Estimating Structural Bond Pricing Models via Simulated Maximum Likelihood

Max Bruche
Financial Markets Group
London School of Economics
Houghton Street
London WC1A 2AE, U. K.
m.e.bruche@lse.ac.uk

This version: April 1, 2005

Abstract

This paper describes how structural bond pricing models can be estimated using a Simulated Maximum Likelihood procedure developed by Durbin and Koopman (1997). The approach has the advantage that price data on any traded claim (such as bonds, equity, and credit default swaps), as well as information about the balance sheet (e.g. accounting data) can be used in the estimation, improving efficiency. Monte Carlo evidence as well as a small application to real data indicates that this approach is superior to both traditional estimation methods and recently proposed versions of Maximum Likelihood Estimation (Ericsson, Reneby 2002).

1 Introduction

In 1974, Merton wrote a seminal paper (Merton 1974) that explained how the then recently presented Black-Scholes model could be applied to the pricing of corporate debt. Many extensions of this model followed. This family of models is sometimes referred to as the family of structural models of corporate bond prices, and views prices of corporate debt and equity as portfolios of options on the fundamental value or asset value of the firm. Extensions of the original model relate to e.g. sub-ordination arrangements, indenture provisions and default before maturity (Black and Cox 1976), coupon bonds (Geske 1977), stochastic interest rates (Shimko, Tejima, and van Deventer 1993, Longstaff and Schwartz 1995) or an optimally chosen capital structure (e.g. Leland 1994), to name but a few.

Structural models have found applications in risk management, (e.g. the KMV EDF^{TM} methodology Crosbie and Bohn 2002), in central banks (Gropp, Vesala, and Vulpes 2002), or in pricing (e. g. the CreditGrades^{TM} model, Finger et al. 2002).

For the pricing of credit derivatives, however, structural models do not seem to be the preferred choice, due to their inability to precisely price instruments. There is evidence, for instance, that the simplest structural model (the original (Merton 1974) model) seems to require implausibly high volatilities to generate reasonable bond prices (cf. e.g. Jones, Mason, and Rosenfeld 1984, Anderson and Sundaresan 2000, Eom, Helwege, and Huang 2001).

So called reduced-form models, pioneered by Jarrow and Turnbull (1995) that simply posit a hazard rate process that describes the instantaneous probability of default seem to match observed bond spreads quite well when implemented (Duffee 1999), possibly because factors like liquidity risk or tax rates are subsumed into the hazard rate estimates. Consequently, they are often the model of choice when practitioners hedge one credit derivative in terms of other credit derivatives. Typically, the derivative chosen to hedge other derivative would be a default swap.

However, in many application where the focus is on both credit-related instruments and the prices of equity, structural models seem a natural and intuitive choice, as they clearly relate both of these via the balance sheet. An example of such an application would be a hedge fund selling a credit default swap, and trying to hedge itself by selling equity. In order to do that, the hedge fund would need to know the "Greeks" or the derivatives of the price of the credit default swap with respect to equity, which a structural model produces quite naturally. Since with very few exceptions (Mamaysky 2002), reduced-form models do not produce an explicit link between the price of equity and the price of the bond, they are not suitable for the above application. Schaefer and Strebulaev (2003) indicate that for this purpose, even the simple Merton model produces hedge parameters that seem to be correct on average (it is not clear whether they claim that in their implementation the Merton model could actually be used for hedging, though).

While part of the poor empirical performance of structural models might be attributable to missing theoretical features of the models or abstracting from non-default risk factors, such as liquidity risk, this paper argues that a large part of the problem is also attributable to poor estimation methodologies. It proposes a likelihood approach based on non-linear filtering.

The approach has the advantage that it can use price data on any traded claim (such as bonds, equity, and credit default swaps), as well as information about the balance sheet (e.g. accounting data), substantially improving efficiency. Also, contrary to alternative approaches, it is not necessary to assume that any asset is priced 100% exactly by the structural model, which is important since it is likely that microstructure effects, liquidity issues, taxes, agency problems etc. (which are not included in typical structural models) add "noise" to asset prices that is unrelated to fundamental asset values. As is shown below, even small amounts of "noise" can have serious consequences for estimation results when they are ignored.

Monte Carlo evidence and a small application to real data indicate that this approach is superior to both traditional estimation methods and recently proposed versions of Maximum Likelihood Estimation (Ericsson and Reneby 2002a, Duan, Gauthier, Simonato, and Zaanoun 2003, Ericsson and Reneby 2005). Hopefully, it is a step into the direction of making structural models useful in contexts such as the one described above.

The rest of this paper is organised as follows: First, some of the fundamental problems of estimating structural models are discussed, and some common estimation approaches examined. A very general description of the estimation problem is then presented, followed by an exposition of how to evaluate the likelihood function in this context. Results of a Monte Carlo experiment are discussed. Finally, the method is applied to real data.

1.1 The inherent problem of estimation

Only in very few special cases is the estimation of structural models straightforward. Gemmill (2002), for instance, picks data on British closed-end funds that issue zero coupon bonds. For this data set, asset values of the funds are readily available (indeed, they are published daily), and the entire debt of the entity consists of one zero coupon bond. This situation exactly matches the assumptions of the Merton model, and direct computation of theoretically predicted bond prices is very simple and raises few problems.

For normal companies, however, asset values are observed very infrequently. Since the asset value of a firm is not typically traded, market prices cannot be observed. Balance sheet information on asset values exists, but it is available at most at a quarterly frequency, and often only at an annual frequency. This would be a problem in the hedging situation described above, as a quarterly or annually rebalanced hedge presumably would be quite useless. For practical purposes, the asset value of a firm is latent or unobserved.

The key distinguishing feature of different estimation approaches is how this problem is dealt with.

1.2 Other issues for estimation

There are a number of other issues which need to be taken into account in estimation.

1.2.1 Data availability

Corporate bonds are typically not listed on exchanges. The parties that collect information on bonds are either market-makers (by market convention, these are often the investment banks that issue bonds on behalf of clients), or organisations that buy a lot of bonds. In either of these cases, since corporate bonds are typically not very liquid, transaction level data is likely to be at a less than daily frequency on average. For the more liquid bonds, there might be several transactions a day, but for most bonds, there might be several transactions a week, or a couple of transactions per month. This means that any time series of bond prices at a higher than monthly frequency is very likely to be irregularly spaced.

In addition, the type of bonds that are traded often have embedded call options, sinking fund clauses and/or double up options. They are much more complicated than the type of bond that Merton originally examined.

1.2.2 Data quality

Prices of claims on asset values of the firm contain information not just about asset values. It is often argued that bond prices contain premia for non-default risk factors, but equity prices as well are probably influenced by market microstructure, the liquidity (Morris and Shin 2004) of market participants and/or the markets, stop-loss limits, tax treatment, agency problems etc. Strictly speaking, prices of claims on the firm (including equity) should be treated as noisy signals about the asset value at best.

1.2.3 Capital structure

In real life, a firm's capital structure is typically much more complicated than that assumed in e.g. the original Merton model. Typically, there will be many layers of debt, of different seniority, with publicly traded bonds as just one of these layers. How to interpret a model that abstracts from complicated capital structures in this context can make a big difference to estimation results.

