

# Robustness of Distance-to-Default

Jessen, Cathrine; Lando, David

*Document Version*

Accepted author manuscript

*Published in:*

Journal of Banking & Finance

*DOI:*

[10.1016/j.jbankfin.2014.05.016](https://doi.org/10.1016/j.jbankfin.2014.05.016)

*Publication date:*

2015

*License*

CC BY-NC-ND

*Citation for published version (APA):*

Jessen, C., & Lando, D. (2015). Robustness of Distance-to-Default. *Journal of Banking & Finance*, 50, 493–505.  
<https://doi.org/10.1016/j.jbankfin.2014.05.016>

[Link to publication in CBS Research Portal](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us ([research.lib@cbs.dk](mailto:research.lib@cbs.dk)) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 04. Jul. 2025



# Robustness of Distance-to-Default

**Cathrine Jessen and David Lando**

Journal article (Post print version)

**CITE:** Robustness of Distance-to-Default. / Jessen, Cathrine; Lando, David. In: Journal of Banking & Finance, Vol. 50, 01.2015, p. 493-505.

DOI: [10.1016/j.jbankfin.2014.05.016](https://doi.org/10.1016/j.jbankfin.2014.05.016)

Uploaded to [Research@CBS](mailto:Research@CBS): December 2016

© 2016. This manuscript version is made available under the CC-BY-NC-ND 4.0 license  
<http://creativecommons.org/licenses/by-nc-nd/4.0/>

# Robustness of distance-to-default\*

Cathrine Jessen & David Lando

`cj.fi@cbs.dk`

`dl.fi@cbs.dk`

Department of Finance

Copenhagen Business School

April 26, 2014

## Abstract

Distance-to-default (DD) is a measure of default risk derived from observed stock prices and book leverage using the structural credit risk model of Merton (1974). Despite the simplifying assumptions that underlie its derivation, DD has proven empirically to be a strong predictor of default. We use simulations to show that the empirical success of DD may well be a result of its strong robustness to model misspecifications. We consider a number of deviations from the Merton model which involve different asset value dynamics and different default triggering mechanisms. We show that, in general, DD is successful in ranking firms' default probabilities, even if the underlying model assumptions are altered. A possibility of large jumps in asset value or stochastic volatility challenge the robustness of DD. We propose a volatility adjustment of the distance-to-default measure that significantly improves the ranking of firms with stochastic volatility, but this measure is less robust to model misspecifications than DD.

## 1 Introduction

'Distance-to-default' is a credit score derived from observed stock prices and book leverage using a structural model of default risk. A version of the measure based on the Merton (1974) model has been shown empirically to perform well when it comes to ranking firms

---

\*Support from the Center for Financial Frictions (FRIC), grant no. DNRF102, is gratefully acknowledged. Cathrine Jessen also thanks the Danish Council for Independent Research | Social Sciences for financial support. We are grateful for comments from our discussant at the International Conference on Financial Regulation and Systemic Risk in Paris, André Lucas, and an anonymous referee.

according to their default risk, see for example Hillegeist, Keating & Cram (2004), Duffie, Saita & Wang (2007) and Bharath & Shumway (2008). Thus, the measure is commonly applied as an alternative to ratings to control for firms' default risk, as in e.g. Chava & Purnanandam (2010), Vassalou & Xing (2004) and Acharya, Lochstoer & Ramadorai (2013). The good ranking performance of a measure based on the Merton model is striking in view of the model's somewhat poorer ability to capture the level of default probabilities, and in view of its simple assumptions on asset dynamics and debt structure. In this paper we investigate whether the success of the distance-to-default can in part be explained by a strong robustness to model misspecifications. In other words, is it the case that the distance-to-default measure performs well even if the observed data are generated using other asset value dynamics or different default triggering mechanisms? As part of this agenda, we focus on understanding which violations of the underlying Merton model that may cause the distance-to-default to fail in its ranking of firms. We base all of our results on simulated samples so that we are able to run tightly controlled experiments.

We find that changing the default triggering mechanism into a model with an exogenous default boundary like the one of Black & Cox (1976) or having an endogenously determined default boundary as in Leland & Toft (1996) has little effect on the robustness of the distance-to-default measure. However, changing asset-value dynamics may have a pronounced effect. Introducing the possibility of large jumps in asset value challenges the robustness of the measure. Interestingly, jumps in themselves are not enough to challenge the robustness. The measure performs well if jumps are frequent but relatively small, such that the majority of the variation in asset value comes from the diffusion part, or the jump-diffusion is well approximated by a diffusion. Having stochastic volatility in asset value also makes the measure significantly less reliable. This happens because the constant volatility estimate may turn out much lower than the realized volatility path, and distance-to-default therefore mistakenly classifies the firm as less risky than it truly is. The opposite effect can be obtained in a model of a firm with dual business lines. If a firm has both high-volatile and low-volatile assets, one may in theory have a firm with zero probability of default for which an estimated distance-to-default implies a non-negligible risk of default. Yet, for realistic parameters in the dual business model, distance-to-default is a robust measure for ranking firms' default risk.

Even if the ranking by distance-to-default turns out poor for the case of stochastic volatility, the procedure by which the distance-to-default is estimated gives remarkably robust estimates of the underlying firm asset value. Consequently, we propose a modified distance-to-default measure that accounts for stochastic volatility. We use the same method for estimating the underlying firm asset value but then in a second step estimate a stochastic volatility specification of the asset value dynamics. Our alternative distance-to-default measure improves the ranking performance for firms with stochastic volatility, but it comes at the cost of being less robust to deviations from this assumption; for both the Merton and jump diffusion models it leads to a slightly poorer ranking performance.

We evaluate distance-to-default's ranking ability using three methods: The first is visual inspection of Cumulative Accuracy Profiles (CAP curves), the second uses a formal test of significance based on Receiver Operator Characteristics and the third is based on a measure of profitability for competing banks that employ models in the decisions to grant loans and whose profitability are a function of the accuracy of their model compared to that of their competitor. With respect to the last method, we find that if firms have stochastic volatility in asset value there is a large potential economic benefit of changing to a more powerful credit scoring model than distance-to-default, whereas the benefit obtained is small for all other violations of Merton's assumptions that we test.

The Merton model is notoriously known to produce default probabilities that are too low, and therefore the common use of distance-to-default is to apply some empirically based transformation which maps the measure into a default probability. This was proposed by Crosbie & Bohn (2003) and Kealhofer (2003) who mapped distance-to-default to a so-called expected default frequency. We will not consider such empirical transformation in this paper, which solely focuses on distance-to-default's ranking performance.

## 2 The distance-to-default

The basic ingredients of the simple version of the Merton model that are used to derive the distance-to-default measure are:

1. The firm's asset value process,  $V$ , follows a geometric Brownian motion and therefore, in particular, has constant volatility and no jumps.
2. The firm's capital structure consists of debt and equity, where debt is issued as a single zero coupon bond. This means that the firm can only default at maturity of debt.
3. All market frictions are ignored.

The default probability predicted by the Merton model is given as  $PD = N(-DD)$ , where  $N$  is the standard normal distribution function and  $DD$  is the distance-to-default:

$$DD = \frac{\ln\left(\frac{V}{P}\right) + \left(\mu - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}. \quad (1)$$

Here  $T$  denotes the maturity of outstanding debt,  $P$  is the face value of debt,  $\mu$  the drift and  $\sigma$  the volatility of the asset value process.

Estimating the primary parameters,  $\sigma$  and  $V$ , that determine  $DD$  is challenging, because the firm asset value,  $V$ , is not directly observable. The classical way to handle this is to apply equity data for the estimation instead. There exist several approaches<sup>1</sup> to estimating

---

<sup>1</sup>Alternatively, one can solve a system of two nonlinear equations with the two unknowns,  $\sigma$  and  $V$ , as e.g. Hillegeist et al. (2004) do using historical equity volatility or apply options implied equity volatility. Maximum likelihood estimation of  $\sigma$  and  $V$  is also possible by viewing the asset value process as a transformation of observed equity data (Duan (1994)).

$\sigma$  and  $V$  from equity data, but we choose the iterative procedure of Vassalou & Xing (2004), which we will refer to as the VX-algorithm and describe below. We consider firm's face value of debt,  $P$ , and maturity,  $T$ , observable from accounting statements. The estimation makes use of the fact that in the Merton model the equity value,  $E$ , equals the value of a call option on the firm's asset value with strike equal to face value of debt:

$$E_t = c(V_t; \sigma, P, T - t, r), \quad (2)$$

where  $c(\cdot)$  is the Black Scholes call option price formula.

Let  $E_{t_0}, E_{t_1}, \dots, E_{t_N}$  be a set of observed equity values over a time-span of  $t_N$  years. For simplicity assume that we work with equidistant observations such that observations are  $\Delta t = \frac{t_N}{N}$  apart. Fix an initial guess of the volatility parameter,  $\hat{\sigma}_0$ , to start the iteration.

The  $n$ 'th step of the iteration updates the asset volatility estimate,  $\hat{\sigma}_{n-1}$ , to an improved estimate,  $\hat{\sigma}_n$ , in two steps. First, calculate estimated asset values  $V_{t_0}(\hat{\sigma}_{n-1}), \dots, V_{t_N}(\hat{\sigma}_{n-1})$  from observed equity values by inverting the Black Scholes formula in (2) using the previous step's volatility estimate,  $\hat{\sigma}_{n-1}$ . Second, find the updated volatility estimate,  $\hat{\sigma}_n$ , as the volatility estimate of the estimated asset value process,  $V_{t_0}(\hat{\sigma}_{n-1}), \dots, V_{t_N}(\hat{\sigma}_{n-1})$ , from step one, where the asset value process is assumed to follow a geometric Brownian motion:

$$\begin{aligned} \hat{\sigma}_n &= \sqrt{\frac{1}{N\Delta t} \sum_{i=1}^N \left( \ln \left( \frac{V_{t_i}(\hat{\sigma}_{n-1})}{V_{t_{i-1}}(\hat{\sigma}_{n-1})} \right) - \bar{\xi} \right)^2} \\ \bar{\xi} &= \frac{1}{N\Delta t} \sum_{i=1}^N \ln \left( \frac{V_{t_i}(\hat{\sigma}_{n-1})}{V_{t_{i-1}}(\hat{\sigma}_{n-1})} \right) = \frac{1}{N\Delta t} \left( \ln V_{t_N}(\hat{\sigma}_{n-1}) - \ln V_{t_0}(\hat{\sigma}_{n-1}) \right). \end{aligned}$$

These two steps are repeated until  $\hat{\sigma}_n$  converges which usually happens after only few iterations. We denote the final volatility estimate by  $\hat{\sigma}$ .

