Maximum Likelihood Estimation of Structural Credit Spread Models - Deterministic and Stochastic Interest Rates

Jin-Chuan Duan, Geneviève Gauthier, Jean-Guy Simonato and Sophia Zaanoun*

(October 2002; incomplete)

Abstract

One difficulty in implementing structural credit spread models is that the underlying asset value cannot be directly observed. Models require the unobserved asset value and the unknown parameter(s) as inputs; for example, asset value and volatility are in practice unknown when the model of Merton (1974) is applied. This paper applies the maximum likelihood principle to the estimation of structural credit spread models. The maximum likelihood estimator for parameter(s), asset value, credit spread and default probability are derived for models with deterministic and stochastic interest rates. Monte Carlo studies are conducted to examine the performance of the maximum likelihood method in finite samples specifically for the structural models of Merton (1974) and Longstaff and Schwartz (1995).

Keywords: Credit risk, maximum likelihood, option pricing, Monte Carlo simulation.

JEL classification code: C22.

---

*Duan, Rotman School of Management, University of Toronto. Gauthier and Simonato, HEC (Montréal). Zaanoun is a Ph.D. candidate at HEC (Montréal). Duan, Gauthier and Simonato acknowledge the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC), Les Fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec (FCAR) and from the Social Sciences and Humanities Research Council of Canada (SSHRC).

1 Introduction

In Merton (1974), a pricing model for corporate liabilities was developed using an option valuation approach. In his setting, the unobserved asset value of the firm is governed by a geometric Brownian motion. Subsequently, many variants of this model have been proposed in the literature. Merton's and its extended models are typically referred to as structural credit spread (or risky bond) models. Examples abound; Longstaff and Schwartz (1995), Madan and Unal (1998) and Collin-Dufresne and Goldstein (2001). This paper develops a maximum likelihood estimation method for this class of models.

As pointed out in Jarrow and Turnbull (2000) among others, there are several limitations associated with the implementation of structural credit spread models. First, the asset value is an unobserved quantity. This in turn creates problems with the estimation of the various required parameters such as the drift and volatility of the asset value process and the correlation among different asset value processes. Other limitations are related to default threshold and stochastic interest rates because the parameter values specific to these features are also unobserved.

In the academic literature, two approaches have been proposed for dealing with the estimation problem when the underlying asset value is unobserved. The first approach, which we will refer to as the implicit estimation method, uses some observed quantities and the corresponding restrictions derived from the theoretical model to extract point estimates for the model parameter(s) and the unobserved asset value. Take the univariate case of Merton's (1974) model as an example. The implicit estimation method relies on two equations: one relating the equity value to the asset value and the other relating the equity volatility to the asset volatility. The two-equation system can then be solved for the two unknown variables: the asset value and volatility. The implicit estimation method has been adopted by Ronn and Verma (1984) to implement the deposit insurance pricing model of Merton (1978) and by Jones, Mason and Rosenfeld (1984) to conduct an empirical study of Merton's (1974) risky bond pricing model. A three-equation extension of the implicit estimation method was used in Duan, Moreau and Sealey (1995) to implement their deposit insurance model with stochastic interest rate where the third equation relates the equity duration to the asset duration.

The second estimation approach was proposed by Duan (1994, 2000). In these papers, a likelihood function based on the observed equity values is derived by employing the transformed data principle in conjunction with the equity pricing equation. With the likelihood function in place, maximum likelihood estimation and statistical inference become straightforward. The maximum likelihood method was applied to Merton's (1978) deposit insurance pricing model in Duan (1994), Duan and Yu (1994), and Laeven (2002). Later, Duan and Simonato (2002) extended the method to deposit insurance pricing under stochastic interest rate. For credit risk, the estimation method has been applied to a strategic corporate bond pricing model by Ericsson and Reneby (2001).

Theoretically, the maximum likelihood estimation method has several advantages relative to the implicit estimation method. First, the maximum likelihood method provides an estimate of the drift of the unobserved asset value process under the physical probability measure. This can in turn be used to obtain an estimate of the default probability of the firm. Such an estimate is not available within the context of the implicit estimation method since the theoretical equity pricing equation typically does not contain the drift of the asset value process under the physical probability measure. The second advantage is associated with the asymptotic properties of the maximum likelihood estimator such as consistency and asymptotic normality, which in turn allows for statistical inference to assess the quality of parameter estimates and/or perform testing on the hypotheses of interest. In contrast, consistency is unattainable with the implicit estimation method because it erroneously forces a stochastic variable to be a constant (see Duan (1994)). Consequently, no reliable statistical inference using the implicit estimation method can be expected.

To our knowledge, the maximum likelihood estimation method, except for Ericsson and Reneby (2001), has not been applied to the credit spread models. Particularly, it has not been applied in the portfolio context of credit risk assessment or to the credit risk models with stochastic interest rates. In the credit risk context, the first important problem is the estimation of correlations among different firms. Indeed, for credit risk it is essential to correctly assess correlations because standard credit risk management methods such as CreditMetrics and KMV are highly sensitive to the correlation coefficients of asset returns (see Crouhy and Mark (1998)). The second problem is related to the unobserved financial distress level at which default is triggered. This issue is inherently critical to the models of Longstaff and Schwartz (1995) and Collin-Dufresne and Goldstein (2001).

This paper adopts the transformed data method of Duan (1994) to come up with a practical maximum likelihood estimation procedure for the structural credit spread model with or without stochastic interest rates. Specifically, we develop a method for the constant interest rate credit spread model of Merton (1974) in a portfolio context, i.e., more than one asset. Moreover, we cast the so-called KMV method (as described in Crouhy *et al.* (2000)), a popular commercial implementation of Merton's structural credit spread model, in relation to the maximum likelihood method and study its performance. We then consider the Longstaff and Schwartz (1995) model which incorporates financial distress. Before proceeding to the full model, we analyze a reduced version of the Longstaff and Schwartz (1995) model by setting the interest rate to a constant. This reduced Longstaff and Schwartz model can also be viewed as a modified Merton's model for it specifically allows for default triggering based on financial distress. Finally, we tackle the full version of the Longstaff and Schwartz (1995) model.

The Longstaff and Schwartz (1995) model has a peculiar property that complicates the estimation problem. Under some parameter values, a risky bond price can increase faster than the asset value of the firm. In other words, one dollar increase in the firm's asset value causes more than one dollar increase in the risky bond value. Conceptually, this implies that injecting equity capital into the firm could actually cause the total equity value to drop. From the estimation point of view, the relationship between the asset and equity values is not always a one-to-one transformation, a critical assumption underlying the approach of Duan (1994). In this paper, we have devised a solution to this estimation problem created by the Longstaff and Schwartz (1995) model.

For both versions of the Longstaff and Schwartz (1995) model, we factor in survivorship, because the typical data is available for survived firms and assessing credit spreads for firms that have survived is really the purpose of credit spread models. These maximum likelihood estimation procedure is implemented empirically on actual data and its finite sample performance is analyzed using a Monte Carlo study.

2 Merton's credit spread model

In the Merton (1974) framework, firms have a very simple capital structure. It is assumed that the ith firm is financed by equity with a market value $S_{i,t}$ at time t and a zero-coupon debt instrument with a face value of F_i maturing at time T_i. Let $V_{i,t}$ be the asset value of the firm and $D_{i,t}(\sigma_{V_i})$ its risky zero-coupon bond value at time t. Let us consider m firms. Naturally, the following accounting identity holds for every time point and for every firm:

$$V_{i,t} = S_{i,t} + D_{i,t}\left(\sigma_{V_i}\right), \text{ for any } t \geq 0 \text{ and } i = 1, ..., m. \tag{1}$$

It is further assumed that the asset values follow geometric Brownian motions; that is, on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t : t \geq 0\}, P)$, we have

$$dV_{i,t} = \mu_i V_{i,t} dt + \sigma_{V_i} V_{i,t}\ dW_{i,t},\ i \in \{1, ..., m\} \tag{2}$$

where μ_i and σ_{V_i} are, respectively, the drift and diffusion coefficients under the physical probability measure P and the m dimensional Brownian motion $W = \{(W_{1,t}, ...W_{m,t}) : t \geq 0\}$ is such that

$$\mathrm{Cov}^P\left[W_{i,t}, W_{j,t}\right] = \rho_{ij} t \text{ for any } t \geq 0. \tag{3}$$

The default-free interest rate r is assumed to be a constant. The default of the ith firm occurs at time T_i if the asset value V_{i,T_i} is below the face value F_i of the debt. Using these assumptions, formulas for the bond value, and default probability can be obtained. These formulas are provided in Appendix A. The credit spread formula follows immediately from the formula for $D_{i,t}\left(\sigma_{V_i}\right)$. Because the default-free interest rate is a constant, the credit spread can be written as

$$C_{i,t}(\sigma_{V_i}) = -\frac{\ln\left[D_{i,t}\left(\sigma_{V_i}\right)/F_i\right]}{T_i - t} - r. \tag{4}$$

To implement this model empirically, one runs into the difficulty of getting the required input variables. Specifically, for the ith firm, the asset value $V_{i,t}$, its drift μ_i and its diffusion coefficient σ_{V_i} are unknown. The correlation coefficient between any two asset values is also unknown. In the next sub-section, we derive the likelihood function which serves as the basis for maximum likelihood estimation.

2.1 The likelihood function

The specific idea for constructing the likelihood function is taken from Duan (1994, 2000) which treats the observed time series of equity prices as a sample of transformed data with the equity pricing equation defining the transformation. Loosely speaking, the resulting likelihood function becomes the likelihood function of the implied asset values multiplied by the Jacobian of the transformation evaluated at the implied asset values.

In general, one observes a time series of equity values for the ith firm corresponding to a known face value of debt over a sample period with a time step of length h. Denote the time series sample up to time t by $\{s_{i,0}, s_{i,h}, s_{i,2h}, ..., s_{i,Nh}\}$ with $t = Nh$ and $i \in \{1, ..., m\}$. The observation period is assumed to be the same for all firms to facilitate the estimation of the correlation coefficients. Let θ denote the vector containing the parameters associated with the m-dimensional geometric Brownian motion process; that is,

$$\theta = [\mu_1, ..., \mu_m, \sigma_{V_1}, ..., \sigma_{V_m}, \rho_{12}, ..., \rho_{1m}, \rho_{23}, ..., \rho_{2m}, ...]. \tag{5}$$

The function defining the critical transformation is the equity pricing equation which is $S_{i,t} = V_{i,t} - D_{i,t}\left(\sigma_{V_i}\right)$. It can be easily shown that $S_{i,t}$ is an invertible function of $V_{i,t}$ for any σ_{V_i}. We denote it by $S_{i,t} = g_i\left(V_{i,t}; t, T, \sigma_{V_i}\right)$.

Theorem 1 *The log-likelihood function corresponding to the stock price sample is*

$$\begin{aligned} & L\left(\mathbf{s}_0, \mathbf{s}_h, \mathbf{s}_{2h}, ..., \mathbf{s}_{Nh}; \theta\right) \\ = \; & -\frac{mN}{2}\ln\left(2\pi\right) - \frac{N}{2}\ln\left|\det\mathbf{\Sigma}\right| - \frac{1}{2}\sum_{k=1}^{N}\mathbf{w}_{kh}^{*\prime}\Sigma^{-1}\mathbf{w}_{kh}^{*} - \sum_{k=1}^{N}\sum_{i=1}^{m}\ln v_{i,kh}^{*} \\ & -\sum_{k=1}^{N}\sum_{i=1}^{m}\ln\Phi\left(d\left(v_{i,kh}^{*}, kh, \sigma_{V_i}\right)\right) \end{aligned} \tag{6}$$

where $v_{i,kh}^{*} = g_i^{-1}\left(s_{i,kh}; kh, T, \sigma_{V_i}\right)$ *is the asset value implied by the equity value,* $\Phi(\cdot)$ *is the standard normal distribution function,* $d(\cdot,\cdot,\cdot)$ *is defined in equation (21) of Appendix A,* $\mathbf{w}_{kh}^{*}$ *is an m-dimensional column vector defined as*

$$\mathbf{w}_{kh}^{*} = \left(\ln v_{i,kh}^{*} - \ln v_{i,(k-1)h}^{*} - \left(\mu_i - \frac{1}{2}\sigma_{V_i}^2\right)h\right)_{m\times 1}$$

and

$$\mathbf{\Sigma} = \left(\sigma_{V_i}\sigma_{Vj}\rho_{ij}h\right)_{i,j=1,\ldots,m} = \begin{bmatrix} \sigma_{V_1}^2 & \cdots & \sigma_{V_1}\sigma_{Vm}\rho_{1m} \\ \vdots & \ddots & \vdots \\ \sigma_{V_1}\sigma_{Vm}\rho_{1m} & \cdots & \sigma_{V_m}^2 \end{bmatrix} h.$$

The proof for the theorem can be found in Appendix B. The first four terms on the right hand side of equation (6) constitute the log-likelihood function if the asset values were observed. The last term in equation (6) corresponds to the Jacobian that accounts for the transformation from the observed equity values to the implied asset values. It is important to note that $\mathbf{v}_{kh}^{*}$ depends on the parameters of the model. If it were not, $\sum_{k=1}^{N}\sum_{i=1}^{m}\ln v_{i,kh}^{*}$ could be dropped from the likelihood function. This would be the case if the asset value could be directly observed.

With the likelihood function in place, one can conduct the maximum likelihood estimation and statistical inference.

