STRUCTURAL MODELS OF CORPORATE BOND PRICES

by

MAX BRUCHE

London School of Economics and Political Science

Thesis submitted in partial fulfilment of the requirements of the degree of Doctor of Philosophy at the University of London

May 2005

Abstract

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. The family of models is sometimes referred to as the family of structural models of corporate bond prices. It has found applications in bond pricing and risk management, but appears to have a more mixed empirical record than the so-called reduced-form models (e.g. Duffie and Singleton 1999), and is often considered imprecise. As a consequence, it is often avoided in pricing and hedging applications.

This thesis examines three possible avenues for improving the performance of structural models:

1. Strategic interaction between debtors: Possible "Coordination failures" - races to recover value that can dismember firms - are a very important form of strategic interaction between debtors that can have a large influence on the value of debt.

2. The econometrics of structural models: The classic "calibration" methodology widely employed in the literature is an ad-hoc procedure that has severe problems from an econometric perspective. This thesis proposes a filtering-based approach instead that is demonstrably superior.

3. The non-default component of spreads: Corporate bond prices most probably do not only represent credit risk, but also other types of risk (e.g. liquidity risk). This thesis attempts to quantify and assess this non-default component.

Acknowledgements

I would like to thank my supervisor Hyun Shin, as well as Jurgen Doornik, Abel Elizalde, Siem Jan Koopman, Oliver Linton, Joel Reneby, Jean-Charles Rochet, Jukka Vesala, and Jean-Pierre Zigrand for some very helpful comments. All errors are my own.

Contents

Abstract 2

Acknowledgements 3

List of Tables 9

List of Figures 10

1 Introduction **11**

1.1 A short survey of structural models 11

1.1.1 Default before maturity 13

1.1.2 Endogenous barriers 13

1.1.3 Coupons 14

1.1.4 Stochastic risk-free rates 14

1.1.5 Reduced information sets 15

1.2 Suitable structural models for bond pricing 16

1.3 The empirical track record of structural models 17

1.4 Improving structural models 19

1.5 A research agenda 20

2 Corporate bond prices and coordination failure **25**

2.1 Introduction 25

2.2 Coordination failure in the context of strong legal frameworks: the US 27
2.2.1 Preference 28
2.2.2 Secured debt 30
2.2.3 Bank loans and covenants 31
2.2.4 Discussion of the incidence of coordination failure 31
2.3 The model 32
2.3.1 The setup 32
2.3.2 The solution 38
2.3.3 Pricing 41
2.4 Discussion of results 46
2.4.1 Comparisons to Merton 1974 and Black-Cox 1976 46
2.4.2 Comparison to the no coordination failure case 47
2.4.3 Comparison to Morris-Shin 2004 48
2.4.4 Thoughts on an empirical test 48
2.5 Concluding Remarks 49

3 Estimating structural models of corporate bond prices 51

3.1 Introduction 51
3.1.1 The inherent problem of estimation 53
3.1.2 Other issues for estimation 54
3.1.3 Estimation approaches 55
3.2 A general setup for estimation 60
3.2.1 The state equation 61
3.2.2 The observation equation 61
3.2.3 The system 62

3.2.4 Restrictions determine feasible estimation procedures 63
3.3 A general procedure for evaluating the likelihood function 65
3.3.1 Approximating the model 68
3.3.2 Conditioning on no default 69
3.3.3 Inference and tests 70
3.3.4 Advantages of the approach 71
3.4 Applications 72
3.4.1 The theoretical model 72
3.4.2 Monte Carlo experiments 74
3.4.3 Application to real data 81
3.5 Concluding Remarks 88

4 The non-default components of corporate yield spreads [1] 90

4.1 Introduction 90
4.2 The model 94
4.2.1 The default component 94
4.2.2 The non-default component 95
4.3 Estimation 95
4.3.1 Setup 95
4.3.2 Evaluating the likelihood function 97
4.4 Data 99
4.4.1 Bond price data 99
4.4.2 Equity prices and accounting data 100
4.4.3 Non-default proxies 100

[1] This chapter is based on joint work with Joel Reneby.

