On the Equivalence of the KMV and Maximum Likelihood Methods for Structural Credit Risk Models

Jin-Chuan Duan, Geneviève Gauthier and Jean-Guy Simonato[*]

June 15, 2005

Abstract

Moody's KMV method is a popular commercial implementation of the structural credit risk model pioneered by Merton (1974). Among the key features of their implementation procedure is an algorithm for estimating the unobserved asset value and the unknown parameters. This estimation method has found its way to the recent academic literature, but has not yet been formally analyzed in terms of its statistical properties. This paper fills this gap and shows that, in the context of Merton's model, the KMV estimates are identical to maximum likelihood estimates (MLE) developed in Duan (1994). Unlike the MLE method, however, the KMV algorithm is silent about the distributional properties of the estimates and thus ill-suited for statistical inference. The KMV algorithm also cannot generate estimates for capital-structure specific parameters. In contrast, the MLE approach is flexible and can be readily applied to different structural credit risk models.

Keywords: Credit risk, transformed data, maximum likelihood, financial distress, EM algorithm.

JEL classification code: C22, G13.

1 Introduction

Credit risk modeling has gained increasing prominence over the years. The interest was in a large part motivated by new regulatory requirements, such as the Basel Accord, which provide strong incentives for financial institutions to quantitatively measure and manage risks of their corporate debt portfolios. Financial institutions have thus either developed their own internal models or relied on third-party software to assess various measures for the credit risks of their portfolios.

There exist two conceptually different modeling approaches to credit risk – structural and reduced-form. The structural approach is rooted in the seminal paper of Merton (1974) in which the firm's asset value is assumed to follow a geometric Brownian motion and the firm's capital structure to consist of a zero-coupon debt and common equity. This structural approach then yields formulas for the value of the risky corporate bond and the default probability of the firm. The structural approach is conceptually elegant but is laden with implementation problems. As pointed out in Jarrow and Turnbull (2000), the firm's asset values are unobservable and the model parameters (the assets' expected return and volatility in the case of Merton's model) are unknown. Unobservability is, in fact, considered by many to be the major drawback of the structural credit risk models.

A popular commercial implementation of the structural credit risk model is Moody's KMV method. This proprietary software builds on a variation of Merton's (1974) credit risk model and prescribes an iterative algorithm which infers, from the firm's equity price time series, the firm's unobserved total asset value and unknown expected return and volatility, which are the quantities required for computing, for example, credit spread and default probability.[1] In addition to its popularity among financial institutions, the KMV estimation method has also found its way into the recent academic literature; for example, Vassalou and Xing (2004) use the computed default probability to study equity returns, Cao, Simin and Zhao (2005) use the computed asset values and volatilities to examine growth options, and Bharath and Shumway (2004) use the computed default probability to study forecasting performance. Albeit its popularity, not much is known about the statistical properties of the KMV estimates. In fact, it is not entirely clear as to whether this method is statistically sound.

---

[1]It is our understanding that the KMV method subjects the estimates to further proprietary calibrations. The KMV estimation method discussed in this paper is the one described in Crosbie and Bohn (2003), which is only restricted to the iterative algorithm.

In the academic literature, there exist at least three other ways of dealing with the unobservability issue. First, a proxy asset value may be computed as the sum of the market value of the firm's equity and the book value of liabilities. Examples of using a market value proxy are Brockman and Turtle (2003) and Eom, *et al* (2004). The second approach is based on solving a system of equations; for example in the case of Merton's model, the equity pricing and volatility formulas form a system of two equations linking the unknown asset value and asset volatility to the equity value and equity volatility. This approach originating from Jones, *et al* (1984) and Ronn and Verma (1986) has been used extensively in the deposit insurance literature. The third approach put forward in Duan (1994) is based on maximum likelihood estimation (MLE) which views the observed equity time series as a transformed data set with the theoretical equity pricing formula serving as the transformation. This transformed-data MLE method has been applied to several structural models in Ericsson and Reneby (2003) and Wong and Choi (2004) and adapted to address survivorship in Duan, *et al* (2003).

The benefits of using the MLE method are well understood in statistics and econometrics. In the context of structural credit risk models, its dominance over the two other approaches discussed in the preceding paragraph has already been documented in Ericsson and Reneby (2003), Duan, *et al* (2003) and Wong and Choi (2004). But what statistical properties does the KMV method possess? This question is of both academic and practical interest because of the method's popularity. We show here that the KMV estimates are identical to the maximum likelihood estimates in the case of Merton's (1974) model. Our theoretical argument relies on characterizing the KMV method as an EM algorithm, a well-known approach for obtaining maximum likelihood estimates. The transformed-data MLE method is a better alternative, however, because statistical inference follows naturally in that framework. The KMV method, on the other hand, only generates point estimates. Perhaps more importantly, the KMV method does not work for structural credit risk models that involve any unknown capital structure parameter(s) such as, for example, the financial distress level in Brockman and Turtle (2003). The transformed-data MLE approach can still be applied to such a model to obtain an estimate for the financial distress level along with other model parameters, as demonstrated in Wong and Choi (2004). In short, the KMV and MLE methods are equivalent only in a limited sense because they lead to the same point estimates for the specific structural model of Merton (1974) but differ for more general structural models.

The paper is organized as follows. Section 2 presents an argument showing the equivalence of

the KMV and MLE estimator in the Merton (1974) context. We examine the equivalence both numerically and theoretically and show that the KMV estimator agrees with the transformed-data MLE estimator but is deficient for the inference purpose. Section 3 discusses the limitations of the KMV approach. The Brockman and Turtle (2003) model is used there to demonstrate that the KMV method does not generate the capital-structure specific parameter. Section 4 concludes.

2 Equivalence of the KMV and maximum likelihood estimates under Merton's (1974) model

2.1 Merton's (1974) model

Merton (1974) derived a risky bond pricing formula using the derivative pricing technology and a simple capital structure assumption. The firm's assets are financed by equity and a zero-coupon debt with a face value of F maturing at time T. The risk-free interest rate r is assumed to be constant. The market value of assets, equity and risky debt at time t are denoted by V_t, S_t and D_t, respectively. Naturally, the following accounting identity holds for every time point :

$$V_t = S_t + D_t. \tag{1}$$

It was further assumed that the asset value follows a geometric Brownian motion :

$$d \ln V_t = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dW_t \tag{2}$$

where μ and σ are, respectively, the expected return and volatility rates, and W_t is a Wiener process. The risky debt in this framework is entitled to a time-T cash flow of $\min \{V_T, F\}$ whose market value prior to time T can be obtained using the standard risk-neutral pricing theory; that is,

$$D_t = Fe^{-r(T-t)} \left(\frac{V_t}{Fe^{-r(T-t)}} \Phi(-d_t) + \Phi \left(d_t - \sigma \sqrt{T-t} \right) \right) \tag{3}$$

where $\Phi(\bullet)$ is the standard normal distribution function and $d_t = \frac{\ln(V_t/F) + \left(r + \frac{1}{2}\sigma^2 \right)(T-t)}{\sigma\sqrt{T-t}}$. In addition, the formulas for the conditional default probability and equity value can be readily derived to be :

$$P_t = \Phi \left(\frac{\ln(F/V_t) - \left(\mu - \frac{1}{2}\sigma^2 \right)(T-t)}{\sigma\sqrt{T-t}} \right), \tag{4}$$

and

$$S_t = V_t \Phi(d_t) - Fe^{-r(T-t)} \Phi \left(d_t - \sigma\sqrt{T-t} \right). \tag{5}$$

4

Operationally, one can implement Merton's (1974) model only if the unobserved asset value, V_t, and the model parameters, μ and σ, can be reasonably estimated.