2.2 The estimation procedure

Although directly optimizing the log-likelihood function given in Theorem 1 appears to be a straightforward way of approaching the estimation problem, it is actually not practical when many firms are in the data sample. The number of parameters that need to be estimated jointly increases rapidly and becomes unmanageable. We therefore adopt the following multi-step estimation procedure knowing that the true optimum may not be obtained this way:

- Step 1: Estimate the Merton (1974) model for each firm separately. For firm i, estimates $\hat{\mu}_i$ and $\hat{\sigma}_{V_i}$ are obtained using the log-likelihood function in (6) by imposing $m = 1$. The Monte Carlo study (see Section 2.3) indicates that the following standard asymptotic results are applicable to the typical application sample size.
  Statement: *The parameter estimates* $(\hat{\mu}_i, \hat{\sigma}_{V_i})$ *are asymptotically normally distributed around the true parameter values with the covariance matrix being approximated by* $\hat{\mathbf{F}}_i^{-1}$ *where*

$$\hat{\mathbf{F}}_i = \begin{pmatrix} -\frac{1}{N}\frac{\partial^2 L\left(s_0,\ldots,s_{Nh};\hat{\mu}_i,\hat{\sigma}_{V_i}\right)}{\partial\mu_i^2} & -\frac{1}{N}\frac{\partial^2 L\left(s_0,\ldots,s_{Nh};\hat{\mu}_i,\hat{\sigma}_{V_i}\right)}{\partial\mu_i\partial\sigma_{V_i}} \\ -\frac{1}{N}\frac{\partial^2 L\left(s_0,\ldots,s_{Nh};\hat{\mu}_i,\hat{\sigma}_{V_i}\right)}{\partial\mu_i\partial\sigma_{V_i}} & -\frac{1}{N}\frac{\partial^2 L\left(s_0,\ldots,s_{Nh};\hat{\mu}_i,\hat{\sigma}_{V_i}\right)}{\partial^2\sigma_{V_i}} \end{pmatrix}.$$

- Step 2: Compute the implied firm value by

$$\hat{V}_{i,t} \equiv V_{i,t}(\hat{\sigma}_{V_i}) = g_i^{-1}\left(S_{i,t}; t, \hat{\sigma}_{V_i}\right). \tag{7}$$

Because $V_{i,t}(\hat{\sigma}_{V_i})$ is a continuously differentiable function of $\hat{\sigma}_{V_i}$, the distribution for the firm's asset value can be approximated by a normal distribution (see Lo (1986) or Rao (1973), page 385). Let $\hat{\nabla}_{V_i} = \left(\frac{\partial V_{i,t}(\hat{\sigma}_{V_i})}{\partial \mu_i}, \frac{\partial V_{i,t}(\hat{\sigma}_{V_i})}{\partial \sigma_{V_i}}\right)$, then[1]

$$V_{i,t}(\hat{\sigma}_{V_i}) - V_{i,t} \sim N\left\{0, \hat{\nabla}_{V_i}\hat{\mathbf{F}}_i^{-1}\hat{\nabla}'_{V_i}\right\}. \tag{8}$$

For the credit spread, recall equation (4). Its point estimate can be computed by

$$C_{i,t}(\hat{\sigma}_{V_i}) = -\frac{\ln\left(\frac{\hat{V}_{i,t}(\hat{\sigma}_{V_i}) - S_{i,t}}{F_i}\right)}{T_i - t} - r \tag{9}$$

and its distribution can be approximated by

$$C_{i,t}(\hat{\sigma}_{V_i}) - C_{i,t}(\sigma_{V_i}) \sim N\left\{0, \hat{\nabla}_{C_i}\hat{\mathbf{F}}_i^{-1}\hat{\nabla}'_{C_i}\right\} \tag{10}$$

where[2] $\hat{\nabla}_{C_i} = \left(\frac{\partial C_{i,t}(\hat{\sigma}_{V_i})}{\partial \mu_i}, \frac{\partial C_{i,t}(\hat{\sigma}_{V_i})}{\partial \sigma_{V_i}}\right)$. The default probability is a function of both μ_i and σ_{V_i} and its expression is given in Appendix A. The point estimate can thus be expressed as

$$P_{i,t}(\hat{\mu}_i, \hat{\sigma}_{V_i}) = \Phi\left(x_{i,t}(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t})\right) \tag{11}$$

where Φ is the standard normal distribution function and

$$x_{i,t}(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t}) = \frac{\ln(F_i) - \ln(V_{i,t}(\hat{\sigma}_{V_i})) - \left(\hat{\mu}_i - \frac{1}{2}\hat{\sigma}^2_{V_i}\right)(T_i - t)}{\hat{\sigma}_{V_i}\sqrt{T_i - t}}.$$

Similarly, the distribution for the default probability estimate can be approximated by

$$P_{i,t}(\hat{\mu}_i, \hat{\sigma}_{V_i}) - P_{i,t}(\mu_i, \sigma_{V_i}) \sim N\left\{0, \hat{\nabla}_{P_i}\hat{\mathbf{F}}_i^{-1}\hat{\nabla}'_{P_i}\right\} \tag{12}$$

---

[1] Note that the equity pricing formula does not depend on μ_i. Thus, $\frac{\partial V_{i,t}(\hat{\sigma}_{V_i})}{\partial \mu_i} = 0$. Moreover,

$$\frac{\partial V_{i,t}(\hat{\sigma}_{V_i})}{\partial \sigma_{V_i}} = \frac{1}{\hat{V}_i(t)\sqrt{T-t}\phi\left(d\left(\hat{V}_{i,t}, t, \hat{\sigma}_{V_i}\right)\right)}$$

where $\phi(\cdot)$ denotes the standard normal density function.

[2] Again, the first component of $\hat{\nabla}_{C_i}$ is zero. The second one is

$$\frac{\partial C_{i,t}(\hat{\sigma}_{V_i})}{\partial \sigma_{V_i}} = -\frac{1}{T_i - t}\frac{1}{V_{i,t}(\sigma_{V_i}) - S_{i,t}}\frac{\partial V_{i,t}(\hat{\sigma}_{V_i})}{\partial \sigma_{V_i}},$$

where $\frac{\partial V_{i,t}(\hat{\sigma}_{V_i})}{\partial \sigma_{V_i}}$ is given in an earlier footnote.

where[3] $\hat{\nabla}_{P_i} = \left(\frac{\partial P_{i,t}(\hat{\mu}_i,\hat{\sigma}_{V_i})}{\partial \mu_i}, \frac{\partial P_{i,t}(\hat{\mu}_i,\hat{\sigma}_{V_i})}{\partial \sigma_{V_i}}\right)$.

- Step 3: Take the sample correlation coefficient between $\ln\left(\hat{V}_{i,kh}/\hat{V}_{i,(k-1)h}\right)$ and $\ln\left(\hat{V}_{j,kh}/\hat{V}_{j,(k-1)h}\right)$ as the estimate for ρ_{ij}. The estimated correlation coefficient is expected to distribute normally around its true value with a variance taken from the corresponding diagonal entry of $\hat{\mathbf{F}}_{ij}^{-1}$ where $\hat{\mathbf{F}}_{ij} = -\frac{1}{N}\left.\frac{\partial^2 L(\bullet)}{\partial\theta_{ij}(k)\partial\theta_{ij}(l)}\right|_{\theta_{ij}=\hat{\theta}_{ij}}$. $\theta_{ij} = \left[\mu_i, \mu_j, \sigma_{V_i}, \sigma_{V_j}, \rho_{ij}\right]$, $\theta_{ij}(k)$ is the kth element of θ_{ij}, $\hat{\theta}_{ij}$ is the estimate for θ_{ij} and $L(\cdot)$ is the joint log-likelihood function of the data sample for the ith and jth firms. Note that the first four entries of $\hat{\theta}_{ij}$ are taken from the individual estimations for the ith and jth firms in Step 1 whereas the correlation coefficient estimate is obtained in this step. For any quantity that is a function of θ_{ij}, the variance of its distribution can be obtained using the whole matrix $\hat{\mathbf{F}}_{ij}^{-1}$ in a way similar to those in Step 2. If one is interested in any quantity that is a function of the parameters for more than two firms, the dimension of $\hat{\mathbf{F}}_{ij}$ can expanded to accommodate the new requirement.

The three-step estimation procedure can be computed fairly quickly; for example, on a standard desktop computer, the completion for two firms usually takes approximately 10 seconds. Parameter estimation for a large portfolio of firms is thus feasible with the three-step estimation procedure. The numerical optimization routine used here is the quadratic hill–climbing algorithm of Goldfeld, Quandt and Trotter (1966) with a convergence criterion based on the absolute values of the changes in parameter values and functional values between successive iterations. We consider convergence achieved when both of these changes are smaller than 1e-5. The three-step estimation procedure uses several simplifications. We need to ascertain its performance by a Monte Carlo study. Such a study is carried out in the next sub-section.

2.3 A Monte Carlo study

In order to assess the quality of the maximum likelihood procedure, we examine how well the normal distribution suggested by the theory approximates the actual distribution for a reasonable sample size. In other words, we verify whether the parameter estimates for a sample size N is well approximated by the distribution given in the preceding subsection. Similarly, we check the distributions for the asset value $V_{i,t}(\hat{\theta}_N)$, credit spread $C_{i,t}(\hat{\theta}_N)$ and default probability $P_{i,t}(\hat{\theta}_N)$.

We consider the case of two firms and simulate the data on a daily basis as follows:

1. Let $V_{i,kh}$ for $i = \{1,2\}$ and $k = \{1,...,N\}$ denote the simulated asset values in accordance

[3]

$$\begin{aligned}
\frac{\partial P_{i,t}(\hat{\mu}_i,\hat{\sigma}_{V_i})}{\partial \mu_i} &= -\phi\left(x_{i,t}(\hat{\mu}_i,\hat{\sigma}_{V_i},\hat{V}_{i,t})\right) \times \frac{\sqrt{T_i - t}}{\hat{\sigma}_{V_i}} \\
\frac{\partial P_{i,t}(\hat{\mu}_i,\hat{\sigma}_{V_i})}{\partial \sigma_{V_i}} &= \phi\left(x_{i,t}(\hat{\mu}_i,\hat{\sigma}_{V_i},\hat{V}_{i,t})\right)\left[\frac{\ln\left(V_{i,t}(\hat{\sigma}_{V_i})\right) - \ln\left(F_i\right) + \left(\hat{\mu}_i + \frac{1}{2}\hat{\sigma}_{V_i}^2\right)(T_i - t)}{\hat{\sigma}_{V_i}^2\sqrt{T_i - t}} - \frac{\frac{\partial V_{i,t}(\hat{\sigma}_{V_i})}{\partial \sigma_{V_i}}}{\hat{\sigma}_{V_i}\sqrt{T_i - t}V_{i,t}(\hat{\sigma}_{V_i})}\right]
\end{aligned}$$

with

$$
\begin{aligned}
V_{1,(k+1)h} &= V_{1,kh} \exp\left(\mu_1 h - \frac{1}{2}\sigma_{V_1}^2 h + \sigma_{V_1}\sqrt{h}\epsilon_{1,k}\right) && (13)\\
V_{2,(k+1)h} &= V_{2,kh} \exp\left(\mu_2 h - \frac{1}{2}\sigma_{V_2}^2 h + \sigma_{V_2}\sqrt{h}\epsilon_{2,k}\right) && (14)
\end{aligned}
$$

where $\left\{(\epsilon_{1,k}, \epsilon_{2,k})' : k \in \{1, ..., N\}\right\}$ is a sequence of independent and identically distributed vectors of standard normal random variables with a correlation coefficient of ρ_{12}. To be consistent with daily data, we set $h = 1/250$. The specific parameter values used in the simulation are given in the table.

2. Use these simulated time series and the equity pricing equation in Appendix B to compute the corresponding series of simulated equity prices.

For each simulated data set, we conduct maximum likelihood estimation and compute the point estimates and their associated variances based on the equity prices. We repeat the simulation run 5000 times to obtain the Monte Carlo estimates for the relevant quantities. Four different cases are examined. Each case is for a combination of high or low debt-to-asset-value ratio ($F_i/V_{i,0}$) with high or low noise-to-signal ratio (σ_{V_i}/μ_i). To conserve space, we have only reported the high-high case in Table 1. In all these experiments, $T = 3, t = 2, r = 0.05, h = 1/250, \rho_{ij} = 0.5$ and $N = 500$. For a given simulation, the values of $V_{i,t}$ and $P_{i,t}(\mu_i, \sigma_{V_i})$ depend on the specific sample path taken by $V_{i,t}$, we thus examine the statistics for the difference between the estimate and the true value.

Table 1 reports the simulation results for the high-high case with $F_i/V_{i,0} = 0.9$ and $\sigma_{V_i}/\mu_i = 3$ for both firms. The results reveal that, for all parameters and the inferred variables, except for the default probabilities, the maximum likelihood estimators are unbiased. The coverage rates[4] indicate that the normal distribution is a very good approximation of the small sample distribution. In this table, we have also reported the results using the implicit estimation method. The description of this method is given in Appendix C. It is clear from this table that the implicit estimation produces poor estimates. The estimates for volatility are significantly biased downward. For asset value, they are biased upward substantially. For the inference on biases, the standard deviations reported in the tables should be divided by a factor of $\sqrt{n}$ to reflect the fact that n simulation runs are used to produce the means. The performance of the implicit estimation method is related to the noise-to-signal ratio of the firm. For high noise-to-signal firms, the biases are very pronounced whereas for low noise-to-signal firms, the estimates (not reported here) may be regarded as acceptable. Also worth noting is that the implicit estimation method cannot yield an estimate for the drift coefficient and consequently it cannot provide an estimate for the default probability. On the efficiency side, the maximum likelihood estimators always show a smaller standard deviation when compared to the implicit estimators. For all parameter estimates and inferred variables, except for the correlation coefficient, the standard deviations of the implicit estimators are approximately 10 times larger than their maximum likelihood counterparts. The correlation estimate based on the implicit estimation method is very good in terms of bias and standard deviation. This indicates that, in the context of Merton's (1974) model, the correlation between the stock price return is a very good estimate of the true correlation between the asset returns. This is perhaps not too surprising because the

[4] A coverage rate is the percentage of the parameter estimates for which the true parameter value is contained in the α confidence interval implied by the asymptotic distribution

quadratic variation process between $\ln S_i$ and $\ln S_j$ is a function of the correlation between the two Wiener processes driving the asset value processes. Normalizing this expression by the sample standard deviation of the stock returns thus gives an estimate that is close to the true correlation.

To our knowledge, the KMV method (as described in Crouhy *et al.* (2000)), a popular commercial implementation of Merton's model, uses an iterative scheme to estimate the parameters. The method assumes a set of parameter values and inverts the observed equity values to yield a sample of implied asset values. These implied asset values are then used in the likelihood function as if they were the observed values to obtain the updated parameter values. The process is repeated until convergence is obtained. This estimation method can be understood better by referring to the log-likelihood function in Theorem 1. This iterative scheme essentially ignores the last term in the log-likelihood function. Interestingly, this term constitutes the Jacobian of the transformation from the equity value to the asset value, a critical component of the maximum likelihood method for the transformed data. We will refer to this iterative scheme as the incomplete likelihood method. The bottom panel of Table 1 reports the results for the incomplete likelihood method. It is clear from Table 1 that ignoring the Jacobian has serious consequences. The parameter estimates are significantly biased downward whereas the estimates for asset values are biased upward. The coverage ratios are quite off as well. The incomplete likelihood method's performance is also related to the noise-to-signal ratio (not reported in the paper to conserve space). Qualitatively, it behaves like the implicit estimation method for the relevant variables.