4.4.4 Assumed and estimated parameter values 100

4.5 Results 101

4.5.1 Out-of-sample prediction performance 101

4.5.2 Estimated parameters 103

4.5.3 Cross-section properties of estimated parameters 105

4.6 Concluding remarks 108

5 Conclusion **110**

A Pricing coordination failure **114**

A.1 Solution of the discrete time model 114

A.1.1 Basic procedure 114

A.1.2 Information 115

A.1.3 Posteriors 115

A.1.4 Critical value of $V_{t+\Delta}$ for which the firm fails 116

A.1.5 Utility 117

A.1.6 Critical value of ρ 118

A.1.7 Equilibrium forced reorganisation 120

A.1.8 Uniqueness 121

A.2 Uncertainty in the limit 122

A.3 Pricing the bond 124

A.3.1 The default process 124

A.3.2 The first passage density 125

A.3.3 Perpetual dollar-in-default claim 125

A.3.4 Survival probabilities 126

A.3.5 The finite-maturity dollar-in-default claim 126

A.3.6 The down-and-out binary option 127
A.3.7 Probability of survival and being in-the-money 127
A.3.8 The price of a down-and-out call 127
A.3.9 Pricing under assumption 2.A: (Bond represents entire debt) . 128
A.3.10 Pricing under assumption 2.B: (Bond represents negligible proportion of debt) 128

B Simulated maximum likelihood estimation of structural models 129

B.1 Sampling from the importance density 129
B.2 Description of selected issuers and bonds 130

C The non-default component of corporate yield spreads 133

C.1 Cross-sectional regressions 133

List of Tables

2.1 Payoffs to creditors in the discrete time game 35

3.1 Monte Carlo results (high quality issuer) 78

3.2 Result of Monte Carlo (low quality issuer) 79

3.3 Prediction errors on real data 86

4.1 Mean parameter estimates (and t-stats) for the different quartiles . . 104

B.1 Summary of chosen bonds 130

B.2 Issuers chosen for empirical analysis 131

B.3 Bonds chosen for empirical analysis 132

C.1 Regression of Constant on cross-sectional variables 133

C.2 Regression of ONOFF2 on cross-sectional variables 134

C.3 Regression of ONOFF30 on cross-sectional variables 134

C.4 Regression of MMMF on cross-sectional variables 135

List of Figures

2.1 Timing assumptions 37

3.1 Alternative asset value paths that could have led to an asset value of 100 in 1991 75

3.2 SMLE and ERMLE actual versus predicted quantiles (high quality issuer) 80

3.3 SMLE and ERMLE actual versus predicted quantiles (high quality issuer) 81

Chapter 1

Introduction

In 1974, Merton wrote a seminal paper (Merton 1974) that explained how the then recently presented Black-Scholes model (Black and Scholes 1973) could be applied to the pricing of corporate debt. Viewing equity and debt as options on the underlying assets of the firm allows the use of the Black-Scholes option pricing calculus for valuation. Many extensions to this model followed. These models, which price bonds with reference to the *structure* of a firm's liabilities are sometimes referred to as structural models of corporate bond prices.

1.1 A short survey of structural models

In the Merton (1974) model, the assumptions are parallel to those in the (Black and Scholes 1973) paper: The underlying (in this case the value of assets) is assumed to follow a geometric Brownian motion:

$$dV = \mu V dt + \sigma V dW \tag{1.1}$$

This asset value is assumed to be traded. There exist two contingent claims on this asset value: The firms' debt is assumed to be a zero coupon bond with maturity T and a face value D. At the maturity of the bond, the firm is wound down. if $V_T < D$, all assets are given to bondholders, if $V_T > D$, the bondholders receive D, and the equity holders (as the residual claimants) receive the rest. If the other assumptions of Black and Scholes hold, equity is essentially a call option on the assets with payoff $\max(0, V_T - D)$, and the bond is essentially the value of assets minus that call (or, equivalently, risk-free debt minus a put) with payoff $\min(V_T, D)$. Hence the equity pricing function is simply

$$F_E(t, V, D, \sigma, T-t) = VN(d_1) - De^{-r(T-t)}N(d_2) \tag{1.2}$$

where

$$\begin{aligned} d_1 &= \frac{log(V/D) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} && (1.3) \\ d_2 &= d_1 = \sigma\sqrt{T-t} && (1.4) \end{aligned}$$

and $N(\cdot)$ is the cumulative normal distribution function.