1.3 Estimation approaches

1.3.1 The direct approach

The first attempt at implementing structural models on corporate bonds was conducted by Jones, Mason, and Rosenfeld (1984). Aside from some interesting ways of dealing with the embedded optionality in bonds that was common at the time, they suggested the following method: First, estimate the asset value (V) as the sum of the value of equity (E), the observed value of traded debt and the estimated value of non-traded debt (assuming that the book to market ratio of traded and non-traded debt is the same). The volatility of the asset value is then calculated directly from the returns of the estimated asset value. They also proposed refining this by using the following relationship (derived from the equity pricing equation F_E using Itôs lemma):

$$\sigma_E = \sigma_V \frac{V}{E} \frac{\partial F_E}{\partial V} \tag{1}$$

An equity volatility is estimated from historical equity returns, and a second estimate of the asset volatility is obtained by plugging this and the first-pass estimate of the asset value into this equation.

The essential feature is that the asset value is estimated by a back-of-the-envelope calculation based on book values and some observed market values of components of the total liabilities. Note that this has no statistical basis, and that it does not involve the assumptions of the model. Although possibly a reasonable educated guess, there is no reason to expect that this method will yield particularly reliable estimates of asset values, asset value volatilities, or to predict bond prices well.

More recently, variants of this technique have been employed by e.g. Lyden and Saraniti (2000) who remain relatively close to the original version, or for example Anderson and Sundaresan (2000) who combine stock and flow accounting data to arrive at a leverage proxy, or most recently Eom, Helwege, and Huang (2004), who simply add the book value of debt to the observed market value of equity to arrive at an estimate of the asset value.

1.3.2 The yield curve approach

Wei and Guo (1997) choose an approach similar to that often used when 'calibrating' models of the risk-free yield curve. For data on the spread between the term structures of Eurodollar and US Treasury debt, for each time period for which they have observations, they choose parameters (where they view the asset value as a parameter as well) to minimise the squared fitting errors. This allows them to back out implied asset values (as well as finding estimates of model parameters). Unfortunately, they only have 5 data points on the spread curve for each date, so they necessarily need models with at most that number of parameters.

While this approach might appeal to fixed-income practitioners, it also ignores the information in the time series of data (how spreads change across time), and focuses solely on the (small) cross-sectional element. It is furthermore impractical for the type of application considered above, as it is practically impossible to obtain a term structure of yields for individual firms, since the number of actively traded corporate bonds per firm is typically very small.

1.3.3 The calibration approach

The most common approach to implementing structural models to date, sometimes termed 'calibration' has been to solve a set of two equations relating the observed price of equity and estimated (i.e. usually historical) equity volatility to asset value and asset value volatility (Ronn and Verma 1986). The equations used for this are the option-pricing equation describing the value of equity as an option on the underlying asset value (F_E), and the equation describing the relationship between equity volatility and asset value volatility derived from the equity pricing equation via Itô's lemma.

$$E = F_E(V, \sigma_V) \tag{2}$$

$$\sigma_E = \sigma_V \frac{V}{E} \frac{\partial F_E}{\partial V}. \tag{3}$$

Once the equity-implied asset value and asset value volatility have been obtained, calculating a theoretical price for a bond is comparatively straightforward, since most other parameters, such as the bond cash flows and the risk-free term structure are observed directly. The theoretical prices, yields or spreads can then be compared to the actual values, to give information on the accuracy of various different models.

Note that when the volatility of equity returns is calculated from historical data, this is typically done assuming that the volatility is constant. Of course, this contradicts equation 3. Since the equity volatility changes as the ratio of the value of equity to the value of assets changes, and as the derivative of the equity pricing function changes, this problem will be especially apparent when the asset value (or the leverage) of the firm changes a lot over the estimation period.

This approach is common in the commercial world (see e.g. the KMV methodology described by Crosbie and Bohn 2002) and is used in academia (e.g. Delianedis and Geske 1999, Delianedis and Geske 2001). It is often the only approach described in major textbooks (Hull 2003). There are variants, such as the approach used by Huang and Huang (2003), who use four target variables (observed default probabilities, leverage ratios, recoveries given default and equity premiums) to match four parameters (asset value, a market price of risk parameter, the asset value volatility and a recovery rate).

1.3.4 The Ericsson and Reneby approach

Ericsson and Reneby (2005) (cf. also Ericsson and Reneby 2002a) demonstrate the biases that result from the calibration technique (which they call the "volatility restriction" method) if leverage is not constant. As an alternative, they propose a Maximum Likelihood method based on a method first proposed by Duan (1994) (This method is also utilised by Duan, Gauthier, Simonato, and Zaanoun (2003)): Typically, structural models start with postulating that the asset value follows a geometric Brownian motion. This of course implies that the changes in the log of the asset value are Gaussian. The density of the log asset value hence takes a simple form. They suggest that to relate this density to something observable, one can simply change variable to the equity price. If the Jacobian of the function relating equity price to the asset value is known, the log-likelihood function of equity prices can be derived, and subsequently maximised using standard techniques. With a Monte Carlo study, Ericsson and Reneby show that using this technique on equity prices is superior to the calibration method described above. Essentially, the calibration approach is unable to disentangle the effects of volatility and leverage, and tends to confuse the two.

There are two related drawbacks to the Ericsson and Reneby Maximum Likelihood Estimation (ERMLE).

1. Ericsson and Reneby implement it using only the price of equity. Ideally, if we have information on one or more bond prices, credit and equity derivatives, accounting information or any other information on the underlying (the fundamental value or asset value of the firm), we would like to include this in the estimation.

2. It is not clear why if structural models price bonds with an error, we should be willing to assume that they would price equity without error. Given that asset prices including equity prices are influenced by market microstructure, agency problems, taxes etc. which are outside the model, it is not reasonable to make this assumption. This paper will demonstrate that making the assumption of zero equity pricing error (or observation error) can induce serious bias in estimates.

Including more information such as e.g. a bond price in the estimation will make it necessary to find a compromise between bond-implied asset value and equity-implied asset value, or in general between the different asset values implied by the different observed variables. The problem of estimation is intimately related to the problem of recovering the value of the latent or unobserved asset value from the observed variables, it is one of calculating posterior densities, or filtering. Since the option-pricing equations are highly non-linear, the problem is one of non-linear filtering.

2 A general setup for estimation

Suppose we have data on e.g. the prices of debt and equity, which our model suggests are functions of a latent unobserved state (e.g. the asset value of the firm in the structural model context). An econometric model then consists of a transition equation for the latent state (an equation describing how e. g. the asset value changes), and some observation equations that describe the functions that map asset values into the observed prices of debt and equity respectively.

2.1 The state equation

Structural models of corporate bond and equity prices typically specify that the process describing the evolution of the value of assets V of the firm (which determines the value of equity and debt) follows a geometric Brownian motion:

$$dV = \mu V dt + \sigma V dW. \tag{4}$$

It is obvious that $\log V$ is Gaussian:

$$\int_t^T d\log V = (\mu - \frac{1}{2}\sigma^2)\int_t^T ds + \sigma\int_t^T dW \sim N\left\{\mu\left(T-t\right), \sigma^2\left(T-t\right)\right\}. \tag{5}$$

Hence we can define $\alpha_t = \int_0^t d\log V$, and write the discrete time form as

$$\alpha_t = d + \alpha_{t-1} + \eta_t, \quad \eta_t \sim NID\left\{0, \sigma_\eta\right\}, \tag{6}$$

where $d = (\mu - \frac{1}{2}\sigma^2)\int_{t-1}^t ds$ and $\eta_t = \sigma\int_{t-1}^t dW$.