2.2 The KMV estimation method

The equity values are observed at regular time points, and we denote the time series of $n + 1$ observations by $\{S_0, S_h, S_{2h}, \cdots, S_{nh}\}$ where h is the length of time, measured in years, between two observations; for example, a time series obtained by sampling on a daily basis over a two-year period will correspond roughly to $h = 1/250$ and $n = 500$.

The key to understanding the KMV method is that the unobserved asset value corresponding to a given observed equity value can be implied out from the equity pricing equation (5) if the asset volatility is known. This is true because the pricing equation forms a one-to-one relationship between the asset value and the equity price. We use function $g(\bullet; \sigma)$ to denote the equity pricing equation recognizing specifically the role played by the unknown asset volatility; that is, $S_t = g(V_t; \sigma)$. Since this function is invertible at any given asset volatility, we can express $V_t = g^{-1}(S_t; \sigma)$. The inversion can be easily performed numerically using, say, a bisection search algorithm.

The KMV method is a simple two-step iterative algorithm which begins with an arbitrary value of the asset volatility and repeats the two steps until achieving convergence. The two steps going from the m-th to $(m + 1)$-th iteration are :

Step 1: Compute the implied asset value time series $\{\hat{V}_0(\hat{\sigma}^{(m)}), \hat{V}_h(\hat{\sigma}^{(m)}), \cdots, \hat{V}_{nh}(\hat{\sigma}^{(m)})\}$ corresponding to the observed equity value data set $\{S_0, S_h, \cdots, S_{nh}\}$ where $\hat{V}_{ih}(\hat{\sigma}^{(m)}) = g^{-1}(S_{ih}; \hat{\sigma}^{(m)})$.

Step 2: Compute the implied asset returns $\{\hat{R}_1^{(m)}, \cdots, \hat{R}_n^{(m)}\}$ where $\hat{R}_i^{(m)} = \ln\left(\hat{V}_{ih}(\hat{\sigma}^{(m)})/\hat{V}_{(i-1)h}(\hat{\sigma}^{(m)})\right)$ and update the asset drift and volatility parameters as follows :

$$\bar{R}^{(m)} = \frac{1}{n}\sum_{k=1}^{n}\hat{R}_k^{(m)}$$

$$\left(\hat{\sigma}^{(m+1)}\right)^2 = \frac{1}{nh}\sum_{k=1}^{n}\left(\hat{R}_k^{(m)} - \bar{R}^{(m)}\right)^2$$

$$\hat{\mu}^{(m+1)} = \frac{1}{h}\bar{R}^{(m)} + \frac{1}{2}\left(\hat{\sigma}^{(m+1)}\right)^2$$

Note that the division by h is to state the parameter values on a per annum basis. In updating the variance, one could divide by $n - 1$ instead of n, which makes no difference to the asymptotic analysis.

Three points are important to note. The procedure outlined above is only the first part of the approach described in Crosbie and Bohn (2003) where they mentioned that the volatility estimate obtained with the above algorithm is then used in a Bayesian procedure to obtain the final estimate. Since we do not have access to the specifics of that Bayesian procedure, we cannot analyze its properties. Second, a model slightly more general than that of Merton (1974) is used by KMV. Finally, Crosbie and Bohn (2003) provided no description as to how the drift parameter estimate is computed. We follow Vassalou and Xing (2004) for they compute this parameter estimate using a sample average of the implied asset returns.

The top panel of Table 1 provides an example of the KMV estimation using a simulated sample of daily equity values using $V_0 = 1$, $\mu = 0.1$, $\sigma = 0.2$, $T = 2$, $F = 0.9$, $r = 0.05, n = 250$ and $h = 1/250$. For this example, the initial value for the volatility is set to $\hat{\sigma}^{(0)} = 0.2$ and the convergence criteria is 1×10^{-6}, that is, we stop the iteration when the difference between two successive estimation of σ is less than 1×10^{-6}. The parameter values corresponding to the final iteration are presented in this table along with the implied asset values for the last 10 data points. The KMV algorithm converges to $\hat{\mu} = -0.025$ and $\hat{\sigma} = 0.177$. The final implied asset value for the last data point, V_{250h}, is 0.9708. This implied asset value along with $\hat{\mu}$ and $\hat{\sigma}$ can be used to compute, for instance, a default probability estimate corresponding to the last data point, which turns out to be 0.420.

2.3 The transformed-data maximum likelihood estimation

The estimation problem associated with unobserved asset values can be naturally cast as a transformed-data maximum likelihood estimation (MLE) problem. Such an approach was first developed in Duan (1994). The obvious advantages are that (1) the resulting estimators are known to be statistically efficient in large samples, and (2) sampling distributions are readily available for computing confidence intervals or for testing hypotheses.

We denote the log-likelihood function of the observed data set under a specified model as $L(\theta; \text{data})$ where θ is the set of unknown parameters under the model. The MLE is to find the value of θ at which the data set has the highest likelihood of occurrence. Under Merton's (1974) model and assuming that one could directly observe the firm's asset values $\{V_0, V_h, \cdots, V_{nh}\}$, the

log-likelihood function could be written as :

$$L^V\left(\mu, \sigma; V_0, V_h, \cdots, V_{nh}\right) = -\frac{n}{2}\ln\left(2\pi\sigma^2 h\right) - \frac{1}{2}\sum_{k=1}^{n}\frac{\left(R_k - \left(\mu - \frac{\sigma^2}{2}\right)h\right)^2}{\sigma^2 h} - \sum_{k=1}^{n}\ln V_{kh} \quad (6)$$

where $R_k = \ln\left(V_{kh}/V_{(k-1)h}\right)$.[2] This log-likelihood is a result of assuming the geometric Brownian motion for the asset value dynamic. If the firm had undergone one or more times of refinancing, survivorship becomes an important issue. Duan, *et al* (2003) offers the survivorship adjustment for Merton's (1974) model.