The price of debt is given by

$$F_B(t, V, D, \sigma, T-t) = V - F_E(t, V, D, \sigma, T-t) \tag{1.5}$$

Of course, in order to apply the Black Scholes calculus, many simplifying assumptions are made. Many extensions have been proposed that deal with relaxing some of the assumptions.

1.1.1 Default before maturity

The model has been extended to take into account e.g. sub-ordination arrangements, indenture provisions and default before maturity (Black and Cox 1976). In the original Merton model, since the aggregate debt of the firm is seen as a zero-coupon bond, default can only occur when a payment is supposed to be made: at the maturity of the zero-coupon bond. One important step towards making the model more realistic is allowing for default before the maturity of aggregate debt. Black and Cox show how to price a bond when there is a net worth covenant which is violated when the asset value falls below a specified level. Since working out default probabilities in this case is dual to calculating first passage times, models incorporating this feature are sometimes called "First Passage Models". In this context, bonds are viewed as a type of barrier option, and the same pricing principles apply. Barriers can be fixed, exponentially growing or stochastic and the corresponding prices can easily be calculated - as long as a transformation of the asset value and the barrier process is a Brownian motion.

1.1.2 Endogenous barriers

In the context of a First Passage Model, the barrier can also be made endogenous. For example, Leland (1994) assumes the firm chooses the barrier to maximise firm value, and calculates the prices of debt and equity in a time-homogenous setting. This was extended by Leland and Toft (1996) to the case where the firm also chooses the maturity of its debt.

An endogenous barrier does not have to be an optimal choice, it can also be the outcome of a game. Anderson and Sundaresan (1996) and Mella-Barral and Perraudin

(1997) for example look at a situation in which equity holders strategically choose when to service a firm's debt. Given bankruptcy costs, it is not necessarily optimal for holders of debt to foreclose as soon coupon payments are (partially) missed. This can be exploited by holders of equity. Anderson and Sundaresan present a binomial model, and Mella-Barral and Perraudin provide closed-form solutions based on a time-homogeneous continuous time model.

1.1.3 Coupons

An extension of the basic model to bonds that pay a continuous coupon is relatively straightforward. However, most bonds pay coupons at discrete times. Black and Scholes noted in passing that valuing coupon bonds is similar to valuing compound options. At a coupon date, equity holders have the option to obtain the future value of dividend streams by paying the coupon. Geske (1977) (cf. also Geske 1979) examined this problem, and showed how the resulting price can be calculated recursively, and suggested a procedure for reducing the dimensionality of the resulting multidimensional integral. It is clear, though, that this method of valuing coupon bonds can become intractable very quickly as the number of coupons rises.

1.1.4 Stochastic risk-free rates

It has also been argued that the assumption of constant flat interest rates as initially made in the Black-Scholes model is especially inappropriate when valuing bonds, and the model has been extended to take into account stochastic interest rates. Kim, Ramaswamy, and Sundaresan (1993) extend a structural model by introducing a mean reverting square root short rate process, whereas Longstaff and Schwartz (1995) and

Briys and de Varenne (1997) use a mean reverting Gaussian or Vasicek process for the short rate. Of the two, Briys and de Varenne (1997) specify a more general version where parameters are time dependent. Kim, Ramaswamy, and Sundaresan (1993) report that the short rate has very little influence on credit spreads, whereas Longstaff and Schwartz (1995) conclude that the 30 year treasury yield and the credit spread are negatively correlated, and that this correlation can be economically and statistically significant. Of course, the level of interest rates could also have an effect on the level of the default barrier, providing a further refinement to the model (Collin-Dufresne and Goldstein 2001b).

1.1.5 Reduced information sets

Lastly, for all structural models in which the underlying asset value as well as the default point is perfectly observed, defaults will be predictable in the sense that the instantaneous default probability over a time interval that shrinks to zero also tends to zero. In the absence of components in bond returns that are not related to default (such as a liquidity risk premia, which are typically outside the scope of structural models), this implies that the credit spreads predicted by these structural models also tend to zero. This does not seem to be the case empirically, so several ways of increasing predicted spreads at the short end of the credit spread term structure have been devised. Duffie and Lando (2001) propose assuming that the asset value process is not perfectly observed by bondholders - this increases the default probabilities (conditioned on the information set of the bondholders) at the short end, and hence raises spreads. Alternatively, the level of the default barrier can be unknown (Finger et al. 2002, Giesecke 2004, Giesecke and Goldberg 2004). To produce a similar effect, one can also incorporate jumps in the asset value process (Zhou 2001, Hilberink and Rogers 2002).