The maximum likelihood estimate of the default probability is biased. On average, it overestimates the true default probability by approximately 4%. This bias is mainly due to the imprecision in the expected asset return estimate, which enters into the default probability formula. The coverage rates are also different from the predicted normal distribution. A further analysis reveals that this departure can be explained by the fact that the default probability involves the normal distribution function evaluated at $x_{i,t}(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t})$. This operation is highly non-linear. Even though $x_{i,t}(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t})$ exhibits the property predicted by the theory (see Table 1), the non-linear transformation distorts the default probability. This observation points to a solution as well. A suitable confidence interval for the default probability can actually be obtained. Specifically, the $1-\alpha$ confidence interval for the default probability $P_{i,t}(\hat{\mu}_i, \hat{\sigma}_{V_i}) = \Phi\left(x_{i,t}(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t})\right)$ can be constructed as:

$$\left[\Phi\left(\underline{b}\left(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t}\right)\right); \Phi\left(\overline{b}\left(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t}\right)\right)\right]$$

where $\underline{b}(\bullet)$ and $\overline{b}(\bullet)$ are the lower and upper bounds of the $1-\alpha$ confidence interval for $x_{i,t}(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t})$. The result can be easily verified as follows:

$$\begin{aligned}
& 1-\alpha \\
= \; & P\left[\underline{b}\left(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t}\right) \leq x_{i,t}(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t}) \leq \overline{b}\left(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t}\right)\right] \\
= \; & P\left[\Phi\left(\underline{b}\left(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t}\right)\right) \leq \Phi\left(x_{i,t}(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t})\right) \leq \Phi\left(\overline{b}\left(\hat{\mu}_i, \hat{\sigma}_{V_i}, \hat{V}_{i,t}\right)\right)\right]
\end{aligned}$$

Table 2 presents the results of this correction in terms of the coverage rates. As shown in this table, the corrected confidence intervals do yield the coverage rates consistent with the theoretical prediction.

2.4 Empirical analysis

In this section, we implement Merton's (1974) model using real data. Although this model assumes a zero-coupon debt, most corporations have much more complex liability structures. Liabilities with different properties such as maturity, seniority and coupon rate must be aggregated into one quantity to implement the model. Obviously, there is no clear-cut solution to these problems. One possible approach to determining debt maturity is to find a "theoretical" zero-coupon bond that has the same duration as the aggregated debt. But doing so fundamentally changes the pattern of cash flows. Another way of addressing the issue is to argue that the annual report on profits and losses is perceived by equity holders as the maturity date of their option, which then leads to a pseudo debt maturity of one year. At the time of the public reporting of the annual profits and losses, debt holders may decide to take control of the firm in case of insolvency. On the other hand, if the firm is solvent, the equity value equals the difference between the asset value and the face value of debts.

Determining the amount of debt for the model is also not an obvious matter. The simplest approach would be to set the face value of debt equal to the total amount of short- and long-term liabilities. However, as argued in Crouhy, *et al.* (2000) where some details regarding the implementation of the KMV method are reported, the probability of the asset value falling below the total face value of liabilities may not be an accurate measure of the actual default probability. As reported in an empirical study by KMV, firms default when the asset value reaches a level somewhere between the face value of total liabilities and the face value of the short-term debt. Moreover, there are unknown undrawn commitments (lines of credit) which can be used in case of financial distress. These considerations lead to an ad-hoc approach adopted by KMV to set the face value of debt equal to 1/2 of the long-term debt plus the full amount of the short-term debt.

In this paper, we opt for simplicity and assume that the maturity of debt equals one year and the face value of debt equals to 1/2 of the long-term debt plus the full amount of the short-term debt. Two companies from the Canadian retailing sector are examined for years 1999-2001: Hudson's Bay Company and Sears Canada Inc. Their stock prices are taken from DataStream and the information regarding the long- and short-term debts is extracted from the Financial Post Historical Reports. The short-term debt is defined as the "Current Liabilities" reported yearly in the consolidated balance sheet. The long-term debt is the account "Long-term Debt, net" which represents the long-term debt net of long-term debt maturing during the current year. The short-term interest rate is set to 5% for all years.

The results from the maximum likelihood estimation are reported in Table 3. The daily stock returns for two years preceding to the day of estimation are used to obtain the estimates for each year. For both firms, the estimates for the drift of the asset value process are rather imprecise and statistically insignificant from zero. The estimates for volatility and asset value are, on the other hand, estimated fairly accurately and are clearly statistically significant from zero.

The estimates for the credit spreads are also estimated with precision. The estimated spreads are small when compared with the typical spreads that are observed for firms in this industry and credit class. The default probabilities are estimated to be small but with a large standard error. This is to be expected since the estimated drift, which shows a large variance, enters this quantity. While the default probability of Hudson's Bay is positive and around 2%, the default probability of Sears is virtually zero. The corrected confidence intervals that can be computed around these probabilities are still tight enough to provide interesting information. For example, the 95% confidence interval for Hudson's Bay at the beginning of 2001 is $[0.0004, 0.2936]$ while the one for Sears is $[3.1 \times 10^{-7},\ 0.012]$.

The estimated correlation coefficients between the asset value return also show small standard errors. As in the Monte Carlo study, these estimated coefficients are very close to correlation coefficients for the stock returns.

3 Longstaff and Schwartz's credit spread model

As in Merton (1974), the Longstaff and Schwartz (1995) credit spread model assumes a geometric Brownian motion for the firm's asset value. To simplify the notation, we only consider one firm. Specifically, the firm's asset value dynamic is

$$dV_t = \mu V_t dt + \sigma_V V_t \; dW_{V,t}. \tag{15}$$

The default-free instantaneous interest rate is assumed to be stochastic and governed by an Ornstein-Uhlenbeck process:

$$dr_t = \kappa(\alpha - r_t)dt + \sigma_r dW_{r,t} \tag{16}$$

where r_t is the instantaneous spot interest rate at time t, κ is the mean reverting speed, α is the long run average and σ_r is the interest rate volatility. The two Brownian motions - $W_{V,t}$ and $W_{r,t}$ - are assumed to have the correlation coefficient ρ. In addition, default is triggered whenever the firm's asset value falls below some exogenously specified threshold level K to reflect financial distress.[5]. At default, the debt's recovery rate is assumed to be $1 - w$.

The default-free and risky zero-coupon bond pricing formulas, denoted respectively by $B(r_t, t, T)$ and $D(V_t, r_t, t, T)$, are given in Appendix D. The default probability is also given in this appendix. Because the specific assumption on how default is triggered, a risky coupon bond can be viewed as a simple portfolio of zero-coupon risky bonds in the same way as for the default-free coupon bond. The credit spread formula requires a different definition for coupon bonds, however. When both default-free and risky bonds are of zero-coupon, the definition is the same as the one in equation 4. For coupon bonds, credit spread should be redefined as the difference in the continuously compounded yields between the risky and default-free bonds.

Similar to Merton's (1974) model, several required inputs to the formulas are unknown. Specifically, one needs the value for r_t, V_t, K, w and the parameters governing the asset value and interest rate dynamics. In the next subsection, we develop the likelihood function for the data sample consists of default-free bond yields and equity values of the firm.

In constructing the likelihood function, it is important to take into account survivorship. Assessing credit spreads is an exercise of attaching a premium to debt instruments of a firm that has survived. Survivorship is not the relevant issue in the case of Merton's (1974) model because in that model, the zero-coupon debt only becomes due at some future time point. In other words, there will be no possibility of default prior to that time point.[6] This is not the same as the bond pricing model with financial distress, because under the financial distress model the firm can default prior to the debt maturity.

[5] In Collin-Dufresne and Goldstein (2001), the triggering threshold level is allowed to change dynamically according to $dK_t = \lambda(\ln V_t - v - K_t)dt$

[6] Precisely speaking, this result is related to the way that we have implemented Merton's model. If the data sample contains points of refinancing, survivorship adjustment will be needed.

3.1 The Likelihood function

As before, there is a time series $\{s_0, s_h, ..., s_{Nh}\}$ for the equity values of the firm up to time $t = Nh$. Since interest rates are stochastic, we also need to use a time series of zero-coupon bond prices with \$1 face value, denoted by $\{b_0, b_h, ..., b_{Nh}\}$. The bonds may have different maturities. Denote the parameter set by

$$\theta = (\mu, \sigma_V, \kappa, \alpha, \sigma_r, \lambda, \rho, K, w)$$

Note that λ is the risk premium parameter not in the system defined by equations (15) and (16). This additional parameter arises from the term structure theory.

The unknown interest rate is tied to a default-free bond price by the bond pricing formula (29) in Appendix D. This bond pricing formula defines a one-to-one transformation for any set of interest rate parameters. Specifically,

$$r_t = \frac{\ln b_t - \beta_{T-t}}{\alpha_{T-t}}.$$

Similar to the earlier treatment of Merton's (1974) model, we relates the equity price to the unobserved asset value of the firm: $S_t = g\,(V_t; t, \theta, r_t) = V_t - D(V_t, r_t, t, T)$. Although this function is typically increasing in V_t, it is not always the case. First note that $g\,(V_t; t, \theta, r_t)$ is only defined for $V_t \geq K$. Under some situations, $g\,(V_t; t, \theta, r_t)$ is first decreasing and then increasing. As a result, we cannot be certain that this function is always invertible, a condition for the transformed data method of Duan (1994). But we can be certain that there are at most two asset values corresponding to one equity value. Non-invertibility happens when the debt has a low recovery rate in financial distress. This non-invertibility is a peculiarity of the Longstaff and Schwartz (1995) model. It implies that in some situation, equity capital injection will cause the total equity value to drop because an amount greater than the injected equity capital has been shifted to benefit the bond holders. Addressing non-invertibility becomes particularly important for data samples of firms close to financial distress. The solution turns out to be rather straightforward in principle. We simply add up all likelihood values that correspond to different combinations of possible solutions. Suppose that there are three data points that have two asset values for each equity value. There are eight possible combinations with each yielding a sample likelihood value. We simply add them up and then apply the logarithmic transformation to obtain the log-likelihood function.

To better understand the estimation issue with survivorship, we use the following diagram to describe the estimation task:

$$\underbrace{0 \quad h \quad ... \quad (N-1)\,h \quad Nh}_{\text{Sample period}}\underbrace{\qquad\qquad\qquad\qquad T}_{\text{Future period}}$$

At time Nh, the firm has not defaulted. We need to have a likelihood function for the observations in the sample period, conditional on no default during the sample period. We address this problem by providing two separate results. First we develop a result by assuming constant interest rate. This restriction amounts to modifying Merton's (1974) model by incorporating financial distress. We consider this restricted case of the Longstaff and Schwartz (1995) model so as to isolate the effect of survivorship. We present the likelihood function for the full model later.

Theorem 2 *Assume constant interest rate (i.e., $\sigma_r = 0$). The log-likelihood function based on a sample of observed equity values with no default in the sample period is,*

$$L\left(s_0, s_h, ..., s_{Nh}; \mu, \sigma_V\right) = \ln \sum_{\left(v_0^*, v_h^*, ..., v_{Nh}^*\right) \in \Psi(\sigma_V)} \exp\left[l\left(v_0^*, v_h^*, ..., v_{Nh}^*; \mu, \sigma_V\right)\right] \tag{17}$$

where

$$\begin{aligned}
& l\left(v_0^*, v_h^*, ..., v_{Nh}^*; \mu, \sigma_V\right) \\
= \ & -\sum_{n=1}^{N} \ln v_{nh}^* - \frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\left(\sigma_V^2 h\right) - \frac{1}{2\sigma_V^2 h}\sum_{n=1}^{N}\left(\ln\frac{v_{nh}^*}{v_{(n-1)h}^*} - \left(\mu - \frac{\sigma_V^2}{2}\right)h\right)^2 \\
& + \sum_{n=1}^{N} \ln\left(1 - \exp\left(-\frac{2}{\sigma_V^2 h}\ln\frac{v_{(n-1)h}^*}{K}\ln\frac{v_{nh}^*}{K}\right)\right) \\
& - \ln\left[\Phi\left(\frac{\left(\mu - \frac{\sigma_V^2}{2}\right)Nh - \ln\frac{K}{v_0}}{\sqrt{Nh}\sigma_V}\right) - \exp^{\frac{2}{\sigma_V^2}\left(\mu - \frac{\sigma_V^2}{2}\right)\ln\frac{K}{v_0}} \Phi\left(\frac{\left(\mu - \frac{\sigma_V^2}{2}\right)Nh + \ln\frac{K}{v_0}}{\sqrt{Nh}\sigma_V}\right)\right] \\
& - \sum_{n=1}^{N} \ln\left(1 - \frac{2Fw}{v_{nh}^*}\exp^{-r(T-t)}\left(\begin{array}{c} \frac{1}{\sqrt{T-t}\sigma_V}\phi\left(\frac{1}{\sqrt{T-t}\sigma_V}\left(\ln\frac{K}{v_{nh}^*} - \left(r - \frac{\sigma_V^2}{2}\right)(T-t)\right)\right) \\ + \frac{1}{\sigma_V^2}\left(r - \frac{\sigma_V^2}{2}\right)\exp^{\frac{2}{\sigma_V^2}\left(r - \frac{\sigma_V^2}{2}\right)\ln\frac{K}{v_{nh}^*}} \\ \times\Phi\left(\frac{1}{\sqrt{T-t}\sigma_V}\left(\ln\frac{K}{v_{nh}^*} + \left(r - \frac{\sigma_V^2}{2}\right)(T-t)\right)\right) \end{array}\right)\right),
\end{aligned}$$

and $\Psi(\sigma_V)$ is a set of $(v_0^*, v_h^*, ..., v_{Nh}^*)$ with $v_{nh}^* > K$ being a solution to $g(v_{nh}^*; nh, T, \sigma_V) = s_{nh}$. If $\Psi(\sigma_V)$ is empty, then $L\left(s_0, s_h, ..., s_{Nh}; \mu, \sigma_V\right) = -\infty$.

Note that $\Psi(\sigma_V)$ is a finite set with at most 2^{N+1} elements because each equity value can at most correspond to two asset values and there are $N+1$ data points in the sample. Usually, $\Psi(\sigma_V)$ does not contain many elements. The set $\Psi(\sigma_V)$ depends on σ_V but not μ because the pricing relationship is independent of μ. It is important to note that $\Psi(\sigma_V)$ can be empty because at an arbitrary value of σ_V, one may not be able to find a solution to $g(v_{nh}^*; nh, T\sigma_V) = s_{nh}$ that is greater than K. Around the true parameter value, however, a solution always exists if the model is correctly specified. Note that the last term on the right-hand side of $l\left(v_0^*, v_h^*, ..., v_{Nh}^*; \mu, \sigma_V\right)$ is the Jacobian associated with the transformation from the observed equity values to the implied asset values. The proof is given in Appendix E.

Now we turn to the full version of the Longstaff and Schwartz (1995) model and provide the log-likelihood function. For a technical reason, we cannot solve the exact log-likelihood function, which involves the joint distribution of the firm value and the interest rate, conditional on no default. The strong solution to the interest rate process over one data period is

$$r_{(n+1)h} = \alpha + (r_{nh} - \alpha)\exp(-\kappa h) + \sigma_r \int_{nh}^{(n+1)h} e^{-\kappa[(n+1)h-u]} dW_{r,u},$$

which does not permit the asset price innovation to be expressed as a linear function of interest rate innovation. Instead, we will use the following Euler approximation to construct the likelihood function:

$$r_{(n+1)h} \cong r_{nh} + \kappa(\alpha - r_{nh})h + \sigma_r\left(W_{r,(n+1)h} - W_{r,nh}\right).$$

We introduce the following intermediate result which will be used later in constructing the log-likelihood function.

