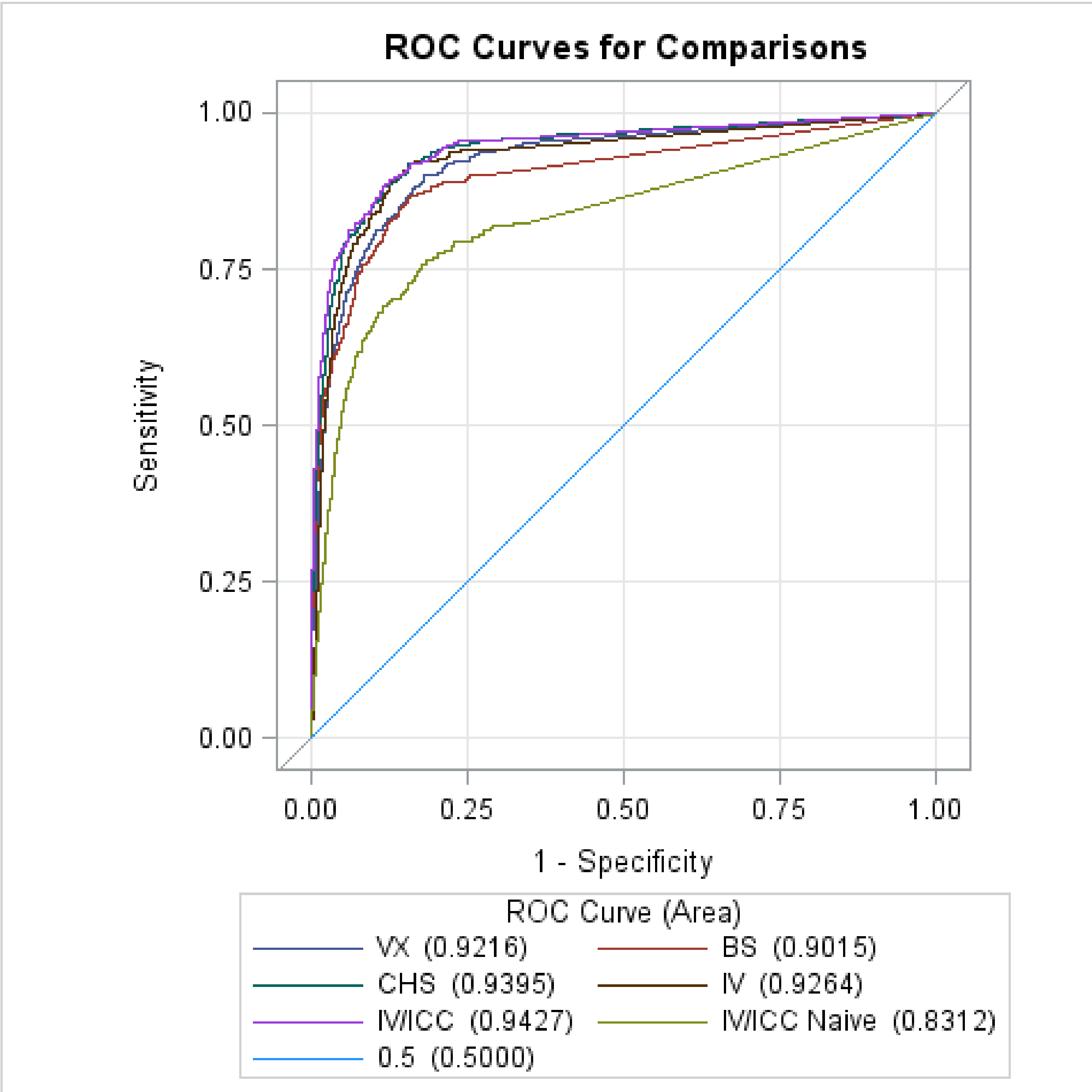
Parm	Original	S^{VX}	S^{BS}	S^{CHS}	S^{M1}	S^{M2}	S^{M3}
Constant	-8.403***	-5.561***	-6.282***	4.105	-5.811***	-2.583**	-6.730***
NIMTAAVG	-24.22***	-22.49***	-20.24***	3.413	-24.22***	-16.98***	-22.65***
TLMTA	2.130***	0.217	0.820**	-1.826*	0.749**	-0.631*	1.211***
EXRETAVG	-24.15***	-18.92***	-21.39***	-11.62***	-18.63***	-18.90***	-23.09***
SIGMA	1.432**	0.546	0.853	-0.440	0.766	-0.001	1.247**
RSIZE	-0.274**	-0.236**	-0.257**	-0.214	-0.199	-0.178*	-0.241**
CASHMTA	-4.004***	-4.161***	-3.319***	-1.420	-4.235***	-3.771***	-3.948***
MB	0.072*	0.078**	0.048	-0.020	0.082**	0.042	0.063*
PRICE	-0.680***	-0.606***	-0.679***	0.000***	-0.619***	-0.408**	-0.793***
S^{VX}		0.177***					
S^{BS}			0.145***				
S^{CHS}				1.057***			
S^{M1}					0.148***		
S^{M2}						0.446***	
S^{M3}							0.082***
Log Likelihood	-1274.5	-1222.2	-1213.2	-1274.5	-1224.4	-1155.1	-1260.2
Pseudo-R^2	0.459	0.475	0.479	0.459	0.481	0.512	0.465
Observations	28341	28341	28341	28341	28341	28341	28341
Bankruptcy	441	441	441	441	441	441	441