1.2 Suitable structural models for bond pricing

Many of the aforementioned models give important qualitative insights. However, the models are typically not easily and consistently implemented empirically. Ericsson and Reneby (2002b) suggest combining several elements to arrive at a very tractable structural model. Given that equity typically does not expire, it makes sense to treat it as a perpetuity. An alternative to modelling the typically very complicated debt structure of a firm, including differences of seniority of different layers of debt, as well as differences in maturities, and covenants, embedded optionality and complicated repayment schedules is to choose a simple model for aggregate debt that treats it as a perpetuity, and treat any one claim that is to be priced as a negligible proportion of aggregate debt. This avoids the compound-optionality problems in pricing (Geske 1977).

These capital structure assumptions greatly simplify pricing, and are also more realistic than assuming that the bond being priced represents the entire debt of a firm as in the original Merton (1974) model. For example, this assumption is particularly troublesome when attempting to price several bonds issued by the same firm.

This kind of model was first proposed by Ericsson and Reneby (2002b), who derive an endogenous default barrier in a manner similar to Leland (1994). Since the exact choice of barrier is often not important for prices, as argued by e.g. Longstaff and Schwartz (1995), their assumptions make for an empirically very tractable version of the first passage model that is likely to produce very similar prices to other first passage models.

1.3 The empirical track record of structural models

Versions of structural models have found important applications in risk management, (e.g. the KMV EDF$^{\mathrm{TM}}$ methodology Crosbie and Bohn 2002). KMV use distances-to-default calculated from structural models (differences between an estimated asset value and a default point, standardised by asset value volatility), and compares these to actual default frequencies. This allows predicting probabilities of default for a given distance-to-default.

There have been attempts to use structural models in pricing and hedging applications, such as the CreditGrades$^{\mathrm{TM}}$ model (Finger et al. 2002), but these attempts usually fare quite badly and produce very large prediction errors and biases (in the author's evaluation this is certainly the case for CreditGrades$^{\mathrm{TM}}$). As a consequence, for the pricing of credit derivatives or bonds, structural models do not seem to be the preferred choice.

Although Anderson and Sundaresan (2000) show that the variation in corporate bond yields seems to be well partly explained by the variables suggested by structural models, and that there is some support for the strategic debt service models over the alternatives, it is not clear whether structural models are appropriate for pricing.

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. Alternatively, for realistic volatilities, the model appears to overpredict bond prices and to underpredict spreads (Jones, Mason, and Rosenfeld 1984).

In their comparison of the Merton (1974), Geske (1977), Longstaff and Schwartz (1995), Leland and Toft (1996) and Collin-Dufresne and Goldstein (2001b) models, Eom, Helwege, and Huang (2004) conclude that only the basic Merton (1974) model underpredicts spreads, whereas most other models overpredict spreads for risky bonds, and underpredict spreads for safe bonds. Crucially, they state that the spread prediction errors are substantial in all cases, presumably precluding their use in pricing applications.

The term "structural models" is often used in contrast to so-called *reduced form* models, pioneered by Jarrow and Turnbull (1995) that simply posit a process describing the evolution of the (instantaneous) probability of default. These models are more tractable empirically speaking, and seem to match observed bond spreads quite well when implemented (Duffee 1999), possibly because they impose less structure, and because possibly factors like liquidity risk or tax rates can be 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.

However, in many applications where the focus is on both credit 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 bank 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 might do rather well, although they do not discuss how to actually implement a working hedge.

In addition, due to the link that structural models provide between equity and credit, they allow estimation of models from equity prices as well as from bond prices. This is especially important if hedging parameters or predicted prices are desired, but there are insufficient bonds or credit derivatives to estimate reduced-form models.

1.4 Improving structural models

The underprediction of spreads by structural models has produced three distinct reactions attempting to explain the failure of structural models to produce reasonable spreads and prices for reasonable volatilities.

