ESTIMATING STRUCTURAL MODELS OF CORPORATE BOND PRICES

Max Bruche

CEMFI Working Paper No. 0610

June 2006

CEMFI
Casado del Alisal 5; 28014 Madrid
Tel. (34) 914 290 551. Fax (34) 914 291 056
Internet: www.cemfi.es

CEMFI Working Paper 0610
June 2006

ESTIMATING STRUCTURAL MODELS OF CORPORATE BOND PRICES

Abstract

One of the strengths of structural models (or firm-value based models) of credit (e.g. Merton, 1974) as opposed to reduced-form models (e.g. Jarrow and Turnbull, 1995) is that they directly link the price of equity to default probabilities, and hence to the price of corporate bonds (and credit derivatives). Yet when these models are estimated on actual data, the existence of data other than equity prices is typically ignored. This paper describes how all available price data (equity prices, bond prices, possibly credit derivative prices) can be used in estimation, and illustrates that using e.g. bond price data in addition to equity price data greatly improves estimates. In this context, the issue of possibly noisy data and/or model error is also discussed.

JEL Codes:
Keywords:

Max Bruche
CEMFI
bruche@cemfi.es

1 Introduction

In 1974, Merton wrote a seminal paper 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 (Longstaff and Schwartz, 1995) or an optimally chosen capital structure (Leland, 1994), to name but a few.

Structural models have found applications in risk management, (e.g. the KMV EDF methodology described by Crosbie and Bohn, 2002), in central banks (Gropp, Vesala, and Vulpes, 2002), or in pricing (e.g. the CreditGrades model described by 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 model) seems to require implausibly high volatilities to generate reasonable bond prices, or that it overpredicts bond prices, or underpredicts spreads for reasonable volatilities (Jones, Mason, and Rosenfeld, 1984). More recently, it has been argued that some structural models that were developed as a reaction to the empirical performance of the simple model actually overpredict spreads, but more crucially, that all structural models appear to make very imprecise predictions (Eom, Helwege, and Huang, 2004; Lyden and Saraniti, 2001).

So called reduced-form models, pioneered by Jarrow and Turnbull (1995) that simply posit a default intensity process that describes the instantaneous probability of default seem to be able to match observed bond spreads quite well when implemented (Duffee, 1999), possibly because factors not considered explicitly in the model, like liquidity risk or tax rates, are subsumed into the hazard rate estimates. Consequently, they are often the model of choice when practitioners price credit derivatives.

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 ("Capital Structure Arbitrage"). 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 reduced-form models do not produce an explicit link between the price of equity and default probabilities, and hence the price of the bond, they are not suitable for the above application. Research by Schaefer and Strebulaev (2003) and Yu (forthcoming) indicates that for this kind of purpose, structural models produce hedge parameters that although they seem to be correct on average across a portfolio, cannot be used for hedging on a firm-by-firm basis. It is possible that this is a consequence of the way these models are implemented, and that enlarging the information set used in estimation (as discussed in this paper) could make firm-by-firm hedging feasible.

While part of the poor empirical performance of structural models might be attributable to misspecified models, this paper argues that a large part of the problem is also attributable to bad estimation practice: In particular, it proposes that to estimate structural models, one would ideally utilize one of their strengths, which is that they price equity *and* debt (or credit derivatives) simultaneously. Typically, models are estimated using equity prices, and the estimates are used to calculate predicted bond prices, which can then be compared to actual bond prices. This paper argues that if bond prices (ideally, also credit derivative prices) are actually available, it is very useful to use these in estimation in addition to equity prices.

A consequence of having more than one actual security price that needs to be matched is that model error or observation errors will have to be considered: For a given set of parameters and a standard structural model, an asset value can be chosen such that at most one observed market price can be matched perfectly. If the model were a correct representation of reality, it should be able to match all prices simultaneously, for the correct asset value. In practice, there are things that are beyond the scope of the model, and not all prices can be matched. It therefore becomes a necessity to consider the possibility that the model does not reproduce all prices exactly (model error). An alternative way of looking at this is to say that price data is “noisy”, and contains the effects of market microstructure, liquidity issues, taxes, agency problems etc., which are not included in typical structural models, and there are observation errors in observing the true (fundamental) prices of debt and equity relevant for the model. An interesting question that arises in this context is whether considering the possibility of noise in equity prices in itself (without utilising bond prices) has an effect on estimation results.

In answering the question whether including bond prices and/or observation errors in equity prices makes a difference for estimation, possible estimation methods need to be discussed. In particular, two maximum likelihood techniques are presented here, Simulated Maximum Likelihood Durbin and Koopman (1997) and its application to the problem of estimating structural models, and an extension of a change-of-variable based maximum likelihood method Duan (1994) that allows using more than one price (this extension is original to this paper). The former has the advantage that it is very general and can deal with a wide variety of estimation problems, whereas the latter is computationally less demanding.

In an application to real pricing data, the paper demonstrates that including bond prices in the estimation reduces the out-of-sample spread prediction errors for three different structural models by up to a third.

The paper also shows that accounting for observation error in equity prices seems to have a small but possibly beneficial effect in terms of reducing errors, although the evidence is not as clear.

The conclusion necessarily is that given that the strength of structural models lies in predicting a tight link between equity and bond prices, including bond prices and not only equity prices in the estimation of these models is necessary.

The rest of this paper is organised as follows: First, some of the fundamental problems of estimating structural models are discussed (section 2), and some common estimation approaches examined (section 3). A very general setup of an estimation is then presented (section 4), which is applied to artificial data (a Monte Carlo experiment) (section 6), and some real data (section 7). Finally, results are discussed and conclusions are drawn.

2 The fundamental estimation problem

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, or, as Jarrow and Turnbull (1995) put it: "to use Merton's model, one must also be able to measure the current value of the firm's assets. This is a difficult task".

The key distinguishing feature of different estimation approaches is how this problem is dealt with. Typically, some form of equity-implied asset values are used.

Once the 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.

3 Previously proposed estimation approaches

The accounting approach The first attempt at implementing structural models on corporate bonds was conducted by Jones, Mason, and Rosenfeld (1984). They initially 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.

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.

