

Predicting corporate bankruptcy using the framework of Leland-Toft: Evidence from U.S.

Chris Charalambous ¹
University of Cyprus

Spiros H. Martzoukos ²
University of Cyprus

Zenon Taoushianis ³
University of Cyprus

June 2019

* We are thankful, for comments and suggestions, to conference participants at the EAA 2019 conference in Paphos, Cyprus, the doctoral consortium of the International Journal of Accounting Symposium 2017 in Limassol, Cyprus, the EFMA 2016 conference in Basel, Switzerland and several anonymous referees.

¹ Professor (Emeritus) of Management Science. Contact address: Department of Business and Public Administration, School of Economics and Management, University of Cyprus, P.O. Box 20537, CY 1678 Nicosia, Cyprus: bachris@ucy.ac.cy (+357 22 892466)

² Associate Professor of Finance. Contact address: Department of Accounting and Finance, School of Economics and Management, University of Cyprus, P.O. Box 20537, CY 1678 Nicosia, Cyprus: baspiros@ucy.ac.cy (+357 22 893615)

³ PhD Candidate and Corresponding Author. Contact address: Department of Accounting and Finance, School of Economics and Management, University of Cyprus, P.O. Box 20537, CY 1678 Nicosia, Cyprus: taoushianis.zenon@ucy.ac.cy (+357 22 893640)

Predicting corporate bankruptcy using the framework of Leland-Toft: Evidence from U.S.

June 2019

Abstract

In this paper, we evaluate an alternative approach for bankruptcy prediction that measures the financial healthiness of firms that have coupon-paying debts. The approach is based on the framework of Leland and Toft (1996), which is an extension of a widely-used model; the Black-Scholes-Merton model. Using U.S. public firms between 1995 and 2014, we show that the Leland-Toft approach is more powerful than Black-Scholes-Merton in a variety of tests. Moreover, extending popular but also contemporary corporate bankruptcy models with the probability of bankruptcy derived from the Leland-Toft model, such as Altman (1968), Ohlson (1980) and Campbell et al. (2008), yields models with improved performance. One of our tests, for example, shows that banks using these extended models, achieve superior economic performance relative to other banks. Our results are consistent under a comprehensive out-of-sample framework.

JEL classification: C51, C52, G13, G33

Keywords: Corporate Bankruptcy, Leland and Toft, Predictive Ability, Bank Performance

1. Introduction

Corporate bankruptcy prediction models are valuable risk management tools to assist bank managers in the decision-making process of identifying firms which are likely to fail and therefore would not be able to pay their obligations. This is because, the consequences arising from bankruptcy are enormous and include, for instance, economic ones such as the loss of the amount lent, impaired profitability for the bank which in certain cases may harm the viability of the bank, the financial system and the economy as well⁴. From the perspective of an investor, economic consequences include the loss of the wealth invested in bankrupt firms but also include non-economic ones, such as the loss of investors' confidence towards the financial markets. For these reasons, among others, it is important for the interested parties to develop and apply reliable corporate bankruptcy prediction models.

While various models have been proposed in the literature, two of the most frequently used by academics and practitioners are Z-score (Altman, 1968) and O-score (Ohlson, 1980). These models mainly use information from the financial statements of the firm to relate past performance with bankruptcy risk. More recently, models with both accounting and market information have been developed. These models have the advantage to incorporate timely market information and thus the likelihood of bankruptcy can be updated more frequently⁵. Studies such as Shumway (2001), Chava and Jarrow (2004), Campbell et al. (2008) and Tinoco and Wilson (2013) show that accounting and market information yield models with improved performance. Another approach is the contingent claims-based approach which is based on the framework of Black and Scholes (1973) and Merton (1974). There, in an options pricing framework, the

⁴ For instance, Papakyriakou et al. (2019) show that the failure of financial institutions from U.S, negatively affect the international stock markets.

⁵ Refer to Agarwal and Taffler (2008) for a discussion between accounting and market information in bankruptcy prediction models.

probability of bankruptcy is the probability that the assets value of the firm will be worth less than its debts, at maturity⁶. Such models, are frequently referred to, as structural models.

In this paper, we evaluate an alternative approach for bankruptcy prediction and we construct with it powerful bankruptcy models, seeking to improve the performance of existing models. Specifically, we evaluate an approach that measures the financial healthiness of firms with coupon-paying debts, using the framework of Leland and Toft (1996). Leland and Toft (1996) belongs to a class of models that extends Merton (1974), to incorporate the effects of taxes and bankruptcy costs to the valuation of equity and a corporate coupon-paying debt with finite maturity. Other significant features of their framework are that, bankruptcy can occur prior to the maturity of the debt but also, they consider the case when the bankruptcy point is determined endogenously. Thus, Leland and Toft is a more appropriate corporate model than Black-Scholes-Merton because it includes a richer information set about the firm which can be useful for bankruptcy prediction⁷.

Several models are considered in this paper. Firstly, we compare the performance of two structural models; Leland-Toft and Black-Scholes-Merton. We believe that the former would outperform the latter since it is an extended version, containing more information for bankruptcy prediction. Next, we compare the performance of three reduced-form models with three hybrid models (i.e reduced-form models augmented with structural models). The first reduced-form model is Ohlson (1980) which is a comprehensive model since it includes various accounting variables such as profitability, liquidity, leverage, cash flows etc. Next, we augment Ohlson model with the probability of bankruptcy derived from Leland-Toft, yielding a hybrid model. We

⁶ See for instance Bharath and Shumway (2008) and Afik et al. (2016) for related literature regarding this approach.

⁷ A strand of the literature, also examines empirically the performance of the structural models in predicting corporate bond prices and spreads and find that they do not accurately predict them (see for instance Lyden and Saraniti, 2000 and Eom et al., 2004 and references therein).

believe that augmenting this comprehensive accounting model with Leland-Toft will improve its performance, yielding a powerful bankruptcy prediction model. Another reduced-form model we examine is an extension of Altman (1968) model, which includes an additional cash flow variable and was suggested by Almamy et al. (2016). The authors find that augmenting Altman's model with a cash flow variable yields improved predictive ability for U.K. firms. We believe that augmenting Altman's model with Leland-Toft will further increase its predictive ability. Next, we use a competent reduced-form model that includes accounting and market variables suggested by Campbell et al. (2008). This model has been examined by Bauer and Agarwal (2014) and was found to outperform other approaches, such as reduced-form models with accounting information as well as the Black-Scholes-Merton structural model. Finally, we seek to improve Campbell et al. (2008), by including Leland-Toft in their model.

For our analysis we use 5460 U.S. public firms with coupon-paying debts between 1995 and 2014. The performance of the models is compared on three dimensions and our results are based on an extensive out-of-sample framework: 1) On their ability to discriminate bankrupt from healthy firms using Receiver Operating Characteristics (ROC) analysis 2) On their ability to predict bankruptcy probabilities close to actual or equivalently on their ability to empirically fit the data using log-likelihood statistics and 3) By measuring the economic performance of banks when they are competing to grant loans to individual firms and each bank uses a corresponding model to evaluate prospective firm-customers. For this last test, we employ the setting of Agarwal and Taffler (2008).

The key findings of the paper are that 1) Leland-Toft approach is more powerful than Black-Scholes-Merton. Sensitivity analysis tests for Leland-Toft shows that its forecasting power is not affected by the choice of parameter values underlying the model. However, none of the structural

models can stand alone in forecasting bankruptcies since they are outperformed by reduced-form (and also hybrid) models, 2) Further increase in predictive ability is achieved when augmenting Altman's model with Leland-Toft rather than a cash-flow variable, 3) Augmenting the comprehensive models of Ohlson (1980) and Campbell et al. (2008) with Leland-Toft yields models with improved performance 4) Reduced-form models augmented with Leland-Toft, outperform reduced-form models augmented with BSM. In fact, the hybrid models which include Leland-Toft are the best performing models in all tests. Most importantly, in our experiment with the competitive loan market we find that banks using these extended models achieve superior economic performance relative to their competitors.

The paper proceeds as follows; Section 2 describes the bankruptcy models and the research hypotheses, section 3 discusses our data, section 4 discusses the methodology, section 5 reports the results and section 6 concludes.

2. Bankruptcy Models and Research Hypotheses

2.1. Structural Models

2.1.1. Black-Scholes-Merton

Merton (1974) shows that the equity value of the firm (E) can be viewed as a European call option underlying the assets of the firm (V) and with strike price the zero-coupon debt of the firm (D). The Black and Scholes (1973) options pricing formula can therefore be used to price the equity of the firm:

$$E = VN(d_1) - De^{-rT}N(d_2) \quad (1)$$

where r is the riskless rate of return, $N(\cdot)$ is the standard normal distribution function, T is the maturity of the debt and d_1, d_2 are defined as follows:

$$d_1 = \frac{\ln(V/D) + (r + 0.5\sigma_V^2)T}{\sigma_V\sqrt{T}} \quad (2)$$

$$d_2 = d_1 - \sigma_V\sqrt{T} \quad (3)$$

and σ_V is the volatility of assets value returns. In the framework of Black-Scholes-Merton, the firm goes bankrupt when $V < D$ and thus the probability of bankruptcy, $prob(V < D)$, is the probability that at debt maturity, the assets value is lower than the debt. The probability of bankruptcy is then given by the Black-Scholes-Merton (BSM hereafter) formula:

$$prob = N\left(-\frac{\ln(V/D) + (\mu - 0.5\sigma_V^2)T}{\sigma_V\sqrt{T}}\right) \quad (4)$$

where μ is the return of assets.