1. Mella-Barral and Perraudin (1997) for example argue that strategic debt service will (ceteris paribus) reduce the value of debt (including high-quality debt), and therefore has the potential for explaining some of the apparent problems of the original model. In essence, any feature that will lower the value of payoffs to bondholders vis-a-vis equity holders or increase the probability of default will mean that a bond being priced in terms of the corresponding equity price will have a higher spread. Other models that have this effect are e.g. the ones proposed by Duffie and Lando (2001) or Zhou (2001). This approach would suggest that the underprediction of the spread is a result of the assumed payoffs being too high, or the calculated default probabilities being to low.

2. Another approach is to argue that the implementation of structural models (e.g. the canonical implementation of Jones, Mason, and Rosenfeld 1984), has so far

been flawed, and that better estimation methods will improve the predicted bond prices and spreads (Duan, Gauthier, Simonato, and Zaanoun 2003, Ericsson and Reneby 2005).

3. Another idea is that maybe structural models produce reasonable predictions for the default-related components of corporate bond spreads, but that there are components of the spread that are not related to default (Delianedis and Geske 2001). These might be related to taxes or liquidity risk, for example.

1.5 A research agenda

In this vein, this thesis attempts to contribute to our understanding of structural models, re-iterate that they are useful in understanding important issues, and improve their specification and estimation to the point where they can be used for applications like the one described above. In particular, it addresses the following issues (in this order):

1. **Strategic interaction between debtors and default**
   Modelling the incidence of default, bankruptcy and liquidation is of course very important to the resulting bond prices. To summarise the above discussion, the following ways of describing the incidence of default and/or bankruptcy and liquidation have been proposed:

   (a) Default can occur only at the time when aggregate debt matures, provided assets are insufficient to pay off bondholders (Merton 1974). This model blurs the boundaries between default, bankruptcy and liquidation by assuming that at the maturity of debt and equity, the firm is liquidated in any case.

(b) A violation of the debt contract occurs when assets fall below a level specified by a net worth covenant (Black and Cox 1976). At this point, presumably debt-holders can attempt to enforce their claims, and the firm will enter bankruptcy. At this point, being controlled by debt-holders, the firm will be liquidated. An alternative to motivating a default barrier in this context is to assume that cash flow is proportional to assets, and assume that the amount needed to service aggregate debt is constant (e.g. proportional to the notional value of aggregate debt). Also, assume that the firm will not be able to issue equity to stave off default. Then default on payments will occur when the asset value falls below the level at which the generated cash flow is insufficient to cover payments. This will again allow debt-holders to attempt to enforce their claims with the same result as above.

(c) The firm can optimally choose a level of assets, such that when that level is reached, it voluntarily goes into bankruptcy. Optimality here could be defined e.g. in terms of maximising firm value (Leland 1994). This is an alternative way of motivating a default barrier.

(d) Equity holders strategically choose whether or not to service debt (Anderson and Sundaresan 1996, Mella-Barral and Perraudin 1997). Given the level of coupons the holders of equity choose to pay, debt-holders can decide whether to enforce claims and to push the firm into bankruptcy or not. Equity holders can pay less than the promised coupon in equilibrium if they have a sufficient amount of bargaining power.

Of these alternatives, the last one can lead to (ceteris paribus) lower bond prices (and hence higher spreads) by lowering the payoffs of bondholders, possibly improving the empirical fit of structural models.

If creditors are dispersed and are unsure about the actions of other creditors, they might not always agree on when to enforce claims once they are given the option to to so, however.

This possibility is explored in chapter 2. Suppose that either a net-worth covenant exists as in Black and Cox (1976) or equivalently, a cash flow covenant exists (and cash flows are proportional to the value of assets as above). Furthermore assume that due to cross-default clauses, if a covenant on a small part of aggregate debt is violated, technically the firm is considered to be in violation of all debt contracts. Often in a situation like this, the firm is not immediately liquidated. Typically, firms keep operating for a period until a sufficiently large number of creditor actually decides to enforce claims, at which point the firm is forced to declare bankruptcy. In these kind of situations, races to grab assets and failures to coordinate between different creditors can exist and are indeed likely. These will have an effect on the point at which bankruptcy (as opposed to default or a covenant violation) actually occurs, and hence will affect the price of debt (Morris and Shin 2004a).