The iterative procedure also provides a drift estimate:  $\hat{\mu} = \frac{1}{\Delta t} \bar{\xi} + \frac{1}{2} \hat{\sigma}^2$ . This estimate, however, has a large standard deviation and is therefore rarely applied in empirical studies.<sup>2</sup> Generally, it is well known that the expected return is hard to estimate with precision as documented by e.g. Merton (1980). Instead we will assume that asset returns satisfy a CAPM-style relation, such that the drift varies with the volatility;  $\mu = r + \lambda \sigma$ . Thus, we estimate the drift by  $\hat{\mu} = r + \lambda \hat{\sigma}$ , assuming  $\lambda$  is known.

The default probabilities that we would estimate using this model on real data would in general be far too small, especially for relatively safe firms. But since the default probability is a monotone function of  $DD$ , this measure can still be used as a measure for ranking firms' default risk. One could proceed and use  $DD$  as an explanatory variable in a statistical model of default. Working with a slightly simpler distance-to-default measure, Crosbie & Bohn (2003) and Kealhofer (2003) translate distance-to-default into an empirical default probability by fitting a firm's distance-to-default non-parametrically to the historical default

---

<sup>2</sup>E.g. Campbell, Hilscher & Szilagyi (2008) set the drift equal to the risk free rate plus some fixed risk premium,  $\bar{\mu} = r + \pi_a$ .

frequency of other firms with the same distance-to-default. Numerous other studies have used  $DD$  as a covariate in hazard regressions, see for example Duffie et al. (2007), Lando & Nielsen (2010), Lando, Medhat, Nielsen & Nielsen (2013), and Shumway (2001). All studies have found a very significant role of  $DD$  in default prediction.

### 3 Experimental design

The success of  $DD$  as a default predictor depends critically on its ability to rank firms according to their default risk, and its empirical performance indicates that this ability is robust to model misspecification. We therefore conduct the following experiment. For six models, five of which violates at least one of the three Merton assumptions, we simulate asset values for a large sample of firms and calculate corresponding equity values, which we then think of as the observed equity values. We use the equity paths to estimate  $DD$  by the iterative procedure of Vassalou & Xing (2004) in order to test the robustness of the distance-to-default to a number of deviations from the simple Merton model.

To evaluate the robustness of distance-to-default, we compare the ranking of firms' default risk according to the estimated  $DD$  to the ranking by the default probability calculated according to the true model specifications. We evaluate the two credit scores' ranking abilities for each model along several dimensions; graphically in terms of cumulative accuracy profiles (CAP curves), statistically in terms of a test of whether the difference between the curves is significant and economically by comparing the returns of two banks using the two credit scores for loan approval. If the performance of the estimated  $DD$ 's ranking is significantly poorer for some model, we conclude that the assumption violated by the model is a potential source of error when employing Merton's distance-to-default as measure of default risk.

#### 3.1 Simulation setup

For each model specification the experiment proceeds as illustrated in Figure 1. We first simulate daily asset values of  $M$  firms over a period of  $T_1$  years, which we will call the estimation period. Using the simulated asset value process we calculate each firm's equity value process according to the given model specification. We treat these computed equity values as observed data and use them for estimating  $DD$ . Consistent with the usage of  $DD$  in practice, we treat the underlying firm asset values as unobservable and use them only for comparing our estimated distance to default to a measure based on the true default probabilities.

Next, we use the VX-algorithm described in Section 2 on the equity data to obtain estimates of the firms' asset volatility,  $\hat{\sigma}$ , drift,  $\hat{\mu}$ , and value,  $V_{T_1}(\hat{\sigma})$ . This allows us to

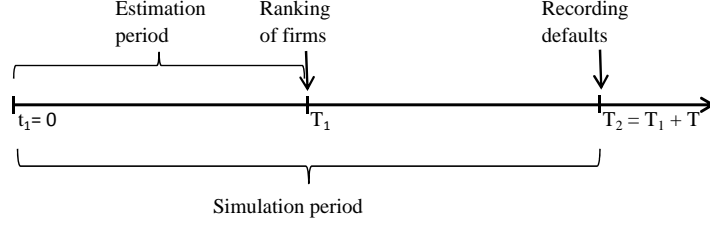


Figure 1: Timeline of experiment.

calculate a VX-estimated distance-to-default,  $DD_{VX}$ , at time  $T_1$ :

$$DD_{VX} = \frac{\ln\left(\frac{V_{T_1}(\hat{\sigma})}{P}\right) + (\hat{\mu} - \frac{1}{2}\hat{\sigma}^2)T}{\hat{\sigma}\sqrt{T}}.$$

$T$  is the default horizon we wish to consider. For comparison, we also calculate the true default probability,  $PD_{\text{true}}$ , at time  $T_1$  according to the underlying model specification using the simulated asset value at time  $T_1$  and the true model parameters applied in the simulation. The true default probability can be transformed into what we will call the 'true distance-to-default' by  $DD_{\text{true}} = -N^{-1}(PD_{\text{true}})$ , where  $N^{-1}$  is the inverse standard normal distribution function. Note that this measure is not computed from estimated asset values and volatilities. Rather it is a transformation of the true default probability into a DD-like measure, which ranks the firms according to their true default probabilities and which is on the same scale as the estimated DD.

Next we rank firms' by their default risk at time  $T_1$ , first as measured by the estimated distance-to-default ( $DD_{VX}$ ) and second, as measured by the true distance-to-default ( $DD_{\text{true}}$ ). To compare these two rankings, we continue the simulation for another  $T$  years and record whether or not each firm defaults during this period. If distance-to-default is a good measure for ranking firms' default risk, then firms that end up in default after  $T$  years ought to be listed among firms with the lowest distance-to-default at the ranking time,  $T_1$ . Note, that even though  $DD_{\text{true}}$  is calculated based on all available information about the true model specification and parameters, the stochastic evolution of asset value between the ranking time,  $T_1$ , and debt maturity,  $T_2$ , prevents  $DD_{\text{true}}$  from providing a perfect ranking.

To explore the relative information contained in the volatility estimate compared to the leverage ratio, we also consider the ranking of firms by leverage,  $L$ . Here, leverage is calculated at time  $T_1$  as the book value of debt relative to the market value of equity plus the book value of debt;  $L_{T_1} = \frac{P}{E_{T_1} + P}$ .

### 3.2 Performance evaluation

To compare the ranking of  $DD_{\text{true}}$  and  $DD_{VX}$  we first calculate Spearman's  $\rho$ . We choose this rank correlation coefficient because it punishes large dislocations harder than e.g.

Kendall's  $\tau$ , and does not assume a linear relationship as does Pearson's correlation coefficient. The rank correlation only measures how the estimated  $DD_{VX}$  relates to  $DD_{\text{true}}$ , and not how the two perform for identifying firms that actually default.

We use Moody's cumulative accuracy profiles (CAP) to evaluate the ability of our three risk scores,  $DD_{\text{true}}$ ,  $DD_{VX}$ , and leverage, to identify firms that actually default within  $T$  years. For  $x \in [0, 1]$ , the CAP curve corresponding to a given scoring method plots the fraction of defaulted firms whose risk score was in the lowest  $x$ -percentile of risk scores in the sample. The CAP curves give a graphical indication of the relative performance of the risk scores, whereas a comparison of the areas under the risk scores' CAP curves, referred to as accuracy ratios (AR), will identify the score with the overall superior ranking performance.

To give a statistical answer to which risk score performs the best, we test whether the areas under the curves differ significantly. Instead of testing for differences in the accuracy ratios we apply another popular but equivalent measure called receiver operating characteristic (ROC), which relates to AR as  $ROC = \frac{1}{2}(AR + 1)$ . The null-hypothesis, that two risk scores perform equally well, is  $H_0 : ROC_1 = ROC_2$ , where  $ROC_i$  is the ROC for risk score  $i$ ,  $i = 1, 2$ . In testing  $H_0$ , we follow Engelmann, Hayden & Tasche (2003), who employ the fact that ROC can be calculated as the test statistic of a Mann-Whitney U-test, which is asymptotically normally distributed. They propose the test-statistic

$$T = \frac{(ROC_1 - ROC_2)^2}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}^2}, \quad (3)$$

where  $\sigma_i^2$  is the variance of risk score  $i$ , and  $\sigma_{12}^2$  is the correlation between the risk scores. The test statistic is  $\chi^2(1)$ -distributed with one degree of freedom.

Strictly speaking, it only makes sense to compare CAP curves and ROC measures for risk scores calculated on the same sample, cf. Section 4.9 in Lando (2004). However, by the nature of our experiment, we create different samples for each model specification. To make the comparison of the relative performance of  $DD_{VX}$  and  $DD_{\text{true}}$  meaningful across models, we impose some homogeneity across the simulated samples by choosing parameters such that initially at time  $t_0$ , the default probabilities for all firms are equal. Over the estimation period the firms' asset values will develop differently and thereby ensure diversity in firm values at the ranking time point  $T_1$ . We need this diversity since even the best risk model cannot separate firms whose differences in characteristics are small compared to the random shocks in future asset values. In summary, we ensure diversity of risks within a sample while preserving homogeneity in the distribution of characteristics across samples.

The statistical test is an improvement over the mere visual inspection of CAP curves, but still it tells us little about the economic significance of using a misspecified model to calculate the credit score. To quantify the potential economic benefit of having a more powerful credit scoring model for ranking firms' default risk, we follow the approach developed by Blöchliger & Leippold (2006) who set up a stylized loan market consisting of two banks

with different credit scoring models available. In a similar fashion, Stein (2005) and Stein & Jordão (2003) also develop an approach for linking power curves and the prices of loans.

Consider a lending market consisting of two banks, both interested in lending to the firms in our sample. The first bank ranks firms by their true default probability and the second by the VX-estimated distance-to-default. Both banks discretize their credit scores such that firms receive a score between 1 and 20 corresponding to each 5%-quantiles. Let  $P(Y = 1|X = x)$  ( $P(Y = 0|X = x)$ ) denote the probability that a firm with default indicator  $Y$  and risk score  $x$  defaults (does not default) after one year. Both probabilities are calculated based on our simulated sample. Let  $LGD$  denote the loss given default and assume it equals 40% for all firms. Then, if we assume a discount rate of  $r$ , a loan with face value 1 to a firm with score  $x$  has a net present value of zero if the spread on the loan,  $s(x)$ , is chosen such that

$$-1 + \frac{(1 + r + s(x))P(Y = 0|X = x) + (1 - LGD)P(Y = 1|X = x)}{1 + r} = 0.$$

We assume that both banks reject the loan application from firms with score 1, i.e. firms ranked among the lowest 5%. Furthermore, the banks charge a fixed fee of 30 bps on top of the spread that gives an NPV of zero. Hence, the rate offered to the remaining 95% firms equals

$$r + s(x) = \frac{P(Y = 1|X = x)}{P(Y = 0|X = x)}LGD + \frac{r}{P(Y = 0|X = x)} + 30bp. \quad (4)$$

Firms that are offered a loan by both banks accept only the cheapest or if the banks offer the same spread, the loan is split between the two. Firms that are offered a loan only by one of the banks accept this regardless of the spread. The bank's realized return,  $R$ , on a loan to firm  $i$  with default indicator  $Y_i$  and risk score  $x$  is

$$R(x) = -1(1 + r) + \mathbf{1}_{\{Y_i=1\}}(1 - LGD) + \mathbf{1}_{\{Y_i=0\}}(1 + r + s(x)). \quad (5)$$

As our final performance evaluation we compare the two banks' average return to measure the economic benefit of having a more powerful credit scoring model than the VX-estimated distance-to-default.