2.1.2. Leland and Toft (1996)

Leland and Toft (1996) extends the framework of Merton (1974) to incorporate the effects of taxes and bankruptcy costs in the valuation of a corporate risky debt with finite maturity. Their framework considers the valuation of debt that pays coupons as opposed to the framework of Merton where the firm issues a zero-coupon debt. In this context, Leland-Toft derive closed-form solutions for the market value of equity, debt and total firm value. Most importantly, they consider the case where bankruptcy is determined endogenously as opposed to Merton (1974) where bankruptcy is determined exogenously. This consideration allows the calculation of an optimal bankruptcy point which is chosen by the management of the firm in favor of shareholders such that the equity value is maximized. When assets value reaches that point, it is

optimal, from shareholders' perspective, for the firm to file for bankruptcy. In contrast, when bankruptcy is determined exogenously, the bankruptcy point is chosen arbitrarily⁸. However, this consideration is suboptimal because firms usually continue operations even when assets value falls below firm's debt and practically there is not an agreed value to use. Eq. (5) shows the calculation of the bankruptcy point, VB_{LT} , underlying Leland-Toft model which is a key determinant of the bankruptcy probability⁹:

$$VB_{LT} = \frac{\left(\frac{C}{r}\right)\left(\frac{A}{rT} - B\right) - A\frac{P}{rT} - \tau\frac{Cx}{r}}{1 + cx - (1 - c)B} \quad (5)$$

where

$$A = 2\alpha e^{-rT}N(a\sigma_V\sqrt{T}) - 2zN(z\sigma_V\sqrt{T}) - \frac{2}{\sigma_V\sqrt{T}}n(z\sigma_V\sqrt{T}) + \frac{2e^{-rT}}{\sigma_V\sqrt{T}}n(a\sigma_V\sqrt{T}) + (z - a)$$

$$B = -\left(2z + \frac{2}{z\sigma_V^2T}\right)N(z\sigma_V\sqrt{T}) - \frac{2}{\sigma_V\sqrt{T}}n(z\sigma_V\sqrt{T}) + (z - a) + \frac{1}{z\sigma_V^2T}$$

$$a = \frac{(r - \delta - 0.5\sigma_V^2)}{\sigma_V^2}, \quad z = \frac{\sqrt{a^2\sigma_V^4 + 2r\sigma_V^2}}{\sigma_V^2}, \quad x = a + z$$

with $N(\cdot)$ and $n(\cdot)$ denoting the cumulative standard normal distribution and standard normal density functions respectively. A closer examination shows that Eq. (5) is a function of eight parameters: risk-free rate (r), tax rate (τ), coupon payments (C), bankruptcy costs (c), volatility of assets (σ_V), debt principal (P), payout yield (δ), and debt maturity (T).

To evaluate bankruptcy risk in discrete points of time t , where $t \leq T$, we need to define a cumulative distribution function. In the framework of Leland-Toft (LT hereafter), the probability

⁸ For example, in the Merton's model the bankruptcy point is the debt of the firm and thus, is determined exogenously.

⁹ Hilberink and Rogers (2002), extend Leland-Toft (1996) to allow for sudden jumps in the asset value, V , and derive a new optimal bankruptcy point. However, the solutions are not explicit and some of the parameters are not straightforward to compute (see Eq. (3.16) and Eq. (3.23) in their paper).

that the current value of firm's assets, V , will fall to the bankruptcy point, VB_{LT} , for the first time at time t is obtained from Leland (2004) and defined as:

$$prob(t) = N(X) + e^Y N(Z) \quad (6)$$

where

$$X = \frac{-\ln\left(\frac{V}{VB_{LT}}\right) - (\mu - \delta - 0.5\sigma_V^2)t}{\sigma_V\sqrt{t}}, Y = \frac{-2\ln\left(\frac{V}{VB_{LT}}\right)(\mu - \delta - 0.5\sigma_V^2)}{\sigma_V^2}$$

$$Z = \frac{-\ln\left(\frac{V}{VB_{LT}}\right) + (\mu - \delta - 0.5\sigma_V^2)t}{\sigma_V\sqrt{t}}$$

Finally, t is the forecasting horizon, which in our case is one year.

2.1.3. Estimating Asset Values and Volatilities

The most important inputs to LT and BSM models are the value of assets and the volatility of assets returns which are not observed. In the context of options pricing, however, the following two non-linear equations can be solved iteratively to obtain the two variables of interest:

$$V = \frac{E - De^{-rT}N(d_2)}{N(d_1)} \quad (7)$$

$$\sigma_V = \frac{E\sigma_E}{VN(d_1)} \quad (8)$$

where σ_E is the volatility of equity returns that is directly estimated from daily equity data. The above iterative procedure, which we use to estimate the two unobserved inputs, is the standard approach for the estimation of asset value and volatility and has also been used by Eom et al.

(2004), Hillegeist et al. (2004), Campbell et al. (2008), while Vassalou and Xing (2004) use a variation of the above iterative process¹⁰.

2.2. Reduced-Form Models

Several reduced-form models are also considered. Ohlson (1980) is a model which relates bankruptcy with a set of accounting-based variables, defined as follows:

$$Ohlson = f(SIZE, TLTA, WCTA, CLCA, D(TL > TA), NITA, CFOTL, D(NI_t + NI_{t-1} < 0), CHINI) \quad (9)$$

Next, we consider Almamy, Aston and Ngwa (2016), which we refer as *AAN*. This model is an extension of Altman's model which incorporates a cash flow variable as additional predictor and it is defined as follows:

$$AAN = f(WCTA, RETA, EBITTA, MVTL, SLTA, CFOTA) \quad (10)$$

Moreover, we consider the model proposed by Campbell, Hilscher and Szilagyi (2008), which we refer to as *CHS*. This model is a mixture of accounting ratios, scaled by the market value of assets, and other market information as predictors, defined as follows:

$$CHS = f(NIMTA, TLMTA, EXRET, SIGMA, RSIZE, CASHMTA, MB, PRICE) \quad (11)$$

The definition of the variables is in table 1.

[Insert Table 1 here]

¹⁰ For other approaches, see Bharath and Shumway (2008), Charitou et al. (2013) and Afik et al. (2016)

2.3. Hybrid Models

Finally, we incorporate the probability of bankruptcy derived from LT as additional predictor in Ohlson, Altman and CHS models, yielding the following hybrid models which we refer to as *OLT*, *ALT* and *CHSLT* respectively:

$$OLT = f(SIZE, TLTA, WCTA, CLCA, D(TL > TA), NITA, CFOTL, D(NI_t + NI_{t-1} < 0), CHINI, LT) \quad (12)$$

$$ALT = f(WCTA, RETA, EBITTA, MVTL, SLTA, LT) \quad (13)$$

$$CHSLT = f(NIMTA, TLMTA, EXRET, SIGMA, RSIZE, CASHMTA, MB, PRICE, LT) \quad (14)$$

2.4. Research Hypotheses

LT is an extended version of Merton's model with less restrictive assumptions and a richer information set about the firm. Therefore:

Hypothesis 1: LT is a better alternative approach than BSM

Prior research suggests that accounting and market information should be included in corporate bankruptcy prediction models since they provide complementary information. For instance, variables such as the volatility of equity and excess equity returns improve the performance of accounting-based models (Chava and Jarrow, 2004; Hillegeist et al., 2004; Agarwal and Taffler, 2008; Tinoco and Wilson, 2013 etc.). Therefore, we expect that including LT in Ohlson model will enhance its performance:

Hypothesis 2: Incorporating LT as additional predictor in Ohlson, yields a model with improved performance.

Hence, the model in Eq. (12) should outperform the model in Eq. (9). An extension of Hypothesis 2, is as follows:

Hypothesis 2a: Ohlson model augmented with LT, will outperform Ohlson model augmented with BSM.

Almamy et al. (2016) suggest that augmenting Altman's model with a cash-flow variable, increases its predictive ability. However, further increase in predictive ability could be obtained when augmenting Altman's model with a predictor that measures the financial healthiness of firms with coupon-paying debts. Hence, the model in Eq. (13) should outperform the model in Eq. (10).

Therefore:

Hypothesis 3: Augmenting Altman's model with LT will further increase predictive ability than a cash-flow variable.

An extension of Hypothesis 3, is as follows:

Hypothesis 3a: Altman's model augmented with LT, will outperform Altman's model augmented with BSM.

Campbell et al. (2008) find that augmenting their model with BSM, doesn't yield improved performance, arguing that all the information incorporated in BSM, such as returns and volatilities, are already included in their model. Since LT is an extension of BSM that includes additional information, we want to investigate if augmenting Campbell et al. (2008) with LT, would improve its performance. This leads to the fourth hypothesis:

Hypothesis 4: Incorporating LT as additional predictor in Campbell et al. (2008), yields a model with improved performance.

Finally, an extension of the fourth hypothesis, is the following:

Hypothesis 4a: Augmenting Campbell et al. (2008) with LT, will outperform Campbell et al. (2008) augmented with BSM

3. Data

3.1. Sample

We analyze a sample of 5460¹¹ U.S. public firms from which 333 filed for bankruptcy in a specific year between the 20-year period 1995-2014. Bankruptcy filings were identified from BankruptcyData¹² and include firms which filed for bankruptcy under Chapter 7 or Chapter 11. To avoid problems related to sample selection bias and increase efficiency of regression estimates, we collect all available observations in the selected period for each bankrupt and healthy firm. This practice increases our sample to 39830 firm-year observations. Furthermore, once a firm files for bankruptcy, future observations for that firm are excluded but past observations for all bankrupt firms are included in our sample (i.e. before a firm file for bankruptcy, it is considered as healthy).

Table 2 presents the distribution of observations across the years.

[Insert Table 2 here]

¹¹ The framework of Leland and Toft (1996) applies for firms with coupon-paying debt. Thus, we keep only firms which have interest payments in their income statements

¹² Available at <http://www.bankruptcydata.com/findabrtop.asp>

In general, bankruptcy rate in all years is less than 1% except for years 1999 (1.493%) and the mid-crisis years 2008 and 2009 with the bankruptcy rate being at its peak (1.190% and 2.133% respectively). The average bankruptcy rate in the sample is 0.836% indicating the fact that bankruptcy is a rare event.

Similar with Bharath and Shumway (2008), Afik et al. (2016) and others, we exclude financial firms (SIC 6000-6799) due to the different nature of their operations and structure of their financial statements relative to other industrial firms. Firms are classified into a specific industry according to the Standard Industrial Classification (SIC) code provided by the United States Department of Labor. Table 3 shows the distribution of observations across industries.

[Insert Table 3 here]

Most of observations (53%) comes from the Manufacturing sector and then from Services, Transportation, Retail and Mining sectors, accounting for 16.42%, 10.36%, 8.41% and 5.87% of the sample respectively, whereas the Wholesale, Construction, Public Administration and Agriculture sectors account for the smallest proportions of the sample (4.03%, 0.95%, 0.62% and 0.35% respectively).

3.2. Variables Construction

To construct the relevant variables used in the structural and reduced-form models, we collect annual financial data from Compustat and daily equity data from CRSP. All variables are constructed at the fiscal year-end, prior to the year of bankruptcy.

To estimate the value of assets and the volatility with the iterative process described earlier, we need the market value of equity and the (annualized) volatility of equity return. For the first, we take the stock price at the fiscal year-end and multiply it with the number of shares outstanding.

For the latter, we follow Campbell et al. (2008) by taking the standard deviation of stock returns for the last 30 days, prior to fiscal year-end. For the face value of debt (D), we follow convention in the literature and we set it equal with short-term debt plus half of long-term debt (also used in Vassalou and Xing, 2004 and Campbell et al., 2008). The prediction horizon is one, thus T for the BSM and t for the Leland-Toft models equal 1. Another input to the structural models, is the assets value returns (μ). Campbell et al. (2008) use an equity risk-premium equal to the riskless rate plus 6% for all firms. However, we believe that using a common return for all firms, would undermine the predictive ability of the structural models. A better alternative, would be to use a firm-specific return. A reasonable proxy, which we use in our study, is the annualized return of equity, also used by Bharath and Shumway (2008). Afik et al. (2016), instead, suggest using the maximum between the riskless rate of return and equity return. However, this specification would overstate the asset drift for firms with negative prospects (i.e. the bankrupt firms), undermining again the predictive ability of the structural models.