Variants of this technique have been employed by e.g. Lyden and Saraniti (2001), or more recently Eom, Helwege, and Huang (2004), who simply add the book value of debt (Total Liabilities) to the observed market value of equity to arrive at an estimate of the asset value.

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 (this method was first proposed in the context of deposit insurance by 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.

$$
\begin{aligned}
E &= F_E(V, \sigma_V) && (1)\\
\sigma_E &= \sigma_V \frac{V}{E} \frac{\partial F_E}{\partial V}. && (2)
\end{aligned}
$$

This approach is often the only one discussed in major textbooks (Hull, 2003).

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 2. 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 cannot utilize the dynamic implications of structural models in estimation.

The Change-of-Variable (CoV) approach Ericsson and Reneby (2005) (cf. also Ericsson and Reneby, 2002a) demonstrate the biases that can 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 as first suggested 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. 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 maximized using standard techniques.

4 More general estimation approaches

The objective of this paper is to present and evaluate approaches that in contrast to the previously mentioned approaches allow for the inclusion of more than one type of asset price (e.g. the simultaneous use of equity prices, bond prices and/ or credit derivative prices). In order to do this, it will be useful to clearly state the estimation problem. One of the key points here is that the asset value needs to be treated as a latent or unobserved variable.

We can set up an econometric model that consists of a transition equation for the latent or unobserved state state (an equation describing how asset value changes), and some observation equations that describe the functions that map asset values into the observed prices of debt and equity respectively (the pricing equations of the structural model). The aim is to arrive at a structure that allows maximum likelihood estimation to be performed in terms of a vector of observed asset prices.

4.1 The state equation

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

$$dV_t = \mu V_t dt + \sigma V_t dW_t. \tag{3}$$

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

$$\int_t^T d\log V_s = (\mu - \frac{1}{2}\sigma^2) \int_t^T ds + \sigma \int_t^T dW_s \sim N\left\{\mu\left(T-t\right), \sigma^2\left(T-t\right)\right\}. \tag{4}$$

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

$$\alpha_t = d + \alpha_{t-1} + \eta_t, \quad \eta_t \sim NID\left\{0, \sigma_\eta\right\}, \tag{5}$$

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

The methods described below would also work for more complicated structural models, providing the transition density can be expressed in closed form.

4.2 The observation equation

A given structural model produces some pricing functions F_{it} taking parameters ψ that map a factor (the asset value) α_t into a vector of market prices ξ_t.

If the number of market prices is larger than the number of states, a given model will not be able to match all prices exactly unless the model is correct. Put differently, if the model were correct, knowing one element of ξ would imply knowing all other elements of ξ, and the variance-covariance matrix of ξ would be singular. Since in practice, this is unlikely to be the case, this stochastic singularity needs to be broken by introducing observation errors - we can assume that at most one price is observed without error.

Essentially, 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 will then be 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 essentially one of non-linear filtering.

Prices could be higher or lower than predicted by the model, but they cannot be negative (due to limited liability). Also, it seems plausible to suggest that the size of the errors to be proportional to the price. A multiplicative log-normal observation error therefore appears to be a simple and natural modelling choice.

$$\xi_{it} = F_{it}\left(\alpha_t; \psi_i\right) e^{\varepsilon_t}, \quad \varepsilon_t \sim NID\left\{0, \Sigma_\varepsilon\right\}. \tag{6}$$

Prices conditional on the unobserved asset value would then be log-normal. Note that this is not a totally innocuous assumption, as it implicitly fixes the type of distance between actual and predicted prices that matters for estimation. Here, we are implicitly looking at something like the squared percentage pricing error. If the focus is more on yields or spreads than prices, for instance, it might be useful to choose an expression that takes as the quantity of interest an error related to yields or spreads.

Also, it would be possible to put different and more elaborate structure on the observation errors or components that are outside the structural model, in particular if one had a view on what causes the divergence between model prices and observed prices. For example, liquidity risk has been suggested as an important component in the pricing of corporate bonds (Longstaff, Mithal, and Neis, 2005), and an ad-hoc specification could include proxy variables for liquidity in the specification of the observation or model error.

4.3 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

$$\alpha_t = \quad d + \alpha_{t-1} + \eta_t, \quad \eta_t \sim NID\left\{0, \sigma_\eta\right\}, \tag{7}$$

$$y_t = \quad Z_t\left(\alpha_t; \psi\right) + \varepsilon_t, \quad \varepsilon_t \sim NID\left\{0, \Sigma_\varepsilon\right\} \tag{8}$$

where e.g. η_t is independent of ε_t, and Σ_ε is diagonal. Identification 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.

4.4 Possible assumptions and estimation approaches

Given the previous discussion, one could consider the following different combinations of estimation with different information sets and acknowledging or ignoring the possibility of observation or model errors in different types of prices:

1. Estimation using equity and bond price data, allowing for observation errors in equity,

2. using equity and bond price data, not allowing for observation errors in equity ($\sigma_E = 0$),

3. using only equity data, allowing for observation errors in equity,

4. using only equity price data, and not allowing for observation errors in equity ($\sigma_E = 0$).

Here, we only consider estimation using equity and bond prices, and only consider ignoring observation error in equity prices for the sake of simplicity.

In the most general case (the first item above), evaluating (and hence maximizing) the likelihood function is difficult if $Z(\cdot)$ is non-linear and complicated, which is very likely for corporate bond pricing models. The exact likelihood can be evaluated numerically without an analytical expression, using a method described by Durbin and Koopman (1997): Simulated Maximum Likelihood Estimation (SMLE). This method does need to be implemented in a way that takes into account some peculiarities of structural models, though. For example, given a first passage model that views equity and debt as barrier options (Black and Cox, 1976, e.g.), it will be necessary to condition on no default (it is not possible to observe equity prices implying asset values below the barrier). For details, see appendix A. This method is relatively computationally intensive.