## 4 Models

We have chosen five different models, besides the Merton model, for the experiment. First, we use a jump-diffusion and a stochastic volatility specification of the asset value process to test the robustness of distance-to-default when the firm asset value does not follow a geometric Brownian motion. A violation of this model assumption is also tested with the dual business model of Arora & Sellers (2004), who study firms that consist of a high-volatile and a low-volatile business part. Second, to test Merton's simple assumption on the firm's capital structure and the default triggering mechanism, we first simulate asset values

from the Black & Cox (1976) model, where firms default the first time an exogenous default boundary is hit. Next, we use the Leland & Toft (1996) model to further investigate whether market frictions such as taxes and dead weight costs of default leading to an endogenous default boundary causes the ranking by  $DD$  to break down.

This section provides a short introduction to each of the models. As we wish to study the default risk under the real-world probability measure,  $\mathbb{P}$ , all model dynamics are specified with respect to  $\mathbb{P}$ . Yet, we do assume arbitrage free markets and the existence of a pricing measure,  $\mathbb{Q}$ , since we need to convert asset values into traded equity prices. For simplicity, we assume a constant risk free rate,  $r$ , in all of the models.

## 4.1 Violations of the Merton asset value specification

### 4.1.1 Jumps and stochastic volatility

First we look at two extensions of the Merton model, one with stochastic volatility and one with jumps in asset value.<sup>3</sup> The model with stochastic volatility specifies the dynamics for the firm asset value  $V$  as follows:

$$\frac{dV_t}{V_t} = \mu dt + \sqrt{v_t} dW_t^1 \quad (6)$$

$$dv_t = \kappa(\theta - v_t)dt + \eta\sqrt{v_t}dW_t^2. \quad (7)$$

$W^1$  and  $W^2$  are Brownian motions with correlation  $\text{cor}(dW_t^1, dW_t^2) = \rho dt$ .

The jump-diffusion model assumes the following dynamics of the firm assets:

$$\frac{dV_t}{V_t} = (\mu - \lambda\mu_J)dt + \theta dW_t + J_t dN_t.$$

Here,  $N$  is a Poisson process with intensity  $\lambda$ .  $J_t$  is the jump size, which is log-normally distributed:  $\ln(1 + J_t) \sim \mathcal{N}(\ln(1 + \mu_J) - \frac{1}{2}\sigma_J^2, \sigma_J^2)$ .

For asset value specifications, the value equals the sum of debt,  $D_t$ , and equity,  $E_t$ :  $V_t = D_t + E_t$ . As in the original Merton model the firm defaults at debt's expiration time,  $T$ , if the firm value is not sufficient to repay the face value of debt;  $V_T < P$ . The value of equity equals the value of a call option on firm assets:  $E_t = E^\mathbb{Q}[e^{-r(T-t)}(V_T - P)^+ | \mathcal{F}_t]$ , where  $E^\mathbb{Q}[\cdot | \mathcal{F}_t]$  denotes the expectation with respect to the pricing measure  $\mathbb{Q}$  given the information available at time  $t$ ,  $\mathcal{F}_t$ . When pricing equity, we allow for a proportional volatility risk premium,  $\pi_v = \kappa^\mathbb{Q} - \kappa$ ,  $\theta^\mathbb{Q} = \frac{\theta\kappa}{\kappa^\mathbb{Q}}$ , and a jump risk premium,  $\pi_J = \lambda\mu_J - \lambda^\mathbb{Q}\mu_J^\mathbb{Q}$ . Both the stochastic volatility and the jump specification are affine processes, so in both models equity values,  $E_t$  and default probability,  $PD_t := \mathbb{P}(V_T < P | \mathcal{F}_t)$ , can be calculated using the transform methods of Duffie, Pan & Singleton (2000).

---

<sup>3</sup>Zhang, Zhou & Zhu (2009) provide empirical justification for a structural model with jumps and stochastic volatility in the firm value process in that this helps explain the credit default swap premium.

#### 4.1.2 Dual business model

Arora & Sellers (2004) introduce a model inspired by the business composition of large financial institutions, which can be thought of as composed of two parts: a high-risk and a low-risk business. To model this idea, let the firm value  $V$  be given as  $V = V^1 + V^2$ , where  $V^1$  and  $V^2$  follow correlated geometric Brownian motions:

$$\begin{aligned}\frac{dV_t^1}{V_t^1} &= \mu_1 dt + \sigma_1 dW_t^1 \\ \frac{dV_t^2}{V_t^2} &= \mu_2 dt + \sigma_2 dW_t^2,\end{aligned}$$

where  $\text{cor}(dW_t^1, dW_t^2) = \rho dt$ . As in the Merton model, the firm defaults if  $V_T < P$  at debt maturity  $T$ .

Instead of following Arora & Sellers (2004) who simplifies the setup in order to calculate the default probability analytically, we will use Monte Carlo simulations for calculating both default probabilities and equity values.

### 4.2 Violations of the Merton default triggering mechanism

#### 4.2.1 Exogenous default barrier

The Black-Cox model is an extension of the Merton model that allows for a more realistic default trigger; the firm defaults the first time the asset value falls below some exogenously given default boundary,  $B$ , and hereby the firm may default any time prior to maturity of debt. This default assumption implies a stochastic default time given by  $\tau = \inf\{t > 0 | V_t \leq B\}$ , where we assume that the default barrier is a constant fraction of face value of debt:  $B = \beta P$ . Like in the Merton model, firm's asset value process follows a geometric Brownian motion.

#### 4.2.2 Endogenous default barrier

The Merton model's assumption on the firm's capital structure is not maintained in the Leland-Toft model, where the firm at each moment has a continuum of bonds outstanding with total principal  $P$  and an aggregate annual coupon payment  $C$ . Each bond has maturity  $T$ , and as bonds mature they are rolled over. Furthermore, the Leland-Toft model violates the assumption on market frictions, which is incorporated in terms of a bankruptcy cost and a tax benefit from issuing debt specified as a marginal tax benefit rate,  $\tau$ . These frictions introduce a trade-off from issuing debt: debt coupons are tax-deductible and therefore carry a tax-benefit over equity but debt also carries a dead-weight cost in the event of default. In the original article, Leland & Toft (1996) allow firm owners to choose leverage to optimize the value of the firm. Here we want to ensure that leverage is in line with the rest of the models, so we consider the face value of debt exogenously given. Still, equity holders

optimally choose a constant default boundary,  $V_B$ , at which it is optimal to default, and in default debt holders recover only a fraction,  $\alpha$ , of the asset value. As in the Merton and Black-Cox models, firm's asset value process follows a geometric Brownian motion.

## 5 Results

In this section we present the results of our simulations for each of the six models. We first discuss our choice of model parameters thoroughly. We then evaluate the ranking performance of the estimated distance-to-default compared to the true one for each model in turn.

### 5.1 Calibration of key parameters

Our choice of key parameters such as leverage and default rate seeks to match observed averages for the period 2002–2011. The empirical sample, whose characteristics we try to match, consists of 13,216 firm years for which we have default data from Moody's, equity data from CRSP and accounting data from Compustat. To match the size of a typical sample, we conduct the experiment for  $M = 10,000$  simulated firms for each model. It is fairly common in the literature to use a 1-year estimation period and a 1-year default horizon, so we choose  $T_1 = 1$  and  $T = 1$ . For all models we set the risk free rate to  $r = 2\%$  matching the observed 10-year average 1-year T-bill rate of 2.0%.

We obtain dispersion in firm characteristics by varying initial leverage ratios,  $\frac{P}{V_0}$ , from 20% to 70% with an average of 45% close to the empirical sample average of 40%. In the empirical sample, the average yearly default rate is 1.3%, and we target this default frequency in our simulated samples to preserve homogeneity across models. First, we fix all parameters, except the diffusion volatility ( $\theta$  in the stochastic volatility specification and  $\sigma_2$  in the dual business model) according to existing literature as laid out in Table 1. Then for each model, we use the remaining free volatility parameter to ensure that each model's  $T_2$ -year default probability equals the empirical target of 1.3%. This results in an average 1-year distance-to-default of 3.1 which is lower than the observed average of 4.1.

The chosen parameters imply an average asset volatility ( $\sigma^M$ ) of 28% for our simulated Merton sample, whereas the empirical sample has an average observed volatility of 35% when estimated by the VX-algorithm. Initially, the average equity volatility in the Merton-sample is 49% and very close to our empirical sample average of 52%. For all models, the asset risk premium is set to  $\pi_a = 3.5\%$  as found by Zhang et al. (2009) for BBB-rated firms. This corresponds to a market price of risk of  $\lambda = 0.132$ . We set the drift parameter to  $\mu_i = r + \lambda\sigma_i^M$  for firm  $i$  in all models, and the drift thereby ranges from 3.7%–8.5% across firms with an average of 5.7%.

Model	Fixed parameters	Parameter justification
JD	$\lambda = 11$ $\mu_J = 0$ $\sigma_J = 4.2\%$ $\pi_J = 1.9\%$	Parameters are from the estimation in Wong & Li (2006). The jump risk premium, $\pi_J$ , is from Zhang et al. (2009) for BBB-rated firms.
SV	$\kappa = 0.21$ $\eta = 10\%$ $\rho = -60\%$ $\pi_v = -1.5\%$	Parameters, $\kappa$ , $\eta$ and $\rho$ are from the estimation in Bu & Liao (2013). The volatility risk premium, $\pi_v$ , is from Zhang et al. (2009).
DB	$A_1 = 70$ $\rho = 40\%$ $\sigma_1 \in [11\%, 42\%]$	Parameters are inspired by Landier, Krüger & Thesmar (2011). Arora & Sellers (2004) only implement the model for financial firms, which is not our focus, so instead we follow Landier et al. (2011), who study firms with operations in more than one industry.
BC	$\beta = 70\%$	The parameter is based on the average default barrier found by Davydenko (2012b).
LT	$\tau = 27\%$ $C = 6$ $\alpha = 60\%$	Parameters are from He & Xiong (2012). These parameters are slightly different than those in Leland & Toft (1996), however, He & Xiong (2012) provide a careful justification for each choice.

Table 1: The table provides an overview of the models included in our experiment, our choice of fixed parameters under the real world measure,  $\mathbb{P}$ , and justifications of parameter choices. In each sample, leverage varies from 20% to 70%, and for each model the volatility parameter ( $\theta$  in the stochastic volatility model and  $\sigma_2$  in the dual business model) varies with leverage and is chosen to ensure an initial  $T_2$ -year default probability of 1.3%.