For the risk-free rate (r), the one-year Treasury Constant Maturity rate is used, obtained from Federal Reserve¹³. For the coupon payments (C) and debt principal (P), the interest expense and the short-term debt plus half of long-term debt are used as proxies respectively and the payout yield (δ) is defined as the sum of coupon payments plus dividends (ordinary and preferred) divided by the market value of assets. For corporate tax rate (τ), bankruptcy costs (c) and maturity of debt (T) we follow Leland (2004) who sets these parameters equal to 15%, 30% and 10 years respectively. However, as shown later and specifically in Appendix A, results are not sensitive with respect to different parameters choices and thus Leland-Toft is stable as far as its

¹³ Available at <http://www.federalreserve.gov/releases/h15/data.htm>

performance is concerned (refer to section 4 about model performance measures). All inputs of the models are summarized in Table 1.

4. Methodology

This section describes the methodology that is used to assess the performance of the bankruptcy prediction models.

4.1. Discriminatory Power

Discriminatory power refers to the ability of a model to discriminate bankrupt from healthy firms. The ROC curve is a graphical representation of the discriminatory power of a bankruptcy prediction model. It plots the true predictions on the vertical axis (the percentage of bankrupt firms correctly classified as bankrupt) against the false predictions on the horizontal axis (the percentage of healthy firms incorrectly classified as bankrupt) according to a pre-determined cut-off value. By performing this classification procedure for multiple cut-off values, we create a set of points which together constitute the ROC curve. Ideally, a perfect model will never make false predictions and will always correctly classify the bankrupt firms, for any level of cut-off point. Hence the ROC curve of a perfect model will pass through the point (0, 1) and in general, the closer the ROC curve towards the top-left corner of the graph, the better the discriminatory power is.

A quantitative assessment of the discriminatory power of a bankruptcy prediction model is the Area Under ROC (AUROC) curve (see for example Hanley and McNeil, 1982 and Sobehart and Keenan, 2001) and calculated as follows:

$$AUROC = \frac{1}{nm} \sum_{i=1}^n \sum_{j=1}^m \psi(p_B^i, p_H^j) \quad (15)$$

where

$$\psi(p_B^i, p_H^j) = \begin{cases} 1, & p_B^i > p_H^j \\ 0.5, & p_B^i = p_H^j \\ 0, & p_B^i < p_H^j \end{cases}$$

and p_B^i is the bankruptcy probability of the i -th bankrupt firm, p_H^j is the bankruptcy probability of the j -th healthy firm, n is the number of bankrupt firms and m is the number of healthy firms in our sample.

We test for statistically significant differences between the AUROCs of two models. The hypothesis is as follows:

$$H0: AUROC 1 - AUROC 2 = 0 \quad Vs \quad H1: AUROC 1 - AUROC 2 \neq 0$$

We use the non-parametric approach of DeLong et al. (1988), which accounts for the correlation of the AUROCs produced by the two models. The construction of the test statistic is described in Appendix B.

4.2. Logit Models

The logistic regression approach is used to estimate the models in Eqs. (9)-(14). Specifically, we estimate the following logit model:

$$P(Y_{i,t+1} = 1|X_{i,t}) = p_{i,t} = \frac{e^{a+\beta'X_{i,t}}}{1 + e^{a+\beta'X_{i,t}}} \quad (16)$$

where $p_{i,t}$ is the probability of bankruptcy at time “ t ”, that the “ i -th” firm will go bankrupt the next year, $Y_{i,t+1}$ denotes the status of the i -th firm (1 if it goes bankrupt at time $t+1$, 0 otherwise), $X_{i,t}$ is the vector of covariates of the i -th firm at time t , β is the vector of coefficient estimates and a is the constant term which expresses the bankruptcy risk in the absence of the covariates.

The model in Eq. (16) represents a multi-period logit model because it includes observations for each firm across time. However, the inclusion of multiple firm-year observations per firm yields understated standard errors because the log-likelihood objective function, which is maximized to estimate the multi-period logit model, assumes that each observation is independent from each other. This is a wrong assumption since financial information of a firm at time $t+1$ cannot be independent from the financial information of the same firm at t . Failing to address this econometric issue, leads to wrong inference regarding the significance of the individual coefficients. Similar with Filipe et al. (2016), we use clustered-robust standard errors to adjust for the number of firms in the sample but also for heteroskedasticity (Huber, 1967 and White, 1980).

To compare the predictive accuracy of various logit models, we test the difference between their log-likelihoods. Hence, the hypothesis takes the following general form:

$$H_0: L_1(k_1) - L_2(k_2) = 0 \quad V_s \quad H_1: L_1(k_1) - L_2(k_2) \neq 0$$

where $L_1(k_1)$ is the log-likelihood of the first model with k_1 parameters and $L_2(k_2)$, is the log-likelihood of the second model, with k_2 parameters and $k_1 > k_2$. The construction of the test statistic for different types of logit models can be found in Appendix C.

4.3. *Economic Analysis of Bankruptcy Models*

The analysis so far addressed the accuracy of our bankruptcy models. But how accurate is economically beneficial for banks? Here, we follow the approach of Agarwal and Taffler (2008) to examine it by assuming a loan market worth \$100 billion and banks competing for granting loans to individual firms. Each bank uses one of our bankruptcy models to evaluate the credit worthiness of their customers.

4.3.1. Calculating Credit Spreads

We estimate the models using data spanning the years 1995-2005 (70% of the sample). We sort firm-customers from this sample in 10 groups of equal size and a credit spread is calculated according to the following rule; Firms in the first group, which are firms with the lowest bankruptcy risk, are given a credit spread, k , and firms in the remaining groups are given a credit spread, CS_i , obtained from Stein (2005) and Blochlinger and Leippold (2006) and it is defined as follows:

$$CS_i = \frac{p(Y = 1|S = i)}{p(Y = 0|S = i)} LGD + k \quad (17)$$

where $p(Y=1/S=i)$ and $p(Y=0/S=i)$, is the average probability of bankruptcy and non-bankruptcy for the i -th group, with $i=2, 3, \dots, 10$ and LGD is the loan loss upon default. Following Agarwal and Taffler (2008), the average probability of bankruptcy for the i -th group, is the actual bankruptcy rate for that group, defined as the number of firms that went bankrupt the following year divided by the number of firms in the group. Furthermore, $k=0.3\%$ and $LGD=45\%$.

4.3.2. Granting Loans and Measuring Economic Performance

To evaluate economic performance, we assume that banks compete to grant loans to prospective firm-customers between the period 2006-2014. Each bank, uses one of our bankruptcy models which have been estimated in the period 1995-2005. The bank sorts those customers according to their riskiness and rejects the bottom 5% with highest risk. The remaining firms are classified in 10 groups of equal size and firms from each group are charged a credit spread that has been obtained from the period 1995-2005. Finally, for each customer the bank that charges the lowest credit spread is granting the loan. Two measures of profitability are

used. The first one, Return on Assets (ROA) is defined as Profits/Assets lent and the second one, Return on Risk-Weighted Assets (RORWA) takes into consideration the riskiness of the assets, defined as Profits/Risk-Weighted Assets. Risk-Weighted Assets are obtained from formulas provided by the Basel Committee on Banking Supervision (2006)¹⁴.

5. Results

5.1. Descriptive Statistics

Table 4 reports mean values of the explanatory variables for the entire sample 1995-2014. As expected, the performance of bankrupt firms is worse than the performance of healthy firms, one year prior to bankruptcy with the differences in mean values being statistically significant in most cases. For example, bankrupt firms; are less profitable (*EBITTA* is lower), have more leverage (*TLTA* is higher), their liquidity is more constrained (*WCTA* is lower) etc. In terms of the market performance, stock price of bankrupt firms is significantly more volatile than healthy firms (*SIGMA* is higher), they underperform the market (*EXRET* is lower) as well stock price is lower (*PRICE* is lower). Our variable of interest, *LT*, is higher for bankrupt firms relative to healthy firms

[Insert Table 4 here]

5.2. Reduced-Form and Hybrid Models Estimation

Table 5 reports estimation results for our models when applying the logistic regression approach on our data.

[Insert Table 5 here]

¹⁴ See for instance the Appendix in Bauer and Agarwal (2014)

Here, the estimation sample spans the years between 1995 and 2005 which accounts for approximately 70% of our sample. Firstly, most of Ohlson and CHS variables are significant and with the correct sign. Noticeable exception is the case of *SIGMA* where in previous studies (Campbell et al., 2008 and Bauer and Agarwal, 2014) was found significant. Based on an analysis we have performed, we conclude that in our case the interaction of *SIGMA* with other market variables in the CHS model is the main cause for this kind of behavior. For example, when we include *SIGMA* individually or in the Ohlson model, is statistically significant and with the correct (positive) sign. Furthermore, in the estimation sample, average *SIGMA* (not tabulated) for bankrupt firms is 1.21 while for healthy firms is 0.60, which excludes possible data collection error. Another relatively odd estimation result is the positive coefficient for the *SIZE* variable¹⁵. According to Galil and Gilat (2018), a positive sign of this variable may hint on a selection bias in the bankruptcy sample toward larger corporates. However, this is not the case with our sample. In the estimation sample (whole sample), average total assets for bankrupt firms are 679.1 million (662.5 million), while for the healthy group, average total assets are 2074.5 million (3455.3 million). We believe that its interaction with other variables is the main cause for this result (including *SIZE* only in a logistic regression, yields a statistically significant coefficient with the correct (negative) sign.

Secondly, most of Altman variables included in AAN are not significant, consistent with Hillegeist et al. (2004). The cash flow ratio enters significantly and with the correct sign. Finally, the predictor of interest, LT, is highly statistically significant ($\alpha=1\%$) when incorporated in Altman (ALT), Ohlson (OLT), and Campbell (CHSLT) models.

¹⁵ Hillegeist et al. (2004) also find a positive *SIZE* coefficient in the Ohlson model

5.3. Evaluating Leland-Toft and BSM (Hypothesis 1)

First, we compare the performance of the two structural models. For consistency, we estimate two logit models; The first includes the probability of bankruptcy derived from Leland-Toft in the period 1995-2005 as predictor and the second includes the probability of bankruptcy derived from BSM. Using these models, we forecast bankruptcies in the out-of-sample period, 2006-2014¹⁶. The performance is reported in Table 6.