In the second case, we would like to use all data, but are imposing the restriction that there are no observation errors in equity prices. The resulting likelihood can be evaluated with the general method (SMLE), but this is computationally intensive. It is also possible to modify the Change-of-Variables technique discussed above to allow estimation in this case: For maximum likelihood estimation, we need the joint density of equity prices and other asset prices, i.e.

$$p(\xi_E, \xi_1, \xi_2, \cdots \xi_n) \tag{9}$$

where ξ_E denotes the market value of equity and ξ_i denotes the observed value of asset i (assume there are $i = 1, \dots, n$ of such other assets). This joint density can be expressed by first performing a change of variable from the distribution of the log asset value to arrive at the distribution of equity to obtain $p(\xi_E)$, and then writing the joint density as

$$p(\xi_E, \xi_1, \cdots \xi_n) = p(\xi_1, \cdots \xi_n | \xi_E) p(\xi_E) \tag{10}$$

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, 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. This method is denoted the extended Change-of-Variables approach below (ext. CoV). As far as the author knows, this method is original.

In the third case, equity observation errors are taken into account, but only equity data is used for estimation. Again, this kind of structure can be estimated via SMLE as in the first case.

The fourth case describes a situation in which only equity prices matter, and there are no observation errors in equity prices. This corresponds to the situation described by Duan (1994), Duan, Gauthier, Simonato, and Zaanoun (2003) and Ericsson and Reneby (2005), and estimation can proceed via the CoV method.

Lastly, if an econometric motivation is deemed unnecessary, one could also use the Calibration technique (as described above) for estimation in a situation where only equity prices matter and there is no observation error in equity prices. Of course, this has some obvious drawbacks, for example it is not straightforward to calculate confidence intervals for predictions, and the method is likely to produce biased results in cases where leverage changes. In the following, it will be examined because it is used frequently in practice.

The aim of this paper is to firstly illustrate that structural models can be estimated given these combinations of relevant data and restrictions, and secondly to show what difference they make in practice. To explore this, maximum likelihood estimation under these four assumptions (plus the calibration method) will be applied to some real data, and the results compared (section 7). First it will be useful, though, to demonstrate with a Monte Carlo experiment that SMLE can be used to estimate systems of this type, given that this method has not been applied to the estimation of structural models before (section 6). Prior to this, even, some structural models to be estimated must be chosen (discussed in the next section).

5 Theoretical models

Before applying the different methods under the different assumptions, a structural model has to be chosen. Here, three models were considered.

5.1 The Merton 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. There are very few situations in which these assumptions are actually a reasonable description of the situation being modelled. In practice, bonds typically pay coupons, and equity is probably more appropriately treated as a perpetuity. Also the structure of aggregate debt is typically complicated, which makes fully modelling the cross-dependency of all claims very complicated, as e.g. pointed out by Jarrow and Turnbull (1995).

Nevertheless, this model represents somewhat of a benchmark. It was implemented here following the approach of Eom, Helwege, and Huang (2004): Coupon bonds are modelled as a portfolio of zero-coupon Merton bonds (this is of course incorrect, because it ignores the dependence of defaults between the bonds in the portfolio).

5.2 A Leland (1994) model

Leland (1994) assumes that aggregate debt and equity are perpetuities. In this model, default occurs endogenously when the asset value hits the level at which shareholders are no longer willing to contribute funds to stave off financial distress.

There is a question as to how to price finite maturity bonds within this framework. Here, what could be called a "quasi reduced-form approach" was chosen: Assuming that the coupon bonds being priced represent a negligible proportion of aggregate debt, we can calculate default probabilities in line with equity and aggregate debt, and use these to price the bonds in question. This has the advantage that it makes it possible to avoid compound optionality issues when pricing coupon bonds (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 (2002b).

5.3 Leland-Toft

Another approach to modelling finite-maturity debt in the context of the previous model is to assume that aggregate debt consists of finite maturity bonds paying a continuous coupon, but that only small proportion expires every period, and is replaced by newly issued bonds, always with the same maturity. This is the model proposed by Leland and Toft (1996).

5.4 Parameter assumptions

Some of these models have many parameters. From a practical perspective, it turns out that likelihood functions are very flat with respect to many of these (e.g. the level of the barrier in a first-passage type structural model, assuming that the value of assets is also estimated), so it is infeasible to estimate all of them. Here, most of them are fixed at plausible levels, which appears to be the standard approach followed by the literature. Given that the focus is on the relative merits of different information sets and restrictions in estimations and not on the models per se, this is deemed acceptable.

In the Merton model, the time to maturity of equity was arbitrarily fixed at 20 years, as an approximation to the infinite maturity of debt. For the Leland model, 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 payout rate was assumed to be 2%. For all models, 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).

More parameter assumptions are made that apply only to the Monte Carlo experiment or to the analysis of the real data as detailed in the respective sections below.

6 Monte Carlo experiment

Given that SMLE has not previously been applied to the estimation of structural models, the aim of the Monte Carlo experiment is to demonstrate that it is possible to use Simulated Maximum Likelihood Estimation (SMLE) to estimate the system without restrictions, and that it is efficient and unbiased if equity and bond prices are produced by the structural model that is being estimated and prices are indeed observed with a small error.

Also, it will be used to illustrate how some previously proposed approaches (i.e. using only equity data, and estimating via either Calibration or CoV) could fail in this kind of situation.

Unfortunately, a Monte Carlo experiment for a numerical method like SMLE is very computationally intensive. For a dataset of 250 observations (corresponding to e.g. a year of daily data) of equity and bond prices, SMLE estimation of the full system takes about half an hour on a modern PC. Although not a constraint for doing a firm-by-firm estimation for 50 companies, this can hit the limits of feasibility for running a Monte Carlo experiment with, say, 10,000 artificial datasets. In view of computational constraints, only one model, the Leland model was used in the experiment.

6.1 Setup of the Monte Carlo experiment

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, 10,000 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), and corresponding prices of equity and one bond were calculated for each path (with observation error). The estimations were run on each artificial 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.

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 (the asset value), whereas the price of equity becomes less responsive, as measured by the respective deltas. One would expect that including data on bond prices in 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 bonds.

In order to investigate the effect of this on estimation, two versions of the Monte Carlo experiment were run, with different risk-free rates (4 and 6%). Since the drift of the asset value is equal to the risk-free rate 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 faster), it is possible to anticipate that the performance of the calibration technique should be worse in this case.