Substantially more complicated setups than the one described here (i. e. in particular non-Gaussian setups and multivariate setups where α is a vector) are also feasible.

2.2 The observation equation

A given structural model (or reduced form model) produces some pricing functions F_i taking parameters ψ that map a factor α into a vector of market prices ξ_i. If there are factors outside the structural model, this needs to be modelled econometrically as an observation error. Since prices could be higher or lower than predicted by the model, but they cannot be negative (due to limited liability), and since one would expect the size of the errors to be proportional to the price, a multiplicative log-normal observation error is a simple and natural modelling choice.

$$\xi_{it} = F_{it}\left(\alpha_t; \psi_i\right) e^{\varepsilon_t}, \quad \varepsilon_t \sim NID\left\{0, \Sigma_\varepsilon\right\}. \tag{7}$$

Prices conditional on the unobserved quantities would then be log-normal. Of course, it would be possible to put different and more elaborate structure on the observation errors.

2.3 The system

Now define y as the vector with typical element $\log \xi_{it}$ and $Z_t\left(\alpha_t; \psi\right)$ as the vector of the log of the pricing functions $F_{it}\left(\alpha_t; \psi_i\right)$. We can write the system as

$$\alpha_t = \quad d + \alpha_{t-1} + \eta_t, \quad \eta_t \sim NID\left\{0, \sigma_\eta\right\}, \tag{8}$$

$$y_t = \quad Z_t\left(\alpha_t; \psi\right) + \varepsilon_t, \quad \varepsilon_t \sim NID\left\{0, \Sigma_\varepsilon\right\} \tag{9}$$

where e.g. η_t is independent of ε_t. Specification issues can arise (depending on the pricing functions Z and the particular error structure) and have to be checked on a case-by-case basis. Possibly, some restrictions will have to be imposed.

If we were interested in modelling several factors or asset values, it is easy to interpret α_t as a vector, with σ_η as the variance-covariance matrix determining the correlations between the asset values. Also, the vector y_t could be extended to include information other than prices, such as accounting information, or maybe more pertinently, the prices of other assets such as credit default swaps.

2.4 Restrictions determine feasible estimation procedures

Denote the variance of the observation error for asset or observed variable i as σ_i (assume there are n of these), and use σ_E to denote the variance of the observation error of equity. Let E denote the market value of equity and let A_i denote the observed value of asset i (with n_i observations). In the following we assume that equity information is in some form special[1] - we could easily designate another 'special' observed variable. We can then distinguish four cases in which a likelihood function can be evaluated.

Different assumptions made about the variances of the observation errors in this system will make different estimation procedures feasible. The Ericsson-Reneby/Duan Maximum Likelihood Estimation (ERMLE) method, for example, can be viewed as a method to estimate this system given that $\sigma_E = 0$, and $\sigma_i = \infty, \quad \forall i \neq E$:

[1] It is certainly the most easy to obtain.

2.4.1 Case 1: Restrict $\sigma_E = 0$, restrict $\sigma_i = \infty, \forall i \neq E$

In this case, no weight would be attached to observed variables other than the equity price. Since the equity observation error is assumed to be zero, there is a direct correspondence between the price of equity and the asset value, and we can perform a change of variables to arrive at the density of the (observed) equity prices and hence a likelihood function. This is utilised by the ERMLE method. The more general method described below can also be used to evaluate this likelihood function.

2.4.2 Case 2: Restrict $\sigma_E = 0$

Now, although equity is still a special observed variable, as it is observed without noise, other variables are still relevant and can be included in the estimation. We can still do a change of variable to arrive at the distribution of equity, $p(E)$, but ideally one would need the joint likelihood, and hence density of $p(E, A_1, A_2, \cdots A_n)$. This joint density can be written as

$$p(E, A_1, A_2, \cdots A_n) = p(A_1, A_2, \cdots A_n|E)p(E) \tag{10}$$

It is obvious that it does not matter whether the joint density of other observed variables is conditioned on the equity price, or the asset value implied by the equity price as long as there is a one-to-one correspondence between asset value and the equity price. In terms of a log-likelihood function, this means that it can simply be written as the sum of the change-of-variable likelihood function described in Case 1, plus a term relating to the joint density of the other observed variables, conditional on an equity-implied asset value. The shape of the extra term will be determined by the choice of distribution for the observation errors. Note that this extension to ERMLE is not considered by Ericsson and Reneby (2005), or by Duan, Gauthier, Simonato, and Zaanoun (2003), but has been utilised by Bruche and Reneby (2005).

2.4.3 Case 3: No restrictions (or $\sigma_i \neq 0, \forall i$)

In this case, the the likelihood can not be evaluated by a change of variables, and a more general procedure will be necessary. The problem of evaluating the likelihood function is closely related to that of non-linear filtering. Evaluating the likelihood function in this general case is discussed in the following section. Note that the procedure described below can of course also be used to estimate the systems described under Case 1 and 2.

3 A general procedure for evaluating the likelihood function

In the more general case (Case 3), direct construction of the likelihood function is difficult if a latent variable α enters the observation equation in a non-linear and potentially very complicated fashion. Even in these cases, however, the exact likelihood can be evaluated numerically without an analytical expression, using a method described by Durbin and Koopman (1997). Essentially, it works as follows (using the original notation as much as possible): Define the likelihood as

$$L(\psi) = p(y|\psi) = \int p(\alpha, y|\psi) d\alpha. \tag{11}$$

Suppose we use importance sampling (c.f. e.g. Ripley 1987) from a density $g(\alpha \mid y,\ \psi)$ to evaluate this density. An obvious choice for the importance density is the density of a linear Gaussian model, since it is straightforward to handle (a method for obtaining a suitable approximating linear Gaussian model is described in section 3.1)[2]. We can now write the likelihood as

$$L(\psi) = g(y) \int \frac{p(a, y|\psi)}{g(a, y|\psi)} g(\alpha|y, \psi) d\alpha \tag{12}$$

$$= L_g(\psi) E_g \left[w(a, y) \right], \tag{13}$$

where L_g is the likelihood of the approximating linear Gaussian model and

$$w(a,\ y) = \frac{p(a, y|\psi)}{g(a, y|\psi)}. \tag{14}$$

We can interpret this likelihood function as consisting of the likelihood of the approximating model, multiplied by a factor to correct for the approximation.

The likelihood function of the approximating model can be calculated using the Kalman filter, and the correction factor $E_g[w(a, y)]$ can easily be evaluated using Monte-Carlo techniques. Note that the correction factor becomes more important the more heavily non-linear the model is. We can see that for firms which are very far away from default, where the delta of equity with respect to the asset value is essentially one, and the delta of the bond is essentially equal to zero, the correction factor is likely to be unimportant, for example, as the model is essentially linear. Note that omitting the correction factor would imply QMLE. Sampling from the distribution $g(a|y)$ can be accomplished in various ways, below it is implemented via the procedure described by Durbin and Koopman (2002).

A procedure for estimating the likelihood would be

1. Calculate the linear Gaussian approximation.
2. Calculate the likelihood of the linear Gaussian approximation.

[2]In certain kind of situations it can be advantageous to use a distribution with fatter tails.