## 5.2 Ranking of firms' default risk

Figure 2 shows the CAP curves for the Merton model. Here we see practically no difference between the curve generated by the true  $DD$  compared to the estimated  $DD$ . The rank correlation between  $DD_{\text{true}}$  and  $DD_{\text{VX}}$  calculated by Spearman's  $\rho$  is 0.99 and supports the observed closeness of the CAP curves. As expected, the ranking produced by the leverage ratio is clearly inferior, and we conclude that in the Merton model, the volatility estimate carries important information about a firm's default risk.

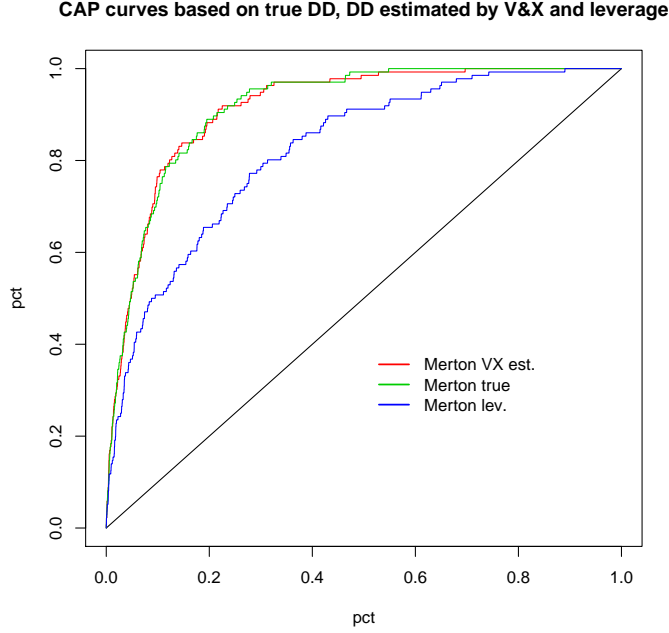


Figure 2: CAP curves for the estimated and true  $DD$  and leverage in the Merton model. Initial leverage ranges from 20% – 70%, and volatility is chosen such that initial 2-year default probability is 1.3%. This implies a volatility  $\sigma$  in the range 48.9% – 13.2%. The risk free rate is  $r = 2\%$  and the market risk premium  $\lambda = 0.132$ .

### 5.2.1 Ranking in the jump-diffusion model

Next we consider deviations from the Merton model assumptions regarding asset value dynamics, and we first evaluate the jump-diffusion specification. The chosen jump parameters,  $\mu_J = 0$ ,  $\sigma_J = 4.2\%$  and  $\lambda = 11$ , are based on the maximum likelihood estimation of Wong & Li (2006) with a downward bias on the jump size (Wong & Li (2006) find an average jump size of 1%). To ensure some diffusion risk for the highest leveraged firms, we impose a lower bound on  $\sigma$  of 6% despite the fact that this results in a default frequency higher than the target of 1.3%. Even though the average jump in asset value is zero, we can on any given day expect to see a negative jump of at least 10% for on average 2.6 of the 10,000 firms in the sample. Wong & Li (2006) assume that the jump distributions under the  $\mathbb{P}$  and  $\mathbb{Q}$ -measure are the same, but we apply the jump risk premium,  $\pi_J = 1.9\%$ , from the calibration in Zhang et al. (2009). This gives us the freedom to choose either  $\mu_J^{\mathbb{Q}}$  or  $\lambda^{\mathbb{Q}}$ , whereafter the other is pinned down by the relation  $\mu_J \lambda = \mu_J^{\mathbb{Q}} \lambda^{\mathbb{Q}} + \pi_J$ . We set  $\mu_J^{\mathbb{Q}} = -0.2\%$  implying  $\lambda^{\mathbb{Q}} = 9.4$ .<sup>4</sup>

With asset values generated by the jump-diffusion specification, the estimated  $DD$  is

<sup>4</sup>Unreported results show, that assuming zero jump risk premium does not change our conclusions.

not able to rank firms as accurately as the true default probability as is shown in figure 3, although the difference between the CAP curves is small still. We also observe a somewhat lower rank correlation of 0.91 compared to the Merton model. The main reason why  $DD_{VX}$  performs fairly well in this model specification is that the majority of the variation in firm asset value comes from the diffusion. With the above parameter specification, the asset value process experiences frequent, but mostly small jumps implying that paths can be well approximated by a geometric Brownian motion and therefore the VX-algorithm still ranks well.

The performance of the ranking based on  $DD$  estimated using the VX-algorithm strongly depends on the jump parameter specification. In the literature, there are not many examples of actual estimations of a Merton model extended to have jumps in asset values, and therefore we cannot claim that the chosen parameters are well established in the literature. Yet, there exist several calibrations of the model, where asset value jump parameters typically are chosen to match some observed characteristics in equity markets. One example is Zhang et al. (2009) who determine jump parameters by simulations that seek to fit the sample average and standard deviation of jumps in equity, and arrive at somewhat different jump parameters:  $\mu_J = 1.2\%$ ,  $\sigma_J = 19\%$  and  $\lambda = 0.16$  for BBB-rated firms. Here jumps are very rare but potentially large in absolute size, and using these parameters in our experiment results in a ranking by  $DD$  which is significantly poorer than the ranking by the true default probability. The Spearman correlation is also low at 81%, and the difference between the areas under the ROC-curves for the VX-estimated versus the true  $DD$ 's rankings is significant as reported in Section 5.3. Part of this conclusion is due to our experimental design, where we hold jump parameters fixed and only use diffusion volatility to ensure a default frequency of 1.3%. This means that the diffusion volatility parameter becomes low for highly leveraged firms.

In summary, our conclusion regarding the jump-diffusion specification is that the estimated  $DD$  performs well as a measure for ranking firms' default risk as long as the majority of the variation in asset value come from the diffusion part (or a process well approximated by a diffusion). However, the ranking will break down if firms have rare, but potentially large jumps and the diffusion part plays a minor role.

### 5.2.2 Ranking in the stochastic volatility model

The parameters for the stochastic volatility model are based on the estimation of Bu & Liao (2013). Performing an estimation of this model is not straight forward since both the asset value process and its volatility are unobservable processes. Bu & Liao (2013) solve this with a particle filtering approach which they test on a sample of 27 Dow Jones firms. We employ their average estimates of  $\eta = 10\%$ ,  $\kappa = 0.21$  and  $\rho = -60\%$  even if based on a relatively small sample. We use the volatility risk premium,  $\mu_v = -1.5\%$ , found by

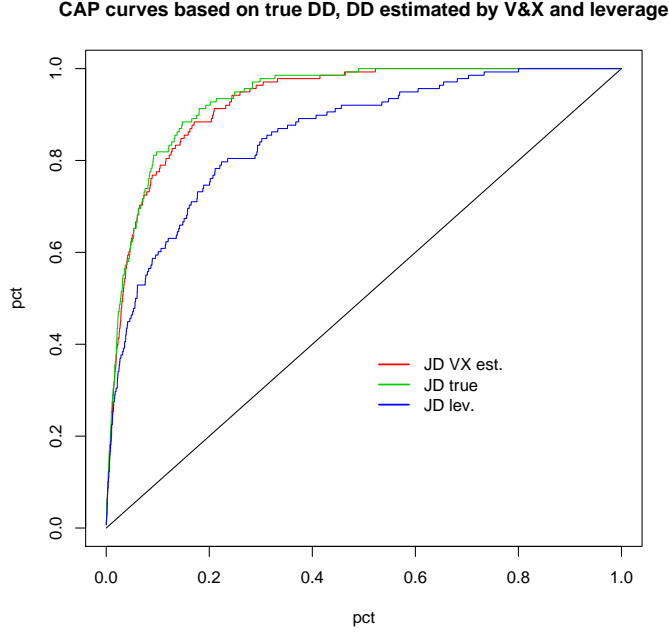


Figure 3: CAP curves for the estimated and true  $DD$  and leverage in the extended Merton model with jumps in firm asset value. Initial leverage ranges from 20% – 70%, and volatility is chosen such that initial 2-year default probability is 1.3%. This implies  $\sigma$  in the range 47% – 6%. The jump intensity is  $\lambda = 11$  ( $\lambda^Q = 9.4$ ), mean jump size  $\mu_J = 0$  ( $\mu_J^Q = -0.2\%$ ) and jump volatility  $\sigma_J = 4.2\%$ . The risk free rate is  $r = 2\%$  and the market risk premium  $\lambda = 0.132$ .

Zhang et al. (2009), since Bu & Liao (2013) implicitly assume this equals zero.<sup>5</sup> As with the jump-diffusion specification, estimations of the stochastic volatility model is very sparse in the literature and we cannot claim that our parameter choices are representative for all firms.

When the asset value dynamics has stochastic volatility, the estimated  $DD$ -measure produces a visibly lower CAP curve than the curve by the true measure as shown in Figure 4. Moreover, the CAP curves show that there is a considerable number of firms with estimated  $DD$  around the median which actually end up in default, but which the estimated  $DD$  cannot identify as high risk firms. This observation of large dislocations in  $DD_{VX}$ 's ranking compared to  $DD_{true}$  is confirmed by a low Spearman's  $\rho$  of 0.91.

$DD_{VX}$  is more often mistaken about the relative default risk of the firms in this model because the  $VX$ -volatility estimate is approximately equal to the mean of the volatility path over the estimation period, which can be far from the true long term mean ( $\theta$ ). The key parameters driving this result are the volatility-of-volatility parameter,  $\eta$ , and the mean reversion speed,  $\kappa$ . In a sample with a low volatility-of-volatility or a high mean reversion

<sup>5</sup>Unreported results show, that assuming zero volatility risk premium does not change our conclusions.

speed, the asset value paths would resemble a geometric Brownian motion, and  $DD$  would in fact be able to rank firms' default risk. In section 6 we return to the performance of distance-to-default for ranking firms with stochastic volatility.

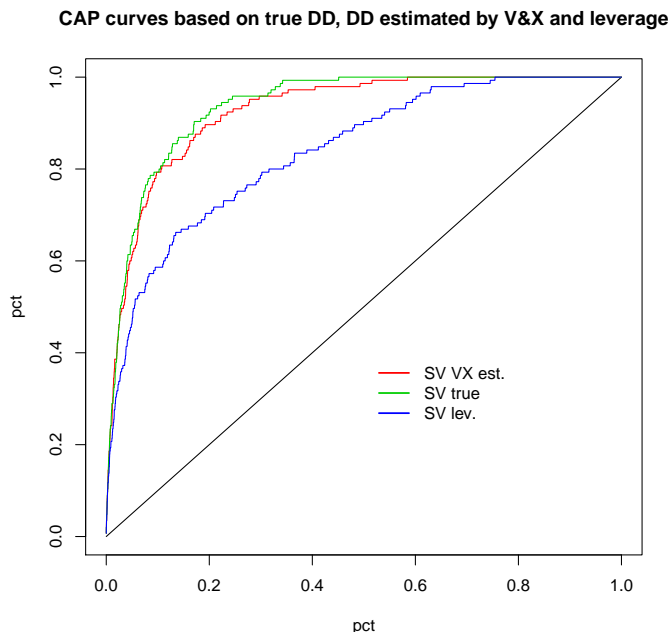


Figure 4: CAP curves for the estimated and true  $DD$  and leverage in the extended Merton model with stochastic volatility of firm asset value. Initial leverage ranges from 20% – 70%, and volatility is chosen such that initial default probability is 1.3%. This implies a long term mean volatility (squared to compare with remaining models),  $\sqrt{\theta}$  in the range 45.2% – 9.6%. The mean reversion speed is  $\kappa = 0.21$  ( $\kappa^Q = 0.19$ ), vol-of-vol  $\eta = 0.10$  and correlation  $\rho = -0.6$ . The risk free rate is  $r = 2\%$  and the market risk premium  $\lambda = 0.132$ .