We, however, do not observe the firm's asset values. Typically, the firm's equity values are available. Recognizing that Merton's (1974) model implicitly provides a one-to-one smooth relationship between the equity and asset values, one can invoke the standard result on differentiable transformations to derive the log-likelihood function solely based on the observed equity data. If we denote the density of the asset value as $f(V)$, the density associated with the equity will be given by $f(V)/\left|\frac{\partial g(V;\sigma)}{\partial V}\right|$. Applying this knowledge yields the following log-likelihood function on the observed equity data :

$$L^S\left(\mu, \sigma; S_0, S_h, \cdots, S_{nh}\right) = L^V\left(\mu, \sigma; \hat{V}_0(\sigma), \hat{V}_h(\sigma), \cdots, \hat{V}_{nh}(\sigma)\right) - \sum_{k=1}^{n}\ln\left(\Phi\left(\hat{d}_{kh}(\sigma)\right)\right) \quad (7)$$

where $\hat{V}_{kh}(\sigma) = g^{-1}\left(S_{kh};\sigma\right)$ and $\hat{d}_{kh}(\sigma) = \frac{\ln\left(\hat{V}_{kh}(\sigma)/F\right) + \left(r + \frac{\sigma^2}{2}\right)(T-kh)}{\sigma\sqrt{T-kh}}$. Note that the inversion does not require μ because the equity pricing equation in (5) is not a function of this parameter. Moreover, $\frac{\partial g(V_{kh};\sigma)}{\partial V_{kh}} = \Phi(d_{kh})$.

One can easily find the ML estimates for μ and σ by numerically maximizing the function in (7). Moreover, one can rely on the ML inference to compute approximate confidence intervals for the estimated parameters – $\hat{\mu}$ and $\hat{\sigma}$ – and other quantities of interest such as the implied asset value $\hat{V}_{nh}(\hat{\sigma})$, the risky bond price $\hat{D}_{nh}(\hat{V}_{nh}(\hat{\sigma}), \hat{\sigma})$ and the default probability $\hat{P}_{nh}(\hat{V}_{nh}(\hat{\sigma}), \hat{\mu}, \hat{\sigma})$. Appendix A provides further discussion on inference.

We take the same equity data sample as in the case of the KMV estimation to compute the ML estimates and report the final results in the bottom panel of Table 1. We have used a convergence criterion of 1×10^{-6} and the initial parameter values of 0.2 for σ and 0.1 for μ to obtain the ML

---

[2] The term $-\sum_{k=1}^{n}\ln V_{kh}$ could be ignored at the optimization stage if the firm's asset values were directly observable since it would simply amount to an irrelevant constant. We must keep this term in order to later deal with the fact that the asset value is not observable. This point was made in Duan (2000).

estimates: $\hat{\mu} = -0.025$ and $\hat{\sigma} = 0.175$ along with their standard errors. The ML estimates are then used to compute the implied asset values and their corresponding standard errors. The results for the final implied asset values and standard errors of the last ten data points are given in the table.

Comparing the two panels, it is clear that the KMV and ML estimates for μ and σ are very close. This result is not a fluke, and its theoretical reason will be given in the next subsection. The final implied asset values are also very close.[3] In contrast to the KMV approach, the MLE method can readily generate standard errors, which are also reported in the table.

2.4 Equivalence

We now proceed to show theoretically that the KMV estimates are the ML estimates for Merton's (1974) model. Our theoretical argument is based on a statistical tool known as the EM algorithm. This algorithm long existed in the statistical literature in different forms. Dempster *et al* (1977) were the first to coin the term and provided a comprehensive treatment of the method. The EM algorithm is essentially an alternative way of obtaining the ML estimate for the incomplete data model, where incomplete data refers to the situation that the model contains some random variable(s) without corresponding observations.

The EM algorithm involves two steps - expectation and maximization - and hence its name. One first writes down the log-likelihood function by assuming the complete data. The expectation of the complete-data log-likelihood function is then evaluated, conditional on the observed data and some assumed parameter values. This completes the expectation step. In the maximization step, one finds the new parameter values that maximize the conditionally expected complete-data log-likelihood function. The updated parameter values are then used to repeat the E and M steps until convergence is achieved. Interestingly, the EM algorithm will converge to the ML parameter estimates based on the incomplete data.[4] Although the EM algorithm is a very powerful tool, its practicality diminishes markedly when the conditional expectation cannot be analytically evaluated and/or the maximization step cannot be solved analytically.

We argue that the KMV method is an EM algorithm in the context of Merton's (1974) model.

---

[3] Although these two approaches can be shown to converge to the same values theoretically, numerical differences can be expected because numerical approximations used in defining the equity pricing function and in differentiations for the optimization purpose; for example, the standard polynomial approximation of the standard normal distribution function used in our program has an accuracy up to the sixth decimal place.

[4] The EM algorithm is not immune to the usual multi-modal problem associated with numerical optimization. In other words, the limit point of the EM algorithm may be a local maximum, just like the typical ML estimation.

A formal proof is given in Appendix B. We provide here a sketch of that proof. In the E-step, one first forms the log-likelihood function associated with the unobserved asset values, i.e., equation (6), which can be viewed as the complete-data log-likelihood using the jargon of the EM algorithm. We then compute its expected value conditional on the observed equity values and some values of μ and σ. Actually, we do not need the value of μ in the E-step because the implied asset value is not a function of μ. Since the relationship between the unobserved asset value and the observed equity value is one-to-one for a given value of σ, this conditional expectation becomes trivial. In fact, it amounts to the log-likelihood function based on the firm's asset values as if they were observed at those implied asset values. The parameter values that maximize this function are those given in Step 2 of the KMV method, which in turn means that the M-step will yield the same estimates for μ and σ as in a KMV update. In short, the KMV method is a degenerate EM algorithm with degeneracy occurring in the E-step.

3   General non-equivalence between the KMV and MLE methods

Merton's (1974) model contains only two unknown parameters and both of which pertain only to the dynamic of the firm's asset value. In general, structural models may contain unknown parameters specific to capital structure. There are numerous examples in the literature. In this section, we focus on the model of Brockman and Turtle (2003) and use it to make an important point on general non-equivalence between the KMV and MLE methods. The Brockman and Turtle (2003) model assumes the same asset value dynamic as that of Merton's (1974). However, unlike Merton's (1974) model, the firm can default at any time prior to the maturity of its debt. A default occurs when the asset value crosses a financial distress level, denoted by K, which can be interpreted as breaching, for example, the debt covenants. The firms' equity can be likened to a European down-and-out call option with the financial distress level serving as the barrier. Since the financial distress level is unknown, their model has three unknown parameters: μ, σ and K. The first two are specific to the asset value dynamic whereas K is related to the capital structure of the firm. The equity pricing formula of this model can be written generically as $S_t = g(V_t; \sigma, K)$, a function of the firm's asset value that depends on two unknown parameters. Its specific expression is given in Appendix C.