6.2 Results of the Monte Carlo experiment

For the CoV technique and calibration, the equity price in all periods is used to estimate the model parameter (the asset value volatility). The bond price in the last period is then predicted, based on the equity-implied asset value in the last period. For the full system estimation without restrictions via SMLE, all equity prices as well as bond prices (except the last bond price, which we are trying to predict) are used 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 cases. This is an out-of-sample prediction (in the cross-sectional, not the time series sense).

The results for the high quality and low quality issuer cases are presented in Table 2 and Table 3 respectively. It can be seen that estimating the full system using SMLE clearly outperforms estimation using only equity prices (and assuming zero equity pricing error) with the other methods if the data is generated according to the structural model being estimated, and if there is indeed a small amount of observation error in prices. Mean spread prediction errors are 1bp or less in size for SMLE for the full system, and the root mean square error (RMSE) is 22 bp, whereas the mean error is in the range of 30bp for the CoV technique on equity data only (a RMSE of about 45bp).

As expected, ignoring an observation error in equity if it is present produces a bias, and not using bond price data decreases efficiency.

The direction of the bias is also as expected: Since the CoV method necessarily ignores the observation error in the equity price by construction, it will in general overestimate the asset value volatility (e.g. for independent observation errors). 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 CoV in terms of the root mean square error (RMSE) in the case of the high quality issuer, since it has a mean spread error of -14bp, and a RMSE of 32bp. For the low quality issuer, changing leverage is more of an issue, and it is indeed clear that in this case Calibration performs much worse, with a mean error of about 100bp, and a RMSE of 157bp (table 3).

We conclude that SMLE of the full system works as expected if the data is generated by the model, and small observation errors are present.

7 Application to real data

The question is of course not whether or not using bond as well as equity prices makes a difference in a Monte Carlo experiment, but whether it makes a difference in practice. To answer this question, the three structural models presented above were estimated under the following assumptions, and with the following techniques on real data:

1. using equity and bond price data, allowing for observation errors in equity (SMLE),
2. using equity and bond price data, not allowing for observation errors in equity (ext. CoV),
3. using only equity price data, allowing for observation errors in equity (SMLE),
4. using only equity price data, and not allowing for observation errors in equity (CoV), and
5. using only equity price data, and not allowing for observation errors in equity (Calibration).

7.1 The data

The data used to test the different estimation procedures and theoretical models came from several sources.

7.1.1 Bond price data

The corporate bond price data is the dataset compiled by Arthur Warga at the University of Houston from the National Association of Insurance Commissioners (NAIC). US regulations stipulate that insurance companies need to report all changes in their fixed income portfolios, 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 models have nothing to say about bonds with embedded optionality, sinking fund 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. Financial issuers are typically excluded when asset values are calculated from accounting data since their balance sheets are very different from the balance sheet of industrial issuers. In our case, it would not really be necessary to exclude them, but it is done here to facilitate comparison of the results with other studies. 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, and most observations occur later in the sample, only issuers for which there were at least 50 bond price observations in 2001 were selected, leaving 50 issuers. The properties of the selected bonds and issuers are described in detail in tables 5 and 4. The issuers rating ranges from AA- (Proctor & Gamble) to B (Lucent Technologies), with BBB being the most common, and equity market capitalization ranging from around 2bn USD (Hertz), to around 300bn USD (Walmart).

There appear to be some outliers in the bond price data, that is to say many cases in which bond prices in the dataset seem implausible. For example, for bond prices at roughly a weekly frequency, prices of 109, 108 and 108 might be followed by a price of 53, which is then followed by 107, 108. It is not possible to tell whether an actual market transaction occurred at that price and that date, or whether there was a data entry error, although it seems somewhat implausible that this is not an error. Since it is not clear, however, that these are errors, they were left in the dataset. If these were errors, this would reduce the value of bond price information, and hence the value of using bond price information in estimation, and we should find the usefulness of bond price data diminished.

7.1.2 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.

7.1.3 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.

7.1.4 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 artificially 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.

7.1.5 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.

7.2 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[1].

The estimations were run separately for each firm, with the individual datasets starting in the beginning of 1999 and going until the date of the last bond price observation (2001-12-31 in most cases, but earlier in some cases). The last bond price observation was excluded from each firm's estimation sample. After estimation, the bond price and spread was predicted and compared to the actual bond price and spread. The accuracy of predic-

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

tions was then compared out-of-sample (in the cross section of 54 bond price observations).

Parameter estimates are reported in tables 6 - 8. It can be seen that including bond price information in the estimation seems to increase the average asset value volatility (σ_η) that is estimated in the case of the Merton and Leland models (from about $0.3 - 0.4$ to about $0.4 - 0.5$). This reflects the fact that these models tend to underpredict spreads on average when estimated using only equity data. When forced to match observed spreads, asset value volatilities have to be higher. For the Leland+Toft model, the asset value volatility estimates are roughly the same across methods (about 0.2). Comparing the average estimated standard deviation of observation errors across bonds and equity, we can see that the standard deviation of bond price observation errors is larger by at least an order of magnitude. In absolute terms, the standard deviation of the equity observation error is in the range of between 3 and 30bp, whereas the standard deviation of the bond observation error is in the range of between 4 and 10%. This implies that apparently it is easier for structural models to match equity prices than it is to match bond prices. This could, for example, be an indication that bond spreads contain a non-default related component (Longstaff, Mithal, and Neis, 2005).

The mean spread prediction error (in bp) and the root mean square error (also in bp) are summarized in table 1. These are convenient measures of the bias, and of the combination of the spread and bias of prediction errors.[2]

mean error	Leland	Leland+Toft	Merton
all data, no restrictions (SMLE)	−50	77	−98
all data, $\sigma_E = 0$ (ext. CoV)	−51	92	−95
equity prices only, no restrictions (SMLE)	−132	61	−177
equity prices only, $\sigma_E = 0$ (CoV)	−127	63	−175
equity prices only, $\sigma_E = 0$ (Calibration)	−128	44	−175
root mean square error	Leland	Leland+Toft	Merton
all data, no restrictions (SMLE)	110	165	156
all data, $\sigma_E = 0$ (ext. CoV)	130	166	167
equity prices only, no restrictions (SMLE)	173	188	233
equity prices only, $\sigma_E = 0$ (CoV)	168	188	233
equity prices only, $\sigma_E = 0$ (Calibration)	173	186	231

Table 1: Spread prediction errors (in bp)

It can be clearly seen that the including bond price data lowers the (root mean square) spread prediction errors for all models, that this effect is less marked for the Leland+Toft model, and that in terms of the RMSE, the Leland model produces the best fit (it produces a minimum of 110bp of RMSE).