### 5.2.3 Ranking in the dual business model

Arora & Sellers (2004), who propose the dual business model, calibrate the model to a small sample of eight financial firms. Because such firms have quite different capital structures compared to firms in other sectors, they are usually excluded from empirical default investigations and they are not the target-firm in our study either. Instead, we use the results of Landier et al. (2011), who compare the characteristics of firms with business in a single industry to conglomerates with businesses in more than one industry. We use their empirical results on conglomerates. First, they find that on average 73% of sales comes from the largest division, which motivates our choice of  $A_1 = 70\%$ . Second, they find the average asset beta for both the core business ( $A_1$ ) and the divisions ( $A_2$ ) to be  $\beta_1 = \beta_2 = 0.55$ . In-

stead of using the average  $\beta$ -estimates, we use the 25th quantile for  $\beta_1 = 0.34$  and the 75th quantile for  $\beta_2 = 0.77$  to obtain a spread between the two. To convert  $\beta$ -values into reasonable estimates of average volatilities,  $\sigma_1$  and  $\sigma_2$ , we use the CAPM-relation  $\beta_i = \frac{\rho}{\sigma_m} \sigma_i$ , where we put market volatility  $\sigma_m = 0.25$  and the industry's correlation with the market to  $\rho = 0.33$ . This results in average volatilities of  $\sigma_1 = 0.26$  and  $\sigma_2 = 0.58$ , which we will target. If we put  $\sigma_1 = 0.85 \cdot \sigma^M \in [11.2\%, 41.6\%]$  (average 24%), and fix initial default probability at 1.3%, this implies  $\sigma_2 \in [27.7; 99\%]$  with an average of 64%. These are not necessarily parameters of the average conglomerate, but if anything they provide a stressed test of the dual business model since a higher difference between  $\sigma_1$  and  $\sigma_2$  represents a model further from the Merton model. We also allow different drift for the two business parts. We set  $\mu_1 = r + \lambda\sigma_1$  and choose  $\mu_2$  such that the weighted (according to size of business parts) average of the two drift parameters equals the drift in the other models;  $\mu = 0.7\mu_1 + 0.3\mu_2$ .

CAP curves for the dual business model are shown in Figure 5. Both CAP curves and a rank correlation of 0.96 show that the estimated *DD* performs well for ranking firms' default risk even in this model, where the asset value specification appear quite different from that in the Merton model. Yet, if we employ an Anderson-Darling test of whether increments in log asset value are normally distributed it is in fact accepted at the 5% level for 91% of the paths compared to 94% for the Merton model. This provides at least a part of the explanation for why *DD* is able to rank the dual business firms' default risk. Yet, we could choose parameters in this model such that *DD*'s ranking is bound to fail. For example, think of a hypothetical firm with  $A_1 = 70$  and a face value of debt of  $P = 70$ . Assume that the low risk part of the firm is completely risk free ( $\sigma_1 = 0$ ), thereby implying that this firm has zero probability of defaulting on its debt. In this case, the VX-algorithm would estimate a strictly positive default probability (given  $\sigma_2 > 0$ ) and thereby overestimate the default risk of such firm. There may of course exist such firms, but this will not characterize the average firm in an empirical sample and we consider the above experiment a realistic test of *DD*'s ranking abilities in the dual business model.

#### 5.2.4 Ranking in the Black Cox model

Now we turn to the robustness of the *DD* when the Merton model's assumption regarding the timing of default is violated, and we start with the exogenous default barrier specification in the Black Cox model. The only parameter in this model, besides the diffusion volatility, is the default barrier. We have specified this as a percentage of the level of debt which varies across firms. Davydenko (2012b) conducts a purely empirical investigation and finds that on average firms default when asset value hits 66% of the face value of debt. Wong & Choi (2009) estimates the Black Cox model and find  $\beta = 74\%$ . We therefore choose  $\beta = 70\%$ . A higher value of  $\beta$  implies a higher hitting probability, i.e. a higher probability of default prior to debt's maturity. In our setup, a higher  $\beta$  therefore means

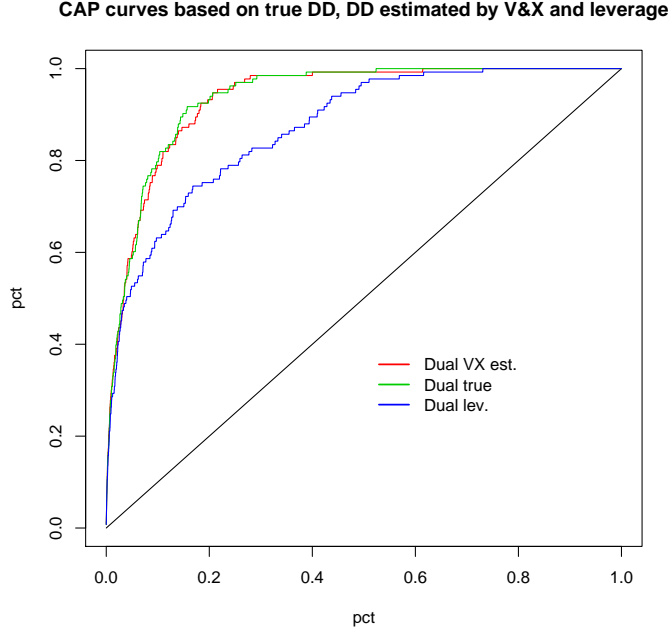


Figure 5: CAP curves for the estimated and true  $DD$  and leverage in the dual business model. Initial leverage ranges from 20% – 70%,  $\sigma_1$  from 41.6% – 11.2%, and  $\sigma_2$  is chosen such that initial default probability is 1.3%. This implies a  $\sigma_2$  of 99.2% – 27.7%. The firm consists of 70% low-volatility business and 30% high-volatility business. The two businesses are correlated  $\rho = 0.4\%$ . The risk free rate is  $r = 2\%$  and the average market risk premium  $\lambda = 0.132$ .

a calibration of the Black Cox model further from the Merton model. Unreported results show that choosing e.g.  $\beta = 80\%$  does not change our conclusions.

Figure 6 shows that the estimated  $DD$ 's ability to rank firms' default risk relative to the true  $DD$  is only slightly weaker in the Black Cox model, and also confirmed by a Spearman rank correlation of 0.99. The good performance of the VX-algorithm can to some extent be explained by the fact that the underlying asset value does indeed follow a geometric Brownian motion. Yet, the CAP curves indicate that the change in default triggering mechanism does not significantly affect the ability of  $DD_{VX}$  to rank firms' default risk. Here we observe that the CAP curve generated by leverage is somewhat closer to the other two CAP curves, and thereby the volatility estimate is relatively less important compared to its importance in the previously analyzed models.

### 5.2.5 Ranking in the Leland-Toft model

Finally, we study the robustness of the  $DD$ -measure for the Leland-Toft model incorporating bankruptcy costs and tax advantage of debt in which firms default the first time an endogenous default boundary is hit. We use the parameters of He & Xiong (2012),

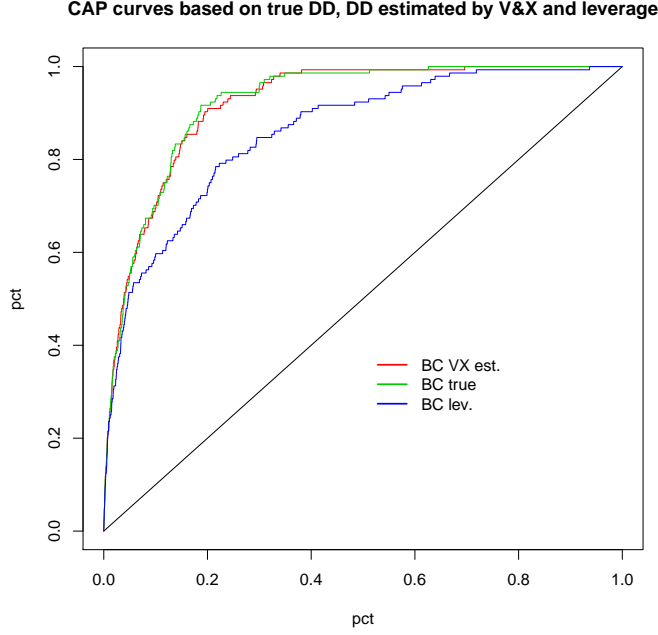


Figure 6: CAP curves for the estimated and true  $DD$  and leverage in the Black Cox model. Initial leverage ranges from 20% – 70%, and volatility is chosen such that initial default probability is 1.3%. This implies a  $\sigma$  of 48.6% – 13.2%. The default barrier is chosen to  $\beta = 70\%$  of face value of debt. The risk free rate is  $r = 2\%$  and the market risk premium  $\lambda = 0.132$ .

who provide a careful discussion of each parameter choice. They employ a slightly higher recovery value,  $\alpha$ , and a slightly lower tax rate,  $\tau$ , than the original article of Leland & Toft (1996).

The CAP curves shown in Figure 7 indicate that the estimated  $DD$  performs well for ranking firms' default risk in this model, and the rank correlation is also high at 0.95. As for the Black Cox model a part of the explanation lies in the fact that the underlying asset value does indeed follow a geometric Brownian motion, and since the endogenous default barrier in this model is even lower than the exogenous barrier applied in the Black Cox model for most firms, defaults prior to debt's maturity will be even less important. Yet, the capital structure in this model is quite different from Merton's assumptions but according to the visual inspection this does not influence the estimated  $DD$ 's ranking significantly.