Brockman and Turtle (2003) conducted an empirical study using a market value proxy for the unobserved asset value of the firm. They concluded that the companies in their large sample of

industrial firms face high financial distress levels (relative to their book values of liabilities). Wong and Choi (2004), however, argued that the high financial distress level is induced by the use of the market value proxy for it is upward biased. They then adopted the transformed-data MLE approach of Duan (1994) to derive the likelihood function for the Brockman and Turtle (2003) model. They proceeded to empirically estimate the model for a large sample of firms and yielded estimates of the financial distress level that are much lower than those based on the market value proxy.

Our purpose here is to use the Brockman and Turtle (2003) model to illustrate the limitation of the KMV method. In short, we will show that the KMV method cannot estimate the capital structure parameter. In addition, the KMV method does not deal with the survivorship issue that is inherent in the estimation of a structural model like Brockman and Turtle (2003). In contrast, the transformed-data MLE approach can handle both.

It is fairly easy to see why the KMV method cannot generate a meaningful estimate for the financial distress level K. Starting with $\hat{\sigma}^{(0)}$ and $\hat{K}^{(0)}$, one can invert the time series of equity values to obtain their corresponding implied asset values; that is, $\hat{V}_{ih}(\hat{\sigma}^{(0)}, \hat{K}^{(0)}) = g^{-1}(S_{ih}; \hat{\sigma}^{(0)}, \hat{K}^{(0)})$ with g defined in Appendix C. Then, apply Step 2 of the KMV method to obtain the updated values: $\hat{\sigma}^{(1)}$ and $\hat{\mu}^{(1)}$. The implied asset values do not, however, provide a mechanism for updating K. Thus, $\hat{K}^{(1)} = \hat{K}^{(0)}$. In short, K will remain at its arbitrary initial value even though μ and σ have been updated.

Again we can use the EM algorithm to better understand the nature of the KMV method in this context. Similar to Merton's model (1974), one first derives the complete-data log-likelihood function, which is expected to be different from equation (6). This is simply due to the fact that if a firm has survived the sample period, it means that the firm's asset values must have stayed above the financial distress level for the entire period. Let $\mathcal{D}_i \equiv \left\{ \inf_{t \in [0,ih]} V_t > K \right\}$ be the event that no default occurred up to time ih. The complete-data log-likelihood function of the Brockman and Turtle (2003) model can be written as

$$
\begin{aligned}
& L_{BT}^V \left(\mu, \sigma, K; V_0, V_h, \cdots, V_{nh} | \mathcal{D}_n \right) \\
= & \ L^V \left(\mu, \sigma; V_0, V_h, \cdots, V_{nh} \right) + \ln P \left(\mathcal{D}_n \, | V_0, V_h, \cdots, V_{nh}; \sigma, K \right) - \ln P \left(\mathcal{D}_n; \mu, \sigma, K \right) \quad (8)
\end{aligned}
$$

The third term in the above expression can be analytically derived using the well-known probability of the Brownian motion infimum. The second term is more complicated and requires the knowledge

10

of the joint probability for the Brownian motion and its infimum. It turns out that the conditional survival probability is not a function of μ. The exact expressions for these two terms are given in Appendix C.

The expected value of the complete-data log-likelihood function conditional on the observed equity values and some parameter values can thus be computed. The result is the log-likelihood function with the firm's asset values being fixed at the implied asset values, say, in the m-th iteration. Thus, this function can be written as $L_{BT}^V\left(\mu, \sigma, K; \hat{V}_0(\hat{\sigma}^{(m)}, \hat{K}^{(m)}), \cdots, \hat{V}_{nh}(\hat{\sigma}^{(m)}, \hat{K}^{(m)})|\mathcal{D}_n\right)$. Obviously, maximizing this function will lead to updated estimates – $\hat{\mu}^{(m+1)}$, $\hat{\sigma}^{(m+1)}$ and $\hat{K}^{(m+1)}$ – that are different from those produced by the KMV updating step. It is now clear that the KMV method is not an EM algorithm for the Brockman and Turtle (2003) model. Thus, there is no reason to expect that the KMV method should yield the ML estimates.

The above discussion provides an EM algorithm for estimating the Brockman and Turtle (2003) model even though the result produced by the EM algorithm no longer coincides with that from the KMV method. The M-step is simply to find $\hat{\mu}^{(m+1)}$, $\hat{\sigma}^{(m+1)}$ and $\hat{K}^{(m+1)}$ that maximize $L_{BT}^V\left(\mu, \sigma, K; \hat{V}_0(\hat{\sigma}^{(m)}, \hat{K}^{(m)}), \cdots, \hat{V}_{nh}(\hat{\sigma}^{(m)}, \hat{K}^{(m)})|\mathcal{D}_n\right)$. Conceptually, this EM algorithm will converge to the maximum likelihood estimates. Operationally, however, it is not an attractive approach mainly because the M-step does not have an analytical solution. The numerical implementation of this EM algorithm thus becomes a two-layer loop where the M-step constitutes the inner loop.

We now use a simulation study to numerically examine the non-equivalence between the KMV and maximum likelihood estimator. Doing so requires the derivation of the log-likelihood function based on the observed equity values under the Brockman and Turtle (2003) model. Its derivation is provided in Appendix C. As discussed earlier, the MLE for this model has been developed in Wong and Choi (2004).[5]

We use the following parameter values to perform the simulation: $r = 0.05, \mu = 0.1, \sigma = 0.3, V_0 = 1.0, F = 1.0, K = 0.8, h = 1/250, n = 250$ and a maturity of $(2 - jh)$ years for the j-th data point of the time series. Since the Brockman and Turtle (2003) model is based on continuous monitoring of financial distress, the simulation of the firm's asset values is conducted by dividing one day into 50 subintervals. If a simulated sample breaches the barrier at any one of the $12,500$ monitoring points, the sample is discarded. We then compute the 251 corresponding daily equity values for any kept sample, and then act as if the firm's asset values were not available. In short,

---

[5]It is our belief that Wong and Choi's (2004) log-likelihood function contains minor errors.

we try to mimic a daily sample of observed equity values of a survived firm.