[Insert Table 6 here]

We find that AUROC of LT is 0.8941 while for BSM is 0.8659, indicating that LT has better out-of-sample discriminating ability. Using the Delong test we find that the difference is statistically significant at $\alpha=1\%$ (test statistic=2.74). Moreover, in Appendix A we perform a sensitivity analysis test to examine whether the AUROC of LT is affected by deviations in the choice of parameter values. In all, results suggest that LT is not sensitive as far as the ordinal ranking (AUROC) is concerned. Further, LT model explains bankruptcy variation better than BSM, according to pseudo- R^2 (19.72% and 17.90% respectively), although differences in their log-likelihoods are not significant (test statistic is 1.38).

However, as it is evident from Table 6, neither LT, nor BSM are sufficient statistics to forecast bankruptcies, since they are outperformed by other models such as Ohlson and CHS (differences in AUROCs and log-likelihoods are significant)¹⁷. Thus, none of the structural models can stand alone.

¹⁶ This adjustment in the bankruptcy probability derived from structural models through a logit regression, is usually referred to as calibration.

¹⁷ This result is also evident by the regression results in table 5, since LT enters significantly in Ohlson, Altman and CHS along with other variables, suggesting that individually, it doesn't capture all the bankruptcy-related information.

Finally, we perform an analysis of the economic benefits when banks use either LT or BSM in evaluating the credit worthiness of prospective customers. We make the paradigm more challenging by using Altman's model as a benchmark. Hence, we assume there are three banks competing for loans; bank 1 uses LT, bank 2 uses BSM and bank 3 uses Altman's model. The results are reported in table 7:

[Insert

Table 7 here]

As can be inferred, the quality of the loan portfolio for bank 1 which uses LT is the best among the three banks, since there are only 10 bankruptcies (0.25%), whereas there are 37 and 38 bankruptcies in the portfolios of banks 2 and 3 respectively, corresponding to 0.79% and 0.76% bankruptcy rate. Most importantly however, is that bank 1 generates superior economic performance relative to its competitors. For instance, on a risk-adjusted basis, bank 1 yields 1.09% return on the capital it has invested while bank 2 generates 0.64% and the return for bank 3 is 0.54%¹⁸.

From the analysis in this section, we conclude that a bank has more gain by using LT rather than BSM which in fact, lends support to our first hypothesis, indicating that LT is a better approach than BSM, due to the richer information set incorporated in LT.

5.4. Reduced-Form and Hybrid Model Performance (Hypotheses 2-4)

In this section, we test the performance of the models using three out-of-sample approaches, as well as the economic benefits when banks adopt the models in a competitive loan market, as outlined below.

¹⁸ Results are robust with respect to various specifications of LGD (0.4-0.7) but k as well (0.002-0.004)

5.4.1. Baseline Approach

Here, we use the models (as estimated in table 5) to forecast bankruptcies in the out-of-sample period which spans the years between 2006 and 2014. Results are reported in table 6.

Panel A reports the performance of the models while panels B and C test for differences in their discriminating ability and predictive accuracy respectively. Firstly, OLT performs better than Ohlson (AUROCs are 0.9449 vs 0.9252 and log-likelihoods are -483.43 vs -535.57). The differences are statistically significant (test statistics are 4.73 and 104.28 respectively) which lends support to our second hypothesis, that extending Ohlson with LT yields improved performance. Secondly, ALT outperforms AAN (AUROCs are 0.9207 vs 0.8597 and log-likelihoods are -519.56 vs -603.81). The differences are statistically significant (test statistics are 4.78 and 6.51 respectively) which is in line with our third hypothesis. That is, augmenting Altman's model with LT, further improves performance relative to a cash flow variable as suggested by Almamy et al. (2016). Including LT in CHS slightly improves discriminating ability (AUROCs are 0.9395 for CHSLT and 0.9332 for CHS), though their differences are not statistically significant (test statistic is 1.47). On the other hand, LT carries incremental information as indicated by their log-likelihoods, meaning that it is a missing variable for the model (log-likelihoods for CHSLT and CHS are -491.41 and -498.85 respectively), with the difference being statistically significant (test statistic is 14.88). Thus, we provide evidence regarding the fourth hypothesis, that augmenting CHS with LT yields improved performance. We complement the aforementioned results, with ROC curves provided in figure 1.

[Insert Figure 1 here]

A related performance statistic with AUROC is the partial AUROC (pAUROC)¹⁹, which is based on a specific region of the area under ROC curve that might be of practical interest (see for instance Dodd and Pepe, 2003)²⁰. Panel A in table 6 reports pAUROCs for the models. We have also tested for differences in pAUROCs, but we do not report the results to save space. Overall, the conclusions are similar with the case of AUROC, giving validity to hypotheses 1-4.

5.4.2. Rolling Window Approach

As a second way to test the models, we re-estimate them yearly based on a rolling window. For instance, we estimate the models using firms between 1995 and 2005 and apply them on firms in 2006. Then we re-estimate the models using firms between 1996 and 2006 and apply them on firms in 2007. This process is repeated until 2014 and we aggregate bankruptcy probabilities obtained from each year to measure the performance of the models. It should be noted that this approach should be used in practice because the models are updated more frequently as new information becomes available. Results are reported in table 8.

[Insert Table 8 here]

We obtain similar results as before and summarized as follows: 1) Augmenting Ohlson with LT yields a model with improved performance as indicated by AUROC statistics (0.9469 for OLT vs 0.9289 for Ohlson) as well as log-likelihood statistics (-470.37 for OLT vs -522.87 for Ohlson). Differences are statistically significant (test statistics are 4.33 and 104.99 for the two performance tests respectively) which is consistent with hypothesis 2, 2) Incorporating LT in Altman's model further improves its performance as opposed to a cash flow variable (AUROC

¹⁹ We thank an anonymous referee for this suggestion.

²⁰ The selection of the partial region under the ROC curve, however, is subjective. We rely on STATA's default specifications to estimate the pAUROCs of the models and to test for significant differences. Specifically, we use the region of the AUROC that starts from false positive rate equal 0 and ends at false positive rate equal 0.5.

for ALT is 0.9253 vs 0.8673 for AAN, while log-likelihoods are -508.25 and -593.86 respectively). Differences in their performance are significant (test statistics for the two tests are 4.62 and 6.28 respectively) which provides evidence to support our third hypothesis and 3) LT incorporates information not included in CHS (log-likelihood for CHSLT is -479.66 while for CHS is -485.94, and test statistic is 12.57) which is in line with the fourth hypothesis. AUROC improvement is not significant.

5.4.3. Five Folds Approach

Here, we divide the whole sample period in five approximately equal-sized sub-samples. We use any four of them to estimate the models and apply them on firms in the left-out sample. This is to break the chronological order of the data, and to consider different periods as well, such as periods before the financial crisis period. Then we aggregate bankruptcy probabilities from each left-out sample to obtain single performance measures. Results are reported in table 9.

[Insert Table 9 here]

As expected, performance according to this approach is lower since we use data from different periods to make predictions, missing therefore potential trends. Despite this fact, we obtain similar insights as with the two previous approaches. Discriminating ability is improved when we include LT as additional predictor in Ohlson (AUROCs are 0.9091 for OLT and 0.8939 for Ohlson) while predictive accuracy is also better (log-likelihoods are -1379.48 for OLT and -1477.06 for Ohlson). Differences are statistically significant (test statistics are 4.56 and 195.16 for the two tests respectively) which is consistent with our second hypothesis regarding the superiority of this extended version of Ohlson's model. Next, ALT outperforms AAN as indicated by AUROC statistics (0.8826 vs 0.8438 respectively) as well as log-likelihood statistics

(-1471.35 vs -1630.45 respectively), meaning that LT further improves performance when included in Altman's model as opposed to a cash flow variable, supporting therefore our third hypothesis. Differences in performance are significant (test statistics are 5.63 and 6.62 for the two tests respectively). Finally, as with the previous tests we find evidence that LT improves the performance of CHS. Specifically, log-likelihood for CHSLT is -1403.72 while for CHS is -1417.16. Difference is statistically significant (test statistic is 26.87) while discriminating ability, measured by AUROC, is only slightly improved.

5.4.4. Economic Benefits

So far, we have considered aspects of model performance such as discriminating ability, measured by AUROC, and empirical fit, measured by log-likelihood statistics. However, a bank is more interested in the economic benefits when using bankruptcy models in the decision-making process of granting loans to individual firms. Here, we show the case of five banks, where bank 1 uses OLT, bank 2 uses CHSLT, banks 3 and 4 use CHS and Ohlson respectively, whereas bank 5 uses a benchmark model such as Altman's model. Table 10 reports information regarding the economic results of these banks.

[Insert Table 10 here]

Clearly, the quality of the loan portfolio for banks 1 and 2 which use OLT and CHSLT respectively, is superior relative to that of banks 3 and 4 which use CHS and Ohlson respectively. This is evident by the lower concentration of bankruptcies in their portfolios (0.11% for bank 1 and 0.16% for bank 2) relative to other banks (0.44% for bank 3 and 0.58% for bank 4).

The most important result, however, is that banks 1 and 2 earn higher returns than the other banks on a risk-adjusted basis (i.e. after adjusting for the riskiness of the assets lent). For instance, for each dollar invested, banks 1 and 2 earn 1.74% and 1.54% respectively on a risk-adjusted basis, whereas the competing banks earn lower returns (1.12% for bank 3 and 1.02% for bank 4). Bank 5 which uses a generic bankruptcy model earns negative returns. It is also worth noting that differences in discriminating ability that we have not found to be statistically significant are reflected in the economic results²¹. For instance, both banks 1 and 2 that use OLT and CHSLT respectively, are more profitable than bank 3 which uses CHS, although there are no significant differences in their AUROCs, as reported in table 6. Thus, banks should take into consideration the economic benefits when judging what bankruptcy model to use.

Based on these results²², we conclude that banks using OLT and CHSLT, can achieve superior economic performance relative to other banks that use, for instance, CHS or Ohlson.

5.5. Augmenting CHS, Ohlson and Altman with LT and BSM (Hypotheses 2a-4a)

Campbell et al. (2008) show that augmenting their model with BSM doesn't yield improved performance based on pseudo-R², indicating that all the information incorporated in BSM is already included in their model. Here, we re-examine this insight and compare it with the case of LT. Table 11 reports the results.

[Insert Table 11 here]

Indeed, augmenting CHS with BSM doesn't improve performance, since volatility and return measures are already included in the model (log-likelihood and pseudo-R² are the same when

²¹ Bauer and Agarwal (2014) also reported that very small differences in AUROCs are shown up in the economic results

²² Results are robust with respect to different parameter specifications, for example, setting LGD = 0.4-0.7 and k=0.002-0.004

compared to CHS. Differences in model performance are not statistically significant according to test-statistics reported in panel A and B. However, this is not the case when we include LT. Specifically, pseudo- R^2 increases to 34.35% as well as log-likelihood (difference with CHS is statistically significant at $\alpha=1\%$), indicating that LT provides additional information not included in CHS. As expected though, we document increase in performance when augmenting Ohlson and Altman with LT and BSM, since the two reduced-form models do not incorporate market information (all test statistics for performance difference are significant at $\alpha=1\%$).