The Merton and Leland model underpredict spreads on average, whereas the Leland+Toft model overpredicts. Interestingly, including bond price seems to increase the bias in the case of the Leland+Toft model, although it decreases the RMSE. It only does so in terms of the predicted spreads,

[2]For the sake of completeness, following Eom, Helwege, and Huang (2004), the means and standard deviations of the percentage pricing error (predicted bond price minus actual price over actual price), the absolute percentage pricing error, the percentage yield error (predicted yield minus actual yield over actual yield), the absolute percentage yield error, the percentage spread error (predicted spread minus actual spread over actual spread) and the absolute percentage spread error are also reported in the appendix (tables 9, 10, 11). Finally, figure 2 illustrates the distribution of spread prediction errors with boxplots.

however. Through the choice of the form of the observation error, a distance criterion is implicitly defined that is taken into account in the maximum likelihood estimation, which is the squared difference of the log predicted and log actual bond prices. If prediction error is measured by a more similar criterion, including bond price information in the estimation actually produces less biased forecasts.

The Leland+Toft model also is particularly imprecise, which parallels the results of Eom, Helwege, and Huang (2004), although here, the largest outliers have an absolute size below 1000bp, whereas they have outliers as large as 4000bp. It also is the model for which adding bond price data has the smallest effect in terms of reducing errors. It seems possible that since the model is less able to generate precise predictions for bond prices, the estimation benefits less from including bond prices.

Allowing for equity observation errors lowers the error if bond prices are included for the Leland and Merton models (10-20bp), but appears to have no effect if only equity price information is used.

It is also not clear that either of the two likelihood methods that is using only equity data (SMLE and CoV) is outperforming calibration.

8 Conclusion

The strength of structural models vis-à-vis reduced form models lies in the fact that they can link not only bond prices do prices of other bonds and credit derivatives, but also to prices of equity. Yet when implemented, even when the focus is on the prediction of bond spreads, typically only equity price data is utilized.

This paper demonstrates that if bond price data is available, it *has* be used in addition to equity price data if accuracy is desired.

It discusses two likelihood-based methods that can be used to estimate these models using all available data (this might include equity prices, bond prices and prices of credit derivatives), Simulated Maximum Likelihood Estimation, which is very general, but computationally intensive, and and extended Change-of-Variable approach (original to this paper), which is less general but less computationally intensive. It illustrates with an application to real data how using bond prices in addition to equity prices substantially reduces out-of-sample root mean square spread prediction errors, by up to a third.

It is also shown that whether or not bond price data is included is much more important than what kind of estimation method is used if only equity price data is used (likelihood based approaches versus calibration), and whether or not possible observation errors in equity prices are taken into account.

As a by-product, it has been shown that out of the theoretical models considered, a version of the Leland model appears to fit the data best.

It is clear that structural models have difficulties in accurately predicting spreads. It is also clear, however, that estimating structural models using only equity prices compounds the problem, and that they are better estimated using all available data.

References

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

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

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.

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, 2004, "Structural Models of Corporate Bond Pricing: An Empirical Analysis," *Review of Financial Studies*, 17, 499–544.

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$^{\text{TM}}$ Technical document," Published on CreditGrades$^{\text{TM}}$ 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.

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.

Leland, H., and K. Toft, 1996, "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure," *Journal of Finance*, 51, 987–1019, Papers and Proceedings of the Fifty-Sixth Annual Meeting of the American Finance Association, San Francisco California, January 5-7, 1996.

Longstaff, F., S. Mithal, and E. Neis, 2005, "Corporate Yield Spreads: Default Risk or Liquidity. New Evidence from the Credit-Default Swap Market," *Journal of Finance*, 60, 2213 – 2253.

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

Lyden, S., and D. Saraniti, 2001, "An Empirical Examination of the Classical Theory of Corporate Security Valuation," SSRN Working Paper.

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

Ripley, B., 1987, *Stochastic Simulation*, John Wiley & Sons Ltd., New York NY 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.

Yu, F., forthcoming, "How Profitable Is Capital Structure Arbitrage?," *Financial Analysts Journal*.

A Maximising the likelihood function in the general case

In the most general case, evaluating (and hence maximizing) the likelihood function is difficult if $Z(\cdot)$ is non-linear and complicated, which is very likely for corporate bond pricing models. 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): The likelihood L, depending on some parameters ψ is the density of the observed variables $p(y)$. We have the joint density of the observed variables y and the unobserved asset variable α. One way of obtaining just the marginal density of the observed variables y and hence the likelihood function is to simply integrate out the α:

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

With some algebra, it can be shown that given that one has an approximation with density $g(\cdot)$, this is equivalent to

$$\begin{aligned} L\left(\psi\right) =& g(y) \int \frac{p(a, y|\psi)}{g(a, y|\psi)} g(\alpha|y, \psi) d\alpha && (12) \\ =& g(y) E_{g(\alpha|y)}\left[w(a, y)\right], && (13) \end{aligned}$$

where

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

This essentially says that given an approximation, the true density or likelihood can be expressed as the approximate density times a correction factor, where this correction factor is related to the distance between the truth p and the approximation g, and can be calculated numerically. If the approximation is sufficiently tractable, this allows the evaluation of the likelihood (and hence maximum likelihood estimation). $g(\cdot)$ can be interpreted as an importance sampler used to integrate out the nuisance variable α (cf. e.g. Ripley, 1987).

The procedure for evaluating the likelihood for a given set of model parameters can be summarised as follows:

1. Calculate a tractable approximation (for details on choosing an approximation, see section A.1).
2. Calculate the likelihood of the tractable approximation.
3. To obtain a correction factor, simulate α from the importance density (for details see section A.2), and numerically calculate the expectation term.

To be concrete, this means maximizing

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

where

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

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

Note that the correction factor w becomes more important the more heavily non-linear the structural 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 also that omitting the correction factor would imply QMLE.