In order to implement the KMV method, one must commit to a specific value of K. The top panel of Table 2 reports the average values of the KMV estimates for μ and σ from $5,000$ simulated samples. In all cases, the initial σ is set to 0.3. When the true value of K is used, i.e., $K = 0.8$, the estimated σ appears to be unbiased. It is not the case for μ, however, and in fact it is substantially upward biased with an average of more than twice the true value. This result is consistent with the survivorship analysis presented in Duan, $et\ al$ (2003) for Merton's (1974) model. They show that failing to adjust for survivorship will result in an upward biased estimate for μ but not for σ. When an incorrect financial distress level is assumed, i.e., $K = 0.7$ or 0.9, the estimate either for μ or σ is of poor quality, a result that is hardly surprising.

The results based on the transformed-data MLE approach on the same $5,000$ simulated samples are reported in the second panel of Table 2.[6] The results show that the average parameter estimates for σ, K and V_{nh} are centered around the correct values. Furthermore, the coverage rates suggest that the actual distributions are reasonably approximated by the asymptotic normal distributions.[7] However, there exists a downward bias for the drift parameter μ for the sample size of 250 daily observations. It is well known in the literature that the drift parameter can only be estimated accurately using a long time span of data. In this transformed-data MLE setting with one-year worth of daily data, it is hardly surprising to experience the same difficulty. The bias is, however, of a much smaller magnitude than the upward bias observed with the KMV approach. To gain a better understanding of the bias, we have performed an additional Monte Carlo experiment with the results being reported in the third panel of Table 2. The parameters of this experiment are identical to those of the second panel except that h is set equal to $1/12$. [8] The bias in the estimated μ clearly becomes smaller as reflected in the fact that the mean is much closer to the true parameter value.

---

[6]As it is often the case in Monte Carlo experiments that involve non-linear optimization, convergence failure may occur in some fraction of cases. Whenever that happens, we add new replications in order to have $5,000$ simulated samples with proper convergence. The numbers of failed convergence are reported in the tables.

[7]A coverage rate represents the proportion of the $5,000$ estimates for which the true parameter value is contained in the α confidence interval implied by the corresponding asymptotic normal distribution.

[8]The simulation of the firm's asset values is conducted by dividing each month into $1,050$ subintervals.

4 Conclusion

The implementation of structural credit risk models requires estimates for asset values and model parameters. We have provided an analysis on the statistical properties of Moody's KMV method for Merton's (1974) model. We have shown theoretically that the KMV estimates are identical to the ML estimates when Merton's (1974) model is employed. This theoretical result relies on casting the KMV method as an EM algorithm. Even under Merton's (1974) model, the KMV method is not equivalent to the MLE method, because the former is limited to generating point estimates whereas the latter has accompanying distributional information ideal for conducting statistical inference. For structural models containing unknown capital structure parameter(s), for example, Brockman and Turtle (2003), we have shown that the KMV method does not generate suitable estimates and fails to address an inherent survivorship issue.

The KMV method is nevertheless an intuitive and numerically efficient way of obtaining point estimates under Merton's (1974) model. Our linking it to the EM algorithm actually opens up ways to modify the KMV method for different structural models. However, such modified schemes are likely to be numerically inefficient because the practicality of the EM algorithm largely rests on one's ability to obtain the analytical solutions in both the E and M steps. For structural models in general, the E step will remain to be a degenerate expectation operation but the M step is unlikely to have an analytical solution.

A   Maximum likelihood inference

Inference is an important part of statistical analysis. The ML inference for the structural credit risk model can be easily implemented using numerical differentiation. For a sufficiently large sample size n, the distribution of the maximum likelihood parameter estimates, stacked in a column vector $\hat{\boldsymbol{\theta}}$, can be well approximated by the following normal distribution :

$$\hat{\boldsymbol{\theta}} \approx N\left(\boldsymbol{\theta}_0, \mathbf{H}_n^{-1}\right)$$

where θ_0 denotes the true parameter value vector and H_n is the negative of the second derivative matrix of the log-likelihood function with respect to θ; that is,

$$\mathbf{H}_n = -\left.\frac{\partial L(\boldsymbol{\theta}; \text{data})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right|_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}}$$

For any differentiable transformation of the ML estimates, $f(\hat{\theta})$, the standard statistical result implies that

$$f(\hat{\theta}) \approx N \left(f(\theta_0), \left[\frac{\partial f(\hat{\theta})}{\partial \theta} \right]' H_n^{-1} \frac{\partial f(\hat{\theta})}{\partial \theta} \right)$$

where $\frac{\partial f(\hat{\theta})}{\partial \theta}$ is a column vector of the first partial derivatives of the function with respect to the elements of θ being evaluated at $\hat{\theta}$. This standard result was, for example, utilized in Lo (1986) in the context of testing the derivative pricing model.

In the context of Merton's (1974) model, $\theta = (\mu, \sigma)'$. For the implied asset value, one needs to recognize that it is a differentiable function of $\hat{\sigma}$. The specific analytical result for the distribution of the implied asset value is available in equation (4.7) of Duan (1994). The expressions for the risky bond and default probability can be derived in a way similar to the deposit insurance value in equation (4.8) of Duan (1994). The derivation requires one to recognize that both risky bond price and default probability are direct as well as indirect functions of the parameter(s) with the indirectness through the implied asset value.

B The KMV method is an EM algorithm

We now describe the EM algorithm for estimating Merton's (1974) model in the framework laid out in Dempster, *et al* (1977). By assuming that both the firm's asset and equity values could be observed, we can express the log-likelihood function of the complete data by

$$L \left(\mu, \sigma; V_0, V_h, \cdots, V_{nh}, S_0, S_h, \cdots, S_{nh} \right) = L^V \left(\mu, \sigma; V_0, V_h, \cdots, V_{nh} \right) + \sum_{i=0}^{n} \ln \mathbf{1}_{\{S_{ih}=g(V_{ih};\sigma_0)\}}$$

where $\mathbf{1}_{\{\bullet\}}$ is the indicator function. The indicator function appears because non-zero likelihood only occurs at the points where the equity and asset values satisfy the equity pricing equation evaluated at the unknown true volatility parameter σ_0. For the EM algorithm, we can in fact drop the sum involving the indicator function, a point that will become clear later in the E step.