Finally, we find that hybrid models with LT outperform hybrid models with BSM. For example, log-likelihood for CHS augmented with LT is -491.41 while for CHS augmented with BSM is -497.96 (Vuong's test statistic is significant at $\alpha=5\%$). AUROC for the first, is slightly higher (test statistic is not significant). Similar is the case with Ohlson and Altman model. Augmenting these models with LT yields models with better performance relative to augmenting them with BSM (DeLong's test for AUROCs and Vuong's test for log-likelihoods are significant at $\alpha=5\%$). In all, we find evidence in favor of hypotheses 2a-4a; hybrid models that incorporate LT, outperform those with BSM.

5.6. Time Robustness

As a final test, we estimate the reduced-form and hybrid models in the period 2006-2014 and forecast bankruptcies in the period 1995-2005. This is to test the performance of the models in a completely new sample, since the previous tests, included data from the recent period to measure performance i.e. 2006-2014. Results are reported in table 12.

[Insert Table 12 here]

In all, the results support the superiority of the hybrid models over the reduced form models, suggesting that the LT is a significant addition to the models²³. For instance, OLT outperforms Ohlson (differences in AUROCs and log-likelihoods are both significant at $\alpha=5\%$ and $\alpha=1\%$ respectively), CHSLT outperforms CHS (differences in log-likelihoods are significant at $\alpha=1\%$) and finally, ALT outperforms AAN (differences in AUROCs and log-likelihoods are significant at $\alpha=1\%$).

6. Summary and Conclusions

In this paper, we examine an alternative approach for bankruptcy prediction that is based on Leland and Toft (1996), which is a model that measures the financial healthiness of firms with coupon-paying debts. This model is an extension of a model widely used in the literature; the BSM model. The Leland-Toft (LT), however, incorporates information not captured by BSM and thus it should be a better one. Based on several tests, we find evidence suggesting that it is a better approach in terms of discriminatory power, predictive accuracy but also in terms of economic performance when a bank implements LT relative to BSM.

Next, we use the probability of bankruptcy derived from LT as additional predictor to extend two widely-used corporate bankruptcy models (Altman, 1968 and Ohlson, 1980) but also, a contemporary model which was found to outperform other approaches for bankruptcy prediction (Campbell et al., 2008). Our objective is to develop powerful models that are practical and easy to implement. Under a comprehensive out-of-sample analysis, we find that augmenting Altman's model with LT further improves its performance as opposed to a cash flow ratio, as suggested by Almamy et al. (2016). The most powerful models, however, are obtained when we augment

²³ We have also performed our test for the economic benefits, and we find that the banks using the hybrid models achieve higher returns relative to banks using the reduced-form models.

Ohlson and CHS, with LT. Further, the models augmented with LT outperform the models augmented with BSM.

However, banks are more interested in the economic performance of such models. Based on our final test we find that banks using OLT and CHSLT earn higher returns than banks which implement other models to evaluate firm-customers in a competitive loan market. We therefore recommend the use of those augmented models as an appropriate risk management tool, that could be economically beneficial for banks.

Future work should emphasize the estimation of bankruptcy costs and debt maturity separately for each firm, rather than using average values, which we think will increase the accuracy of LT and its contribution to OLT and CHSLT.

Appendix A: Sensitivity Analysis of Leland-Toft

Consider the following three vectors with different parameter values for tax rate, τ , debt maturity, T and bankruptcy costs, c :

$$\tau = \{0.15, 0.20, 0.25, 0.30, 0.35, 0.40\}$$

$$T = \{8, 10, 12, 14, 16, 18, 20\}$$

$$c = \{0.15, 0.20, 0.25, 0.30, 0.35, 0.40\}$$

Each scenario has as input the triplet $\{\tau_i, T_j, c_k\}$ where $i=k=1, \dots, 6$ and $j=1, \dots, 7$. For each input scenario, AUROC of Leland-Toft is obtained and a histogram is constructed for the 252 scenarios, as shown in the following figure.

[Insert Figure A.1 here]

As can be inferred, discriminating ability measured by AUROC, is not sensitive at all with respect to the different scenarios since it ranges between 0.8828 to 0.8848, with an average value of 0.8840. Thus, performance is not affected significantly by deviations in the choice of parameter values.

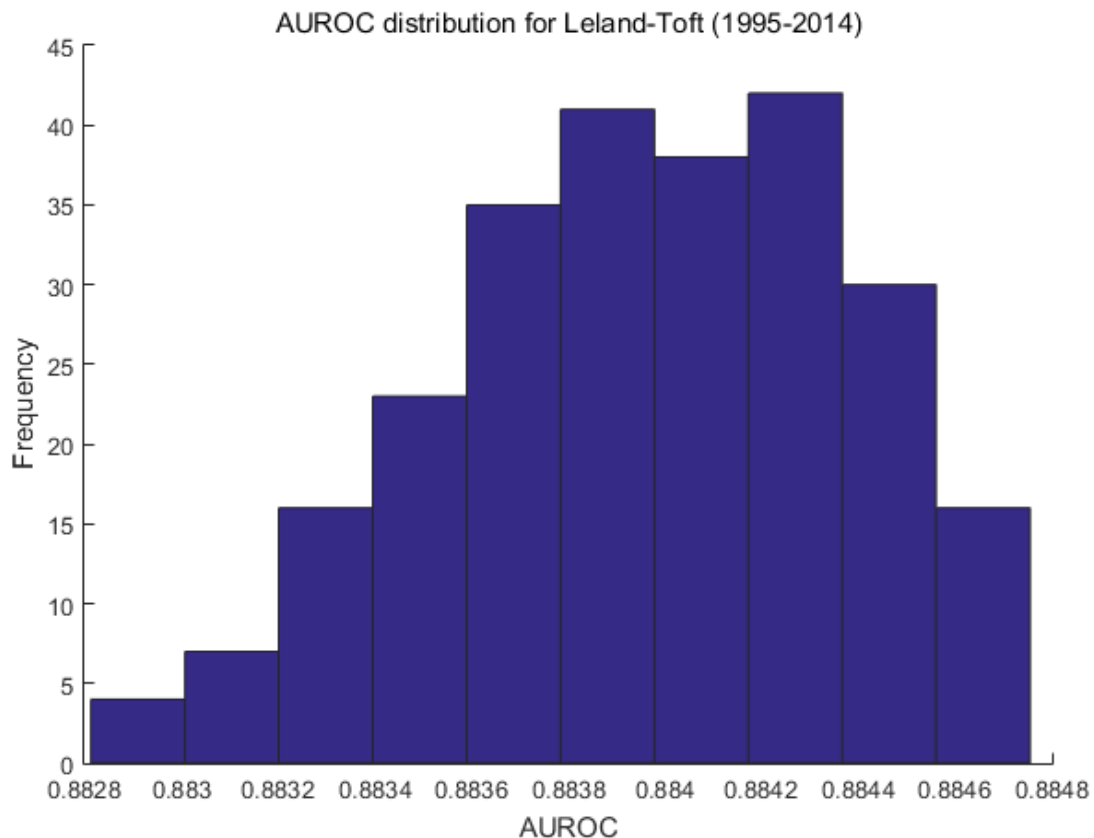


Figure A.1: This figure shows the distribution of AUROC produced by Leland-Toft model under different scenarios of its input parameters τ (the tax rate), T (debt maturity) and c (bankruptcy costs) during the period 1995-2014.

Appendix B: Discriminatory Power Test Statistic

The key element for the estimation of the test statistic is the covariance matrix of the AUROCs produced by the two models. Following DeLong et al. (1988), the covariance matrix is estimated as follows:

1) For each bankrupt firm calculate the AUROC:

$$\widehat{AUROC}(p_B^i) = \frac{1}{m} \sum_{j=1}^m \psi(p_B^i, p_H^j), \quad (i = 1, 2, \dots, n) \quad (\text{B.1})$$

2) For each healthy firm calculate the AUROC:

$$\widehat{AUROC}(p_H^j) = \frac{1}{n} \sum_{i=1}^n \psi(p_B^i, p_H^j), \quad (j = 1, 2, \dots, m) \quad (\text{B.2})$$

3) Define the 2x2 symmetric matrix S_{10} with $(k,r)^{th}$ element defined as:

$$s_{10}^{k,r} = \frac{1}{n-1} \sum_{i=1}^n [\widehat{AUROC}_k(p_B^i) - \widehat{AUROC}_k][\widehat{AUROC}_r(p_B^i) - \widehat{AUROC}_r] \quad (\text{B.3})$$

4) Define the 2x2 symmetric matrix S_{01} with $(k,r)^{th}$ element defined as:

$$s_{01}^{k,r} = \frac{1}{m-1} \sum_{j=1}^m [\widehat{AUROC}_k(p_H^j) - \widehat{AUROC}_k][\widehat{AUROC}_r(p_H^j) - \widehat{AUROC}_r] \quad (\text{B.4})$$

5) Then the covariance matrix of the two AUROCs is defined as:

$$S = \frac{1}{n} S_{10} + \frac{1}{m} S_{01} \quad (\text{B.5})$$

Finally, the z-statistic which is standard-normally distributed is calculated as follows:

$$z = \frac{\widehat{AUROC}_1 - \widehat{AUROC}_2}{(s^{1,1} - 2s^{1,2} + s^{2,2})^{1/2}} \quad (\text{B.6})$$

with $s^{1,1}$ and $s^{2,2}$ being the variances of AUROCs of the two models under comparison and $s^{1,2}$ their covariance, all obtained from Eq. (B.5).

Appendix C: Predictive Accuracy Test Statistic

There are two distinct types of logit models; non-nested and nested models. In the case of non-nested models where the k_2 parameters in model 2 are not subset of the k_1 parameters in model 1, the Vuong (1989) test is used. The z-statistic in this case is standard-normally distributed and it is defined as follows:

$$z = \frac{2(L_1 - L_2) - (k_1 - k_2)\ln(N)}{2\sqrt{N}\omega_N} \quad (\text{C.1})$$

Here, N is the number of observations and ω_N is the sample standard deviation of the individual log-likelihoods produced by each model, l_i , which is defined as follows:

$$l_i = \ln \left[\frac{p_{1,i}^{y_i}(1 - p_{1,i})^{(1-y_i)}}{p_{2,i}^{y_i}(1 - p_{2,i})^{(1-y_i)}} \right] \quad (\text{C.2})$$

where $p_{1,i}$ and $p_{2,i}$ are the bankruptcy probabilities for the i -th firm produced by models 1 and 2 respectively (time index “ t ” is dropped for simplicity). Furthermore, y_i indicates whether the firm is bankrupt ($y_i = 1$) or healthy ($y_i = 0$). Rejection of the null hypothesis indicates significant difference between the predictive accuracy of the two models.