It is possible to derive a bond price for a firm in this kind of situation in the context of a structural model, as shown in chapter 2. Comparisons to other structural models, in particular the Merton and Black Cox models can be made. Since coordination failures mean that bankruptcy is not triggered at the socially optimal point, bond prices will be lower than in the absence of coordination failures. If bankruptcy costs are increased by legal wrangling over claims by dispersed creditors, then clearly this will reinforce the conclusion of a lower price.

2. **The econometrics of structural models**

   In practical applications, it is of course not only important to have good theoretical models, but that the implementation is done in an efficient manner. Some shortcomings in existing methods for estimating structural models are apparent. Although the proposed methods often serve well to obtain a reasonable first approximation, they often lack an econometric basis, i.e. there is no reason to believe that they are particularly efficient. Actually, most are demonstrably inefficient to a degree that it is not surprising that structural models are not chosen in the types of hedging applications described above. A filtering-based approach that can provide more precise estimates, more reliable predicted densities and can include more types of information (such as prices of several assets, accounting information, etc), as well as giving a natural framework for easily dealing with non-default related components in asset prices is proposed in chapter 3.

3. **The non-default component of spreads**

   Lastly, as argued above, the underprediction of corporate bond spreads (especially for short maturity bonds), which seems to occur for many models even with efficient estimation methods could be a result of non-default related risk factors that have an impact on corporate bond prices (Delianedis and Geske 2001, Longstaff, Mithal, and Neis 2004).

   This is an alternative explanation of the underprediction - it could either be a result of existing models underpredicting the default component of the spread, or it could be a result of models correctly predicting the default component, but ignoring the fact that there are other (e.g. liquidity) components in spreads.

A very simple empirical approach for integrating a non-default related component is proposed in chapter 4. It is shown that including a simple linear non-default component improves out-of-sample predictions substantially, that the non-default component is related to various variables associated with market-wide liquidity, and that the relationship of these variables to the non-default component appears to vary cross-sectionally.

All these points can be seen as an attempt to improve the empirical performance of structural models, and are therefore a step in the direction of making structural models more appropriate in hedging and pricing applications.

A summary of conclusions and an outlook for future research is given in chapter 5.

Chapter 2

Corporate bond prices and coordination failure

2.1 Introduction

Morris and Shin (2004a) (cf. also Morris and Shin 2000) argue that coordination failure among creditors can have an effect on the price of debt. The problem of coordination failure is akin to the problem faced by depositors of a bank which is vulnerable to a run. Even if it is not efficient to foreclose on a loan or grab assets, e.g. when the debtor is fundamentally viable, fear that other creditors may foreclose or grab assets can lead to preemptive action and inefficient foreclosure or to a viable firm being pushed into bankruptcy.

In general, coordination failures can arise among creditors in a context where it is possible for individual creditors to improve their position vis-a-vis the firm at the cost of other creditors (i.e. in a situation with strategic complementarities between the different lenders). This is the case for example if creditors can foreclose individually, leaving other creditors exposed to a firm with lower liquidity, or for instance if some creditors will be able to grab assets in the case of financial difficulties at the expense of other creditors.

Of course, bankruptcy codes will generally try to prevent coordination failures (cf. e.g. Baird and Jackson 1990, Jackson 1986). There are many countries in which bankruptcy codes are not developed or not legally enforced, such that coordination failures arise naturally. But even in countries with developed and enforced bankruptcy codes and systems, coordination failures can occur. Since the US bankruptcy system is an important example of a country with a developed and enforced bankruptcy system, the possible incidence of coordination failures in the US will be discussed below.

Morris and Shin (2004a) model coordination failure as a function of a fundamental variable in a simple static model. The fundamental variable could easily be interpreted as the asset value of a firm. This naturally suggests using a structural model (as pioneered by Merton 1974) to price the debt as derivatives on the underlying asset value, in order to allow comparisons to other structural models.

As mentioned in chapter 1, game theoretic considerations have been integrated into corporate bond pricing in the context of a situation in which equity holders strategically choose when to service a firm's debt (Anderson and Sundaresan 1996, Mella-Barral and Perraudin 1997). In these models, holders of debt act in a coordinated manner to decide whether or not to push the firm into bankruptcy, given the offer of equity holders.