E-step: Compute

$$\text{E}\left[L\left(\mu,\sigma;V_0,V_h,\cdots,V_{nh},S_0,S_h,\cdots,S_{nh}\right)\Big|S_0,S_h,\cdots,S_{nh},\hat{\mu}^{(m)},\hat{\sigma}^{(m)}\right]$$

$$= L^V\left(\mu,\sigma;\hat{V}_0(\hat{\sigma}^{(m)}),\hat{V}_h(\hat{\sigma}^{(m)}),\cdots,\hat{V}_{nh}(\hat{\sigma}^{(m)})\right) + \sum_{i=0}^{n}\ln \mathbf{1}_{\{S_{ih}=g(\hat{V}_{ih}(\hat{\sigma}^{(m)});\hat{\sigma}^{(m)})\}}$$

$$= L^V\left(\mu,\sigma;\hat{V}_0(\hat{\sigma}^{(m)}),\hat{V}_h(\hat{\sigma}^{(m)}),\cdots,\hat{V}_{nh}(\hat{\sigma}^{(m)})\right)$$

$$= -\frac{n}{2}\ln\left(2\pi\sigma^2 h\right) - \frac{1}{2}\sum_{k=1}^{n}\frac{\left(\hat{R}_k^{(m)} - \left(\mu - \frac{\sigma^2}{2}\right)h\right)^2}{\sigma^2 h} - \sum_{k=1}^{n}\ln\hat{V}_{kh}(\hat{\sigma}^{(m)})$$

where $\hat{R}_k^{(m)} = \ln\left(\hat{V}_{kh}(\hat{\sigma}^{(m)})/\hat{V}_{(k-1)h}(\hat{\sigma}^{(m)})\right)$ and $\hat{V}_{kh}(\hat{\sigma}^{(m)}) = g^{-1}(S_{kh};\hat{\sigma}^{(m)})$.

M-step: Maximize the conditional log-likelihood function obtained in the E-step to yield

$$\left(\hat{\sigma}^{(m+1)}\right)^2 = \frac{1}{nh}\sum_{k=1}^{n}\left(\hat{R}_k^{(m)} - \bar{R}^{(m)}\right)^2$$

$$\hat{\mu}^{(m+1)} = \frac{1}{h}\bar{R}^{(m)} + \frac{1}{2}\left(\hat{\sigma}^{(m+1)}\right)^2$$

where $\bar{R}^{(m)} = \frac{1}{n}\sum_{k=1}^{n}\hat{R}_k^{(m)}$.

It is now clear that the KMV method as described in the text is indeed an EM algorithm. It is easy to verify $L^V\left(\mu,\sigma;\hat{V}_0(\hat{\sigma}^{(m)}),\hat{V}_h(\hat{\sigma}^{(m)}),\cdots,\hat{V}_{nh}(\hat{\sigma}^{(m)})\right)$ satisfies the conditions in Theorem 4 of Dempster, *et al* (1977) so that the KMV method for Merton's model yield the ML estimates for μ and σ.

C   The Brockman and Turtle (2003) model and the corresponding likelihood function

In Brockman and Turtle (2003), the firm's equity is viewed as a European down-and-out call option with the strike price F, barrier K and maturity T. Thus, the equity pricing equation is:

$$
\begin{aligned}
S_t &= V_t\Phi(a_t) - Fe^{-r(T-t)}\Phi\left(a_t - \sigma\sqrt{T-t}\right) - V_t\left(K/V_t\right)^{2\eta}\Phi\left(b_t\right) \\
&\quad + Fe^{-r(T-t)}\left(K/V_t\right)^{2\eta-2}\Phi\left(b_t - \sigma\sqrt{T-t}\right)
\end{aligned}
$$

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

$$
a_t = \frac{\ln \frac{V_t}{\max(F,K)} + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}
$$

$$
b_t = a_t - \frac{2\ln\frac{V_t}{K}}{\sigma\sqrt{T-t}}
$$

$$
\eta = \frac{r}{\sigma^2} + \frac{1}{2}.
$$

Note that there is a rebate rate in Brockman and Turtle's (2003) original formula. In their implementation, however, the rebate rate was set to zero. The formula above reflects the special case of a zero rebate rate.

Denote the above equity pricing equation by $g(V_t; \sigma, K)$ to reflect the fact that it is a function of the firm's asset value and depends on two unknown parameters. We can apply the transformed-data MLE method of Duan (1994) to obtain the log-likelihood function of a discretely sampled equity values on a firm that survives the entire sample period. It turns out to be the log-likelihood function of the firm's asset value conditional on survival (equation (8)) evaluated at the implied asset values plus a term related to the Jacobian of the transformation. Thus,

$$
\begin{aligned}
& L_{BT}^S\left(\mu, \sigma, K; S_0, S_h, ..., S_{nh}|\mathcal{D}_n\right) \\
={}& L_{BT}^V\left(\mu, \sigma, K; \hat{V}_0(\sigma,K), \hat{V}_h(\sigma,K), \cdots, \hat{V}_{nh}(\sigma,K)|\mathcal{D}_n\right) - \sum_{j=1}^{n} \ln\left|\frac{\partial g(\hat{V}_{jh}(\sigma,K); \sigma, K)}{\partial V_{jh}}\right| \\
={}& L^V\left(\mu, \sigma; \hat{V}_0(\sigma,K), \hat{V}_h(\sigma,K), \cdots, \hat{V}_{nh}(\sigma,K)\right) \\
& + \ln P\left(\mathcal{D}_n|\hat{V}_0(\sigma,K), \hat{V}_h(\sigma,K), \cdots, \hat{V}_{nh}(\sigma,K); \sigma, K\right) \\
& - \ln P(\mathcal{D}_n; \mu, \sigma, K) - \sum_{j=1}^{n} \ln\left|\frac{\partial g(\hat{V}_{jh}(\sigma,K); \sigma, K)}{\partial V_{jh}}\right|.
\end{aligned}
$$

where the first derivative of the equity value with respect to the asset value can be derived to be

$$
\begin{aligned}
\frac{\partial g(V_t; \sigma, K)}{\partial V_t} ={}& \Phi(a_t) + \frac{1}{\sigma\sqrt{T-t}}\left[\phi(a_t) - \frac{Fe^{-r(T-t)}}{V_t}\phi\left(a_t - \sigma\sqrt{T-t}\right)\right] \\
& + \left(\frac{K}{V_t}\right)^{2\eta-1}\left[(2\eta-1)\frac{K}{V_t}\Phi(b_t) + \frac{Fe^{-r(T-t)}}{K}(2-2\eta)\Phi\left(b_t - \sigma\sqrt{T-t}\right)\right] \\
& + \frac{1}{\sigma\sqrt{T-t}}\left(\frac{K}{V_t}\right)^{2\eta-1}\left[\frac{K}{V_t}\phi(b_t) - \frac{Fe^{-r(T-t)}}{K}\phi\left(b_t - \sigma\sqrt{T-t}\right)\right]
\end{aligned}
$$

with $\phi(\cdot)$ representing the standard normal density function.