Lemma 3 *Conditional on no default in the sample period, the log-likelihood function based on the asset values and the interest rates, is, for any $v_0, ..., v_{Nh} > K$,*

$$\begin{aligned}
& G\left(v_0, r_0, v_h, r_h, ..., v_{Nh}, r_{Nh}; \theta\right) \\
\cong\ & -\sum_{n=1}^{N} \ln v_{nh} - N \ln \sigma_r - N \ln \sigma_V - N \ln h - \frac{N}{2} \ln(2\pi) - \frac{1}{2} \sum_{n=1}^{N} \phi_{n-1,n}^2 \\
& + \sum_{n=1}^{N} \ln \left(\begin{array}{c} \frac{\rho}{1-\rho^2} \phi_{n-1,n} \left[\Phi\left(\lambda_{n-1,n}^-\right) + \exp\left(\theta_{n-1}\right) \Phi\left(\lambda_{n-1,n}^+\right) \right] \\ + \frac{1}{\sqrt{1-\rho^2}} \phi\left(\lambda_{n-1,n}^-\right) \left[1 - \exp\left(-\frac{2}{\sigma_V^2 h} \ln \frac{v_{nh}}{K} \ln \frac{v_{(n-1)h}}{K} \right) \right] \\ + \frac{\rho}{1-\rho^2} \phi_{n-1,n} \left[\Phi\left(\lambda_{n-1,n}^-\right) + \exp\left(\theta_{n-1}\right) \Phi\left(\lambda_{n-1,n}^+\right) \right] \end{array} \right) \\
& - \ln \left[\begin{array}{c} \Phi\left(\frac{1}{\sqrt{Nh}\sigma_V} \left(\left(\mu - \frac{\sigma_V^2}{2} \right) Nh - \ln \frac{K}{v_0} \right) \right) \\ - \exp\left(\frac{2}{\sigma_V^2} \left(\mu - \frac{\sigma_V^2}{2} \right) \ln \frac{K}{v_0} \right) \Phi\left(\frac{1}{\sqrt{Nh}\sigma_V} \left(\ln \frac{K}{v_0} + \left(\mu - \frac{\sigma_V^2}{2} \right) Nh \right) \right) \end{array} \right].
\end{aligned}$$

where

$$\begin{aligned}
\phi_{n-1,n} &= \frac{\ln \frac{v_{(n-1)h}}{v_{nh}} + \left(\mu - \frac{\sigma_V^2}{2} \right) h}{\sigma_V \sqrt{h}} \frac{\rho}{\sqrt{1-\rho^2}} + \frac{\left(r_{nh} - r_{(n-1)h} \right) - \kappa(\alpha - r_{(n-1)h})h}{\sigma_r \sqrt{h}} \sqrt{\frac{1}{1-\rho^2}}, \\
\lambda_{n-1,n}^+ &= \eta^+ - \frac{\ln \frac{v_{nh}}{K}}{\sigma_V \sqrt{h}} = \frac{-\ln \frac{v_{nh}}{K} - \ln \frac{v_{(n-1)h}}{K} + \left(\mu - \frac{\sigma_V^2}{2} \right) h}{\sigma_V \sqrt{h}}, \\
\lambda_{n-1,n}^- &= \eta^- + \frac{\ln \frac{v_{nh}}{K}}{\sigma_V \sqrt{h}} = \frac{\ln \frac{v_{nh}}{K} - \ln \frac{v_{(n-1)h}}{K} - \left(\mu - \frac{\sigma_V^2}{2} \right) h}{\sigma_V \sqrt{h}}, \\
\text{and } \theta_{n-1} &= -\frac{2}{\sigma_V^2} \left(\mu - \frac{\sigma_V^2}{2} \right) \ln \frac{v_{(n-1)h}}{K}.
\end{aligned}$$

The proof of this lemma is in Appendix F. Again, the real estimation problem is for the situation in which the asset values and interest rates are not observable. We need to replace the asset values and the interest rates by their implied values v_{nh}^* and r_{nh}^*. But we must take into account the Jacobian for the transformation. The log-likelihood function based on the observed data is given in the following theorem.

Theorem 4 *The log-likelihood function based on a sample of observed equity values (with no default in the sample period) and default-free bond prices is*

$$L\left(s_0, b_0, ..., s_{Nh}, b_{Nh}; \theta\right) = \ln \sum_{\left(v_0^*, v_h^*, ..., v_{Nh}^*\right) \in \Psi(\theta)} \exp\left[l\left(v_0^*, r_0^*, ..., v_{Nh}^*, r_{Nh}^*; \theta\right) \right] \tag{18}$$

where

$$\begin{aligned}
& l\left(v_0^*, r_0^*, ..., v_{Nh}^*, r_{Nh}^*; \theta\right) \\
=\ & G\left(v_0^*, r_0^*, ..., v_{Nh}^*, r_{Nh}^*; \theta\right) - \sum_{n=0}^{N} \ln \left| \alpha_{t_{nh}-nh} b_{nh} \left(1 + w b_{nh} \frac{\partial}{\partial v} \mathcal{Q}_{nh}^T \left(v_{nh}^*, r_{nh}^*, nh, T\right) \right) \right|,
\end{aligned} \tag{19}$$

$r^*_{nh} = \frac{\ln b_{nh} - \beta_{T_{nh}-nh}}{\alpha_{T_{nh}-nh}}$ *($T_{nh} - nh$ is the maturity for the default-free bond price b_{nh}) and $v^*_{nh} > k$ being a solution to the equity valuation equation $g\left(v^*_{nh}; r^*_{nh}, nh, \theta\right) = s_{nh}$. $\mathcal{Q}^T_t\left(V_t, r_t, t, T\right)$ and α_{T-t} are defined in Appendix D and $G\left(v^*_0, r^*_0, ..., v^*_{Nh}, r^*_{Nh}; \theta\right)$ is given in Lemma 3. $\Psi(\theta)$ is a set of $(v^*_0, v^*_h, ..., v^*_{Nh})$. If $\Psi(\theta)$ is empty, then $L\left(s_0, b_0, ..., s_{Nh}, b_{Nh}; \theta\right) = -\infty$.*

This theorem is proved in Appendix G.

3.2 The estimation procedure

Although directly optimizing the log-likelihood function given in Theorem 4 appears to be a sensible way of approaching the estimation problem, it is actually not a good approach in practice. First, the likelihood function is defined for the data set comprising one specific equity value series and the common bond price series. If there are more than one firm, there will be more than one set of parameter values governing the common interest rate dynamic. One can, of course, think of expanding the log-likelihood function to include all firms in the sample to conduct the joint estimation. It is not practical, however, when there are many firms in the sample. The second problem is the difference in the time horizons for bond pricing and equity valuation. The Vasicek (1977) bond pricing model reflects the mean reversion in interest rates. The tendency for interest rates to return to their average is meant to be a long-run phenomenon. It is therefore reasonable to expect that the mean-reversion parameter can only be pinned down using a relatively long interest rate data series. The bond valuation model, on the other hand, depends on the firm's asset volatility parameter. Since the variation of the asset value under the diffusion specification is large, the volatility parameter can usually be estimated with precision using a relatively short time series. To allow for, say, year-to-year changes in the asset volatility, the use of an equity value time series shorter than the interest rate series may be desirable.

We thus devise a two-step estimation procedure. The first step estimates, through maximizing the likelihood function, the Vasicek (1977) bond pricing model parameters using interest rate data only. The likelihood function for this model is given in Duan (1994). The second step estimates the asset value parameters with the likelihood function in Theorem 4 by fixing the interest rate parameters at the values obtained from the first step. This two-step procedure ensures the interest rate parameter estimates are the same for all firms in the sample. Moreover, it allows us to use a longer time series of interest rates to pin down the mean-reversion parameter for the interest rate dynamic. Since the parameters governing the asset value dynamic do not enter the bond pricing model, this two-step procedure continues to yield consistent parameter estimates. However, the estimation of the standard errors of the asset value parameters will be affected. Using the bond pricing parameter estimates as if they were the true parameter values does not account for the additional sampling errors brought about by the their estimation errors in the first step. To account for these sampling errors, the standard errors of the asset value parameters will be taken from the usual information matrix $\mathbf{F}_N = -\frac{1}{N}\frac{\partial^2 L(\bullet)}{\partial\theta_i \partial\theta_j}$, where $\hat{\theta}_n$ is set to the parameter values obtained in the first and second step of the two-step estimation procedure. The statistical properties of our two-step procedure will be studied later using a Monte Carlo analysis.

3.3 A Monte Carlo study

In order to assess the quality of the maximum likelihood method, we examine how well the normal distribution suggested by the theory approximates the actual distribution for a reasonable sample size. In other words, we check to see whether the parameter estimates for a sample size N is well approximated by the distribution given in the preceding subsection. Similarly, we check the distributions for the asset value $V_t(\hat{\theta}_N)$, credit spread $C_t(\hat{\theta}_N)$ and default probability $P_t(\hat{\theta}_N)$.

We consider the case of a single firm and simulate the data on a frequency of 50 points per day and sample one every 50 points to mimic the situation of analyzing daily data. Let $h = 1/250$. Specifically,

- Let $r_{kh/50}$ and $V_{kh/50}$ for $k = \{1, ..., 50N\}$ denote the simulated values of the instantaneous interest rates and asset values. Using the transition density functions corresponding to equations (15) and (16). Specifically, $r(kh)$ and $V(kh)$ are simulated according to the dynamics:

$$\begin{aligned} r_{(k+1)h/50} &= \alpha + \left(r_{kh/50} - \alpha\right) e^{-\kappa(h/50)} + \sqrt{\frac{\sigma_r^2}{2\kappa}(1 - e^{-2\kappa(h/50)})}\epsilon_{r,k}, \\ V_{(k+1)h} &= V_{kh} \exp\left(\mu(h/50) - \frac{1}{2}\sigma_V^2(h/50) + \sigma_V\sqrt{h/50}\epsilon_{V,k}\right) \end{aligned}$$

  where $\left\{(\epsilon_{r,k}, \epsilon_{V,k})' : k \in \{1, ..., 50N\}\right\}$ is a sequence of independent and identically distributed vectors of standard normal random variables with a correlation coefficient of ρ. The specific parameter values used in the simulation are given in the tables.

- Use these simulated time series and the equity pricing equation to compute the corresponding series of simulated equity prices.

To be completed later.

3.4 Empirical analysis

In this section, we implement the Longstaff and Schwartz (1995) models using real data. In this model, unknown triggering level for financial distress and recovery rate are introduced to the model. In a way, it has simplified the default consideration. Nevertheless, one still have to determine the debt size and maturity by aggregating debts of different maturities.

To be completed later.

4 Conclusion

To be completed later.

A Formulas for debt and default probability in Merton's (1974) model

In Merton (1974), $D_{i,T_i}(\sigma_{V_i}) = \min\{V_{i,T_i}, F_i\}$. For valuation, it is well-known that we can use the risk-neutral asset price dynamic to evaluate the discounted expected payout, where the risk-neutral dynamic has the risk-free rate as the drift term but the same diffusion term. Consequently, the bond value at time t is

$$D_{i,t}(\sigma_{V_i}) = F_i e^{-r(T_i-t)} \left(\frac{V_{i,t}}{F_i e^{-r(T_i-t)}} \Phi\left(-d(V_{i,t}, \sigma_{V_i})\right) + \Phi\left(d(V_{i,t}, \sigma_{V_i}) - \sigma_{V_i}\sqrt{T_i - t}\right) \right) \tag{20}$$

where $\Phi(\bullet)$ is the standard normal distribution function and

$$d(V_{i,t}, t, \sigma_{V_i}) = \frac{\ln(V_{i,t}) - \ln(F_i) + \left(r + \frac{1}{2}\sigma_{V_i}^2\right)(T_i - t)}{\sigma_{V_i}\sqrt{T_i - t}}. \tag{21}$$

Furthermore, Merton's model implies the following default probability under measure P at time t:

$$P_{i,t}(\mu_i, \sigma_{V_i}) = \mathrm{Prob}_t\left[V_i(T_i) < F_i\right] = \Phi\left(\frac{\ln(F_i) - \ln(V_{i,t}) - \left(\mu_i - \frac{1}{2}\sigma_{V_i}^2\right)(T_i - t)}{\sigma_{V_i}\sqrt{T_i - t}} \right). \tag{22}$$

where $\mathrm{Prob}_t[\bullet]$ denotes for the conditional probability taken at time t under the measure P. The joint probability of default for several firms can be expressed using the multivariate normal cumulative distribution function $N_{\mathbf{0},\rho} : R^m \to [0,1]$ with with mean $0_{m\times 1}$ and covariance matrix $\rho \equiv (\rho_{ij})_{i,j\in\{1,\ldots,m\}}$ and relying on the following quantity:

$$\mathrm{Prob}_t\left[V_{1,T_i} < \alpha_1, ..., V_{m,T_i} < \alpha_m\right] = \mathbf{N}_{\mathbf{0},\rho}(\beta_1, ..., \beta_m) \tag{23}$$

where

$$\beta_i = \frac{\ln(\alpha_i) - \ln(V_{i,t}) - \left(\mu_i - \frac{1}{2}\sigma_{V_i}^2\right)(T_i - t)}{\sigma_{V_i}\sqrt{T_i - t}}.$$

This gives rise to the joint default probability of the firms $i_1, \ldots, i_k$ where $k \leq m$ by setting

$$\alpha_i = F_i \text{ for any } i \in \{i_1, ..., i_k\} \text{ and } \alpha_i \to \infty \text{ for all } i \notin \{i_1, ..., i_k\}$$

in equation (23).

B The likelihood function for Merton's (1974) model

In this section we derive the likelihood function for the multivariate version of Merton's (1974) model. As stated in the text, θ denote the vector containing all parameters associated with the m-dimensional geometric Brownian motion process. Assume for a moment that we are able to observe $v_0, v_h, v_{2h}, ..., v_{Nh}$ with

$$\mathbf{v}_{kh} = \begin{pmatrix} v_{1,kh} \\ \vdots \\ v_{m,kh} \end{pmatrix}.$$

By the property of the geometrical Brownian motion, the joint density function is

$$f_{\mathbf{V}_0,\mathbf{V}_h,\mathbf{V}_{2h},...,\mathbf{V}_{Nh}}\left(\mathbf{v}_0,\mathbf{v}_h,\mathbf{v}_{2h},...,\mathbf{v}_{Nh};\theta\right)=\prod_{k=1}^{N} f_{\mathbf{V}_{kh}|\mathbf{V}_{(k-1)h}}\left(\mathbf{v}_{kh}\left|\mathbf{v}_{(k-1)h}\right.;\theta\right)$$

where $f_{\mathbf{V}_{kh}|\mathbf{V}_{(k-1)h}}\left(\mathbf{v}_{kh}\left|\mathbf{v}_{(k-1)h}\right.\right)$ denotes the conditional density function of v_{kh} given $v_{(k-1)h}$.