The conclusion from the graphical inspections in this section is that for most violations of the Merton model's simplifying assumptions, the distance-to-default performs well for ranking firms' default risk. Neither a more complex capital structure nor market frictions seem to visibly reduce  $DD$ 's ability to rank firms' default risk compared to the true default probability. However, the estimated  $DD$ 's ranking is less accurate when the underlying asset value does not follow a geometric Brownian motion, in particular stochastic volatility

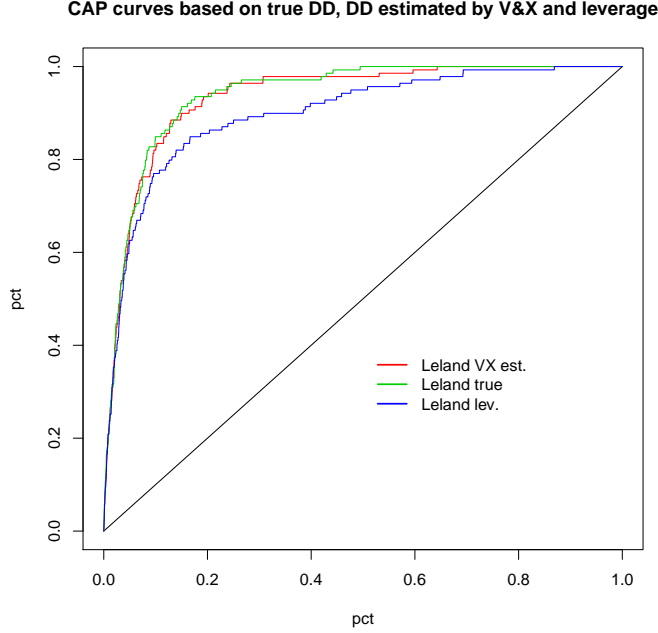


Figure 7: CAP curves for the estimated and true  $DD$  and leverage in the Leland-Toft model. Initial leverage ranges from 20% – 70%, and volatility is chosen such that initial default probability is 1.3%. This implies a  $\sigma$  of 45.5% – 10.9%. The marginal tax rate is  $\tau = 0.27$ , yearly coupon rate is  $C = 6$ , and the debt holders recover a fraction  $\alpha = 60\%$  of firm value in bankruptcy. The risk free rate is  $r = 2\%$  and the market risk premium  $\lambda = 0.132$ .

and jumps in asset value may challenge the ranking by distance-to-default.

Not surprisingly, the performance of leverage as risk score is generally poorer than distance-to-default. However, for models where firms default the first time their asset values hit a default boundary related to leverage, leverage CAP curves are closer to  $DD$  CAP curves, which indicates that the volatility estimate is relatively less important in these models.

### 5.3 Statistical test of ranking performance

In this section we use the test statistic in (3) to test the hypothesis that the ROC of the true  $DD$ 's ranking equals the ROC of the estimated  $DD$ 's ranking. Table 2 shows the results of the test for each of the six model specifications. As we would expect, the test is accepted for the Merton model; the area under the ROC curve generated by VX-estimated distance-to-default is not significantly different from the area under the curve generated by the true default probability. For the dual business and Black Cox models, we also clearly accept the hypothesis, whereas the p-values of the test for the jump-diffusion and Leland-

Toft model specifications are only just over the 5% acceptance level. The test clearly rejects the hypothesis when asset value has stochastic volatility as was already clear from the CAP curves. For a jump specification with very rare and potentially large jumps ( $\lambda = 0.16$ ,  $\mu_J = 1.2\%$  and  $\sigma_J = 19\%$ ) the test is also clearly rejected with a  $p$ -value of 0.1%. Overall, the formal statistical test conducted here confirms the conclusions of the visual inspection of the CAP curves.

	$ROC_{\text{true}}$	$ROC_{\text{VX}}$	$T$	$p$ -value (%)
Mer.	0.922	0.920	0.76	38.5
JD	0.942	0.936	2.2	14.1
SV	0.944	0.933	9.0	0.26
Dual	0.944	0.941	2.3	13.1
BC	0.929	0.927	1.2	26.7
LT	0.946	0.941	3.4	6.4

Table 2: For each model, the table reports the area,  $ROC_{\text{true}}$ , under the ROC-curve generated by the true distance-to-default and the area,  $ROC_{\text{VX}}$ , under the ROC-curve generated by the VX-estimated distance-to-default. We also report the test statistic,  $T$ , in (3) for testing the hypothesis that  $ROC_{\text{true}} = ROC_{\text{VX}}$  together with its  $p$ -value.

## 5.4 Economic benefit of powerful ranking method

Here we calculate the economic benefit of a bank that knows the firms' ranking by the true default probability compared to a bank that grants loans based on an estimate of the distance-to-default. Both banks offer loans to firms they perceive as healthy and calculate spreads based on their credit scoring model, but firms only accept the loan from the bank offering the lowest spread. To obtain a robust estimate of the banks' average returns we follow Stein & Jordão (2003) and draw 1000 samples each consisting of 200 firms from our originally simulated pool of firms. For each sample, we assign risk scores to all firms according to the true default probability and to the estimated distance-to-default, and from equation (4) we then calculate the spread offered to each firm by each bank. Next, for each bank we find the return according to equation (5) for the loans they grant. Results for the two banks' average returns, market shares and shares of defaults are provided in Table 3. Note, that since not every firm is offered a loan, the sum of the banks' market shares and their shares of defaults are below one.

Table 3 confirms our previous findings. In the original Merton, dual business and Black Cox models the additional return earned by ranking firms by their true default probability is very small. For the Leland-Toft model the difference is slightly higher at 6 bp. For firms with frequent but mostly small jumps (parameters from Table 1) the difference in returns is also moderate as reported in Table 3, whereas in a sample of firms with very rare and

	$R_{\text{true}}$ (bp)	$R_{\text{VX}}$ (bp)	$MS_{\text{true}}$ (%)	$MS_{\text{VX}}$ (%)	$DS_{\text{true}}$ (%)	$DS_{\text{VX}}$ (%)
Mer.	21	21	50.2	45.7	30.3	21.3
JD	22	16	53.2	43.1	18.0	23.4
SV	24	15	56.6	39.5	20.9	19.7
Dual	22	20	46.4	49.6	13.7	28.8
BC	23	21	50.1	45.8	24.9	21.4
LT	23	17	51.8	44.2	14.7	25.9

Table 3: For each model, the table reports the return ( $R$ ), market share ( $MS$ ) and share of defaults ( $DS$ ) for the bank ranking firms by their true default probabilities (bank 1) and for the bank ranking firms by their VX-estimated distance-to-defaults (bank 2). E.g. for the sample of firms generated by the Merton model, bank 1 will on average earn a return of 21 bp on their lending activities, and also the competing bank 2 will earn a return of 21 bp on average. Bank 1 lends out to 50.2% of the firms in the sample, 45.7% of the firms choose loans from bank 2, whereas the remaining firms are refused loans from both banks. Of the firms, that end up in default, 30.3% had a loan in bank 1, whereas 21.3% had a loan in bank 2.

potentially large jumps ( $\lambda = 0.16$  and  $\sigma_J = 19\%$ ) the bank ranking by the true default probability would have a greater advantage ( $R_{\text{true}} = 23$  bp vs.  $R_{\text{VX}} = 15$  bp). For firms with stochastic volatility in asset value a bank that ranks firms by  $DD_{\text{VX}}$  can potentially increase its return by up to 9 bp by shifting to a more powerful model.

The low return for the bank ranking firms by their estimated distance-to-default can be caused either because the bank is unable to set competitive spreads and therefore loses a large market share, which is the primary problem in the stochastic volatility model and jump diffusion model, or because it ends up lending to many firms that end up defaulting as is the case in the dual business and Leland-Toft model specifications. For the Merton model, and to some extent also the Black-Cox model, the bank using  $DD_{\text{VX}}$  as risk score is surprisingly successful in refusing loans to firms that later end up in default, but since it is unable to set competitive spreads it does not perform better than the bank having the true default probability as default score.

## 6 Ranking of stochastic volatility firms

Our results so far indicate that jumps and in particular stochastic volatility in the firm asset dynamics pose the biggest challenges to the robustness of the distance-to-default as credit score. Here we study the case where firms have stochastic volatility in more detail and propose an adjustment to the  $DD$  measure that accounts for stochastic volatility.

The VX-estimation can be mistaken about the ranking of firms with stochastic volatility both because it may estimate firm asset value incorrectly, and because the VX-volatility

estimate is a constant, that cannot capture possible future changes in volatility. It turns out, that the asset value estimate is surprisingly close to the true value (see Table 4) even when the VX-estimated volatility is far from the realized volatility. Therefore, the robustness of  $DD$  is primarily weakened because of its constant volatility assumption. From Table 4 we see that the VX-ranking makes the biggest mistakes (i.e. estimates  $DD_{VX}$  relatively high at time  $T_1$  for firms that actually end up in default) for firms whose estimated volatility is far from the realized volatility at the ranking time point,  $T_1$ , and in particular, far from the realized volatility path over the period  $t \in [T_1, T_2]$ . Therefore, when ranking firms' default risk according to the estimated  $DD$ , we risk mis-judging the riskiness of firms whose volatility path during the estimation period turn out significantly lower than after the ranking is done.

	$(v_t)_{t \in [0, T_1]}$	$(v_t)_{t \in [T_1, T_2]}$	$v_{T_1}$	$\hat{\sigma}_{VX}$	$\frac{\hat{V}_{VX}(T_1) - V_{\text{true}}(T_1)}{V_{\text{true}}(T_1)}$ (abs)
total sample	24.4%	24.0%	24.1%	24.4%	0.07% (0.07%)
survivors	24.3%	23.8%	24.0%	24.3%	0.06% (0.06%)
defaults	28.5%	33.5%	31.4%	28.1%	0.85% (0.87%)
defaults with $DD_{VX} > 2$	22.2%	28.0%	24.3%	21.6%	0.17% (0.17%)

Table 4: The table provides mean values of the volatility paths during the estimation period,  $(v_t)_{t \in [0, T_1]}$ , after the ranking,  $(v_t)_{t \in [T_1, T_2]}$ , the mean volatility at the ranking time point,  $v_{T_1}$  and the VX-algorithm's volatility estimate,  $\hat{\sigma}_{VX}$ . The table also reports average (absolute) relative firm value estimation errors,  $\frac{\hat{V}_{VX}(T_1) - V_{\text{true}}(T_1)}{V_{\text{true}}(T_1)}$  ( $|\frac{\hat{V}_{VX}(T_1) - V_{\text{true}}(T_1)}{V_{\text{true}}(T_1)}|$ ).

The reason that the VX-algorithm is fairly precise in its estimation of firm value is that it is the observed path of equity values, not the volatility parameter, that mainly determines the asset value path in the VX-algorithm's inversion of the Black-Scholes formula. For a given equity value, a change of one percentage point in the volatility parameter merely leads to a change in asset value of less than 0.2, whereas for a given volatility parameter a change of 1 in equity value leads to a change of at least 1 in asset value. Furthermore, for fixed volatility and equity values the Black Scholes inversion and the corresponding inversion in the stochastic volatility model produce similar results in most cases; the asset values produced by the two inversion formulas will differ significantly (up to 10%) only for low equity values combined with a VX-estimated volatility far from the realized volatility.