The problem of evaluating the likelihood function is closely related to that of non-linear filtering. If we had the filtered density $p(\alpha|y)$, the density of y (and hence the likelihood) could be computed from the joint density via Bayes' theorem. The filtered density would be the best importance sampler.

Here, this estimation technique was implemented using Ox Version 3.32 (Doornik, 2002) and SsfPack Version 3.0 beta 2 (Koopman, Shephard, and Doornik, 1999).

A.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 linearize the observation and state equations, which delivers an approximating linear Gaussian model with the same mode as the true model.

A linear Gaussian approximation has the obvious advantage that the likelihood of the approximation L_g can easily be calculated via the Kalman filter.

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).

Unfortunately, the approximation method described above does not work well for models with a default barrier if applied naively, since 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, and above the discontinuity), we iterate until convergence. If convergence is achieved, we will have found the mode of the (truncated) posterior density of $\alpha|y$. This is chosen as the mean and mode of the approximating linear Gaussian and non-truncated 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.

A.2 Sampling from the importance density

There are several different ways of sampling from the importance density, the approach described by Durbin and Koopman (2002) was chosen here, as it is simple and efficient.

For a model specified as

$$\alpha_t = d + \alpha_{t-1} + \eta_t$$
$$\tilde{y}_t = \dot{Z}_t(\tilde{\alpha}_t)\alpha_t + \varepsilon_t \tag{27}$$

the approach works as follows (define $w = (\varepsilon_1', \eta_1', \ldots, \varepsilon_n', \eta_n'))$:

1. Draw a random vector w^+ from a joint Gaussian density, and generate α^+ and y^+ by means of recursion 27 with w replaced by w^+.

2. Compute $\hat{\alpha} = E(\alpha|y)$ and $\hat{\alpha}^+ = E(\alpha^+|y^+)$ by means of standard Kalman filtering and disturbance smoothing

3. The draw $\alpha_1 = \hat{\alpha} - \hat{\alpha}^+ + \alpha^+$ is a draws from the desired density. An antithetic draw balanced for location is $\alpha_1 = \hat{\alpha} + \hat{\alpha}^+ - \alpha^+$

B Tables and figures

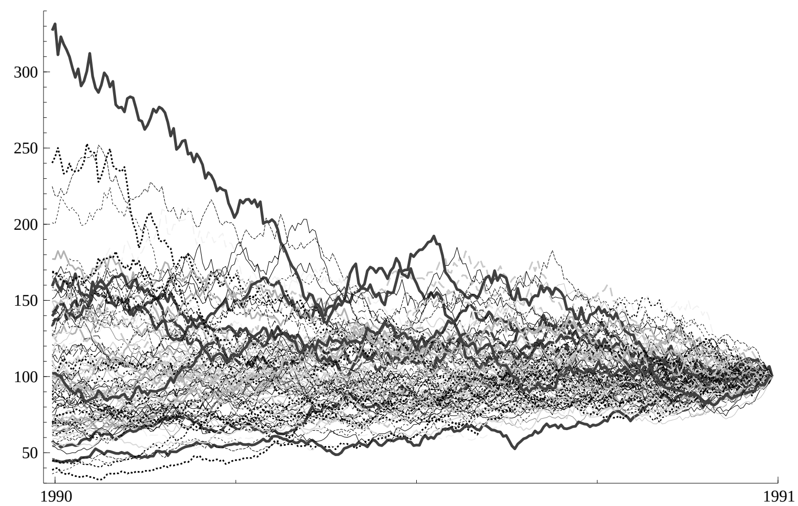


Figure 1: Alternative asset value paths that could have led to an asset value of 100 in 1991

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

Table 2: Monte Carlo results (high quality issuer)

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

Table 3: Result of Monte Carlo (low quality issuer)

	Notional ($m)	Coupon	Time to mat.*
Min.	15	5.250	0.622
1st Qu.	150	6.563	2.528
Median	250	7.200	5.125
Mean	403	7.279	7.669
3rd Qu.	500	7.875	8.401
Max.	1500	10.750	27.190

* Time to maturity in years, at prediction date

Table 4: Summary of chosen bonds

Name	Rating*	Market Cap†
Anadarko Pete Corp	BBB+	4,168
Anheuser Busch Cos Inc	A+	32,570
Archer Daniels Midland Co	A+	7,300
AT & T Corp	BBB	170,551
Bellsouth Telecommunications Inc	A+	86,586
Boeing Co	A	37,907
Charter Communications Hldgs Llc	CCC+	5,009
Coca Cola Enterprises Inc	A	8,482
COLUMBIA / Hca Healthcare Corp	BBB-	16,369
Corning Inc	BB+	29,817
Csx Corp	BBB	6,754
Dayton Hudson Corp	A+	31,576
Deere & Co	A-	10,019
Delta Air Lines Inc Del	B-	6,643
Disney Walt Co	BBB+	61,612
Dow Chem Co	A-	29,084
Du Pont E I De Nemours & Co	AA-	67,803
Emerson Elec Co	A	24,878
Enron Corp		31,084
Ford Mtr Co Del	BBB-	55,606
Gte Corp	A+	66,385
Hasbro Inc	BBB-	3,572
Hertz Corp	BBB-	2,015
Honeywell Intl Inc	A	44,739
International Business Machs Corp	A+	208,426
Lockheed Martin Corp	BBB	8,018
Lowes Cos Inc	A	21,314
Lucent Technologies Inc	B	242,171
Marriott Intl Inc	BBB+	7,476
Masco Corp	BBB+	11,166
May Dept Stores Co	BBB+	10,359
Motorola Inc	BBB	90,755
Norfolk Southn Corp	BBB	7,660
Northrop Grumman Corp	BBB	3,674
Owens Ill Inc	BB-	3,546
Penney J C Inc	BB+	5,127
Philip Morris Cos Inc		55,456
Praxair Inc	A-	7,917
Procter & Gamble Co	AA-	140,862
Ralston Purina Co		8,296
Raytheon Co	BBB-	6,188
Royal Caribbean Cruises Ltd	BB+	8,327
Tenet Healthcare Corp	B	7,492
Union Pac Corp	BBB	10,703
United Technologies Corp	A	29,921
Viacom Inc	A-	31,325
Wal Mart Stores Inc	AA	297,843
Weyerhaeuser Co	BBB	16,476
Worldcom Inc		147,416
Xerox Corp	BB-	16,066