Since the conditional distribution of v_{kh} given $v_{(k-1)h}$ is lognormal, we obtain the log-likelihood function

$$\begin{aligned}
& L\left(\mathbf{v}_0,\mathbf{v}_h,\mathbf{v}_{2h},...,\mathbf{v}_{Nh};\theta\right) \\
= & \ln f_{\mathbf{V}_0,\mathbf{V}_h,\mathbf{V}_{2h},...,\mathbf{V}_{Nh}}\left(\mathbf{v}_0,\mathbf{v}_h,\mathbf{v}_{2h},...,\mathbf{v}_{Nh};\theta\right) \\
= & \sum_{k=1}^{N}\ln f_{\mathbf{V}_{kh}|\mathbf{V}_{(k-1)h}}\left(\mathbf{v}_{kh}\left|\mathbf{v}_{(k-1)h}\right.;\theta\right) \\
= & -\frac{mN}{2}\ln(2\pi)-\frac{N}{2}\ln\left(\left|\det\mathbf{\Sigma}\right|\right)-\frac{1}{2}\sum_{k=1}^{N}\mathbf{w}_{kh}^{'}\Sigma^{-1}\mathbf{w}_{kh}-\sum_{k=1}^{N}\sum_{i=1}^{m}\ln v_{i,kh} \qquad (24)
\end{aligned}$$

where $\Sigma\equiv\left(h\sigma_{V_i}\sigma_{V_j}\rho_{ij}\right)_{i,j\in\{1,...,m\}}$, $\left|\det\mathbf{\Sigma}\right|$ is the determinant of Σ and w_{kh} is the column vector

$$\mathbf{w}_{kh}\equiv\left(\ln v_{i,kh}-\ln v_{i,(k-1)h}-\left(\mu_i-\frac{1}{2}\sigma_{V_i}\right)h\right)_{m\times 1}.$$

We of course do not observe $v_0,v_h,v_{2h},...,v_{Nh}$. Instead, we have a time series of the equity values $s_0,s_h,s_{2h},...,s_{Nh}$ where

$$\mathbf{s}_{kh}=\begin{bmatrix} s_{1,kh}\\ \vdots \\ s_{m,kh}\end{bmatrix}$$

The equity pricing equation is

$$S_{i,t}=V_{i,t}\Phi\left(d\left(V_{i,t},t,\sigma_{V_i}\right)\right)-F_ie^{-r(T_i-t)}\Phi\left(d\left(V_{i,t},t,\sigma_{V_i}\right)-\sigma_{V_i}\sqrt{T_i-t}\right) \qquad (25)$$

where

$$d\left(V_{i,t},t,\sigma_{V_i}\right)=\frac{\ln\left(V_{i,t}\right)-\ln\left(F_i\right)+\left(r+\frac{1}{2}\sigma_{V_i}^2\right)\left(T_i-t\right)}{\sigma_{V_i}\sqrt{T_i-t}} \qquad (26)$$

and $\Phi(\bullet)$ is the cumulative standard normal distribution function. It defines the transformation between $S_{i,t}$ and $V_{i,t}$. We denote it by $S_{i,t}=g_i\left(V_{i,t};t,\sigma_{V_i}\right)$. For any $0\le t<T_i$, this function is invertible in the sense that for any fixed t and σ_{V_i}, $V_{i,t}=g_i^{-1}\left(S_{i,t};t,\sigma_{V_i}\right)$. In fact, $\Delta_i(t)\equiv\frac{\partial g_i(V_{i,t};t,\sigma_{V_i})}{\partial V_i}=\Phi(d\left(V_{i,t},t,\sigma_{V_i}\right))$ is strictly positive, implying that g_i is a strictly increasing function of $V_{i,t}$. Defining v_{kh}^* as the implied asset prices at time kh obtained from the observed equity prices and a given set of parameters $(\sigma_{V_1},...,\sigma_{V_m})$, that is

$$\mathbf{v}_{kh}^*=\begin{bmatrix} v_{1,kh}^*\\ \vdots \\ v_{m,kh}^*\end{bmatrix}=\begin{bmatrix} g_1^{-1}\left(s_{1,kh};kh,\sigma_{V_1}\right)\\ \vdots \\ g_m^{-1}\left(s_{m,kh};kh,\sigma_{V_m}\right)\end{bmatrix}.$$

We can express the log-likelihood function for the sample of equity values as

$$\ln f_{\mathbf{S}_0,\mathbf{S}_h,\mathbf{S}_{2h},...,\mathbf{S}_{Nh}}\left(\mathbf{s}_0,\mathbf{s}_h,\mathbf{s}_{2h},...,\mathbf{s}_{Nh};\theta\right)=\ln f_{\mathbf{V}_0,\mathbf{V}_h,\mathbf{V}_{2h},...,\mathbf{V}_{Nh}}\left(\mathbf{v}_0^*,\mathbf{v}_h^*,...,\mathbf{v}_{Nh}^*;\theta\right)+\ln\left|\det\mathcal{J}\right|$$

where the Jacobian J of the transformation is a block diagonal matrix $J=\left(\mathcal{J}_{nh}\right)_{n\in\{1,...,N\}}$ and the sub-matrix J_{kh} is the $m\times m$ matrix

$$\mathcal{J}_{kh}=\left(\frac{\partial v_{i,kh}^*}{\partial s_{j,kh}}\right)_{i,j\in\{1,...,m\}}.$$

One can show

$$\frac{\partial v_{i,kh}^*}{\partial s_{j,kh}}=\begin{cases}\frac{1}{\Phi\left(d\left(v_{i,kh}^*,,kh,\sigma_{V_i}\right)\right)} & \text{if } i=j\\ 0 & \text{otherwise}\end{cases}. \tag{27}$$

Therefore

$$\ln\left|\det\mathcal{J}\right|=\ln\prod_{k=1}^{N}\prod_{i=1}^{m}\frac{1}{\Phi\left(v_{i,kh}^*\right)}=-\sum_{k=1}^{N}\sum_{i=1}^{m}\ln\Phi\left(d\left(v_{i,kh}^*,kh,\sigma_{V_i}\right)\right).$$

Thus, the log-likelihood function based on the sample of observed equity values is

$$\begin{aligned}&L\left(\mathbf{s}_0,\mathbf{s}_h,\mathbf{s}_{2h},...,\mathbf{s}_{Nh};\theta\right)\\ =\ &-\frac{mN}{2}\ln\left(2\pi\right)-\frac{N}{2}\ln\left(\left|\mathbf{\Sigma}\right|\right)-\frac{1}{2}\sum_{k=1kh}^{N}\mathbf{w}_{kh}^{*\prime}\Sigma^{-1}\mathbf{w}_{kh}^*-\sum_{k=1}^{N}\sum_{i=1}^{m}\ln v_{i,kh}^*\\ &-\sum_{k=1}^{N}\sum_{i=1}^{m}\ln\Phi\left(d\left(v_{i,kh}^*,kh,\sigma_{V_i}\right)\right)\end{aligned}$$

where w_{kh}^* is an m-dimensional column vector defined as

$$\mathbf{w}_{kh}^*=\left(\ln v_{i,kh}^*-\ln v_{i,(k-1)h}^*-\left(\mu_i-\frac{1}{2}\sigma_{V_i}\right)h\right)_{m\times 1}.$$

C The implicit estimation method

Following Jones, et al (1984) and Ronn and Verma (1984), an equation relating the diffusion coefficient of the stock price process to that of the asset value process can be obtained because stock price is a function of the asset value. Formally, $S_{i,t}=g_i(V_{i,t};t,\sigma_{V_i})$. Applying Itô's lemma gives rise to

$$\begin{aligned}d\ln S_{i,t}\ =\ &\left(\frac{\mu_iV_{i,t}\Delta_i\left(t\right)+\theta_i\left(t\right)+\frac{1}{2}\sigma_{V_i}^2V_i^2\left(t\right)\Gamma_i\left(t\right)}{S_{i,t}}-\frac{\sigma_{V_i}^2V_i^2\left(t\right)\Delta_i^2\left(t\right)}{2S_i^2\left(t\right)}\right)dt\\ &+\frac{\sigma_{V_i}V_{i,t}\Delta_i\left(t\right)}{S_{i,t}}\,dW_{i,t}.\end{aligned} \tag{28}$$

where

$$\begin{aligned}
\Delta_i(t) &\equiv \frac{\partial g_i(V_{i,t};t,\sigma_{V_i})}{\partial v} = \Phi(d(V_{i,t},\sigma_{V_i})), \\
\Gamma_i(t) &\equiv \frac{\partial^2 g_i(V_{i,t};t,\sigma_{V_i})}{\partial v^2} = \frac{\phi(d(V_{i,t},\sigma_{V_i}))}{\sigma_{V_i} V_{i,t}(T_i-t)^{1/2}}, \\
\theta_i(t) &\equiv \frac{\partial g_i(V_{i,t};t,\sigma_{V_i})}{\partial t}
\end{aligned}$$

and ϕ and Φ denotes respectively the density function and the cumulative function of a standard normal random variable. The diffusion coefficient of the stock return process can thus be written as:

$$\sigma_{S_i}(t) = \frac{\sigma_{V_i} V_{i,t} \Delta_i(t)}{S_{i,t}}.$$

This coefficient is time dependent and stochastic. Although it is inconsistent with the model, the implicit estimation method assumes that the sample standard deviation of stock returns sampled over a time period prior to time t is a good estimate of $\sigma_{S_{i,t}}$. This relationship in conjunction with $S_{i,t} = g_i(V_{i,t};t,\sigma_{V_i})$ can be used to solve for two unknowns - $V_{i,t}$ and σ_{V_i} - using the observed value for $S_{i,t}$ and the estimate for $\sigma_{S_{i,t}}$.

In the multi-firm context, the covariance between two stock returns is $\sigma_{S_i}(t)\,\sigma_{S_j}(t)\,\rho_{ij}$ because

$$\begin{aligned}
d\langle \ln S_i, \ln S_j\rangle_t &= \frac{\sigma_{V_i} V_{i,t}\Delta_i(t)}{S_{i,t}} \frac{\sigma_{V_j} V_j(t)\,\Delta_j(t)}{S_j(t)} d\langle W_i, W_j\rangle_t \\
&= \frac{\sigma_{V_i} V_{i,t}\Delta_i(t)}{S_{i,t}} \frac{\sigma_{V_j} V_j(t)\,\Delta_j(t)}{S_j(t)} \rho_{ij}\ dt \\
&= \sigma_{S_i}(t)\,\sigma_{S_j}(t)\,\rho_{ij}\ dt.
\end{aligned}$$

Using the sample covariance yields an estimate for ρ_{ij}. Note that this estimation procedure is again inconsistent with the model because the quadratic variation process between $\ln S_i$ and $\ln S_j$ is also time dependent and stochastic.

D Formulas for debt and default probability in the Longstaff and Schwartz (1995) model

As given in Vasicek (1977), the time t value $B(r_t,t,T)$ of a zero-coupon bond that pays F dollars at time T is

$$\begin{aligned}
B(r_t,t,T) &= F\exp^{\alpha_{T-t} r_t + \beta_{T-t}}, \qquad (29) \\
\alpha_{T-t} &= \kappa^{-1}\left(\exp^{-\kappa(T-t)} - 1\right), \\
\beta_{T-t} &= \frac{\sigma_r^2}{4\kappa_3}\left(1 - e^{-2\kappa(T-t)}\right) + \frac{1}{\kappa}\left(\alpha - \frac{\lambda}{\kappa} - \frac{\sigma_r^2}{\kappa^2}\right)\left(1 - e^{-\kappa(T-t)}\right) - \left(\alpha - \frac{\lambda}{\kappa} - \frac{\sigma_r^2}{\kappa^2}\right)(T-t).
\end{aligned}$$

The risky zero-coupon bond is assumed to pay, at time T, F dollars if the default has not occur and $F(1-w)$ otherwise. Default is triggered when the asset value V_t for the first time crosses below a threshold level K. The crossing time is denoted by τ. The time-t value of this risky bond is

$$D(V_t,r_t,t,T)\mathbf{1}_{\{\tau>t\}} = \mathbf{1}_{\{\tau>t\}} B(r_t,t,T)\left(1 - w\mathcal{Q}_t^T\left[\tau \le T\right]\right) \qquad (30)$$

where $Q_t^T[\tau \leq T]$ is the conditional probability, under the forward measure Q_t^T using the zero-coupon default-free bond as the numeraire [7], that the default occurs after time t but before or at time T. The formula for $Q_t^T[\tau \leq T]$ given in Longstaff and Schwartz's (1995) Proposition 1 is actually incorrect. This has been noted in Muller (2000) and Collin-Dufresne and Goldstein (2001). We refer readers to these two papers for two different ways of computing $Q_t^T[\tau \leq T]$.

The default probability under the physical probability measure, as compared to that under measure Q_t^T, is much simpler. It equals

$$\begin{aligned}
& 1 - P\left[\inf_{t\leq s\leq T} V_s > K\right] \\
= \; & 1 - P\left[\inf_{t\leq s\leq T} V_t \exp\left(\left(\mu - \frac{\sigma_V^2}{2}\right)(s-t) + \sigma_V\left(W_{V,s} - W_{V,t}\right)\right) > K\right] \\
= \; & 1 - P\left[\inf_{t\leq s\leq T} \left(\left(\mu - \frac{\sigma_V^2}{2}\right)(s-t) + \sigma_V\left(W_{V,s} - W_{V,t}\right)\right) > \ln K - \ln V_t\right] \\
= \; & \Phi\left(\frac{\ln\frac{K}{V_t} - \left(\mu - \frac{\sigma_V^2}{2}\right)(T-t)}{\sqrt{T-t}\,\sigma_V}\right) + \exp\left(\frac{2}{\sigma_V^2}\left(\mu - \frac{\sigma_V^2}{2}\right)\ln\frac{K}{V_t}\right)\Phi\left(\frac{\ln\frac{K}{V_t} + \left(\mu - \frac{\sigma_V^2}{2}\right)(T-t)}{\sqrt{T-t}\,\sigma_V}\right).
\end{aligned}$$

E The likelihood function for the constant interest rate version of the Longstaff and Schwartz (1995) model

We first assume that it is possible to observe the firm values. Let $U_{nh} = \inf_{(n-1)h\leq t\leq nh} \ln V_t$. Note that the event of "no default during the time interval $[0, Nh]$" can be expressed as

$$U_{nh} > \ln K \text{ for any } n \in \{1, 2, ..., N\}.$$

The log-likelihood function conditional on no default during the sample period is

$$\ln f_{V_0,...,V_{Nh}|\inf_{0\leq t\leq Nh} V_t > K}\left(v_0, v_h..., v_{Nh}\right) = \ln \frac{f_{V_0,...,V_{Nh},\inf_{0\leq t\leq Nh}\ln V_t > \ln K}\left(v_0, ..., v_{Nh}\right)}{\mathrm{P}\left[\inf_{0\leq t\leq Nh} \ln V_t > \ln K\right]} \tag{31}$$

where $f_{V_0,...,V_{Nh}|\inf_{0\leq t\leq Nh}\ln V_t > \ln K}$ is the joint density function of the firm values $V_0, ..., V_{Nh}$ and no default in the sample period.