It remains to derive the expressions for $P(\mathcal{D}_n; \mu, \sigma, K)$ and $P(\mathcal{D}_n | V_0, V_h, \cdots, V_{nh}; \sigma, K)$. First recall the following result for the infimum of a Brownian motion with drift:

$$P\left\{ \inf_{0 \leq s \leq t} W_s + as > y \right\} = \begin{cases} \Phi\left(\frac{at-y}{\sqrt{t}}\right) - \exp(2ay)\Phi\left(\frac{at+y}{\sqrt{t}}\right) & \text{if } y \leq 0 \\ 0 & \text{otherwise} \end{cases}$$

Thus,

$$P\left(\mathcal{D}_n; \mu, \sigma, K\right) = \Phi\left(\frac{(\mu - \frac{\sigma^2}{2})nh - \ln\frac{K}{V_0}}{\sqrt{nh}\sigma}\right) - \exp\left(\frac{2(\mu - \frac{\sigma^2}{2})}{\sigma^2}\ln\frac{K}{V_0}\right)\Phi\left(\frac{(\mu - \frac{\sigma^2}{2})nh + \ln\frac{K}{V_0}}{\sqrt{nh}\sigma}\right).$$

The joint distribution of the Brownian motion with drift and its infimum is known to be

$$P\left\{ \inf_{0 \leq s \leq t} W_s + as \leq y \ \& \ W_t + at \geq x \right\} = \exp(2ay)\Phi\left(\frac{2y - x + at}{\sqrt{t}}\right) \quad \text{for } y \leq \min(x, 0).$$

As a result, the distribution of the infimum conditional on two end points becomes

$$P\left\{ \inf_{0 \leq s \leq t} W_s + as \leq y \ \middle| \ W_0 = 0, W_t = b \right\} = \frac{\exp(2ay)\phi\left(\frac{2y - b}{\sqrt{t}}\right)}{\phi\left(\frac{b}{\sqrt{t}}\right)}$$

for $y \leq \min(b + at, 0)$. This result can be used to characterize the distribution of the infimum up to time ih while conditioning on $V_{(i-1)h}$, V_{ih} and the infimum up to time $(i-1)h$. Hence,

$$\begin{aligned} H_i(v) &\equiv P\left\{ \inf_{0 \leq s \leq ih} V_s \leq v \ \middle| \ V_{(i-1)h}, V_{ih}, \inf_{0 \leq s \leq (i-1)h} V_s \right\} \\ &= 1_{\{\inf_{0 \leq s \leq (i-1)h} V_s \leq v\}} + 1_{\{\inf_{0 \leq s \leq (i-1)h} V_s > v\}} P\left\{ \inf_{(i-1)h \leq s \leq ih} V_s \leq v | V_{(i-1)h}, V_{ih} \right\} \end{aligned}$$

Let $W_{s^*}^* \equiv W_s - W_{(i-1)h}$ where $s^* \equiv s - (i-1)h$ for $s \geq (i-1)h$. For $v < \min(V_{(i-1)h}, V_{ih})$, we can compute

$$\begin{aligned} &P\left\{ \inf_{(i-1)h \leq s \leq ih} V_s \leq v \ \middle| \ V_{(i-1)h}, V_{ih} \right\} \\ &= P\left\{ \inf_{0 \leq s^* \leq h} W_{s^*}^* + \frac{\mu - \sigma^2/2}{\sigma}s^* \leq \frac{1}{\sigma}\ln\frac{v}{V_{(i-1)h}} \ \middle| \ W_0^* = 0, W_h^* = \frac{1}{\sigma}\ln\frac{V_{ih}}{V_{(i-1)h}} - \frac{(\mu - \sigma^2/2)h}{\sigma} \right\} \\ &= \frac{\exp\left(\frac{2(\mu - \sigma^2/2)}{\sigma^2}\ln\frac{v}{V_{(i-1)h}}\right)\phi\left(\frac{\ln(v/V_{(i-1)h}) + \ln(v/V_{ih}) + (\mu - \sigma^2/2)h}{\sigma\sqrt{h}}\right)}{\phi\left(\frac{\ln(V_{ih}/V_{(i-1)h}) - (\mu - \sigma^2/2)h}{\sigma\sqrt{h}}\right)} \\ &= \exp\left(\frac{2\ln\frac{V_{ih}}{v}\ln\frac{v}{V_{(i-1)h}}}{\sigma^2 h}\right). \end{aligned}$$

17

For either $v \geq V_{(i-1)h}$ or $v \geq V_{ih}$, it is obvious that $P\left\{\inf_{(i-1)h \leq s \leq ih} V_s \leq v \middle| V_{(i-1)h}, V_{ih}\right\} = 1$. Thus, for $v < \min(V_{(i-1)h}, V_{ih}, \inf_{0 \leq s \leq (i-1)h} V_s)$,

$$H_i(v) = \exp\left(\frac{2 \ln \frac{V_{ih}}{v} \ln \frac{v}{V_{(i-1)h}}}{\sigma^2 h}\right).$$

Otherwise, $H_i(v) = 1$.

We are now in a position to compute $P\left(\mathcal{D}_n \middle| V_0, V_h, \cdots, V_{nh}; \sigma, K\right)$. Note that $K < \inf_{0 \leq s \leq ih} V_s$ is in the set $\mathcal{D}_i$. As a result, there is no need to know the value of $\inf_{0 \leq s \leq ih} V_s$ in computing $H_{i+1}(K)$ as long as it is restricted to $\mathcal{D}_i$.

$$
\begin{aligned}
P\left(\mathcal{D}_1 \middle| V_0, V_h; \sigma, K\right) &= 1 - H_1(K) \\
P\left(\mathcal{D}_2 \middle| V_0, V_h, V_{2h}; \sigma, K\right) &= P\left(\mathcal{D}_1 \middle| V_0, V_h; \sigma, K\right) - P\left(\mathcal{D}_1 \middle| V_0, V_h; \sigma, K\right) H_2(K) \\
&= [1 - H_1(K)][1 - H_2(K)] \\
P\left(\mathcal{D}_3 \middle| V_0, V_h, \cdots, V_{3h}; \sigma, K\right) &= P\left(\mathcal{D}_2 \middle| V_0, V_h, V_{2h}; \sigma, K\right) - P\left(\mathcal{D}_2 \middle| V_0, V_h, V_{2h}; \sigma, K\right) H_3(K) \\
&= [1 - H_1(K)][1 - H_2(K)][1 - H_3(K)].
\end{aligned}
$$

In general, we have

$$
\begin{aligned}
P\left(\mathcal{D}_n \middle| V_0, V_h, \cdots, V_{nh}; \sigma, K\right) &= \prod_{i=1}^{n}[1 - H_i(K)] \\
&= \prod_{i=1}^{n}\left[1 - \exp\left(\frac{2 \ln \frac{V_{ih}}{K} \ln \frac{K}{V_{(i-1)h}}}{\sigma^2 h}\right)\right].
\end{aligned}
$$

Note that the above conditional survival probability turns out not to be a function of μ.

References

[1] Brockman, P. and H. Turtle, 2003, A Barrier Option Framework for Corporate Security Valuation, *Journal of Financial Economics* 67, 511-529.