* S&P long term domestic issuer rating according to Compustat

† Market capitalisation on 2000-01-03 in $m

Table 5: Issuers chosen for empirical analysis

	Merton	Leland	Leland+Toft
all data, no restrictions	0.4777184	0.4148535	0.2383047
all data, $\sigma_E = 0$	0.5331905	0.4475787	0.2511613
equity data, $\sigma_E \neq 0$	0.3800668	0.2773889	0.2141082
equity data, $\sigma_E = 0$	0.3987499	0.2907822	0.2216615
Calibration	0.3991525	0.2942223	0.2107958

Table 6: cross-sectional mean of estimates of σ_η

	Merton	Leland	Leland+Toft
all data, no restrictions	0.09696077	0.05155426	0.04705640
all data, $\sigma_E = 0$	0.09610188	0.05123680	0.04718293

Table 7: cross-sectional mean of estimates of σ_B

	Merton	Leland	Leland+Toft
all data, no restrictions	0.001864530	0.0006890793	0.002646254
equity data, $\sigma_E \neq 0$	0.002958444	0.0034060808	0.002932257

Table 8: cross-sectional mean of estimates of σ_E

	% pricing err	abs % pricing err	% yield err	abs % yield err	% spread err	abs % spread err
1. mean	2.04	3.53	−9.21	12.61	−37.36	49.94
1. sd	4.99	4.06	13.97	10.93	50.36	37.63
2. mean	1.80	3.42	−11.65	15.86	−44.54	55.93
2. sd	5.00	4.05	18.51	14.99	52.49	39.88
3. mean	6.12	6.83	−20.69	22.84	−77.28	81.53
3. sd	7.16	6.48	16.04	12.73	37.03	26.13
4. mean	5.74	6.45	−20.15	22.31	−75.48	79.74
4. sd	6.62	5.92	16.15	12.93	37.28	26.76
5. mean	5.88	6.58	−20.17	22.43	−74.89	80.06
5. sd	7.35	6.72	16.45	13.14	39.82	27.74

1. all data, no restrictions (SMLE)
2. all data, $\sigma_E = 0$ (ext. CoV)
3. equity prices only, no restrictions (SMLE)
4. equity prices only, $\sigma_E = 0$ (CoV)
5. equity prices only, $\sigma_E = 0$ (Calibration)

Table 9: prediction errors, Leland model

	% pricing err	abs % pricing err	% yield err	abs % yield err	% spread err	abs % spread err
1. mean	−1.29	4.21	21.49	25.68	60.28	118.43
1. sd	5.94	4.36	53.46	51.54	235.16	211.45
2. mean	−1.61	4.34	21.88	25.83	60.29	116.00
2. sd	5.88	4.25	52.67	50.82	232.72	210.14
3. mean	−0.44	5.11	19.52	26.91	53.44	118.16
3. sd	8.50	6.78	54.13	50.80	234.23	208.69
4. mean	−0.64	4.96	19.81	26.76	54.79	117.63
4. sd	8.23	6.57	54.08	50.94	234.65	209.82
5. mean	0.53	5.00	17.53	26.39	48.72	115.88
5. sd	7.91	6.11	54.49	50.72	234.91	209.58

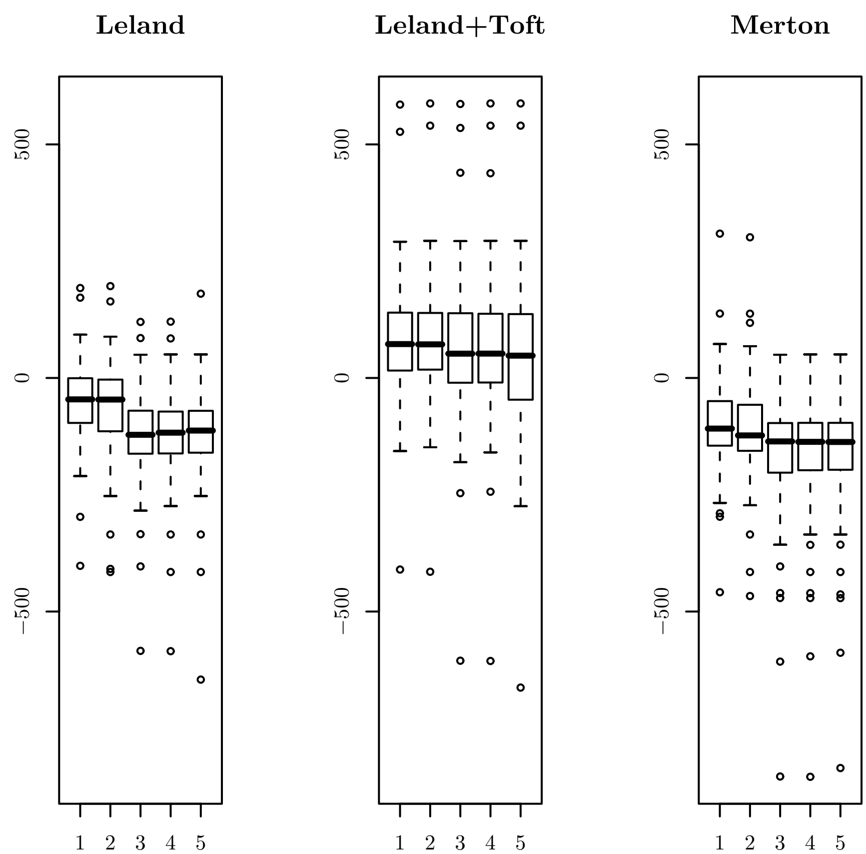
1. all data, no restrictions (SMLE)
2. all data, $\sigma_E = 0$ (ext. CoV)
3. equity prices only, no restrictions (SMLE)
4. equity prices only, $\sigma_E = 0$ (CoV)
5. equity prices only, $\sigma_E = 0$ (Calibration)