On the other hand, to compare predictive accuracy between nested models where the k_2 parameters in model 2 are subset of the k_1 parameters in model 1, the standard Likelihood Ratio (LR) test is used. The statistic in that case is the following:

$$LR = -2[L(k_2) - L(k_1)] \quad (\text{C.3})$$

and follows a chi-squared distribution with $k_1 - k_2$ degrees of freedom. Rejection of the null hypothesis indicates that predictive accuracy of the two models is not equivalent. Therefore, at

least one of the extra $k_1 - k_2$ parameters in model 1 carry additional explanatory power about bankruptcy risk.

References

- Afik, Z., Arad, O., & Galil, K. (2016). Using Merton model for default prediction: An empirical assessment of selected alternatives. *Journal of Empirical Finance*, 35, 43-67.
- Agarwal, V., & Taffler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking and Finance*, 32, 1541-1551.
- Almamy, J., Aston, J., & Ngwa, L. N. (2016). An evaluation of Altman's z-score using cash flow ratio to predict corporate failure amid the recent financial crisis: Evidence from the UK. *Journal of Corporate Finance*, 36, 278-285.
- Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23, 589-609.
- Bauer, J., & Agarwal, V. (2014). Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. *Journal of Banking and Finance*, 40, 432-442.
- Basel Committee on Banking Supervision. (2006). International convergence of capital measurement and capital standards: A revised framework.
- Bharath, S. T., & Shumway, T. (2008). Forecasting Default with Merton Distance to Default Model. *The Review of Financial Studies*, 21, 1339-1369.
- Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81, 637-654.
- Blochlinger, A., & Leippold, M. (2006). Economic benefit of powerful credit scoring. *Journal of Banking and Finance*, 30, 851-873.
- Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. *The Journal of Finance*, 63, 2899-2939.
- Charitou, A., Dionysiou, D., Lambertides, N., & Trigeorgis, L. (2013). Alternative Bankruptcy Prediction Models Using Option-Pricing Theory. *Journal of Banking and Finance*, 37, 2329-2341.
- Chava, S., & Jarrow, R. A. (2004). Bankruptcy prediction with industry effects. *Review of Finance*, 8, 537-569.

- Delong, R., Delong, M., & Clarke-Pearson, L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A non-parametric approach. *Biometrics*, 44, 837-845.
- Dodd, L. E., & Pepe, M. S. (2003). Partial AUC estimation and regression. *Biometrics*, 59, 614-623.
- Eom, Y.H., Helwege, J., & Huang, Z.J. (2004). Structural models of corporate bond pricing: An empirical analysis. *The Review of Financial Studies*, 17, 499-544.
- Filipe, S. F., Grammatikos, T., & Michala, D. (2016). Forecasting distress in European SME portfolios. *Journal of Banking and Finance*, 64, 112-135.
- Galil, K., & Gilat, N. (2008). Predicting default more accurately: To proxy or not to proxy for default? Available at SSRN: <https://ssrn.com/abstract=2618190>.
- Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristics (ROC) curve. *Radiology*, 143, 29-36.
- Hilberink, B., & Rogers, L.C.G. (2002). Optimal capital structure and endogenous default. *Finance and Stochastics*, 6, 237-263.
- Hillegeist, S. A., Keating, E. K., Cram, D. P., & Lundstedt, K. G. (2004). Assessing the probability of bankruptcy. *The Review of Financial Studies*, 9, 5-34.
- Huber, P. J. (1967). The behavior of maximum likelihood estimates under non-standard conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 221-233.
- Leland, H. (2004). Predictions of default probabilities in structural models of debt. *Journal of Investment Management*, 2, 5-20.
- Leland, H., & Toft, K. B. (1996). Optimal capital structure, endogenous bankruptcy and the term structure of credit spreads. *Journal of Finance*, 51, 987-1019.
- Lyden, S., & Saraniti, D. (2001). *An Empirical Examination of the Classical Theory of Corporate Security Valuation*. Available at SSRN: <http://papers.ssrn.com/sol3/papers.cfm?19>.
- Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29, 449-470.
- Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18, 109-131.

- Papakyriakou, P., Sakkas, A., & Taoushianis, Z. (2019). Financial firm bankruptcies, international stock markets, and investor sentiment. *International Journal of Finance and Economics*, 24, 461-473.
- Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business*, 74, 101-124.
- Sobehart, J., & Keenan, S. (2001). Measuring default accurately. *Risk*, 31-33.
- Stein, R. M. (2005). The relationship between default prediction and lending profits: Integrating ROC analysis and loan pricing. *Journal of Banking and Finance*, 29, 1213-1236.
- Tinoco, M. H., & Wilson, N. (2013). Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. *International Review of Financial Analysis*, 30, 394-419.
- Vassalou, M., & Xing, Y. (2004). Default risk in equity returns. *Journal of Finance*, 59, 831-868.
- Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested Hypotheses. *Econometrica*, 57, 307-333.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48, 817-838.

Tables

Table 1: Model definition and estimation of the variables

	Variable	Estimation
AAN (2016)	WCTA	Working capital/total assets
	RETA	Retained earnings/total assets
	EBITTA	Earnings before interests and taxes/total assets
	MVTL	Market value of equity/total liabilities
	SLTA	Net sales/total assets
	CFOTA	Cash flows from operations/total assets
Ohlson (1980)	SIZE	Log (Total assets/GNP price level index)
	TLTA	Total liabilities/total assets
	CLCA	Current liabilities/current assets
	D(TL>TA)	1 if TL>TA and 0 otherwise, where TL are total liabilities and TA are total assets
	NITA	Net income/total assets
	CFOTL	Cash flows from operations/total liabilities
	D(NI _t +NI _{t-1} <0)	1 if the cumulative net income in two consecutive years is negative and 0 otherwise
	WCTA	Working Capital/total assets
	CHNI	$(NI_t - NI_{t-1}) / (NI_t + NI_{t-1})$, is the change in net income (takes values between -1 and 1)
CHS (2008)	NIMTA	Net income/market value of assets, where market value of assets is the sum of market value of equity and total liabilities
	TLMTA	Total liabilities/market values of assets
	EXRET	Log excess return of each firm's equity relative to the S&P 500 index, over the previous 12 months
	SIGMA	Annualized standard deviation of daily stock returns, over the previous 3 months
	RSIZE	Relative size, defined as Market value of equity/market value of S&P 500 index
	CASHMTA	Cash and short-term investments/market value of assets
	MB	Market value of equity/book value of equity
	PRICE	Log (stock price)

	σ	Annualized volatility of asset returns, obtained by solving Eqs. (7) and (8)
	μ	Annualized return on assets, proxied by the annualized return on equity
	V	Market value of assets, obtained by solving Eqs. (7) and (8)
	r	Riskless rate of return, proxied by the one-year treasury constant maturity rate
	C	Coupon payments, proxied by the interest expenses in the income statement
	P	Principal value of debt, proxied by short-term debt plus half of long-term debt
Leland and Toft (1996)	δ	Payout yield, which is the sum of coupons and dividends (ordinary and preferred) divided by the market value of assets
	τ	Corporate tax rate, 15% as in Leland (2004)
	c	Bankruptcy costs, 30% as in Leland (2004)
	T	Maturity, 10 years as in Leland (2004)

This table describes the input variables of four models: Almamy et al. (2016), denoted as AAN, Ohlson (19980), Campbell et al. (2008), denoted as CHS and Leland and Toft (1996). All variables are constructed using financial and market information one year prior to bankruptcy filing. The second column shows how variables are entered in the models and the third column shows how they are calculated.

Table 2: Distribution of observations per year

Bankruptcy year	Bankrupt firms	Healthy firms	Bankruptcy rate
1995	15	2749	0.543
1996	16	2804	0.567
1997	13	2933	0.441
1998	21	2186	0.952
1999	31	2045	1.493
2000	21	2572	0.810
2001	23	2425	0.940
2002	15	2206	0.675
2003	18	2045	0.873
2004	13	1919	0.673
2005	15	1865	0.798
2006	10	1796	0.554
2007	15	1738	0.856
2008	20	1661	1.190
2009	34	1560	2.133
2010	7	1508	0.462
2011	9	1431	0.625
2012	13	1388	0.928
2013	12	1353	0.879
2014	12	1313	0.906
Total	333	39497	0.836

This table reports the number of observations across the years 1995-2014. The first column shows the year of bankruptcy, the second and third columns show the number of bankrupt and healthy firms respectively and the last column shows the annual bankruptcy rate defined as bankrupt firms / (bankrupt firms + healthy firms).

Table 3: Distribution of observations per industry

Industry	Number of observations	Percentage
Agriculture, Forestry and Fishing	138	0.35
Mining	2340	5.87
Construction	379	0.95
Manufacturing	21109	53.00
Transportation, Communications, Electric, Gas, and Sanitary Services	4126	10.36
Wholesale Trade	1604	4.03
Retail Trade	3349	8.41
Services	6539	16.42
Public Administration	246	0.62

This table shows the distribution of observations per industry. Each observation is classified to an industry, according to SIC codes. Column 2 shows the number of observations that belong to each industry and column 3 shows the percentage of sample belonging to each industry calculated as industry observations / total observations.

Table 4: Summary statistics

Variable	Mean values		t-statistic for differences
	Bankruptcies	Non-Bankruptcies	
SIZE	0.410	1.391	8.56
TLTA	0.888	0.523	26.00
WCTA	-0.009	0.227	18.77
CLCA	1.389	0.658	25.75
D(TL>TA)	0.312	0.034	27.32
NITA	-0.414	-0.024	31.99
CFOTL	-0.302	0.102	13.08
D (NI _t +NI _{t-1} <0)	0.913	0.322	23.06
CHINI	-0.260	0.013	8.83
RETA	-1.736	-0.188	22.37
EBITTA	-0.238	0.029	25.03
MVTL	3.103	38.257	3.99
SLTA	1.250	1.172	1.72
CFOTA	-0.147	0.047	21.67
NIMTA	-0.270	-0.008	36.02
TLMTA	0.694	0.412	19.12
EXRET	-0.909	-0.056	26.79
SIGMA	1.219	0.609	26.01
RSIZE	-12.933	-10.820	18.46
CASHMTA	0.075	0.097	2.85
MB	1.459	1.416	0.24
PRICE	0.496	2.422	28.93
LT	0.441	0.042	41.46

This table reports mean value differences for the explanatory variables, between the bankrupt and non-bankrupt firms and t-tests for the statistical significance of the differences. The definition of variables can be found in Table 1.