## 6.1 Volatility adjustment of distance-to-default

The relatively poor performance of the VX-algorithm for ranking the default risk of firms with stochastic volatility leads us to consider an adjustment of the traditional distance-to-default measure that takes stochastic volatility into account. The key observation from Table 4 is that mainly the deviation of volatility at the ranking time point,  $v_{T_1}$ , from its

long term mean,  $\theta$ , leads to errors in the ranking and therefore we need a volatility measure that incorporates this.

Our volatility adjustment is inspired by Packham, Schlögl & Schmidt (2013) who consider a first passage time model, where the firm defaults the first time,  $\tau$ , its credit quality process,  $X$ , hits a barrier  $b < 0$ :  $\tau = \inf\{t > 0 \mid X_t \leq b\}$ . The credit quality is specified by  $X_t = \int_0^t \sigma_s dW_s$ , where  $W_t$  is a Brownian motion and  $\sigma_t$  a strictly positive, càdlàg process independent of  $W_t$ . Packham et al. (2013) show that in this model the  $T$ -year default probability can be calculated as

$$\mathbb{P}(\tau \leq T) = 2E^{\mathbb{P}} \left[ N \left( \frac{b - X_0}{\sqrt{QV_T}} \right) \right],$$

where  $QV_T$  is the  $T$ -year quadratic variation of  $X$ . For a credit quality process with deterministic quadratic variation, this translates to a "distance-to-default" given by  $\frac{X_0 - b}{\sqrt{QV_T}}$ .

The stochastic volatility model considered in our experiment does not quite fit into this setup, firstly because default happens only at maturity, not the first time the barrier  $b$  is crossed, secondly because the stochastic volatility specification in (6)–(7) does not fit into the format of Packham et al. (2013), and thirdly because our "credit quality process" (i.e. log firm asset value) has drift. Nevertheless, we propose to rank firms by the following volatility adjusted distance-to-default measure:<sup>6</sup>

$$DD_V = \frac{\ln V - \ln P}{\sqrt{E^{\mathbb{P}}[QV_T]}}. \quad (8)$$

The quadratic variation takes both the long term level of volatility as well as the current spot volatility into account, and therefore we expect that the quadratic variation is better able to address the problem of ranking firms whose realized volatility at the time of ranking is far from its long term volatility as illustrated in Table 4.

In order to calculate  $DD_V$  we need an estimate of the expected quadratic variation of the asset value, which is considered an unobservable variable. However, since Table 4 indicates that the VX-algorithm is able to estimate the firms' assets fairly accurately, we will use the VX-estimated asset value path over the period  $[0, T_1]$  to estimate the quadratic variation. It is commonly accepted that volatility follows a mean reverting process, and given that this is diffusion driven the expected value of the  $t$ -year quadratic variation as seen from time  $t = 0$  is given by

$$E^{\mathbb{P}}[QV_t] = \theta t + \frac{1}{\kappa}(\theta - v_0)(e^{-\kappa t} - 1) \quad (9)$$

In our experiment we assume that the stochastic volatility follows a CIR-process (7), which has exactly this expected quadratic variation. Our experiment would be biased if we apply

---

<sup>6</sup>We also conducted this experiment with  $DD_V = \frac{\ln(\frac{V}{P}) + \mu T - \frac{1}{2} E^{\mathbb{P}}[QV_T]}{\sqrt{E^{\mathbb{P}}[QV_T]}}$ , however the simpler formulation proved more robust in the simulation experiment.

information about the underlying model in our estimation, however, the expected quadratic variation in (9) applies for a large class of stochastic volatility models.

To estimate the parameters,  $\theta$  and  $\kappa$ , and the state variable,  $v_0$ , in (9) we turn to the literature on stochastic volatility models for stock prices, where different methods are proposed. ? derive a maximum likelihood procedure to estimate all model parameters simultaneously under both the  $\mathbb{P}$  and  $\mathbb{Q}$ -measure, whereas e.g. Reno (2006) and ? use realized volatility to proxy for the spot volatility process and estimate the volatility parameters directly from this. We will use the realized volatility approach in our estimation.

Whereas volatility estimation in stock pricing models commonly use intraday data to calculate the realized variance on a daily basis, we face the challenge of having only daily "observations" of the asset value process available. Therefore, we use daily data over the period of two weeks ( $\Delta t = \frac{1}{26}$ ) to calculate the realized variance,  $RV$ , on a biweekly basis:

$$RV_t = \frac{1}{\Delta t} \sum_{\substack{i=2 \\ s_i \in [t-\Delta t; t]}}^{14} (\ln V_{s_i} - \ln V_{s_{i-1}})^2.$$

This gives us 26 observations of the realized variance if considering an estimation period of one year ( $T_1 = 1$ ) as we have done up until now. To obtain a longer time series, we extend the estimation period to two years ( $T_1 = 2$ ) such that we have 52 observations to estimate  $\kappa$  and  $\theta$ . The spot volatility at the ranking time point,  $v_{T_1}$ , is set equal to the realized volatility over the preceding two weeks.

We follow the estimation approach of Reno (2006) who uses the realized volatility path to obtain a nonparametric estimate of the linear drift function of the volatility process. We then find the parameters as the slope ( $-\hat{\kappa}$ ) and interception ( $\hat{\kappa}\hat{\theta}$ ) of the drift function. For our simulated sample we find an average mean reversion estimate of  $\hat{\kappa} = 1.7$  compared to the true value  $\kappa = 0.21$ ,<sup>7</sup> and an average long term volatility estimate of  $\hat{\theta} = 0.079$  compared to the true average of  $\theta = 0.073$ . The average spot volatility at time  $T_1$  is  $v_{T_1} = 0.074$ , whereas the average realized volatility is  $\hat{v}_{T_1} = 0.066$ . The mean estimated quadratic variation is 0.073, close to that calculated with the true model parameters, which is 0.074.

In Figure 8 we compare the ranking ability of  $DD_V$  to that of the true default probability and the VX-estimated  $DD$  in the stochastic volatility model. Notice that with a 2-year estimation period the CAP-curves are closer and the ROC-numbers higher than in our original experiment in Figure 4. The primary explanation lies in our experimental design, where we start out all firms with the same initial asset value at time  $t_0$ , and over a 2-year estimation period their asset values spread out more than over a 1-year period. This makes it easier for our credit scoring models to distinguish good from bad firms and

---

<sup>7</sup>It is well known from the stock pricing literature, that especially the volatility's mean reversion parameter is hard to estimate precisely. Since we use low-frequency estimates of the asset value process (which is itself an estimated process) compared to the stock pricing literature's high-frequency data, we are satisfied with this estimation precision.

therefore improves the ranking compared to our original experiment. Moreover, with a longer estimation period the VX-algorithm produces a more precise estimate of the average volatility level. Still, the CAP curves show a clear improvement in the ranking, when using  $DD_V$  instead of  $DD_{VX}$ . This is also confirmed by the receiver operating characteristics in Table 5. The ROC-test now accepts the hypothesis that  $ROC_{\text{true}} = ROC_V$ , while still  $ROC_{\text{true}} = ROC_{VX}$  is rejected, although not as firmly as in Table 2. Yet, the ROC-test also show that  $ROC_V$  is not significantly different from  $ROC_{VX}$ . The volatility adjusted distance-to-default is highly correlated with the VX-estimated distance-to-default with a Spearman's  $\rho$  of 95%, yet its correlation with the true default probability is slightly higher than the correlation between  $DD_{VX}$  and  $DD_{\text{true}}$ , see Table 6.

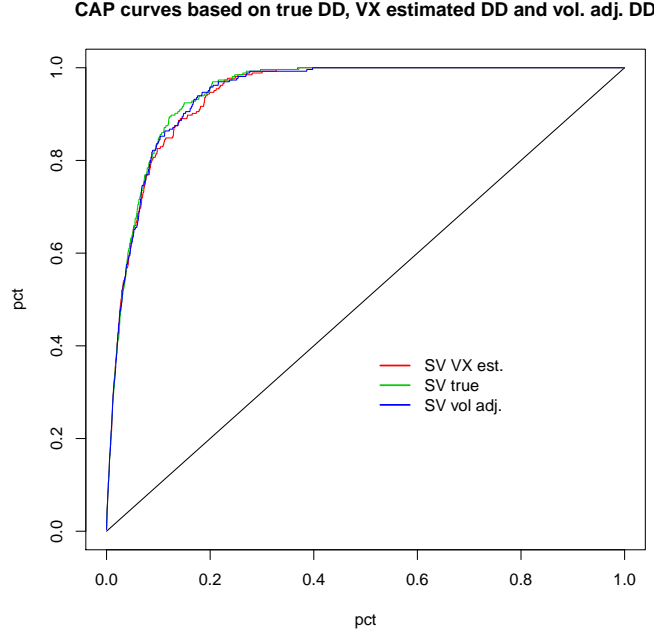


Figure 8: CAP curves for the VX-estimated, the true  $DD$  and the volatility adjusted  $DD_V$  in the stochastic volatility model. Parameters are as in Table 1, except here the estimation period is extended to 2 years ( $T_1 = 2$ ).

Even if the volatility adjustment clearly improves the ranking performance in the stochastic volatility model, we see that the performance of  $DD_V$  is poorer than  $DD_{VX}$  for the Merton and in particular for the jump diffusion model as shown in figures 9 and 10. This is also confirmed by the ROC-numbers in Table 5, where we see that difference between  $ROC_{\text{true}}$  and  $ROC_V$  is borderline significant at the 5% level for the Merton model and highly significant for the jump-diffusion model.

In conclusion, we see a significant improvement in the ranking by the volatility adjusted distance-to-default compared to the original measure when firms have stochastic volatility,

but the volatility adjustment is less robust to model misspecifications compared to the VX-estimated distance-to-default.

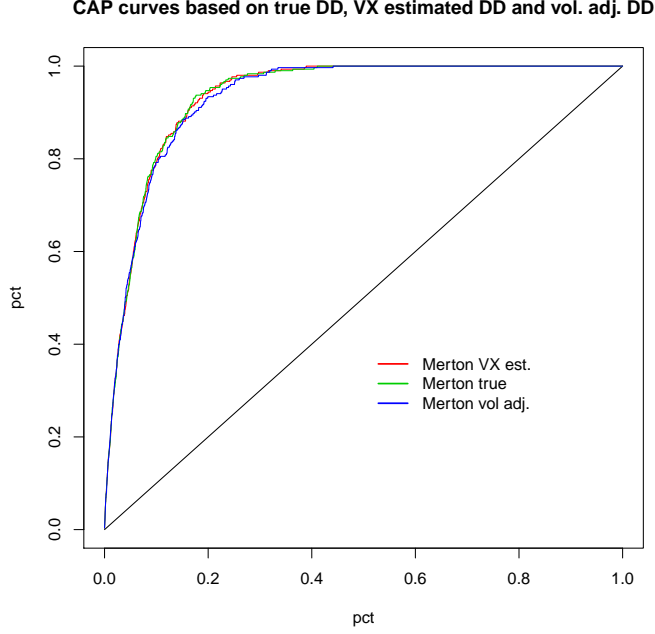


Figure 9: CAP curves for the VX-estimated, the true  $DD$  and the volatility adjusted  $DD_V$  in the Merton model. Parameters are as in Table 1, except here the estimation period is extended to 2 years ( $T_1 = 2$ ).