3. To obtain a correction factor, simulate α from the importance density, and numerically calculate the expectation term.

We can write

$$\hat{L}(\psi) = L_g(\psi)\bar{w} \tag{15}$$

where

$$\bar{w} = \frac{1}{M}\sum_{i=1}^{M} w_i, \quad w_i = \frac{p\left(\alpha^i, y|\psi\right)}{g\left(a^i,\ y \mid \psi\right)}, \tag{16}$$

and α^i is drawn from the importance density. The accuracy of this numerically evaluated likelihood only depends on M, the size of the Monte Carlo simulation. In practice, the log transformation of the likelihood is used. This introduces a bias for which a modification has been suggested that corrects for terms up to order $O\left(M^{-3/2}\right)$ (cf. Shephard and Pitt 1997, Durbin and Koopman 1997):

$$\log \hat{L}(\psi) = \log L_g(\psi) + \log \bar{w} + \frac{s_w^2}{2M\bar{w}^2}, \tag{17}$$

with $s_w^2 = (M-1)^{-1}\sum_{i=1}^{M}(w_i - \bar{w})^2$.

The technique described here allows the numerical evaluation of the likelihood function and hence its numerical maximisation. Note that many of the assumptions made above serve to simplify the resulting model, but are not necessary for the methodology to be applicable. In particular, errors in the state and observation equation need not be Gaussian or independent.

This approach is generally more computationally intensive than ERMLE, but more general (it can evaluate the likelihood of the restricted version of the system necessary for ERMLE).

3.1 Approximating the model

Durbin and Koopman (2001) describe several methods for deriving an appropriate approximate model. The basic idea of the appropriate method for this case is to iteratively linearise the observation and state equations, which delivers an approximating linear Gaussian model with the same mode as the true model. Starting with an initial guess of α which we call $\tilde{\alpha}$, linearise the observation equation around this guess:

$$Z_t(\alpha_t) \cong Z_t(\tilde{\alpha}_t) + \dot{Z}_t(\tilde{\alpha}_t)(\alpha_t - \tilde{\alpha}_t), \tag{18}$$

where

$$\dot{Z}_t(\tilde{\alpha}_t) = \left.\frac{\partial Z_t(\alpha_t)}{\partial {\alpha_t}'}\right|_{\alpha_t=\tilde{\alpha}_t}. \tag{19}$$

Defining

$$\tilde{y}_t = y_t - Z_t(\tilde{\alpha}_t) + \dot{Z}_t(\tilde{\alpha}_t)\tilde{\alpha}_t, \tag{20}$$

we can approximate the observation equation by

$$\tilde{y}_t = \dot{Z}_t\left(\tilde{\alpha}_t\right)\alpha_t + \varepsilon_t. \tag{21}$$

In most structural models, the state equation is already linear and Gaussian (through the assumption of geometric Brownian motion for the asset value). So the approximating model is:

$$\alpha_t = d + \alpha_{t-1} + \eta_t \tag{22}$$

$$\tilde{y}_t = \dot{Z}_t(\tilde{\alpha}_t)\alpha_t + \varepsilon_t \tag{23}$$

where

$$\eta_t \sim NID\left\{0, \sigma_\eta\right\} \tag{24}$$

$$\varepsilon_t \sim NID\left\{0, \Sigma_\varepsilon\right\} \tag{25}$$

and

$$\eta_t \perp \varepsilon_t. \tag{26}$$

If we start with a guess of α, obtain our guess of y, and then smooth to obtain our next guess of α, and iterate this until convergence, the linear model in the last step is the one that has the same conditional mode as the actual model. For a proof, consult the references cited by e.g. Durbin and Koopman (2001).

3.2 Conditioning on no default

Unfortunately, the approximation method described above does not work well for models with a default barrier if applied naively. Firstly, in this case the Z function is not necessarily monotonic, and hence the true posterior distribution being approximated is possibly bimodal, although the two maxima will be located close to each other for reasonable parameter values. Secondly, the density will not in general be continuous or continuously differentiable at the default barrier. It is not clear what the "best" approximating density is in this case, and the algorithm described above can break down when an iteration reaches the region below the barrier.

In order to reliably find a reasonable approximating model/ density, the following algorithm was implemented: Starting with a large guess for α (ensuring that the starting value is above the mode(s), and above the discontinuity), we iterate until convergence. If convergence is achieved, we will have found a (possibly local) maximum of the posterior density of $\alpha|y$. This is chosen as the mean and mode of the approximating linear Gaussian density. For situations where the mode of the true density is equal to the truncation point, the truncation point is chosen as the mode of the approximating density.

This is a somewhat arbitrary choice of importance density, which implies that convergence might be slow. In order to ensure that convergence is reasonable, the number of simulations is set to 1,000, which is a number for which good convergence occurs in Monte Carlo experiments for various different parameter combinations. Different choices of importance densities might provide efficiency gains, and are an issue for future research.

3.3 Inference and tests

Simulated Maximum Likelihood Estimation (SMLE) allows for classical as well as Bayesian inference techniques. The posterior density of the state given the data can be calculated via simulation (Durbin and Koopman 2001). This allows for calculating expected values of functions of the state (such as bond prices and spreads), as well as for their posterior densities. Both classical as well as Bayesian versions of the calculation can be implemented. For the classical versions, parameters are assumed fixed at their estimates, and since the bias is $O\left(\frac{1}{T}\right)$ as MLE estimators are root-T consistent, it is ignored. For the Bayesian version, they are treated as having posterior distributions. The classical inference version can produce confidence intervals that are too tight in finite samples, so the Bayesian calculations are performed here.

Standard Maximum Likelihood tests could be applied. Consider the case of testing various nested models, or imposing a parameter restriction. It is possible to simply estimate the model under the null and the alternative, compute the likelihoods associated with the two hypotheses, and conduct a Likelihood Ratio test. Wald and Score tests are equally feasible. Diagnostics can be based on the one-step-ahead prediction errors for the observed variables.

3.4 Advantages of the approach

The main advantages of the approach are the following:

- The approach can utilise all data that contains information about the asset value of a firm, including prices of equity, different bonds and credit derivatives. This improves the efficiency of the estimation as shown below.
- It is not necessary to assume that the model prices equity (or any other asset) with 100% accuracy - assets are assumed to be priced with an observation error as described above. Given that we would expect market microstructure effects, agency problems, liquidity issues, time horizon issues, etc. to show up in equity returns (and other asset returns), it is unreasonable to assume that any asset is priced perfectly by a structural model. In fact, this approach allows separating out non-asset value related components in asset prices via the (estimated) observation errors.

The approach also has the following advantages:

- Since we are in a filtering framework, an extension to asynchronously and irregularly spaced data or sporadic missing observations is trivial, which is important especially since corporate bond markets are typically not very liquid meaning that some bonds are not traded on a very frequent basis, as well as allowing the use of e.g. annual balance sheet information together with daily price data.

- It can easily be extended to the case of a multivariate asset value process - allowing for example for the estimation of default correlations based on not only prices of equity (as is common practice) but on prices of debt and equity.

- Estimating a model connecting different observed variables allows for a rigorous and theory-based analysis of relationships between these variables, e.g. the lead-lag behaviour between equity and debt markets.