Table 10: prediction errors, Leland+Toft model

	% pricing err	abs % pricing err	% yield err	abs % yield err	% spread err	abs % spread err
1. mean	5.82	7.15	−15.35	18.73	−61.26	73.61
1. sd	8.33	7.20	15.04	10.42	52.12	31.81
2. mean	5.00	6.73	−18.73	23.08	−74.37	85.39
2. sd	8.01	6.60	18.86	13.05	51.56	29.42
3. mean	9.68	9.72	−26.13	27.29	−95.45	95.50
3. sd	10.52	10.49	16.12	14.02	16.99	16.73
4. mean	9.49	9.65	−25.88	27.27	−94.11	95.21
4. sd	10.59	10.44	16.59	14.15	22.53	17.20
5. mean	9.48	9.66	−25.86	27.28	−94.06	95.36
5. sd	10.46	10.29	16.48	13.95	22.64	16.20

1. all data, no restrictions (SMLE)
2. all data, $\sigma_E = 0$ (ext. CoV)
3. equity prices only, no restrictions (SMLE)
4. equity prices only, $\sigma_E = 0$ (CoV)
5. equity prices only, $\sigma_E = 0$ (Calibration)

Table 11: prediction errors, Merton model



1. all data, no restrictions (SMLE)
2. all data, $\sigma_E = 0$ (ext. CoV)
3. equity prices only, no restrictions (SMLE)
4. equity prices, $\sigma_E = 0$ (CoV)
5. equity prices, $\sigma_E = 0$ (Calibration)

Figure 2: Boxplots of the spread prediction errors (in bp)

CEMFI WORKING PAPERS

0401 *Andres Almazan, Javier Suarez and Sheridan Titman*: "Stakeholders, transparency and capital structure".

0402 *Antonio Diez de los Rios*: "Exchange rate regimes, globalisation and the cost of capital in emerging markets".

0403 *Juan J. Dolado and Vanessa Llorens*: "Gender wage gaps by education in Spain: Glass floors vs. glass ceilings".

0404 *Sascha O. Becker, Samuel Bentolila, Ana Fernandes and Andrea Ichino*: "Job insecurity and children's emancipation".

0405 *Claudio Michelacci and David Lopez-Salido*: "Technology shocks and job flows".

0406 *Samuel Bentolila, Claudio Michelacci and Javier Suarez*: "Social contacts and occupational choice".

0407 *David A. Marshall and Edward Simpson Prescott*: "State-contingent bank regulation with unobserved actions and unobserved characteristics".

0408 *Ana Fernandes*: "Knowledge, technology adoption and financial innovation".

0409 *Enrique Sentana, Giorgio Calzolari and Gabriele Fiorentini*: "Indirect estimation of conditionally heteroskedastic factor models".

0410 *Francisco Peñaranda and Enrique Sentana*: "Spanning tests in return and stochastic discount factor mean-variance frontiers: A unifying approach".

0411 *F. Javier Mencía and Enrique Sentana*: "Estimation and testing of dynamic models with generalised hyperbolic innovations".

0412 *Edward Simpson Prescott*: "Auditing and bank capital regulation".

0413 *Víctor Aguirregabiria and Pedro Mira*: "Sequential estimation of dynamic discrete games".

0414 *Kai-Uwe Kühn and Matilde Machado*: "Bilateral market power and vertical integration in the Spanish electricity spot market".

0415 *Guillermo Caruana, Liran Einav and Daniel Quint*: "Multilateral bargaining with concession costs".

0416 *David S. Evans and A. Jorge Padilla*: "Excessive prices: Using economics to define administrable legal rules".

0417 *David S. Evans and A. Jorge Padilla*: "Designing antitrust rules for assessing unilateral practices: A neo-Chicago approach".

0418 *Rafael Repullo*: "Policies for banking crises: A theoretical framework".

0419 *Francisco Peñaranda*: "Are vector autoregressions an accurate model for dynamic asset allocation?"

0420 *Ángel León and Diego Piñeiro*: "Valuation of a biotech company: A real options approach".

0421 *Javier Alvarez and Manuel Arellano*: "Robust likelihood estimation of dynamic panel data models".

0422 *Abel Elizalde and Rafael Repullo*: "Economic and regulatory capital. What is the difference?".

0501 C*laudio Michelacci and Vincenzo Quadrini*: "Borrowing from employees: Wage dynamics with financial constraints".

0502 *Gerard Llobet and Javier Suarez*: "Financing and the protection of innovators".

0503 *Juan-José Ganuza and José S. Penalva*: “On information and competition in private value auctions”.

0504 *Rafael Repullo*: “Liquidity, risk-taking, and the lender of last resort”.

0505 *Marta González and Josep Pijoan-Mas*: “The flat tax reform: A general equilibrium evaluation for Spain”.

0506 *Claudio Michelacci and Olmo Silva*: “Why so many local entrepreneurs?”.

0507 *Manuel Arellano and Jinyong Hahn*: “Understanding bias in nonlinear panel models: Some recent developments”.

0508 *Aleix Calveras, Juan-José Ganuza and Gerard Llobet*: “Regulation and opportunism: How much activism do we need?”.

0509 *Ángel León, Javier Mencía and Enrique Sentana*: “Parametric properties of semi-nonparametric distributions, with applications to option valuation”.

0601 *Beatriz Domínguez, Juan José Ganuza and Gerard Llobet*: “R&D in the pharmaceutical industry: a world of small innovations”.

0602 *Guillermo Caruana and Liran Einav*: “Production targets”.

0603 *Jose Ceron and Javier Suarez*: “Hot and cold housing markets: International evidence”.

0604 *Gerard Llobet and Michael Manove*: “Network size and network capture”.

0605 *Abel Elizalde*: “Credit risk models I: Default correlation in intensity models”.

0606 *Abel Elizalde*: “Credit risk models II: Structural models”.

0607 *Abel Elizalde*: “Credit risk models III: Reconciliation reduced – structural models”.

0608 *Abel Elizalde*: “Credit risk models IV: Understanding and pricing CDOs”.

0609 *Gema Zamarro*: “Accounting for heterogeneous returns in sequential schooling decisions”.

0610 *Max Bruche*: “Estimating structural models of corporate bond prices”.