The joint density function of the firm values $V_0, ..., V_{Nh}$ and the infimums $U_0, ..., U_{Nh}$ is

$$f_{V_0,U_h,V_h...,U_{Nh},V_{Nh}}\left(v_0, u_h, v_h..., u_{Nh}, v_{Nh}\right) = \prod_{n=1}^{N} f_{U_{nh},V_{nh}|V_{(n-1)h}}\left(u_{nh}, v_{nh}|v_{(n-1)h}\right)$$

[7] Under this measure, asset prices relative to $B(r_t, t, T)$ are martingales with respect to Q_t^T. See Madan and Unal (1999).

because V is Markovian. Therefore, the joint density function is

$$\begin{aligned}
& f_{V_0,\dots,V_{Nh},\inf_{0\le t\le Nh}\ln V_t>\ln K}\left(v_0,\dots,v_{Nh}\right) \\
= & \int_{\ln K}^{\infty}\dots\int_{\ln K}^{\infty} f_{V_0,U_h,V_h\dots,U_{Nh},V_{Nh}}\left(v_0,u_h,v_h\dots,u_{Nh},v_{Nh}\right)du_h\dots du_{Nh} \\
= & \int_{\ln K}^{\infty}\dots\int_{\ln K}^{\infty}\prod_{n=1}^{N} f_{U_{nh},V_{nh}\left|V_{(n-1)h}\right.}\left(u_{nh},v_{nh}\left|v_{(n-1)h}\right.\right)du_h\dots du_{Nh} \\
= & \prod_{n=1}^{N}\int_{\ln K}^{\infty} f_{U_{nh},V_{nh}\left|V_{(n-1)h}\right.}\left(u_{nh},v_{nh}\left|v_{(n-1)h}\right.\right)du_{nh}.
\end{aligned}$$

The next step is to compute $f_{U_{nh},V_{nh}\left|V_{(n-1)h}\right.}$. It is based on the the joint law of the arithmetic Brownian motion and its infimum. Let $X_{nh}=\ln V_{nh}$. Then

$$f_{U_{nh},V_{nh}\left|V_{(n-1)h}\right.}\left(u_{nh},v_{nh}\left|v_{(n-1)h}\right.\right)=\frac{1}{v_{nh}}f_{U_{nh},X_{nh}\left|X_{(n-1)h}\right.}\left(u_{nh},\ln v_{nh}\left|\ln v_{(n-1)h}\right.\right)$$

and

$$\begin{aligned}
& f_{V_0,\dots,V_{Nh},\inf_{0\le t\le Nh}\ln V_t>\ln K}\left(v_0,\dots,v_{Nh}\right) \\
= & \prod_{n=1}^{N}\frac{1}{v_{nh}}\int_{\ln K}^{\infty} f_{U_{nh},X_{nh}\left|X_{(n-1)h}\right.}\left(u_{nh},\ln v_{nh}\left|\ln v_{(n-1)h}\right.\right)du_{nh} \\
= & \begin{cases} \prod_{n=1}^{N}\left[\begin{array}{c} \frac{1}{v_{nh}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_V\sqrt{h}}\exp\left(-\frac{1}{2\sigma_V^2 h}\left(\ln\frac{v_{nh}}{v_{(n-1)h}}-\left(\mu-\frac{\sigma_V^2}{2}\right)h\right)^2\right) \\ \times\left(1-\exp\left(-\frac{2}{\sigma_V^2 h}\ln\frac{v_{(n-1)h}}{K}\ln\frac{v_{nh}}{K}\right)\right)\end{array}\right] & v_0,\dots,v_{Nh}>K \\ 0 & \text{otherwise.}\end{cases}
\end{aligned} \tag{32}$$

The denominator of equation (31) can be derived as

$$\begin{aligned}
\mathrm{P}\left[\inf_{0\le t\le Nh}\ln V_t>\ln K\right] = & \Phi\left(\frac{1}{\sqrt{Nh}\sigma_V}\left(\left(\mu-\frac{\sigma_V^2}{2}\right)Nh-\ln\frac{K}{v_0}\right)\right) \\
& -\exp\left(\frac{2}{\sigma_V^2}\left(\mu-\frac{\sigma_V^2}{2}\right)\ln\frac{K}{v_0}\right)\Phi\left(\frac{1}{\sqrt{Nh}\sigma_V}\left(\left(\mu-\frac{\sigma_V^2}{2}\right)Nh+\ln\frac{K}{v_0}\right)\right)
\end{aligned}$$

provided that $v_0>K$.

Since we only observe equity values, it is necessary to express the likelihood function in terms of equity values. we need to establish the relation that links together the asset and the equity values. On the set $\tau>t$, the time-t value of the risky zero-coupon bond is

$$\begin{aligned}
& D_t\left(V_t,r,T\right) \\
= & \exp^{-r(T-t)}F\left(1-wQ\left[\tau\le T\left|\tau>t\right.\right]\right) \\
= & \exp^{-r(T-t)}F\left(1-w\left[\begin{array}{c}\Phi\left(\frac{1}{\sqrt{T-t}\sigma_V}\left(\ln\frac{K}{V_t}-\left(r-\frac{\sigma_V^2}{2}\right)(T-t)\right)\right) \\ +\exp\left(\frac{2}{\sigma_V^2}\left(r-\frac{\sigma_V^2}{2}\right)\ln\frac{K}{V_t}\right)\Phi\left(\frac{1}{\sqrt{T-t}\sigma_V}\left(\ln\frac{K}{V_t}+\left(r-\frac{\sigma_V^2}{2}\right)(T-t)\right)\right)\end{array}\right]\right)
\end{aligned}$$

where Q is the risk-neutral measure, F is the face value of the debt, T is the maturity of the debt, $1-w$ is the recovery rate at default. Therefore, using the relation (1), the equity value is

$$\begin{aligned}
& S_t \\
= \; & V_t - D_t\left(V_t, r, T\right) \\
= \; & V_t - \exp^{-r(T-t)} F \left(1 - w \left[\begin{array}{c} \Phi\left(\frac{1}{\sqrt{T-t}\sigma_V}\left(\ln\frac{K}{V_t} - \left(r - \frac{\sigma_V^2}{2}\right)(T-t)\right)\right) \\ + \exp\left(\frac{2}{\sigma_V^2}\left(r - \frac{\sigma_V^2}{2}\right)\ln\frac{K}{V_t}\right) \\ \times\Phi\left(\frac{1}{\sqrt{T-t}\sigma_V}\left(\ln\frac{K}{V_t} + \left(r - \frac{\sigma_V^2}{2}\right)(T-t)\right)\right) \end{array} \right] \right) \\
= \; & g\left(V_t; t, T, \sigma_V\right).
\end{aligned}$$

Although function g is not a one-to-one mapping from V_t to S_t, we can consider the inverse function locally around a solution. Let v_{nh}^* be a solution. The relevant logarithm of the Jacobian becomes

$$\begin{aligned}
& -\sum_{n=1}^{N} \ln\left|\frac{\partial g\left(v_{nh}^*; nh, T, \sigma_V\right)}{\partial V}\right| \\
= \; & -\sum_{n=1}^{N} \ln\left| 1 - \frac{2Fw}{v_{nh}^*}\exp^{-r(T-t)} \left(\begin{array}{c} \frac{1}{\sqrt{T-t}\sigma_V}\phi\left(\frac{1}{\sqrt{T-t}\sigma_V}\left(\ln\frac{K}{v_{nh}^*} - \left(r - \frac{\sigma_V^2}{2}\right)(T-t)\right)\right) \\ + \frac{1}{\sigma_V^2}\left(r - \frac{\sigma_V^2}{2}\right)\exp^{\frac{2}{\sigma_V^2}\left(r - \frac{\sigma_V^2}{2}\right)\ln\frac{K}{v_{nh}^*}} \\ \times\Phi\left(\frac{1}{\sqrt{T-t}\sigma_V}\left(\ln\frac{K}{v_{nh}^*} + \left(r - \frac{\sigma_V^2}{2}\right)(T-t)\right)\right) \end{array} \right) \right|.
\end{aligned}$$

Putting all together gives rise to the result.

F The proof for Lemma 3

The conditional likelihood function, conditional on no default in the sample period, is

$$\begin{aligned}
& f_{V_0, r_0, \ldots, V_{Nh}, r_{Nh} | \inf_{0 \le t \le Nh} \ln V_t > \ln K}\left(v_0, r_0, v_h, r_h, \ldots, v_{Nh}, r_{Nh}\right) \\
= \; & \frac{f_{V_0, r_0, \ldots, V_{Nh}, r_{Nh}, \inf_{0 \le t \le Nh} \ln V_t > \ln K}\left(v_0, r_0, v_h, r_h, \ldots, v_{Nh}, r_{Nh}\right)}{\mathrm{P}\left[\inf_{0 \le t \le Nh} \ln V_t > \ln K\right]}
\end{aligned} \tag{33}$$

where $f_{V_0, r_0, \ldots, V_{Nh}, r_{Nh}, \inf_{0 \le t \le Nh} \ln V_t > \ln K}$ is the joint density function of the firm values $V_0, \ldots, V_{Nh}$, the interest rates $r_0, \ldots, r_{Nh}$ and no default in the sample period.

The joint density function of the firm values V_{nh}, the rates r_{nh} and the infimums $U(nh) = \inf_{(n-1)h \le t \le nh} \ln V_t$ is

$$\begin{aligned}
& f_{V_0, r_0, U_0, V_h, r_h, \ldots, U_{Nh}, V_{Nh}, r_{Nh}}\left(v_0, r_0, u_h, v_h, r_h, \ldots, u_{Nh}, v_{Nh}, r_{Nh}\right) \\
= \; & \prod_{n=1}^{N} f_{U_{nh}, V_{nh}, r_{nh} | v_{(n-1)h}, r_{(n-1)h}}\left(u_{nh}, v_{nh}, r_{nh}\right)
\end{aligned}$$

because the bivariate stochastic process (V, r) is Markovian. Therefore, the joint density function of the V_{Nh} and the absence of default is

$$\begin{aligned} & f_{V_0, r_0, \ldots, V_{Nh}, r_{Nh}, \inf_{0 \leq t \leq Nh} \ln V_t > \ln K} (v_0, r_0, \ldots, v_{Nh}, r_{Nh}) \\ = & \prod_{n=1}^{N} \frac{1}{v_{nh}} \int_{\ln K}^{\infty} f_{U(nh), V_{Nh}, r_{Nh} | v_{(n-1)h}, r_{(n-1)h}} (u_{nh}, \ln v_{nh}, r_{nh}) \, du_{nh} \end{aligned} \tag{34}$$

To complete the proof, we need the conditional density function $f_{U_{nh}, V_{Nh}, r_{Nh} | v_{(n-1)h}, r_{(n-1)h}}$. This can be obtained by first computing the conditional distribution function (Lemma 5). To complete the proof, it suffices to evaluate the definite integral in equation 34. These last two steps are straightforward but tedious. The details are available upon request.

Lemma 5 *Let $P_{(n-1)h}[\bullet]$ denotes the distribution conditional on $(v_{(n-1)h}, r_{(n-1)h})$. Assume $\rho \neq 0$. If $v_{(n-1)h} > K$ and $x > \ln K$, then*

$$\begin{aligned} & P_{(n-1)h} \left[U(nh) > \ln K,\ X(nh) \leq x \text{ and } r_{Nh} \leq r \right] \\ \cong & \Phi(\phi_1(r, x)) \left[\Phi\left(\eta^- + \frac{x - \ln K}{\sigma_V \sqrt{h}} \right) + \exp(\theta) \Phi\left(\eta^+ - \frac{x - \ln K}{\sigma_V \sqrt{h}} \right) \right] \\ & - sign(\rho) \times \Phi(\phi_2(r)) \left[\Phi(\eta^-) + \exp(\theta) \Phi(\eta^+) \right] \\ & + \int_{\phi_1(r,x)}^{\phi_2(r)} \left[\Phi\left(\eta^- + \frac{\sqrt{1-\rho^2}}{\rho} (\phi_2(r) - z) \right) + \exp(\theta) \Phi\left(\eta^+ - \frac{\sqrt{1-\rho^2}}{\rho} (\phi_2(r) - z) \right) \right] \phi(z) \, dz \end{aligned}$$

and $P_{(n-1)h}[U(nh) > \ln K,\ X(nh) \leq x$ and $r_{Nh} \leq r] = 0$ otherwise, where $\Phi(\bullet)$ and $\phi(\bullet)$ are the distribution and density function of a standard normal random variable,

$$\begin{aligned} \theta &= -\frac{2}{\sigma_V^2} \left(\mu - \frac{\sigma_V^2}{2} \right) \ln \frac{v_{(n-1)h}}{K}, \\ \eta^+ &= \frac{-\ln \frac{v_{(n-1)h}}{K} + \left(\mu - \frac{\sigma_V^2}{2} \right) h}{\sigma_V \sqrt{h}}, \\ \eta^- &= \frac{-\ln \frac{v_{(n-1)h}}{K} - \left(\mu - \frac{\sigma_V^2}{2} \right) h}{\sigma_V \sqrt{h}} \\ \phi_1(r, x) &= \frac{\ln v_{(n-1)h} - x + \left(\mu - \frac{\sigma_V^2}{2} \right) h}{\sigma_V \sqrt{h}} \frac{\rho}{\sqrt{1-\rho^2}} + \frac{\left(r - r_{(n-1)h} \right) - \kappa(\alpha - r_{(n-1)h}) h}{\sigma_r \sqrt{h}} \sqrt{\frac{1}{1-\rho^2}}, \\ \phi_2(r) &= \phi_1(r, x) + \frac{x - \ln K}{\sigma_V \sqrt{h}} \frac{\rho}{\sqrt{1-\rho^2}}. \end{aligned}$$

Proof of Lemma 5. Assume that $\rho > 0$. The case $\rho < 0$ can be proved in a similar manner.