4 Applications

In order to get an impression of the usefulness of the SMLE technique, it was compared to ERMLE and the calibration technique in an application to simulated data (a Monte Carlo experiment), as well as in an application to real bond pricing data (a large part of this was implemented in Ox Version 3.32 (Doornik 2002) using SsfPack Version 3.0 beta 2 (Koopman, Shephard, and Doornik 1999)).

4.1 The theoretical model

The Merton model is not a model that can be appropriately applied to real bond pricing data, because it makes the assumptions that bonds do not pay coupons, and that the bond represents the entire debt of the firm. As mentioned above, there are very few situations in which these assumptions are actually a reasonable description of the situation being modelled. In practise, bonds typically pay coupons, and equity and aggregate debt are probably more appropriately treated as perpetuities. Also the structure of aggregate debt is typically complicated, which makes fully modelling the cross-dependency of all claims very complicated.

Leland (1994) assumes that aggregate debt and equity are perpetuities. As it turns out, pricing coupon bonds in the context of this model is relatively easy if it is assumed that they represent a negligible proportion of aggregate debt. This has the advantage it makes it possible to avoid compound optionality issues when pricing coupon bonds (described in Geske 1977), as well as providing a reasonable approximation to a possibly very complicated debt structure that it would otherwise be impossible to model. This was first proposed by Ericsson and Reneby (1998) (. Ericsson and Reneby 2002b)

Since Ericsson and Reneby have shown that this model can produce very good out-of-sample performance, and in order to facilitate comparisons with their estimation method (ERMLE), which they test on it, the model is chosen here[3].

In the model, default occurs when the asset value hits the level at which shareholders are no longer willing to contribute funds to stave off financial distress. Ericsson and Reneby note that the barrier could be determined as the outcome of a strategic game. They also allow the aggregate level of debt to grow over time. To simplify the model and to make it essentially the same as the model proposed by Leland (1994), the growth rate of aggregate debt is restricted to zero in the application here. Furthermore, several other parameters were fixed: The recovery to equity was set to 5% (this is a deviation from absolute priority). The recovery to aggregate debt was set to 80%, the corporate tax rate was assumed to be 20% and the per-bond recovery (recovery fraction of principal) was taken to be 50%. These numbers could be made more realistic, for example by choosing a bond recovery fraction according to a table such as the one given in the paper by Altman and Kishore (1996).

4.2 Monte Carlo experiments

The aim of the Monte Carlo experiment is to demonstrate that using Simulated Maximum Likelihood Estimation (SMLE) and including bond price information (as well as information on the price of equity) produces efficient and unbiased estimates of the parameters. It will also aim to demonstrate that Ericsson and Reneby Maximum Likelihood Estimation (ERMLE) as well as the traditional calibration approach on the same dataset produces biased and inefficient estimates.

4.2.1 Setup of the Monte Carlo experiments

Since the calibration technique only allows for the estimation of the asset value volatility parameter and the asset value, these were the only quantities estimated by all the estimation approaches to facilitate a comparison.

To generate the Monte Carlo data, the following parameter assumptions were made: The standard deviation of the observation errors was assumed to be 1%, the asset value volatility was set to 30% p.a. and the market price of risk was set to 0.5 (to produce a reasonable positive drift on average in the simulated asset values).

For each Monte Carlo experiment, 1000 asset value paths of 250 periods each (to represent trading days) were simulated using the same set of parameters, all ending up at the same asset value (figure 1 illustrates this concept),

[3] Note that the Monte Carlo experiment was also conducted with the Merton model, producing very similar results. These are not reported here, however.

and corresponding prices of equity and one bond were calculated for each path (with observation error). The estimations were run on each artifical data set corresponding to one asset value path, on all equity prices and all bond prices excluding the last bond price. This last bond price, the corresponding asset value and spread was then predicted.

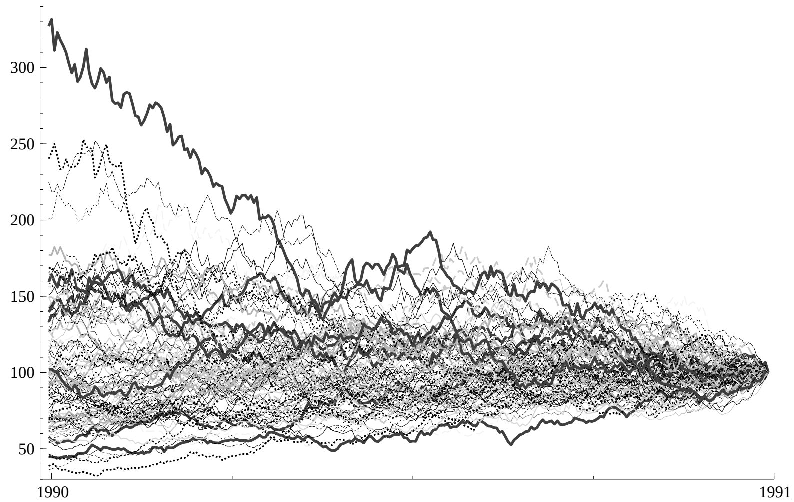

Figure 1: Alternative asset value paths that could have led to an asset value of 100 in 1991

The final debt/equity ratio and the bond parameters were loosely based on the situation of K-Mart in Dec 2001: The final asset value is 12.5b \$, and the face value of aggregate debt in the final period is 12b \$ (the firm is highly leveraged). At the beginning of the sample, the bond has 6 years left to maturity, and pays a semiannual coupon of 5%.

As a firm comes closer to default, its bond prices become more responsive to the underlying financial situation of the firm (asset value), whereas the price of equity becomes less responsive. One would expect that including data on bond prices into the estimation would make a difference in particular for firms which are close enough to default for the default risk to be reflected in the price of the bonds. For firms which are not risky, the price of equity is likely to contain most of the relevant information, as the bonds are essentially priced as risk-free.

In order to investigate the effect of this on estimation, two Monte Carlo experiments were run, with different risk-free rates (4 and 6%). Since the drift of the asset value is equal to the risk-free plus a risk premium (which is the same in both cases), the drift is higher in the case where the risk-free rate is set to 6%. Given that in both cases, the paths end up at the same point, on average, the starting point will be lower for the case with the greater drift. The default probability over the 250 trading days (one year) is 0.52% for the greater drift, and 0.03% for the smaller drift. Average one year default probabilities for 1920 - 2004 as reported by Moodys are 0.06% for Aa rated issuers, 0.07% for A rated issuers, 0.30% for Baa rated issuers, and 1.31% for Ba rated issuers. It can be seen that the case of $r = 6\%$ corresponds roughly to an issuer which is at the border between investment grade and non-investment grade in terms of its rating (low quality issuer), and the case of $r = 4\%$ represents the case of an issuer with a quality higher than an Aa rated issuer (high quality issuer).

Also, choosing these two cases with different drifts will highlight the difficulties that the calibration technique faces when leverage changes. Since the drift in the low quality issuer case is higher (leverage changes more on average), it is possible to anticipate that the performance of the calibration technique should be worse in this case.