	$ROC_{\text{true}}$	$ROC_V$	$ROC_{VX}$	$H_1$ p-value	$H_2$ p-value	$H_3$ p-value
Merton	0.952	0.949	0.952	3.2%	25%	4.0%
JD	0.948	0.940	0.945	0.1%	9.2%	2.4%
SV	0.960	0.959	0.956	65%	5.6%	23.4%

Table 5: The first three columns in the table provide receiver operating characteristics for the ranking by the true default probability,  $DD_V$  and  $DD_{VX}$  for the Merton, jump-diffusion and stochastic volatility models. The three last columns reports the p-values for tests of the hypotheses  $H_1 : ROC_{\text{true}} = ROC_V$ ,  $H_2 : ROC_{\text{true}} = ROC_{VX}$  and  $H_3 : ROC_V = ROC_{VX}$ .

## 7 Conclusion

We have found that distance-to-default is a robust measure for ranking firms according to their default risk under most violations of the assumptions underlying the Merton model

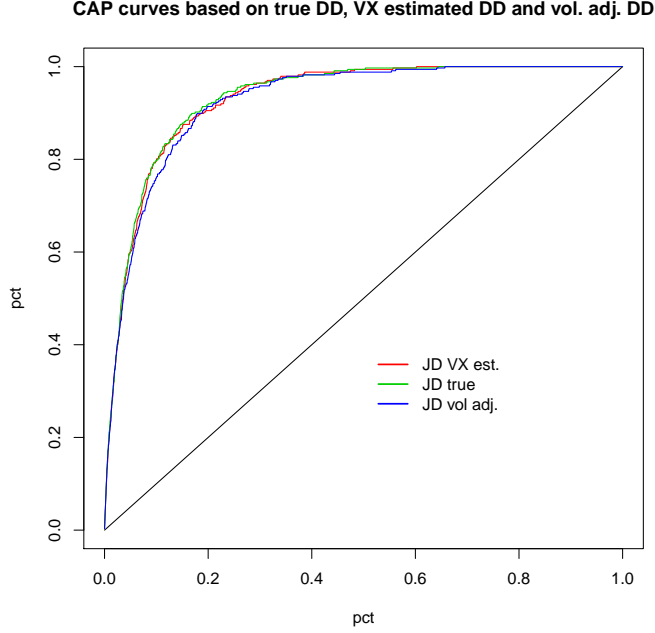


Figure 10: CAP curves for the VX-estimated, the true  $DD$  and the volatility adjusted  $DD_V$  in the jump diffusion model. Parameters are as in Table 1, except here the estimation period is extended to 2 years ( $T_1 = 2$ ).

Rank correlation	$\rho(DD_{\text{true}}, DD_V)$	$\rho(DD_{\text{true}}, DD_{VX})$	$\rho(DD_V, DD_{VX})$
Merton	95.6%	99.7%	95.9%
JD	89.6%	97.0%	89.9%
SV	89.1%	88.5%	95.1%

Table 6: The table provides pairwise Spearman rank correlations for  $DD_{\text{true}}$ ,  $DD_V$  and  $DD_{VX}$  for the Merton, jump diffusion and stochastic volatility models.

that we consider. However, when the asset value process may experience large jumps or has stochastic volatility, we see a significantly poorer ranking performance documented by both a lower CAP curve and by less profitable lending decisions for a bank using distance-to-default for credit scoring.

Ignoring stochastic volatility will lead to a significant underestimation of the credit risk of firms that happen to experience low asset volatility during the estimation period but which are in fact at risk of a significant increase in future asset volatility. We propose a volatility adjustment to the traditional distance-to-default, that accounts for the potential difference between the spot volatility of assets and the long term level. This significantly improves the ranking of the firms' default risk when the true model has stochastic volatility, but it shows less robustness than the traditional  $DD$  measure when the true model has

constant volatility or jumps in asset value. Hence, the increase in precision comes at the cost of lower robustness.

Clearly, there may be other covariates than DD that are relevant for predicting firm defaults. For example, the empirical literature (e.g. Campbell et al. (2008) and Bharath & Shumway (2008)) documents that measures of a firm’s profitability such as net income have strong explanatory power for the firm’s default risk. Others, such as Davydenko (2012*a*) show empirically that a significant fraction of defaults happens because of illiquidity rather than insolvency. Gryglewicz (2011) introduces cash holdings to the Leland-Toft model and thereby combines liquidity and solvency concerns of the firm. He & Xiong (2012) extend the Leland-Toft model to include liquidity shocks to the corporate bond market. For this particular model, unreported results show that distance-to-default’s ranking performance is very similar to that in the Leland-Toft model.

As a final caveat, note that we have only checked robustness within reasonable parameter ranges. We are able to set up samples of extreme firms in each model (i.e. firms with extreme parameters) for which distance-to-default is unable to provide a precise ranking of the firms’ default risk. Yet, the aim of our study is to draw conclusions for a simulated sample of firms whose characteristics are empirically plausible.

## References

- Acharya, V. V., Lochstoer, L. A. & Ramadorai, T. (2013), ‘Limits to arbitrage and hedging: Evidence from commodity markets’, *Journal of Financial Economics* **109**, 441–465.
- Arora, N. & Sellers, M. (2004), ‘Financial EDF Measures: A new Model of Dual Business Lines’, *Moody’s KMV*.
- Aït-Sahalia, Y. & Kimmel, R. (2007), ‘Maximum likelihood estimation of stochastic volatility models’, *Journal of Financial Economics* **83**, 413–452.
- Bharath, S. & Shumway, T. (2008), ‘Forecasting Default with the Merton Distance to Default Model’, *Review of Financial Studies* **21**, 1339–1369.
- Black, F. & Cox, J. (1976), ‘Valuing corporate securities: some effects of bond indenture provisions’, *Journal of Finance* **31**, 351–367.
- Blöchliger, A. & Leippold, M. (2006), ‘Economic benefit of powerful credit scoring’, *Journal of Banking & Finance* **30**, 851–873.
- Bu, D. & Liao, Y. (2013), Structural Credit Risk Model with Stochastic Volatility: A Particle-filter approach. Working paper.
- Campbell, J., Hilscher, J. & Szilagyi, J. (2008), ‘In Search of Distress Risk’, *Journal of Finance* **63**, 2899–2939.
- Chava, S. & Purnanandam, A. (2010), ‘Is Default Risk Negatively Related to Stock Returns?’, *Review of Financial Studies* **23**, 2523–2559.

- Crosbie, P. & Bohn, J. (2003), ‘Modelling Default Risk’, *Moody’s KMV*.
- Davydenko, S. A. (2012a), Insolvency, Illiquidity, and the Risk of Default. Working paper.
- Davydenko, S. A. (2012b), When do firms default? A study of the default boundary. Working paper.
- Duan, J.-C. (1994), ‘Maximum likelihood estimation using price data of the derivative contract’, *Mathematical Finance* **4**, 155–167.
- Duffie, D., Pan, J. & Singleton, K. (2000), ‘Transform analysis and asset pricing for affine jump-diffusions’, *Econometrica* **68**, 1343–1376.
- Duffie, D., Saita, L. & Wang, K. (2007), ‘Multi-period corporate default prediction with stochastic covariates’, *Journal of Financial Economics* **83**, 635–665.
- Engelmann, B., Hayden, E. & Tasche, D. (2003), ‘Testing rating accuracy’, *RISK* **16**, 82–86.
- Gryglewicz, S. (2011), ‘A theory of corporate financial decisions with liquidity and solvency concerns’, *Journal of Financial Economics* **99**, 365–384.
- He, Z. & Xiong, W. (2012), ‘Rollover Risk and Credit Risk’, *Journal of Finance* **67**, 391–430.
- Hillegeist, S. A., Keating, E. K. & Cram, D. P. (2004), ‘Assessing the Probability of Bankruptcy’, *Review of Accounting Studies* **9**, 5–34.
- Kealhofer, S. (2003), ‘Quantifying Credit Risk I: Default Prediction’, *Financial Analysts Journal* **January/February**, 30–44.
- Kristensen, D. (2010), ‘Nonparametric filtering of the realized spot volatility: A kernel-based approach’, *Econometric Theory* **26**, 60–93.
- Landier, A., Krüger, P. & Thesmar, D. (2011), The WACC Fallacy: The Real Effects of Using a Unique Discount Rate. Working paper.
- Lando, D. (2004), *Credit Risk Modeling*, Princeton.
- Lando, D., Medhat, M., Nielsen, M. S. & Nielsen, S. F. (2013), ‘Additive Intensity Regression Models in Corporate Default Analysis’, *Journal of Financial Econometrics* **11**, 443–485.
- Lando, D. & Nielsen, M. S. (2010), ‘Correlation in Corporate Defaults: Contagion or Conditional Independence?’, *Journal of Financial Intermediation* **19**, 355–372.
- Leland, H. & Toft, K. B. (1996), ‘Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads’, *Journal of Finance* **51**, 987–1019.
- Merton, R. (1974), ‘On the pricing of corporate debt: the risk structure of interest rates’, *Journal of Finance* **29**, 449–470.
- Merton, R. (1980), ‘On estimating the expected return on the market’, *Journal of Financial Economics* **8**, 323–361.

- Packham, N., Schlögl, L. & Schmidt, W. (2013), ‘Credit gap risk in a first passage time model with jumps’, *Quantitative Finance* **forthcoming**.
- Reno, R. (2006), ‘Nonparametric estimation of stochastic volatility models’, *Economics Letters* **90**, 390–395.
- Shumway, T. (2001), ‘Forecasting Bankruptcy more accurately: A simple hazard model’, *Journal of Business* **74**, 101–124.
- Stein, R. (2005), ‘The relationship between default prediction and lending profits: Integrating ROC analysis and loan pricing’, *Journal of Banking & Finance* **29**, 1213–1236.
- Stein, R. & Jordão, F. (2003), ‘What is a more powerful model worth?’, *Moody’s KMV*.
- Vassalou, M. & Xing, Y. (2004), ‘Default Risk in Equity Returns’, *Journal of Finance* **59**(2), 831–868.
- Wong, H. Y. & Choi, T. W. (2009), ‘Estimating default barriers from market information’, *Quantitative Finance* **9**(2), 187–196.
- Wong, H. Y. & Li, C. P. (2006), Estimating Jump Diffusion Structural Credit Risk models. Working paper.
- Zhang, B. Y., Zhou, H. & Zhu, H. (2009), ‘Explaining Credit Default Swap Spreads with the Equity Volatility and Jump Risks of Individual Firms’, *Review of Financial Studies* **22**(12), 5099–5131.