Table 5: Reduced-form and hybrid models estimation, 1995-2005

Ohlson										
Constant	SIZE	TLTA	WCTA	CLCA	D(TL>TA)	NITA	CFOTL	D (NI _t +NI _{t-1} <0)	CHINI	
-7.280***	0.093***	1.461***	-1.997***	0.040	-0.047	-0.451*	-0.416***	2.120***	-0.545***	
(0.368)	(0.035)	(0.391)	(0.553)	(0.155)	(0.361)	(0.275)	(0.130)	(0.253)	(0.124)	
AAN										
Constant	WCTA	RETA	EBITTA	MVTL	SLTA	CFOTA				
-4.470***	-3.628***	0.084	-1.420***	-0.015	0.119*	-1.390***				
(0.136)	(0.363)	(0.057)	(0.474)	(0.012)	(0.066)	(0.563)				
CHS										
Constant	NIMTA	TLMTA	EXRET	SIGMA	RSIZE	CASHMTA	MB	PRICE		
-3.916***	-3.206***	2.566***	-0.545***	-0.073	0.140***	-4.186***	-0.00	-0.531***		
(0.604)	(0.331)	(0.333)	(0.135)	(0.220)	(0.047)	(0.917)	(0.015)	(0.105)		
ALT										
Constant	WCTA	RETA	EBITTA	MVTL	SLTA	LT				
-5.222***	-2.485***	0.004	-1.955***	-0.008	0.157**	2.520***				
(0.148)	(0.331)	(0.055)	(0.316)	(0.007)	(0.067)	(0.181)				
CHSLT										
Constant	NIMTA	TLMTA	EXRET	SIGMA	RSIZE	CASHMTA	MB	PRICE	LT	
-3.528	-3.129***	1.955***	-0.30***	-0.150	0.146***	-4.195***	-0.002	-0.529***	1.174***	
(0.607)	(0.342)	(0.391)	(0.143)	(0.212)	(0.047)	(0.890)	(0.015)	(0.104)	(0.298)	
OLT										
Constant	SIZE	TLTA	WCTA	CLCA	D(TL>TA)	NITA	CFOTL	D (NI _t +NI _{t-1} <0)	CHINI	LT
-7.09***	0.030	1.202***	-1.809***	-0.102	-0.016	-0.447*	-0.341***	1.844***	-0.369***	2.034***
(0.357)	(0.038)	(0.375)	(0.552)	(0.159)	(0.345)	(0.267)	(0.127)	(0.252)	(0.132)	(0.194)

This table reports estimation results for six models; Ohlson (1980), Almamy et al., (2016), referred to as AAN, Campbell et al. (2008), referred to as CHS and three extended versions of Altman, Ohlson and Campbell et al. (2008) models, which include LT as additional predictor (referred to as ALT, OLT and CHSLT respectively). The sample includes 25950 firm-year observations, from which 201 went bankrupt in a year between 1995 and 2005. The predictor variables are constructed one year prior to bankruptcy. For the definition of variables refer to table 1. ***, ** and * indicate statistical significance at $\alpha=1\%$, $\alpha=5\%$ and $\alpha=10\%$ respectively. In parentheses clustered robust standard errors are reported, that take into consideration the panel

character of our data.

Table 6: Model performance and test for differences, baseline approach

Panel A: Out-of-sample performance, baseline approach (2006-2014)							
Model	AUROC	pAUROC	Log-Likelihood	Pseudo-R ² (%)			
Structural							
Leland-Toft	0.8941	0.3953	-600.93	19.72			
BSM	0.8659	0.3750	-614.55	17.90			
Hybrid							
OLT	0.9449	0.4477	-483.43	35.41			
CHSLT	0.9395	0.4426	-491.41	34.35			
ALT	0.9207	0.4288	-519.56	30.59			
Reduced-Form							
CHS (2008)	0.9332	0.4392	-498.85	33.35			
AAN (2016)	0.8597	0.3754	-603.81	19.33			
Ohlson (1980)	0.9252	0.4280	-535.57	28.45			

Panel B: Test-statistics for differences in AUROCs							
Model	OLT	CHSLT	ALT	CHS	AAN	Ohlson	Leland-Toft
CHSLT	0.64						
ALT	2.58	1.63					
CHS	1.12	1.47	-0.92				
AAN	5.36	5.30	4.78	4.51			
Ohlson	4.73	1.65	-0.41	0.8127	-4.41		
Leland-Toft	3.59	3.15	1.67	2.46	-1.64	1.96	
BSM	6.02	4.14	3.03	3.58	-0.28	3.28	2.74

Panel C: Test-statistics for differences in log-likelihoods							
Model	OLT	CHSLT	ALT	CHS	AAN	Ohlson	Leland-Toft
CHSLT	0.75						
ALT	4.38	2.38					
CHS	1.23	14.88	-1.48				
AAN	7.51	6.16	6.51	5.86			
Ohlson	104.28	2.98	1.22	2.58	-6.21		
Leland-Toft	8.07	6.93	6.73	5.76	-0.16	3.61	
BSM	8.01	7.21	6.56	6.51	0.63	4.47	1.38

This table reports out-of-sample performance for the two structural models (Leland-Toft and BSM), the three hybrid models (OLT, CHSLT and ALT) as well as the three reduced-form models (Ohlson,1980; Campbell et al., 2008, referred to as CHS and Alamy et al., 2016, referred to as AAN). For the definition of the models, refer to table 1. Panel A reports discriminating ability, measured by AUROC and partial AUROC, as well as predictive accuracy, measured by log-likelihood (and pseudo-R²). For the partial AUROC, we use the region of AUROC that starts from false positive rate equal to 0 and ends at false positive rate equal to 0.5. Panel B reports test statistics for differences in the discriminating ability between various models, using Delong et al. (1988). Panel C reports test statistics for differences in predictive accuracy between various models using likelihood ratio tests or Vuong (1989) test. The results are based on a baseline approach, where the models are estimated on the period 1995-2005 and applied on the period 2006-2014. In the case of structural models, for consistency, we estimate two logistic regression models where the first contains the probability of bankruptcy derived from Leland-Toft and the second the probability of bankruptcy derived from BS as predictors.

Table 7: Economic performance of banks using different bankruptcy models (LT vs BSM)

	Bank 1	Bank 2	Bank 3
	LT	BSM	Altman
Credits	4037	4667	5022
Market Share (%)	29.09	33.62	36.18
Bankruptcies	10	37	38
Bankruptcies/Credits (%)	0.25	0.79	0.76
Average Spread (%)	0.38	0.49	0.45
Revenues (\$M)	110.71	163.22	163.89
Loss(\$M)	26.41	97.70	100.34
Profit(\$M)	84.30	65.52	63.55
Return on Assets (%)	0.29	0.19	0.018
Return on RWA (%)	1.09	0.64	0.54

This table reports economic results for three banks in a competitive loan market worth \$100 billion. Bank 1 uses LT, bank 2 uses BSM and bank 3 uses Altman. For the definition of the models, see table 1. The banks sort prospective customers (2006-2014) and reject the 5% of firms with the highest risk. The remaining firms are classified in 10 groups of equal size and for each group, a credit spread is calculated, as described in the main text (section 4.3). The bank that classifies the firm to the group with the lowest spread is finally granting the loan. Market share is the number of loans given divided by the number of firm-years, Revenues = market size*market share*average spread, Loss=market size*prior probability of bankruptcy*share of bankruptcies*loss given default. Profit=Revenues-Loss. Return on Assets is profits divided by market size*market share and Return on Risk-Weighted-Assets is profits divided by Risk-Weighted Assets, obtained from formulas provided by the Basel Accord (2006). The prior probability of bankruptcy is the bankruptcy rate for firms between 1995-2005 and equals 0.77%. Loss given default is 45%.

Table 8: Model performance and test for differences, rolling window approach

Panel A: Out-of-sample performance, rolling approach (2006-2014)					
Model	AUROC	Log-Likelihood	Pseudo-R ² (%)		
Hybrid					
OLT	0.9469	-470.37	37.00		
CHSLT	0.9438	-479.66	35.76		
ALT	0.9253	-508.25	31.93		
Reduced-Form					
CHS (2008)	0.9372	-485.94	34.92		
AAN (2016)	0.8673	-593.86	20.46		
Ohlson (1980)	0.9289	-522.87	29.97		
Panel B: Test-statistics for differences in AUROCs					
Model	OLT	CHSLT	ALT	CHS	AAN
CHSLT	0.41				
ALT	2.31	1.58			
CHS	1.00	1.80	-0.88		
AAN	5.24	5.12	4.62	4.42	
Ohlson	4.33	1.94	-0.34	0.91	-4.38
Panel C: Test-statistics for differences in log-likelihoods					
Model	OLT	CHSLT	ALT	CHS	AAN
CHSLT	0.79				
ALT	4.57	2.20			
CHS	1.13	12.57	-1.46		
AAN	7.61	6.17	6.28	5.91	
Ohlson	104.99	2.85	1.06	2.49	-6.48

This table reports out-of-sample performance for the three hybrid models (OLT, CHSLT and ALT) as well as the three reduced-form models (Ohlson,1980; Campbell et al., 2008, referred to as CHS and Alamy et al., 2016, referred to as AAN). For the definition of the models refer to table 1. Panel A reports discriminating ability, measured by AUROC as well as predictive accuracy, measured by log-likelihood (and pseudo-R²). Panel B reports test statistics for differences in the discriminating ability between various models, using Delong et al. (1988). Panel C reports test statistics for differences in the predictive accuracy between various models using likelihood ratio tests or Vuong (1989) test. The results are based on a rolling window approach, where the models are updated yearly and used to predict bankruptcies next year. For instance, the models are estimated between 1995 and 2005 and apply them on firms in 2006. Then we re-estimate the models between 1996 and 2006 and apply them on firms in 2007. This process is repeated up to 2014. Bankruptcy probabilities for each year are aggregated to obtain single performance measures.