4.2.2 Results of the Monte Carlo experiments

ERMLE uses the equity price in all periods to estimate the asset value volatility. Its estimate of the bond price in the last period is based on the equity-implied asset value (implied by the last equity price). For the calibration, the historical equity volatility is calculated utilising the entire sample of equity prices (250 periods), and the predicted bond price in the last period is based on the asset value implied by the equity price and the historical volatility. SMLE uses all equity prices as well as bond prices (except the last bond price, which we are trying to predict) to estimate the asset value volatility and the asset value in the last period, on which the estimate of the bond price in the last period is based.

We can then compare the estimated asset value volatility, the predicted asset value, bond price and spread for the three estimation methods. In order to see whether the posterior densities for the yields produced by SMLE and ERMLE are reasonable, a plot similar to a Q-Q plot is produced (the calibration approach cannot produce posterior densities, of course). In the case of ERMLE, the posterior density is approximated by the asymptotic distribution of yields via the Delta Method (Ericsson and Reneby 2005), in the case of SMLE, the (Bayesian) posterior densities are calculated via simulation.

The results for the high quality and low quality issuer cases are presented in Table 1 and Table 2 respectively. It can be seen that SMLE clearly outperforms the other methods.

Since ERMLE ignores the observation error in the equity price, it will in general overestimate the asset value volatility. Contrary to the assumption of the method, not all the volatility in equity prices comes from volatility in

	Mean error	Median error	RMSE
σ_η			
SMLE	0.0002	0.0003	0.0021
ERMLE	0.0394	0.0391	0.0447
Calibration	−0.0174	−0.0166	0.0285
Asset Value ($m)			
SMLE	3	0	68
ERMLE	−514	−514	583
Calibration	231	239	395
Spread (bp)			
SMLE	0	1	22
ERMLE	36	35	46
Calibration	−14	−14	32
Bond Price ($ per 100$ face value)			
SMLE	0.00	−0.02	0.94
ERMLE	−1.50	−1.49	1.95
Calibration	0.59	0.60	1.38

Table 1: Monte Carlo results (high quality issuer)

the asset value. This in turn means that asset values will be underestimated. Hence default probabilities are overstated, bond prices underpredicted, and the resulting spread predictions are too high.

For the parameter values chosen here, it is actually clear that the calibration technique outperforms ERMLE in terms of RMSE in the case of the high quality issuer. Since we know that equity volatility is an increasing function of asset value, and since the asset value is increasing (on average) in the sample, we know that the equity volatility that the approach uses in the final period to forecast the bond price is understated. The direction of the bias is not clear in general, although given that the problem is caused by ignoring that equity volatility varies as a function of the asset value, we would expect the calibration technique to perform worse in the case of a larger change of leverage over the sample (i.e. the low quality issuer case for this Monte Carlo), as mentioned above. For the low quality issuer, it is indeed clear that the calibration technique performs much worse than ERMLE (Table 2).

Both SMLE and ERMLE allow the calculation of posterior densities for spread or yield predictions. Figures 2 and 3 produce a Q-Q plot of actual versus theoretical quantiles of tests as follows: For each run, the Bayesian predicted density around the point forecast of the yield is obtained. The quantile according to the posterior density for the actual yield is calculated. The quantiles according to the posterior density can then be plotted against the quantiles of

	Mean error	Median error	RMSE
σ_η			
SMLE	0.0000	0.0000	0.0012
ERMLE	0.0262	0.0257	0.0363
Calibration	0.0948	0.0699	0.1465
Asset Value ($m)			
SMLE	0	2	57
ERMLE	−313	−312	433
Calibration	−1017	−851	1519
Spread (bp)			
SMLE	0	−1	23
ERMLE	29	27	46
Calibration	101	77	157
Bond Price (bp) ($ per 100$ face value)			
SMLE	0.01	0.03	0.87
ERMLE	−1.08	−1.01	1.70
Calibration	−3.58	−2.84	5.49

Table 2: Result of Monte Carlo (low quality issuer)

the actual empirical distribution of the yields. If the predicted densities are broadly reasonable, one would expect the plots to lie on the 45 degree line. This allows a quick visual inspection of the quality of the posterior densities produced by the methods.

It is clear that if equity is observed with error, this will mean that the ERMLE confidence intervals for yield predictions (based on asymptotic theory) will be too small, which is indeed confirmed by the plots.

4.3 Application to real data

In order to get an impression of the usefulness of the SMLE technique, it was also applied to real data. Again, a comparison was made to ERMLE and the calibration technique.

4.3.1 The data

The data used to test the different estimation procedures and theoretical models came from several sources. The corporate bond price data is the dataset compiled by Arthur Warga at the University of Houston from the National Association of Insurance Companies (NAIC). US regulations stipulate that insurance companies need to report all changes in their fixed income portfolios,

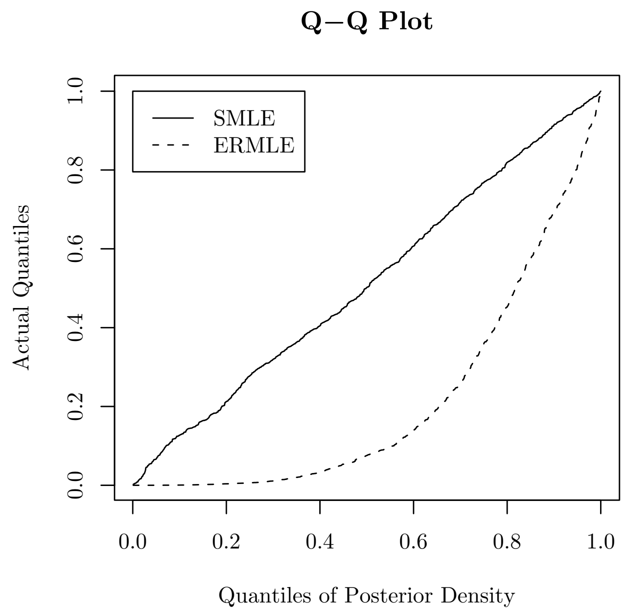

Figure 2: SMLE and ERMLE actual versus predicted quantiles (high quality issuer)

including prices at which fixed income instruments where bought and sold. Insurance companies are some of the major investors in fixed income instruments and the data is therefore reasonably comprehensive. Also, the reported prices are actual transaction prices, and not matrix prices or quotes. The descriptive information on the bonds is obtained from the Fixed Income Securities Database (FISD) marketed by Mergent, Inc., which is a comprehensive database containing all fixed income securities issued after 1990 or appearing in NAIC-recorded bond transactions.