$$
\begin{aligned}
& P_{(n-1)h}\left[U(nh) > \ln K,\ X(nh) \le x \text{ and } r_{Nh} \le r\right] \\
\cong\ & P_{(n-1)h}\left[\begin{array}{c}
\ln v_{(n-1)h} + \inf_{(n-1)h \le t \le nh}\left(\begin{array}{c}\left(\mu - \frac{\sigma_V^2}{2}\right)t \\ +\sigma_V\left(W_V(t) - W_V((n-1)h)\right)\end{array}\right) > \ln K, \\
\ln v_{(n-1)h} + \left(\mu - \frac{\sigma_V^2}{2}\right)h + \sigma_V\left(W_V(nh) - W_V((n-1)h)\right) \le x \\
\text{and } r_{(n-1)h} + \kappa(\alpha - r_{(n-1)h})h + \sigma_r\rho\left(W_V(nh) - W_V((n-1)h)\right) \\
+\sigma_r\sqrt{1-\rho^2}\left(W_V^*(nh) - W_V^*((n-1)h)\right) \le r
\end{array}\right] \\
=\ & P_{(n-1)h}\left[\begin{array}{c}
\inf_{(n-1)h \le t \le nh}\left(\begin{array}{c}\left(\mu - \frac{\sigma_V^2}{2}\right)t \\ +\sigma_V\left(W_V(t) - W_V((n-1)h)\right)\end{array}\right) > \ln \frac{K}{v_{(n-1)h}}, \\
\left(\mu - \frac{\sigma_V^2}{2}\right)h + \sigma_V\left(W_V(nh) - W_V((n-1)h)\right) \le x - \ln v_{(n-1)h} \\
\text{and } \left(\mu - \frac{\sigma_V^2}{2}\right)h + \sigma_V\left(W_V(nh) - W_V((n-1)h)\right) \\
\le \left(\mu - \frac{\sigma_V^2}{2}\right)h + \frac{\sigma_V}{\sigma_r\rho}\left(r - r_{(n-1)h} - \kappa(\alpha - r_{(n-1)h})h\right) \\
-\sigma_V\frac{\sqrt{1-\rho^2}}{\rho}\left(W_V^*(nh) - W_V^*((n-1)h)\right)
\end{array}\right]
\end{aligned}
$$

$$
\begin{aligned}
=\ & \int_{-\infty}^{\infty} P_{(n-1)h}\left[\begin{array}{c}
\inf_{(n-1)h \le t \le nh}\left(\begin{array}{c}\left(\mu - \frac{\sigma_V^2}{2}\right)t \\ +\sigma_V\left(W_V(t) - W_V((n-1)h)\right)\end{array}\right) > \ln \frac{K}{v_{(n-1)h}}, \\
\left(\mu - \frac{\sigma_V^2}{2}\right)h + \sigma_V\left(W_V(nh) - W_V((n-1)h)\right) \le x - \ln v_{(n-1)h} \\
\text{and } \left(\mu - \frac{\sigma_V^2}{2}\right)h + \sigma_V\left(W_V(nh) - W_V((n-1)h)\right) \\
\le \left(\mu - \frac{\sigma_V^2}{2}\right)h + \frac{\sigma_V}{\sigma_r\rho}\left(r - r_{(n-1)h} - \kappa(\alpha - r_{(n-1)h})h\right) - \sigma_V\frac{\sqrt{1-\rho^2}}{\rho}\sqrt{h}z
\end{array}\right]\phi(z)\,dz \\
=\ & \int_{-\infty}^{\infty} P_{(n-1)h}\left[\begin{array}{c}
\inf_{(n-1)h \le t \le nh}\left(\begin{array}{c}\left(\mu - \frac{\sigma_V^2}{2}\right)t \\ +\sigma_V\left(W_V(t) - W_V((n-1)h)\right)\end{array}\right) > \ln \frac{K}{v_{(n-1)h}} \\
\text{and } \left(\mu - \frac{\sigma_V^2}{2}\right)h + \sigma_V\left(W_V(nh) - W_V((n-1)h)\right) \\
\le \min\left(\begin{array}{c} x - \ln v_{(n-1)h}\ ,\left(\mu - \frac{\sigma_V^2}{2}\right)h \\ +\frac{\sigma_V}{\sigma_r\rho}\left(r - r_{(n-1)h} - \kappa(\alpha - r_{(n-1)h})h\right) - \sigma_V\frac{\sqrt{1-\rho^2}}{\rho}\sqrt{h}z\end{array}\right)
\end{array}\right]\phi(z)\,dz.
\end{aligned}
$$

Note that

$$
\begin{aligned}
& x - \ln v_{(n-1)h} \le \left(\mu - \frac{\sigma_V^2}{2}\right)h + \frac{\sigma_V}{\sigma_r\rho}\left(r - r_{(n-1)h} - \kappa(\alpha - r_{(n-1)h})h\right) - \sigma_V\frac{\sqrt{1-\rho^2}}{\rho}\sqrt{h}z \\
\Leftrightarrow\ & z \le \underbrace{\frac{1}{\sqrt{1-\rho^2}}\left(\frac{\rho}{\sigma_V\sqrt{h}}\left(\left(\mu - \frac{\sigma_V^2}{2}\right)h - x + \ln v_{(n-1)h}\right) + \frac{1}{\sigma_r\sqrt{h}}\left(r - r_{(n-1)h} - \kappa(\alpha - r_{(n-1)h})h\right)\right)}_{=\phi_1(r,x)}
\end{aligned}
$$

Hence

$$
\begin{aligned}
& P_{(n-1)h}\left[U_{nh} > \ln K,\ X_{nh} \leq x \text{ and } r_{nh} \leq r\right] \\
\cong\ & \int_{-\infty}^{\phi_1(r,x)} P_{(n-1)h}\left[\begin{array}{c} \inf_{(n-1)h \leq t \leq nh}\left(\begin{array}{c}\left(\mu - \frac{\sigma_V^2}{2}\right)t \\ +\sigma_V\left(W_V(t) - W_V((n-1)h)\right)\end{array}\right) > \ln \frac{K}{v_{(n-1)h}} \\ \text{and } \left(\mu - \frac{\sigma_V^2}{2}\right)h + \sigma_V\left(W_V(nh) - W_V((n-1)h)\right) \leq x - \ln v_{(n-1)h}\end{array}\right]\phi(z)\,dz \\
& + \int_{\phi_1(r,x)}^{+\infty} P_{(n-1)h}\left[\begin{array}{c} \inf_{(n-1)h \leq t \leq nh}\left(\begin{array}{c}\left(\mu - \frac{\sigma_V^2}{2}\right)t \\ +\sigma_V\left(W_V(t) - W_V((n-1)h)\right)\end{array}\right) > \ln \frac{K}{v_{(n-1)h}} \\ \text{and } \left(\mu - \frac{\sigma_V^2}{2}\right)h + \sigma_V\left(W_V(nh) - W_V((n-1)h)\right) \\ \leq \left(\mu - \frac{\sigma_V^2}{2}\right)h + \frac{\sigma_V}{\sigma_r\rho}\left(r - r_{(n-1)h} - \kappa(\alpha - r_{(n-1)h})h\right) - \sigma_V\frac{\sqrt{1-\rho^2}}{\rho}\sqrt{h}z\end{array}\right]\phi(z)\,dz \\
=\ & \Phi\left(\phi_1(r,x)\right)P_{(n-1)h}\left[\begin{array}{c} \inf_{(n-1)h \leq t \leq nh}\left(\left(\mu - \frac{\sigma_V^2}{2}\right)t + \sigma_V\left(W_V(t) - W_V((n-1)h)\right)\right) > \ln \frac{K}{v_{(n-1)h}} \\ \text{and } \left(\mu - \frac{\sigma_V^2}{2}\right)h + \sigma_V\left(W_V(nh) - W_V((n-1)h)\right) \leq x - \ln v_{(n-1)h}\end{array}\right] \\
& + \int_{\phi_1(r,x)}^{+\infty} P_{(n-1)h}\left[\begin{array}{c} \inf_{(n-1)h \leq t \leq nh}\left(\begin{array}{c}\left(\mu - \frac{\sigma_V^2}{2}\right)t \\ +\sigma_V\left(W_V(t) - W_V((n-1)h)\right)\end{array}\right) > \ln \frac{K}{v_{(n-1)h}} \\ \text{and } \left(\mu - \frac{\sigma_V^2}{2}\right)h + \sigma_V\left(W_V(nh) - W_V((n-1)h)\right) \\ \leq \left(\mu - \frac{\sigma_V^2}{2}\right)h + \frac{\sigma_V}{\sigma_r\rho}\left(r - r_{(n-1)h} - \kappa(\alpha - r_{(n-1)h})h\right) - \sigma_V\frac{\sqrt{1-\rho^2}}{\rho}\sqrt{h}z\end{array}\right]\phi(z)\,dz
\end{aligned}
$$

We apply the identity

$$
\begin{aligned}
& P\left[\inf_{0 \leq s \leq t}(\sigma W_s + \alpha s) \geq y \text{ and } \sigma W_t + \alpha t \leq x\right] \\
=\ & \begin{cases} \Phi\left(\frac{x-\alpha t}{\sigma\sqrt{t}}\right) - \Phi\left(\frac{y-\alpha t}{\sigma\sqrt{t}}\right) - \exp\left(2\frac{\alpha y}{\sigma^2}\right)\Phi\left(\frac{y+t\alpha}{\sigma\sqrt{t}}\right) + \exp\left(2\frac{\alpha y}{\sigma^2}\right)\Phi\left(\frac{2y-x+t\alpha}{\sigma\sqrt{t}}\right) & \text{if } y \leq \min(x,0) \\ 0 & \text{otherwise,}\end{cases}
\end{aligned}
$$

to both terms of the last expression.

For the first term, the condition $y \leq \min(x,0)$ becomes $\exp(x) \geq K$ and $v_{(n-1)h} \geq K$. For the second term, the condition $y \leq \min(x,0)$ becomes

$$
z \leq \underbrace{\frac{1}{\sqrt{1-\rho^2}}\left(\frac{\rho}{\sigma_V\sqrt{h}}\left(\left(\mu - \frac{\sigma_V^2}{2}\right)h - \ln\frac{K}{v_{(n-1)h}}\right) + \frac{1}{\sigma_r\sqrt{h}}\left(r - r_{(n-1)h} - \kappa(\alpha - r_{(n-1)h})h\right)\right)}_{=\phi_2(r)}
$$

and $v_{(n-1)h} \geq K$, which implies that the integral only goes from $\phi_1(r,x)$ to $\phi_2(r)$. Note that

$$
\phi_2(r) = \phi_1(r,x) + \underbrace{\frac{\rho}{\sqrt{1-\rho^2}}\frac{x - \ln K}{\sigma_V\sqrt{h}}}_{\geq 0}.
$$

If $\exp(x) \geq k$ and $v_{(n-1)h} \geq K$, then

$$
\begin{aligned}
& P_{(n-1)h}\left[U_{nh} > \ln K,\ X_{nh} \leq x \text{ and } r_{nh} \leq r\right] \\
\cong \quad & \Phi\left(\phi_1(r,x)\right) \underbrace{\left(\begin{array}{c} \Phi\left(\frac{x - \ln v_{(n-1)h} - \left(\mu - \frac{\sigma_V^2}{2}\right)h}{\sigma_V \sqrt{h}} \right) - \Phi\left(\eta^-\right) \\ - \exp(\theta)\, \Phi\left(\eta^+\right) + \exp(\theta)\, \Phi\left(\frac{2 \ln \frac{K}{v_{(n-1)h}} - x + \ln v_{(n-1)h} + \left(\mu - \frac{\sigma_V^2}{2}\right)h}{\sigma_V \sqrt{h}} \right) \end{array} \right)}_{\Phi\left(\eta^- + \frac{x - \ln K}{\sigma_V \sqrt{h}}\right) - \Phi(\eta^-) - \exp(\theta)\Phi(\eta^+) + \exp(\theta)\Phi\left(\eta^+ - \frac{x - \ln K}{\sigma_V \sqrt{h}}\right)} \\
& + \int_{\phi_1(r,x)}^{\phi_2(r)} \underbrace{\left(\begin{array}{c} \Phi\left(\frac{\frac{\sigma_V}{\sigma_r \rho}\left(r - r_{(n-1)h} - \kappa(\alpha - r_{(n-1)h})h\right) - \sigma_V \frac{\sqrt{1-\rho^2}}{\rho}\sqrt{h} z}{\sigma_V \sqrt{h}} \right) - \Phi\left(\eta^-\right) - \exp(\theta)\, \Phi\left(\eta^+\right) \\ + \exp(\theta)\, \Phi\left(\frac{2 \ln \frac{K}{v_{(n-1)h}} - \frac{\sigma_V}{\sigma_r \rho}\left(r - r_{(n-1)h} - \kappa(\alpha - r_{(n-1)h})h\right) + \sigma_V \frac{\sqrt{1-\rho^2}}{\rho}\sqrt{h} z}{\sigma_V \sqrt{h}} \right) \end{array} \right)}_{= \Phi\left(\eta^- + \frac{\sqrt{1-\rho^2}}{\rho}(\phi_2(r) - z)\right) - \Phi(\eta^-) - \exp(\theta)\Phi(\eta^+) + \exp(\theta)\Phi\left(\eta^+ - \frac{\sqrt{1-\rho^2}}{\rho}(\phi_2(r) - z)\right)} \phi(z)\, dz \\
= \quad & \Phi\left(\phi_1(r,x)\right) \left[\Phi\left(\eta^- + \frac{x - \ln K}{\sigma_V \sqrt{h}}\right) + \exp(\theta)\, \Phi\left(\eta^+ - \frac{x - \ln K}{\sigma_V \sqrt{h}}\right) \right] \\
& - \Phi\left(\phi_2(r)\right) \left[\Phi\left(\eta^-\right) + \exp(\theta)\, \Phi\left(\eta^+\right) \right] \\
& + \int_{\phi_1(r,x)}^{\phi_2(r)} \left[\Phi\left(\eta^- + \frac{\sqrt{1-\rho^2}}{\rho}\left(\phi_2(r) - z\right)\right) + \exp(\theta)\, \Phi\left(\eta^+ - \frac{\sqrt{1-\rho^2}}{\rho}\left(\phi_2(r) - z\right)\right) \right] \phi(z)\, dz. \ \square
\end{aligned}
$$

G The likelihood function for the full version of the Longstaff and Schwartz (1995) model

The instantaneous interest rate r_t is unobserved. We relate it to the zero-coupon bond price $B(r_t, t, T)$ using equation (29). The second mapping is the equity pricing equation which links the unobserved asset value to the observed equity value. Using equations (1) and (30) gives rise to

$$
\begin{aligned}
S_t &= V_t - D(V_t, r_t, t, T) \\
&= V_t - B(r_t, t, T) + wB(r_t, t, T)\mathcal{Q}_t^T\left(V_t, r_t, t, T\right) \\
&= g\left(V_t; r_t, t, T, \theta\right)
\end{aligned}
$$

where $Q_t^T\left(V_t, r_t, t, T\right)$ is the conditional default probability under the forward measure Q_t^T.