The model presented here examines a different situation. It focuses on creditors, and assumes that they are dispersed, and are uncertain about the actions of other debtholders. It examines the resulting strategic situation and its implications for bond pricing in a continuous time structural model.

A simple discrete time model will be set up that exhibits coordination failures. The model will be solved, and the continuous time limit will be taken. Pricing will be explored under two different assumptions about the capital structure: Firstly that the bond represents the entire debt of the firm (this is the original Merton assumption), and under the assumption that the bond represents a negligible proportion of overall debt (this yields an empirically more tractable model as discussed in chapter 1).

It will be shown that under the first assumption, the model has the Merton (1974) model and the Black and Cox (1976) model as limiting cases, and that the model can produce lower bond prices[1].

2.2 Coordination failure in the context of strong legal frameworks: the US

On a casual examination, it would appear that in the US, for holders of unsecured debt, it is not in general possible to seize assets individually once formal bankruptcy proceedings (like chapter 7 or chapter 11) have commenced, since these proceedings are associated with an automatic stay. However, outside of formal bankruptcy, state laws govern what creditors can or cannot do. As opposed to formal bankruptcy, these often allow asset grabs, which can push firms into formal bankruptcy. Co-ordination failures can therefore occur before bankruptcy. There is a so-called preference system in place that is meant to prevent these situations, but it is not clear at all that it is effective at preventing coordination failures. On top of this, coordination failures can

[1]As mentioned in chapter 1, this is important in the context of empirical research (e.g. Jones, Mason, and Rosenfeld 1984) that seems to indicate that simple structural models tend to overprice bonds.

occur within formal bankruptcy due to secured creditors. They can also occur between banks that lend as a syndicate. It is clear that in many situations, coordination failure will affect the price of debt. These issues are discussed below in more detail.

2.2.1 Preference

The intersection of state laws and formal bankruptcy is governed by section 547 of the Bankruptcy Code: If a creditor manages to grab assets 90 days before the commencement of formal bankruptcy proceedings, the actions of the creditor fall under state law, and it will be almost impossible for other creditors to make claims against the recovered money. Assets that are grabbed within 90 days of the commencement of formal bankruptcy proceedings are governed by section 547 of the Bankruptcy Code. This is e.g. described by Neustadter (2002):

> In the state system, with some exceptions, the creditor first to obtain an interest in the debtor's property (by obtaining payment or by obtaining a judicial, statutory or consensual lien) is entitled to satisfy as much of its claim as the value of the property permits before any other creditor may satisfy a different claim from the same property. To the winner of the race (to the holder of a lien) go the spoils.
>
> In contrast, the bankruptcy system treats unsecured creditors collectively and, with some exceptions, equally, without regard to the position that the unsecured creditor held in the race in the state system. The two systems are, therefore, fundamentally inconsistent with one another.
>
> Bankr. Code 547 stands at the intersection of the two systems. It provides, among other things, that an unsecured creditor who had won

> a race to an interest in the debtor's property using the state remedies system within 90 days of the filing of the bankruptcy petition may have to forfeit its winnings (without compensation for any expenses it may have incurred in winning the race) for the benefit of all unsecured creditors. The section therefore prevents certain creditors from being preferred over others (hence, section 547 of the Bankruptcy Code is titled "Preferences").

Preferential transfers are transfers of e.g. cash or right to collateral (a lien) to a subgroup of creditors in the 90 days prior to the commencement of formal bankruptcy. The trustee can demand that these are returned to the bankruptcy pool, to be shared between all creditors. These demands are frequently made, but just as frequently challenged, especially in business cases. There are doubts as to whether the preference law that stipulates that such demands may be made prevents asset grabs and hence coordination failures.

In 1997, the American Bankruptcy Institute conducted a number of surveys of preference law practitioners as well as credit providers to evaluate the effectiveness of preference law in achieving its two purposes:

1. to "facilitate the prime bankruptcy policy of equality of distribution among creditors of the debtor".

2. to discourage creditors "from racing to the courthouse to dismember the debtor during his slide into bankruptcy".

See H.R. Rep. No. 595, 95th Cong., 1st Sess. 177-78 (1977). Of course, both of these questions are directly related to the possible incidence of coordination failure.