Over the period of 1994-2001, NAIC reports a total of 1,117,739 transactions. First, all trades with unreasonable counterparties (such as trades with the counterparty "ADJUSTMENT" etc.) are eliminated, leaving 866,434 transactions, representing about 43,330 bonds. Since often, one insurance company will buy a bond from another insurance companies, the same price can enter twice into the database, once for each side of the transactions. In order to prevent double counting, all prices for transactions for the same issue on the same date are averaged, to yield a maximum of one bond price observation per issue per date. This leaves 562,923 observations. Since the selected structural model has nothing to say about bonds with embedded optionality, sinking fund

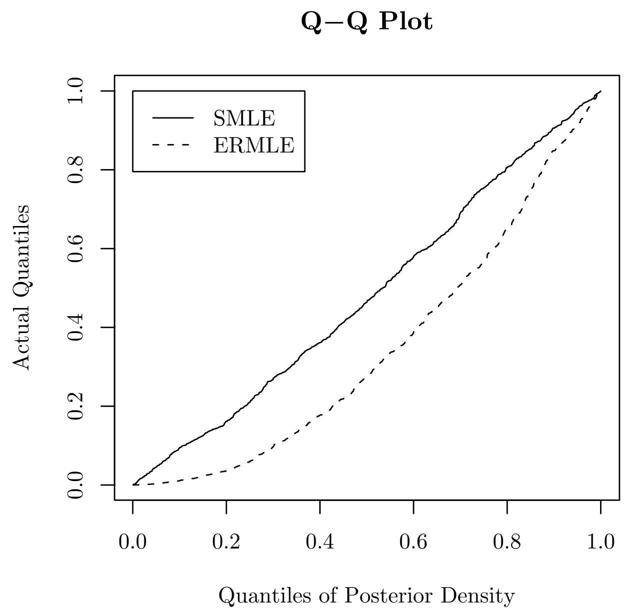

Figure 3: SMLE and ERMLE actual versus predicted quantiles (high quality issuer)

provisions and non-standard redemption agreements, all these bonds were eliminated, leaving 156,837 observations of 8,234 bonds for 1,332 issuers. Finally, government and agency bonds are eliminated, as well as bonds of financial issuers (since the balance sheet of financial issuers are very different from the balance sheet of industrial issuers). This leaves 88,243 observations of 3,907 bonds, for 817 issuers.

Since the point of the methodology presented here is that including additional information improves estimation, only issuers for which there were at least 50 bond price observations in 2001 were selected, leaving 39 issuers.

Equity and accounting data For the selected issuers, the market value of equity was obtained from CRSP, and the value of total liabilities exclusive of shareholder equity (the notional value of aggregate debt) was obtained from Compustat. Issuers for which the data could not be unambigously matched up were ignored. This left 33 issuers, detailed in table 3. Bond price data for which no equity price data was obtainable was ignored.

4.3.2 Risk-free rate data

Implementing corporate bond price models necessitates using risk-free interest rate data. Although most structural models assume constant interest rates (including the one utilised here), this is patently a simplifying assumption which will create problems if used to implement the model. In the implementation, a distinction was made between two types of interest rates: Those to discount individual bond payments, and those used to calculate default probabilities.

Discounting corporate bonds A risk-free curve (zero coupon constant maturity fitted yields) was obtained from Lehman Brothers (calculated from treasury market data). For each corporate bond, the price of a risk-free bond with the same payments was constructed artifically using the risk-free curve. The yield of this artificial risk-free bond was computed. This yield was used to discount the corporate bond in the pricing calculations. This procedure ensures that if any of the corporate bonds were (almost) risk-free, their price would equal the price of a risk-free bond with all payments discounted by rates given by the risk-free curve.

The interest rate in the default probability formulas There were many bonds for any one particular firm, and each of these would have a separate risk-free yield associated with it. The default probabilities formula takes a risk-free rate. If the same default probability is desired for all bonds, a single risk-free rate for use in the formula has to be obtained. A simple arithmetic average of all rates for one firm was chosen. Other possibilities were explored, but the results were not sensitive to the choice of this interest rate parameter.

4.3.3 Results of the application to real data

Since the likelihood function turned out to be not very sensitive to the drift of the asset value, the market price of (asset value) risk was fixed at 0[4].

The estimation was run on data from the beginning of 1999 until the date of the last bond price observation (2001-12-31 in most cases, but earlier in some cases). This last bond price observation was excluded from the estimation sample. After estimation, the bond price and spread was predicted and compared to the actual bond price and spread.

The standard deviation of observation errors could be estimated separately for each asset, but given that for this dataset, the number of bond observations ranges from 20% to 50% of the equity price observations, and given that the standard deviation of the observation errors determines the influence of the various series in the likelihood function, this would imply that the estimation would automatically attach more weight to the asset prices which are more frequently observed.

[4] Changing the market price of risk to 0.5 had no effect on the estimates.

As a simple workaround, it was again fixed at 1% and not estimated here. If anything, this should make it harder for SMLE to do well in comparison to the other methods.

The results are reported in Table 4. As can be seen, the SMLE estimates are substantially less biased (a mean error of -12 bp for SMLE and -77 and -79 bp for ERMLE and Calibration in terms of spreads respectively), and and have a lower RMSE for the same model (88 bp for SMLE and 125 and 128 bp for ERMLE and Calibration respectively), even though a comparatively small amount of bond pricing information was used. The performance of SMLE is good, considering the fact that apart from the asset value, only a single parameter was estimated, and can be seen to be substantially less biased than ERMLE and the calibration technique. ERMLE seems to only slightly outperform the calibration technique for this dataset.

5 Concluding Remarks

This paper suggests applying Simulated Maximum Likelihood Estimation as proposed by Durbin and Koopman (1997) to the estimation of bond prices. This estimation technique allows estimating a system relating several observed variables (e.g. price of equity, price of debt) to an unobserved factor (the asset value). Monte Carlo experiments as well as an application to real data seems to indicate that large gains in efficiency can be achieved over existing techniques that utilise only the price of equity to estimate bond pricing models.

The approach has two main advantages: It can utilise all data that contains information about the asset value of a firm, including prices of equity, different bonds and credit derivatives, improving the efficiency of the estimation as shown above. Also, given that more than one asset price is utilised, it is not necessary to assume that the model prices equity (or any other asset) with 100% accuracy - assets are assumed to be priced with an observation error. Given that market microstructure effects, agency problems, liquidity issues, time horizon issues, etc. are likely to influence equity returns (and other asset returns), and given that these effects are not usually included in structural models, it is unreasonable to assume that any asset is priced perfectly by a structural model. In fact, the approach allows for non-asset value related components in asset prices implicitly via the observation errors.

The approach also easily deals with asynchronously and irregularly spaced data and missing observation, which is important especially since corporate bond markets are typically not very liquid meaning that some bonds are not traded on a very frequent basis.

Furthermore, the approach would facilitate a rigorous and theory-based examination of the relationships between various asset prices used in estimation , for example to examine lead-lag relationships between equity and bond prices. It could easily be extended to a multivariate scenario, where asset value processes are correlated across firms, leading to estimates of portfolio credit risk. All these are topics for future research.

References

Altman, E., and V. Kishore, 1996, "Almost everything you wanted to know about recoveries on defaulted bonds," *Financial Analysts Journal*, pp. 57–64.

Anderson, R., and S. Sundaresan, 2000, "A Comparative Study of Structural Models of Corporate Bond Yields: An Exploratory Investigation," *Journal of Banking & Finance*, 24, 255–269.

Black, F., and J. Cox, 1976, "Valuing Corporate Securities: Some Effects of Bond Indenture Provisions," *Journal of Finance*, 31, 351–367.

Bruche, M., and J. Reneby, 2005, "Joint Estimation of Default and Non-Default Components of Corporate Bond Spreads," unpublished, mimeo.

Crosbie, P., and J. Bohn, 2002, "Modeling Default Risk," Published on Moody's KMV website: http://www.kmv.com, Release Date: 15 Nov 1993, Revised: 14 Jan 2002.

Delianedis, G., and R. Geske, 1999, "Credit Risk And Risk Neutral Default Probabilities: Information about Rating Migrations and Defaults," UCLA Anderson GSM Working Paper.