[2] Bharath, T. and T. Shumway, 2004, Forecasting Default with the KMV-Merton Model, University of Michigan working paper.

[3] Cao, C., Simin, T. and J. Zhao, 2005, Do Growth Options Explain the Trend in Idiosyncratic Risk?, Pensylvania State working paper.

[4] Crosbie, P. and J. Bohn, 2003, Modeling Default Risk, Moody's KMV technical document.

[5] Dempster, A., Laird, N., and D. Rubin, 1977, Maximum-likelihood from Incomplete Data Via the EM Algorithm, *Journal of the Royal Statistical Society B* 39, 1-38.

[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., G. Gauthier, J.G. Simonato and S. Zaanoun, 2003, Estimating Merton's Model by Maximum Likelihood with Survivorship Consideration, University of Toronto working paper.

[9] Eom, Y., J. Helwege and J. Huang, 2004, Structural Models of Corporate Bond Pricing: An Empirical Analysis, *Review of Financial Studies* 17, 499-544.

[10] Ericsson, J. and J. Reneby, 2003, Estimating Structural Bond Pricing Models, to appear in *Journal of Business*.

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

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

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

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

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

[16] Vassalou, M. and Y. Xing, 2004, Default Risk in Equity Returns, *Journal of Finance* 59, 831-68.

[17] Wong, H. and T. Choi, 2004, The Impact of Default Barrier on the Market Value of Firm's Asset, Chinese University of Hong Kong working paper.

Table 1: Comparisons of the KMV and maximum
likelihood estimates
for the Merton (1974) model

KMV

i	S_{ih}	$T - ih$	$\hat{V}_{ih}^{(KMV)}$	s.e.$(\hat{V}_{ih}^{(KMV)})$
241	0.1377	1.036	0.9689	N.A.
242	0.1377	1.032	0.9691	N.A.
243	0.1352	1.028	0.9662	N.A.
244	0.1469	1.024	0.9814	N.A.
245	0.1652	1.020	1.0039	N.A.
246	0.1600	1.016	0.9979	N.A.
247	0.1610	1.012	0.9994	N.A.
248	0.1531	1.008	0.9900	N.A.
249	0.1598	1.004	0.9984	N.A.
250	0.1372	1.000	0.9708	N.A.

$$\hat{\mu}^{(KMV)} = -0.025 \qquad \text{s.e.}(\hat{\mu}^{(KMV)}) = \text{N.A.}$$
$$\hat{\sigma}^{(KMV)} = 0.177 \qquad \text{s.e.}(\hat{\sigma}^{(KMV)}) = \text{N.A.}$$

Maximum Likelihood

i	S_{ih}	$T - ih$	$\hat{V}_{ih}^{(MLE)}$	s.e.$(\hat{V}_{ih}^{(MLE)})$
241	0.1377	1.036	0.9695	0.0051
242	0.1377	1.032	0.9697	0.0051
243	0.1352	1.028	0.9668	0.0052
244	0.1469	1.024	0.9819	0.0047
245	0.1652	1.020	1.0043	0.0041
246	0.1600	1.016	0.9983	0.0043
247	0.1610	1.012	0.9999	0.0042
248	0.1531	1.008	0.9905	0.0044
249	0.1598	1.004	0.9989	0.0042
250	0.1372	1.000	0.9713	0.0050

$$\hat{\mu}^{(MLE)} = -0.025 \qquad \text{s.e.}(\hat{\mu}^{(MLE)}) = 0.174$$
$$\hat{\sigma}^{(MLE)} = 0.175 \qquad \text{s.e.}(\hat{\sigma}^{(MLE)}) = 0.014$$

The last ten data points of a sample of 250 daily equity prices. S_{ih} is the i-th equity value corresponding to a debt maturity of $T - ih$. The true values for μ and σ are respectively 0.1 and 0.2. The face value of debt, the risk-free rate and the time step are $F = 0.9$, $r = 0.05$ and $h = 1/250$, respectively. The implied asset values are computed with a Gauss-Newton algorithm.

Table 2: A Monte Carlo study of the
KMV and MLE methods for the
Brockman and Turtle (2003) model

KMV with $h = 1/250$

	$\hat{\mu}$	$\hat{\sigma}$	$\hat{K}$	$\hat{V}_{nh} - V_{nh}$
True	0.100	0.300	0.800	0.000
Mean ($K = 0.700$)	0.308	0.342	N.A.	-0.024
Mean ($K = 0.800$)	0.260	0.299	N.A.	0.000
Mean ($K = 0.900$)	0.213	0.254	N.A.	0.026

MLE with $h = 1/250$

	$\hat{\mu}$	$\hat{\sigma}$	$\hat{K}$	$\hat{V}_{nh} - V_{nh}$
True	0.100	0.300	0.800	0.000
Mean	0.044	0.296	0.806	0.001
Median	0.101	0.294	0.810	0.001
Std	0.422	0.043	0.085	0.027
25 % cvr	0.207	0.243	0.247	0.241
50 % cvr	0.485	0.497	0.502	0.493
75 % cvr	0.831	0.736	0.751	0.735
95 % cvr	0.957	0.934	0.942	0.923

MLE with $h = 1/12$

	$\hat{\mu}$	$\hat{\sigma}$	$\hat{K}$	$\hat{V}_{nh} - V_{nh}$
True	0.100	0.300	0.800	0.000
Mean	0.086	0.298	0.816	0.001
Median	0.098	0.298	0.807	0.000
Std	0.094	0.018	0.137	0.022
25 % cvr	0.258	0.252	0.252	0.259
50 % cvr	0.521	0.493	0.498	0.511
75 % cvr	0.787	0.749	0.751	0.786
95 % cvr	0.965	0.945	0.947	0.962

True is the parameter value used in the Monte Carlo simulation. Mean, Median and Std are the sample statistics of the estimates from the 5,000 simulations. cvr is the coverage rate defined as the proportion of the 5,000 parameter estimates for which the true parameter value is contained in the α confidence interval implied by the asymptotic distribution. Values used in simulation: $V_0 = 1.0$, $F = 1.0$, $K = 0.8$, $r = 0.05$, $n = 250$, and a maturity of $2 - ih$ years for the i-th data point of the time series of equity values. For the case MLE with $h = 1/250$, 603 failure to converge where observed among 5,603 simulated paths. For the case MLE with $h = 1/12$, 213 failure to converge where observed among the 5,213 simulated paths. For each MLE case, a convergence criterion based on the maximum absolute difference in both parameter and functional values between two successive iterations was used with a threshold of 1×10^{-6}. A threshold of

1×10^{-4} for the gradient norm was also used to determine the validity of a solution obtained with the hill climbing algorithm.