Let Y_t be the vector of observed variables at time t; that is, $Y_t = [B(r_t, t, T_t), S_t]'$, $\mathbf{Y} = [\mathbf{y}_0', \mathbf{y}_h', ..., \mathbf{y}_{Nh}']'$ denotes the sample of all observed variables; that is, $\mathbf{y}_{nh}' = [b_{nh}, s_{nh}]'$. We want to establish the mapping that links together the observed and unobserved variables at time t. For interest rates, the relationship is straightforward from equation (29):

$$
r_t = \frac{\ln b_t - \beta_{T_t - t}}{\alpha_{T_t - t}}.
$$

For the implied firm value, we find the solution to $s_t = g(v_t; r_t, t, \theta)$. Due to a peculiar feature of the Longstaff and Schwartz (1995) model, the solution may not be unique but there will be at most two. For each solution $(v_0^*, r_0^*, v_h^*, r_h^*, ..., v_{Nh}^*, r_{Nh}^*)$, we can apply the transformed data method of Duan (1994, Theorem 2.1) to yield the log-likelihood:

$$L(\mathbf{Y};\theta) = G\left(v_0^*, r_0^*, v_h^*, r_h^*, ..., v_{Nh}^*, r_{Nh}^*; \theta\right) + \ln\left(\left|\det\left(\mathbf{D}_{\mathbf{Y}}^{-1}\right)\right|\right),$$

where $\mathbf{D_Y}$ is a matrix of partial derivatives computed from the transformation relating the unobservable variables to the observed and is evaluated at $(v_0^*, r_0^*, v_h^*, r_h^*, ..., v_{Nh}^*, r_{Nh}^*)$. More precisely, the matrix $\mathbf{D_Y}$ is a block diagonal matrix:

$$\mathbf{D}_Y = \begin{bmatrix} \mathbf{D}_{\mathbf{Y}_0} & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \mathbf{D}_{\mathbf{Y}_{Nh}} \end{bmatrix}$$

with

$$\mathbf{D}_{\mathbf{Y}_t} = \begin{bmatrix} \frac{\partial B(r_t,t,T_t)}{\partial r} & \frac{\partial B(r_t,t,T_t)}{\partial V} \\ \frac{\partial S_t}{\partial r} & \frac{\partial S_t}{\partial V} \end{bmatrix}.$$

The individual elements in each of the $N+1$ matrices are given by

$$\begin{aligned}
\frac{\partial B(r_t,t,T_t)}{\partial r} &= \alpha_{T_t-t} B(r_t,t,T_t), \\
\frac{\partial B(r_t,t,T_t)}{\partial V} &= 0, \\
\frac{\partial S_t}{\partial r} &= \left(-\alpha_{T-t} + w\alpha_{T-t}\mathcal{Q}_t^T(v_t,r_t,t,T) + w\frac{\partial \mathcal{Q}_t^T(v_t,r_t,t,T)}{\partial r}\right) B(r_t,t,T) \\
\frac{\partial S_t}{\partial V} &= 1 + wB(r_t,t,T)\frac{\partial \mathcal{Q}_t^T(v_t,r_t,t,T)}{\partial V}.
\end{aligned}$$

The determinant of $\mathbf{D}_{\mathbf{Y}_t}$ is found to be

$$\det\left(\mathbf{D}_{\mathbf{Y}_t}\right) = \alpha_{t_T-t} B(r_t,t,T_t)\left(1 + wB(r_t,t,T)\frac{\partial}{\partial V}\mathcal{Q}_t^T(v_t,r_t,t,T)\right).$$

Since the determinant of the inverse of the block diagonal matrix $\mathbf{D_Y}$ is simply

$$\det\left(\mathbf{D}_{\mathbf{Y}}^{-1}\right) = \prod_{n=0}^{N} \frac{1}{\det\left(\mathbf{D}_{\mathbf{Y}_{nh}}\right)},$$

The Jacobian term in the log-likelihood function thus becomes

$$\begin{aligned}
\ln\left(\left|\det\left(\mathbf{D}_{\mathbf{Y}}^{-1}\right)\right|\right) &= -\sum_{n=0}^{N} \ln\left|\det\left(\mathbf{D}_{\mathbf{Y}_{nh}}\right)\right| \\
&= -\sum_{n=0}^{N} \ln\left|\alpha_{T_{nh}-nh} b_{nh}\left(1 + wb_{nh}\frac{\partial}{\partial v}\mathcal{Q}_{nh}^T(v_{nh}^*, r_{nh}^*, nh, T)\right)\right|.
\end{aligned}$$

Summing over the likelihood functions for all possible combinations gives rise to the result.

References

[1] Black, F. and M. Scholes, 1973, The Pricing of Options and Corporate Liabilities, Journal of Political Economy 81, 637–659.

[2] Brys, E. and F. de Varenne, 1997, Valuing Risky Fixed Rate Debt: An Extension, Journal of Financial and Quantitative Analysis 32, 239-248.

[3] Collin-Dufresne, P. and R. Goldstein, 2001, Do Credit Spreads Reflect Stationary Leverage Ratios, Journal of Finance 56, 1929-1957.

[4] Crouhy, M., and R. Mark, 1998, Model Risk, Journal of Financial Engineering, forthcoming.

[5] Crouhy, M., Galai, D. and R. Mark, 2000, A Comparative Analysis of Current Credit Risk Models, Journal of Banking and Finance 24, 59-117

[6] Duan, J.C., 1994, Maximum Likelihood Estimation Using Price Data of the Derivative Contract, Mathematical Finance 4, 155-167.

[7] Duan, J.C., 2000, Correction: "Maximum Likelihood Estimation Using Price Data of the Derivative Contract, Mathematical Finance 10, 461-462.

[8] Duan, J.C., A. Moreau and C.W. Sealey, 1995. Deposit Insurance and Bank Interest Rate Risk: Pricing and Regulatory Implications, Journal of Banking and Finance 19, 1091-1108.

[9] Duan, J.C. and J.G. Simonato, 2002, Maximum Likelihood Estimation of Deposit Insurance Value with Interest Rate Risk, Journal of Empirical Finance 9, 109-132.

[10] Duan, J.C. and M.T. Yu, 1994, Assessing the Cost of Taiwan's Deposit Insurance, Pacifice Basin Finance Journal 2, 73-90.

[11] Ericsson, J. and J. Reneby, 2001, Estimating Structural Bond Pricing Models, McGill University working paper.

[12] Goldfeld, S., Quandt, R. and H. Trotter, 1966, Maximization by Quadratic Hill-Climbing, Econometrica 34, 541-551

[13] Jarrow, R. and S. Turnbull, 2000, The Intersection of Market and Credit Risk, Journal of Banking and Finance 24, 271-299

[14] Jones, E., Mason, S. and E. Rosenfeld, 1984, Contingent Claims Analysis of Corporate Capital Structures: An Empirical Investigation, Journal of Finance 39, 611-627.

[15] Laeven, L., 2002, Bank Risk and Deposit Insurance, World Bank Economic Review 16, 109-137.

[16] Lo, A., 1986, Statistical Tests of Contingent-Claims Asset-Pricing Models - A New Methodology, Journal of Financial Economics 17, 143-173.

[17] Longstaff, F. and Schwartz, E., 1995, A Simple approach to Valuing Risky Fixed and Floating Rate Debt, Journal of Finance 49, 1213-1252.

[18] Madan, D., and H. Unal, 1998, Pricing the Risks of Default, Review of Derivatives Research 2, 121-160.

[19] Madan, D., and H. Unal, 2000, A Two-Factor Hazard-Rate Model for Pricing Risky Debt and the Term Structure of Credit Spread, Journal of Financial and Quantitative Analysis 35, 43-65.

[20] Merton, R.C., 1974, On the Pricing of Corporate Debt: the Risk Structure of Interest Rates, Journal of Finance 28, 449-470.

[21] Merton, R.C., 1977, An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees, Journal of Banking and Finance 1, 3–11.

[22] Muller, C., 2000, A Simple Multi-Factor Model of Corporate Bond Prices, working paper, University of Wisconsin – Madison.

[23] Rao, C. R. 1973, Linear Statistical Inference and Its Application, second edition, Wiley, New York.

[24] Ronn, E.I. and A.K. Verma, 1986, Pricing Risk–Adjusted Deposit Insurance: An Option-Based Model, Journal of Finance 41, 871–895.

[25] Saunders, A., 1999, Credit Risk Measurement: New Approaches to Value at Risk and Other Paradigms, Wiley.

Table 1: Simulation results for Merton's model in the case of high leverage-high noise to signal

Maximum Likelihood

	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_{V_1}$	$\hat{\sigma}_{V_2}$	$\hat{\rho}_{12}$	$\hat{V}_{1t_0}$-V_{1t_0}	$\hat{V}_{2t_0}$-V_{2t_0}	$\hat{C}_{1t_0}$-C_{1t_0}	$\hat{C}_{2t_0}$-C_{2t_0}	$\hat{x}_{1t_0}$-x_{1t_0}	$\hat{x}_{2t_0}$-x_{2t_0}	$\hat{P}_{1t_0}$-P_{1t_0}	$\hat{P}_{2t_0}$-P_{2t_0}
True	0.100	0.100	0.300	0.300	0.500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Mean	0.101	0.095	0.300	0.300	0.500	-0.784	1.853	0.000	-0.000	-0.010	0.012	0.048	0.049
Median	0.102	0.098	0.299	0.299	0.501	0.155	0.118	-0.000	-0.000	-0.003	0.008	-0.000	0.001
Std	0.209	0.208	0.018	0.018	0.033	110.522	116.660	0.020	0.021	0.711	0.702	0.080	0.080
25 % cvr	0.258	0.251	0.250	0.255	0.244	0.252	0.255	0.252	0.255	0.260	0.259	0.246	0.246
50 % cvr	0.514	0.516	0.506	0.504	0.497	0.506	0.509	0.507	0.509	0.512	0.512	0.433	0.438
75 % cvr	0.751	0.756	0.754	0.749	0.757	0.752	0.750	0.753	0.750	0.747	0.759	0.568	0.573
95 % cvr	0.951	0.955	0.947	0.942	0.953	0.934	0.933	0.934	0.932	0.952	0.955	0.679	0.685

Implicit

	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_{V_1}$	$\hat{\sigma}_{V_2}$	$\hat{\rho}_{12}$	$\hat{V}_{1t_0}$-V_{1t_0}	$\hat{V}_{2t_0}$-V_{2t_0}	$\hat{C}_{1t_0}$-C_{1t_0}	$\hat{C}_{2t_0}$-C_{2t_0}	$\hat{x}_{1t_0}$-x_{1t_0}	$\hat{x}_{2t_0}$-x_{2t_0}	$\hat{P}_{1t_0}$-P_{1t_0}	$\hat{P}_{2t_0}$-P_{2t_0}
True	–	–	0.300	0.300	0.500	0.000	0.000	0.000	0.000	–	–	–	–
Mean	–	–	0.230	0.228	0.492	612.955	632.409	-0.086	0.023	–	–	–	–
Median	–	–	0.243	0.240	0.493	134.507	141.436	-0.016	-0.122	–	–	–	–
Std	–	–	0.119	0.120	0.036	1003.317	1022.175	0.153	0.755	–	–	–	–

Incomplete Maximum Likelihood

	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_{V_1}$	$\hat{\sigma}_{V_2}$	$\hat{\rho}_{12}$	$\hat{V}_{1t_0}$-V_{1t_0}	$\hat{V}_{2t_0}$-V_{2t_0}	$\hat{C}_{1t_0}$-C_{1t_0}	$\hat{C}_{2t_0}$-C_{2t_0}	$\hat{x}_{1t_0}$-x_{1t_0}	$\hat{x}_{2t_0}$-x_{2t_0}	$\hat{P}_{1t_0}$-P_{1t_0}	$\hat{P}_{2t_0}$-P_{2t_0}
True	0.100	0.100	0.300	0.300	0.500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Mean	0.085	0.080	0.128	0.129	0.499	442.768	453.016	-0.064	-0.066	-0.274	-0.247	0.015	0.017
Median	0.066	0.060	0.219	0.218	0.500	189.563	199.785	-0.023	-0.024	-0.355	-0.337	-0.007	-0.006
Std	0.189	0.186	0.175	0.174	0.033	586.419	591.893	0.099	0.100	0.990	0.985	0.088	0.089
25 % cvr	0.159	0.156	0.001	0.001	0.158	0.001	0.001	0.001	0.001	0.134	0.130	0.128	0.126
50 % cvr	0.327	0.335	0.001	0.001	0.336	0.001	0.001	0.001	0.001	0.284	0.276	0.249	0.241
75 % cvr	0.518	0.525	0.005	0.005	0.531	0.003	0.003	0.003	0.003	0.466	0.467	0.351	0.353
95 % cvr	0.747	0.751	0.020	0.018	0.768	0.010	0.011	0.010	0.011	0.742	0.747	0.462	0.465

True is the parameter value used in the Monte Carlo simulation; **Mean**, **Median** and **Std** are the sample statistics computed with the 5000 estimated parameter values; **cvr** is the coverage rate defined as the percentage of the 5000 parameter estimates for which the true parameter value is contained in the α confidence interval implied by the asymptotic distribution; $V_{0,1} = 10000$, $V_{0,2} = 10000$, $F_1 = 9000$, $F_2 = 9000$, $T = 3.00$, $t_0 = 2.00$, $r = 0.05$ and $N = 500$.

Table 2

Corrected confidence intervals for the default probability

	$\hat{P}_{1t_0}$-P_{1t_0}	$\hat{P}_{2t_0}$-P_{2t_0}
25 % cvr	0.260	0.259
50 % cvr	0.512	0.512
75 % cvr	0.747	0.759
95 % cvr	0.952	0.955

cvr is the coverage rate defined as the percentage of the 5000 parameter estimates for which the true parameter value is contained in the α confidence interval implied by the asymptotic distribution; cvr is the coverage rate defined as the percentage of the 5000 parameter estimates for which the true parameter value is contained in the α confidence interval implied by the asymptotic distribution; $V_{0,1} = 10000$, $V_{0,2} = 10000$, $F_1 = 9000$, $F_2 = 9000$, $T = 3.00$, $t_0 = 2.00$, $r = 0.05$ and $N = 500$.

Table 3

Maximum Likelihood Estimation (Merton's model)

Year	$\hat{\mu}$	$\hat{\sigma}$	$\hat{V}_{t_0}$ \$000's	$\hat{C}_{t_0}$	$\hat{P}_{t_0}$	Correlation $\hat{\rho}$
			Hudson's Bay			
1999	0.011	0.204	2,851,034	2.075e-005	0.001	0.173
	(0.142)	(0.002)	(4)	(2.969e-006)	(0.002)	(0.040)
2000	-0.007	0.233	3,046,716	9.955e-004	0.026	0.157
	(0.178)	(0.002)	(140)	(7.912e-005)	(0.047)	(0.039)
2001	-0.040	0.226	2,488,371	6.637e-004	0.026	0.120
	(0.161)	(0.002)	(72)	(5.095e-005)	(0.043)	(0.044)
			Sears Canada			
1999	0.193	0.190	3,317,074	1.729e-007	0.000	
	(0.134)	(0.004)	(0)	(8.080e-008)	(0.000)	
2000	0.261	0.206	5740209	1.987e-012	0.000	
	(0.147)	(0.005)	(0)	(0.000e+000)	(0.000)	
2001	0.113	0.235	4089644	2.226e-005	0.000	
	(0.165)	(0.004)	(11)	(6.240e-006)	(0.000)	

The maximum likelihood estimates at the begining of Year are computed using the previous two years of daily time series data; $r =$ is set to the yield to maturity of a representative one year canadian government bond while the maturity of debt is set equal to 1.0 year at all point in time; For each firm, the face value of debt is set equal to $0.5\times$ long term debt plus short term debt obtained from the yearly annual report.