Delianedis, G., and R. Geske, 2001, "The Components of Corporate Credit Spreads: Default, Recovery, Tax, Jumps, Liquidity and Market Factors," UCLA Anderson GSM Working Paper.

Doornik, J. A., 2002, *Objected-Oriented Matrix Programming Using Ox*, Timberlake Consultants Press, London, UK, 3 edn.

Duan, J., 1994, "Maximum likelihood estimation using price data of the derivative contract," *Mathematical Finance*, 4.

Duan, J., G. Gauthier, S. Simonato, and S. Zaanoun, 2003, "Estimating Merton's Model by Maximum Likelihood with Survivorship Consideration," EFA 2004 Maastricht Meetings Paper No. 4190.

Duffee, G., 1999, "Estimating the Price of Default Risk," *Review of Financial Studies*, 12, 197–226.

Durbin, J., and S. J. Koopman, 1997, "Monte Carlo Maximum Likelihood Estimation for Non-Gaussian State Space Models," *Biometrika*, 84, 669–684.

Durbin, J., and S. J. Koopman, 2001, *Time Series Analysis by State Space Methods*, Oxford University Press, Oxford, UK.

Durbin, J., and S. J. Koopman, 2002, "A simple and efficient simulation smoother for state space time series analysis," *Biometrika*, 89, 603–615.

Eom, Y., J. Helwege, and J. Huang, 2001, "Structural Models of Corporate Bond Pricing: An Empirical Analysis," Ohio State University Working Paper.

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

Ericsson, J., and J. Reneby, 1998, "A Framework for Valuing Corporate Securities," *Applied Mathematical Finance*, 5, 143–163.

Ericsson, J., and J. Reneby, 2002a, "Estimating Structural Bond Pricing Models," Working Paper.

Ericsson, J., and J. Reneby, 2002b, "The Valuation of Corporate Liabilities: Theory and Tests," McGill University Working Paper.

Ericsson, J., and J. Reneby, 2005, "Estimating Structural Bond Pricing Models," *Journal of Business*, 78.

Finger, C., et al., 2002, "CreditGrades^{TM} Technical document," Published on CreditGrades^{TM} website: http://www.creditgrades.com.

Gemmill, G., 2002, "Testing Merton's Model for Credit Spreads on Zero Coupon Bonds," Cass Business School Working Paper.

Geske, R., 1977, "The Valuation of Corporate Liabilities as Compound Options," *Journal of Financial and Quantitative Analysis*, 12, 541–552.

Gropp, R., J. Vesala, and G. Vulpes, 2002, "Equity and Bond Market Signals as Leading Indicators of Bank Fragility," European Central Bank Working Paper No. 150.

Huang, J., and M. Huang, 2003, "How Much of the Corporate-Treasury Yield Spread is Due to Credit Risk?," 14th Annual Conference on Financial Economics and Accounting (FEA); Texas Finance Festival.

Hull, J., 2003, *Options, Futures, and Other Derivatives*, Pearson Education, Upper Saddle River NJ USA.

Jarrow, R., and S. Turnbull, 1995, "Pricing derivatives on financial securities subject to credit risk," *Journal of Finance*, 51, 53–85.

Jones, E., S. Mason, and E. Rosenfeld, 1984, "Contingent Claims Analysis of Corporate Capital Structures: An Empirical Investigation," *Journal of Finance*, 39, 611–625.

Koopman, S. J., N. Shephard, and J. A. Doornik, 1999, "Statistical Algorithms for Models in State Space Using SsfPack 2.2," *Econometrics Journal*, 2, 113–166.

Leland, H., 1994, "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure," *Journal of Finance*, 49, 1213–1252.

Longstaff, F., and E. Schwartz, 1995, "A Simple Approach to Valuing Risky Debt," *Journal of Finance*, 50, 789–821.

Lyden, S., and D. Saraniti, 2000, "An Empirical Examination of the Classical Theory of Corporate Security Valuation," Barclays Global Investors, San Francisco, CA.

Mamaysky, H., 2002, "A Model for Pricing Stocks and Bonds with Default Risk," Yale ICF Working Paper No. 02-13.

Merton, R., 1974, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 29, 449–470.

Morris, S., and H. Shin, 2004, "Liquidity Black Holes," *Review of Finance*, 8, 1–18.

Ripley, B., 1987, *Stochastic Simulation*, Wiley, New York, USA.

Ronn, E., and A. Verma, 1986, "Pricing risk-adjusted deposit insurances: An option-based model," *Journal of Finance*, 41, 871–895.

Schaefer, S., and I. Strebulaev, 2003, "Structural Models of Credit Risk are Useful: Evidence from Hedge Ratios on Corporate Bonds," LBS IFA Working Paper.

Shephard, N., and M. K. Pitt, 1997, "Likelihood Analysis of Non-Gaussian Measurement Time Series," *Biometrika*, 84, 653–667.

Shimko, D., N. Tejima, and D. van Deventer, 1993, "The Pricing of Risky Debt When Interest Rates are Stochastic," *The Journal of Fixed Income*, 3, 58–65.

Wei, D., and D. Guo, 1997, "Pricing Risky Debt: An Empirical Comparison of the Longstaff Ad Schwartz and Merton Models," *Journal of Fixed Income*, 7, 8–28.

Name	Rating*	Market Cap†
Adelphia Communications Corp		3,473
Archer Daniels Midland Co	$A+$	7,300
Boeing Co	A	37,907
Charter Communications Hldgs Llc	$CCC+$	5,010
Coca Cola Enterprises Inc	A	8,483
Consolidated Edison Co NY Inc	A	7,332
CSX Corp	BBB	6,755
Dayton Hudson Corp (target corp)	$A+$	31,576
Deere & Co	$A-$	10,019
Disney Walt Co	$BBB+$	61,612
Dow Chem Co	$A-$	29,084
Emerson Elec Co	A	24,878
Enron Corp		31,084
Ford Mtr Co Del	$BBB-$	45,863
General Mtrs Corp	BBB	47,776
GTE Corp	$A+$	66,386
Hertz Corp	$BBB-$	2,015
Honeywell Intl Inc	A	44,740
International Business Machs Corp	$A+$	208,426
Lockheed Corp	BBB	8,018
Lowes Cos Inc	A	21,314
Lucent Technologies Inc	B	242,171
Norfolk Southn Corp	BBB	7,660
Penney J C Inc	$BB+$	5,127
Philip Morris Cos Inc		55,456
Procter & Gamble Co	$AA-$	140,863
Ralston Purina Co		8,296
Raytheon Co	$BBB-$	6,188
Union Pac Corp	BBB	10,704
Viacom Inc	$A-$	31,325
Wal Mart Stores Inc	AA	297,844
Weyerhaeuser Co	BBB	16,476
Worldcom Inc		147,417

* S&P long term domestic issuer rating according to Compustat
† Market capitalisation on 2000-01-03 in $m

Table 3: Issuers chosen for empirical analysis

	Mean error	Median error	RMSE
Spread (bp)			
SMLE	-12	-12	88
ERMLE	-77	-73	125
Calibration	-79	-68	128
Bond price (bp) ($ per 100$ face value)			
SMLE	-0.22	0.23	6.71
ERMLE	4.08	2.82	9.71
Calibration	4.26	2.89	10.13

Table 4: Yield/ spread prediction errors on real data