Panel B: Likelihood Ratio Test of Model Difference

	χ^2
S^{VX} versus Original	61.29***
S^{BS} versus Original	79.28***
S^{CHS} versus Original	0.00
S^{M1} versus Original	100.18***
S^{M2} versus Original	238.84***
S^{M3} versus Original	28.51***

Figure 1: Daily Levels of S&P 500 Moments

Project/Default Risk Hong/FR/Figure1.png



Appendix A: Computation of Implied Cost of Capital

We follow Chava and Purnanandam (2010) to compute the Implied Cost of Capital (ICC) using the discounted cash flow framework to value equity. The model relies on explicitly forecasting cash flows for the next $T = 15$ years and incorporates the effect of subsequent cash flows using a terminal value calculation. The expected free cash flow to equity, $FCFE$,of firm i in year $t+k$ is given by:

$$E_t\left[FCFE_{i,t+k}\right] = FE_{i,t+k} \times (1 - b_{i,t+k}), \tag{21}$$

where, $FE_{i,t+k}$ is the earning estimate of firm i in year $t+k$ and $b_{i,t+k}$ is the earnings plow-back or retention rate.

1. We compute the terminal value as a perpetuity:

$$TV_{i,t+T} = \frac{FE_{i,t+T+1}}{r_{i,e}}. \tag{22}$$

2. Collecting all the terms, we get the following:

$$P_{i,t} = \sum_{k=1}^{T} \frac{FE_{i,t+k} \times (1 - b_{i,t+k})}{(1 + r_{i,e})^k} + \frac{FE_{i,t+T+1}}{r_{i,e}(1 + r_{i,e})^T}. \tag{23}$$

3. Solving the equation for $r_{i,e}$ gives us the Implied Cost of Capital – Expected Return.

Estimating Earnings:

$FE_{i,t+k}$ is estimated using the earnings forecast available from I/B/E/S (Institutional Brokers' Estimate System) database. We use one-year and two-year ahead consensus (median) forecasts as proxies for $FE_{i,t+1}$ and $FE_{i,t+2}$, respectively. $FE_{i,t+3} = FE_{i,t+2} \times (1 + LTG)$, where, LTG is the consensus long-term growth forecast from I/B/E/S. We forecast earning from year $t+4$ to $t+T+1$ by mean-reverting the year $t+3$ earnings growth rate to a steady state growth rate of the firm's earnings which is assumed to be the ten year moving average of gross domestic product (GDP) growth rate (g).[7] The growth rate for year $t+k$ ($k>3$) is

$$g_{i,t+k} = g_{i,t+k-1} \times \exp\left(\frac{\ln(g/LTG)}{T-2}\right). \tag{24}$$

Using this, we compute earning as follows:

$$FE_{i,t+k} = FE_{i,t+k-1} \times (1 + g_{i,t+k}). \tag{25}$$

Estimating Earnings Plowback Rate:

The most recent year payout is defined as the sum of dividends and share repurchases minus any issuance of new equity. This data is obtained from COMPUSTAT. The payout ratio is the payout divided by net income. Plow-back rate is one minus the payout ratio. We use the above plow-back ratio for the first year and mean-revert it to a steady state value by year $t+T+1$. At the steady state $g = ROI \times b$. Set ROI for new investments to r_e. That is, we assume that the presence of industry competition would drive the return on a firm's new investments to be equal to its cost of equity. Thus, the plow-back rate for year $t+k$ ($2 \leq k \leq T+1$) is:

$$b_{i,t+k} = b_{i,t+k-1} - \frac{b_{i,t+1} - b_i}{T}, \tag{26}$$

$$b_i = \frac{g}{r_{i,e}}. \tag{27}$$

Expected risk-premium: We subtract the prevailing one-year risk-free rate from the ICC to obtain the expected risk-premium.

[7]GDP growth data is taken from Bureau of Economic Analysis.