Table 9: Model performance and test for differences, five folds approach

Panel A: Out-of-sample performance, five folds approach (1995-2014)					
Model	AUROC	Log-Likelihood	Pseudo-R ² (%)		
Hybrid					
OLT	0.9091	-1379.48	28.49		
CHSLT	0.9057	-1403.72	27.23		
ALT	0.8826	-1471.35	23.73		
Reduced-Form					
CHS (2008)	0.9014	-1417.16	26.54		
AAN (2016)	0.8438	-1630.45	15.48		
Ohlson (1980)	0.8939	-1477.06	23.43		
Panel B: Test-statistics for differences in AUROCs					
Model	OLT	CHSLT	ALT	CHS	AAN
CHSLT	0.50				
ALT	3.95	2.86			
CHS	1.02	1.96	-2.05		
AAN	7.09	6.14	5.63	5.42	
Ohlson	4.56	1.54	-1.38	0.94	-5.75
Panel C: Test-statistics for differences in log-likelihoods					
Model	OLT	CHSLT	ALT	CHS	AAN
CHSLT	1.25				
ALT	5.20	3.07			
CHS	1.68	26.87	-2.11		
AAN	7.63	6.41	6.62	6.07	
Ohlson	195.16	2.79	0.22	2.34	-5.53

This table reports out-of-sample performance for the three hybrid models (OLT, CHSLT and ALT) as well as the three reduced-form models (Ohlson,1980; Campbell et al., 2008, referred to as CHS and Almany et al., 2016, referred to as AAN). For the definition of the models refer to table 1. Panel A reports discriminating ability, measured by AUROC as well as predictive accuracy, measured by log-likelihood (and pseudo-R²). Panel B reports test statistics for differences in the discriminating ability between various models, using Delong et al. (1988). Panel C reports test statistics for differences in the predictive accuracy between various models using likelihood ratio tests or Vuong (1989) test. The results are based on a five-fold cross-validation approach, where we divide the whole sample into five equal sub-samples. Any four of them are used to estimate the models and apply them on firms in the left-out sub-sample. Bankruptcy probabilities for each left-out sub-sample are aggregated to obtain single out-of-sample performance measures

Table 10: Economic performance for five banks when using different bankruptcy models

	Bank 1	Bank 2	Bank 3	Bank 4	Bank 5
	OLT	CHSLT	CHS	Ohlson	Altman
Credits	3571	2561	2758	1886	2935
Market Share (%)	25.73	18.45	19.87	13.59	21.15
Bankruptcies	4	4	12	11	48
Bankruptcies/Credits (%)	0.11	0.16	0.44	0.58	1.64
Average Spread (%)	0.35	0.35	0.43	0.51	0.59
Revenues (\$M)	90.14	65.35	85.42	69.41	125.65
Loss(\$M)	10.56	10.56	31.69	29.05	126.75
Profit(\$M)	79.58	54.79	53.73	40.36	-1.10
Return on Assets (%)	0.31	0.30	0.27	0.30	-0.00
Return on RWA (%)	1.74	1.54	1.12	1.02	-0.00

This table reports economic results for five banks in a competitive loan market worth \$100 billion. Bank 1 uses OLT, bank 2 uses CHSLT, bank 3 uses CHS, bank 4 uses Ohlson and bank 5 uses Altman. For the definition of the models, see table 1. The models are estimated using data from 1995-2005. The banks sort prospective customers (2006-2014) and reject the 5% of firms with the highest risk. The remaining firms are classified in 10 groups of equal size and for each group, a credit spread is calculated, as described in the main text (section 4.3). The bank that classifies the firm to the group with the lowest spread is finally granting the loan. Market share is the number of loans given divided by the number of firm-years, Revenues = market size*market share*average spread, Loss=market size*prior probability of bankruptcy*share of bankruptcies*loss given default. Profit=Revenues-Loss. Return on Assets is profits divided by market size*market share and Return on Risk-Weighted-Assets is profits divided by Risk-Weighted Assets, obtained from formulas provided by the Basel Accord (2006). The prior probability of bankruptcy is the bankruptcy rate for firms between 1995-2005 and equals 0.77%. Loss given default is 45%.

Table 11: Extending Campbell et al. (2008), Ohlson (1980) and Altman (1968) with BSM and LT

Panel A: Out-of-sample performance, 2006-2014					
Model	AUROC	LL	Pseudo-R ² (%)		
CHS models					
CHS (2008)	0.9332	-498.85	33.35		
CHS with BSM	0.9343	-497.96	33.47		
CHS with LT	0.9395	-491.41	34.35		
Ohlson models					
Ohlson (1980)	0.9252	-535.57	28.45		
Ohlson with BSM	0.9383	-497.71	33.51		
Ohlson with LT	0.9449	-483.43	35.41		
Altman models					
AAN (2016)	0.8597	-603.81	19.33		
Altman with BSM	0.9109	-538.87	28.01		
Altman with LT	0.9207	-519.56	30.59		
Panel B: Test statistics for differences in AUROC's					
CHS vs CHS with BSM	0.82	Ohlson vs Ohlson with BSM	3.54	AAN vs Altman with BSM	4.13
CHS vs CHS with LT	1.49	Ohlson vs Ohlson with LT	4.71	AAN vs Altman with LT	4.78
CHS with LT vs CHS with BSM	1.47	Ohlson with LT vs Ohlson with BSM	2.37	Altman with LT vs Altman with BSM	1.97
Panel C: Test statistics for differences in log-likelihoods					
CHS vs CHS with BSM	1.78	Ohlson vs Ohlson with BSM	75.72	AAN vs Altman with BSM	5.74
CHS vs CHS with LT	14.88	Ohlson vs Ohlson with LT	104.28	AAN vs Altman with LT	6.51
CHS with LT vs CHS with BSM	2.09	Ohlson with LT vs Ohlson with BSM	2.36	Altman with LT vs Altman with BSM	2.52

This table reports out-of-sample performance for the three reduced-form models (Campbell et al., 2008, referred to as CHS, Ohlson, 1980 and Alamy et al., 2016, referred to as AAN) and several hybrid models that augment the reduced-form models with the LT or BSM. The models are estimated using data from 1995-2005 and the results are based on the out-of-sample period, 2006-2014. Panel A reports discriminating ability and predictive accuracy as measured by AUROC and log-likelihood respectively. Panel B reports test statistics for differences in the discriminating ability between various models, using Delong et al. (1988). Panel C reports test statistics for differences in the predictive accuracy between various models using likelihood ratio tests or Vuong (1989) test.

Table 12: Model performance and test for differences, baseline approach (time-robustness)

Panel A: Out-of-sample performance, baseline approach (1995-2005)					
Model	AUROC	Log-Likelihood	Pseudo-R ² (%)		
Hybrid					
OLT	0.8854	-960.51	18.72		
CHSLT	0.8838	-937.03	20.71		
ALT	0.8716	-1029.06	12.92		
Reduced-Form					
CHS (2008)	0.8811	-951.09	19.52		
AAN (2016)	0.8344	-1116.46	5.53		
Ohlson (1980)	0.8741	-1014.27	14.17		
Panel B: Test-statistics for differences in AUROCs					
Model	OLT	CHSLT	ALT	CHS	AAN
CHSLT	0.16				
ALT	1.47	1.14			
CHS	0.42	1.28	-0.81		
AAN	3.88	3.49	3.98	3.25	
Ohlson	2.41	0.87	-0.22	0.62	-3.18
Panel C: Test-statistics for differences in log-likelihoods					
Model	OLT	CHSLT	ALT	CHS	AAN
CHSLT	-1.20				
ALT	2.36	3.02			
CHS	-0.43	28.12	-2.41		
AAN	3.13	3.62	3.07	3.33	
Ohlson	107.52	3.14	-0.43	2.56	-2.10

This table reports out-of-sample performance for the three hybrid models (OLT, CHSLT and ALT) as well as the three reduced-form models (Ohlson,1980; Campbell et al., 2008, referred to as CHS and Almayy et al., 2016, referred to as AAN). For the definition of the models, refer to table 1. Panel A reports discriminating ability, measured by AUROC as well as predictive accuracy, measured by log-likelihood (and pseudo-R²). Panel B reports test statistics for differences in the discriminating ability between various models, using Delong et al. (1988). Panel C reports test statistics for differences in the predictive accuracy between various models using likelihood ratio tests or Vuong (1989) test. The results are based on a baseline approach, where the models are estimated on the period 2006-2014 and applied on the period 1995-2005, in order to test the time robustness of our models.

Figures

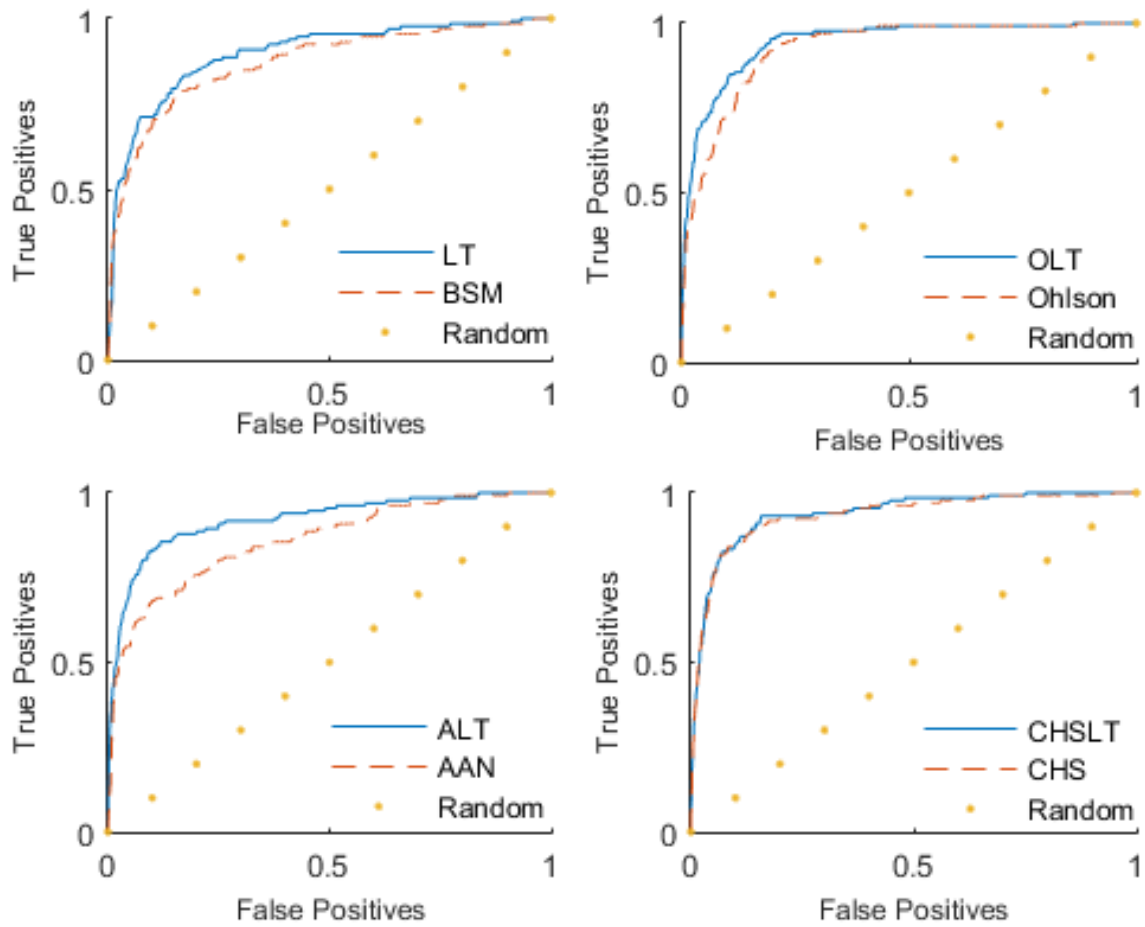


Figure 1: This figure provides graphical representation of the discriminatory power of various bankruptcy prediction models through the ROC curves. The ROC curves are constructed for the out-of-sample period 2006-2014.