Out of the credit providers, more than half (52.1%) answered that equitable redistribution occurred not at all or only a little, and 84.3% said that the preference laws penalise providers who attempt to work with a debtor (presumably as opposed to attempting to recover the maximum amount in the shortest possible time). Almost three-fourths (73.7 %) said that the objective of "helping a debtor deal with its creditors prior to a bankruptcy filing" was either not met at all or only a little. Answers from law practitioners were only slightly less pessimistic.

In the light of this survey, it is not clear that preference law ensures that some creditors are not preferred over others, and that creditors will not race to grab assets to dismember a financially distressed debtor before formal bankruptcy.

On top of this, in situations were assets are oversecured, holders of the lien may assume more debt that can be netted against the collateral. Since this does not constitute a transfer in the legal sense (if the loan was oversecured to begin with), this cannot legally constitute a preferential transfer and can therefore not even be challenged under section 547. It is in fact a perfectly legal way to grab assets.

2.2.2 Secured debt

Note that even once the firm is in formal bankruptcy, secured creditors can of course apply for a relief from the stay, and can attempt to seize collateral or to force the debtor to make payments as a condition to continuing the stay. Otherwise, their collateral would essentially be meaningless. This can affect other creditors in that crucial assets can be taken from the firm, or it can affect the debtors ability to service other debt.

2.2.3 Bank loans and covenants

Lastly, most loans that banks make to firms are governed by very restrictive covenants that give banks the ability to stop lending or foreclose in typically a very wide range of circumstances. For reasons relating to informational asymmetries, banks typically choose covenants over interest rates as a way of controlling their risky loans. From a practical perspective, covenants allow banks to foreclose at the slightest sign of financial trouble. Almost all companies will depend on bank credit of some form in a crucial way, even if they depend on markets to raise the bulk of their funds. Firms that raise short term cash in the commercial paper market, for instance, will almost always have commercial paper backup lines of credit, to have an emergency supply of liquidity in situations where they cannot raise money on the commercial paper markets (i.e. the kind of situation we are interested in), and indeed banks are uniquely positioned to lend in exactly these kind of situations (Gatev and Strahan 2003). This refusing to roll over a loan does not constitute an asset grab, it cannot be challenged in formal bankruptcy, but can produce coordination failures, in the sense that the refusal of one bank to roll over can spell the end of the firm, and can adversely affect the value of loans held by other banks/ creditors.

2.2.4 Discussion of the incidence of coordination failure

To conclude this section, it appears that coordination failures are possible even in the US, which is often given as an example of a country with a strong legal framework with respect to bankruptcy and creditor protection. Given the realities of the preference system, they can occur if assets are grabbed before formal bankruptcy, and can indeed trigger bankruptcy. They can also occur in the context of secured debt (even during formal bankruptcy), and between several banks that are syndicated lenders to a firm.

2.3 The model

2.3.1 The setup

As in the Merton (1974) model, the bond will be priced as a function of the asset value process. First, a discrete time game will be set up, where in every period in which the firm is vulnerable, creditors have to decide whether or not to attempt to grab assets. It will then be possible to derive a critical point for the asset value for which the firm will just be forced into bankruptcy (the 'trigger point'). The asset value changes between periods, such that when the continuous time limit is taken, the process will turn out to be a geometric Brownian motion. This will then allow us to price the bond as a combination of barrier options on the asset value using standard techniques, where the barrier is given by the trigger point.

Capital structure

As discussed in chapter 1, contrary to the original Merton assumption, a bond typically represents only a very small part of the total liabilities of a firm, and capital structures are very complex. In a model, there is a limit to the complexity that can be represented successfully. Merton originally assumed that the (zero coupon) bond was the only debt of the firm. This makes pricing easy, but makes introducing coupons rather hard, raises issues such as what the maturity of equity is, etc. Since many models make this assumption, it could be considered a starting point to facilitate comparison. However, another assumption that allows for simple pricing but is maybe more realistic is to assume that the bond represents a negligible amount of total debt. In the following, prices will be derived for both assumptions:

Assumption 2.A. *The bond represents the entire debt of the firm. At maturity of the bond, the firm will be liquidated (if it has not been liquidated early due to bankruptcy).*

This is keeping with the original Merton assumption. The bond can be priced assuming that it represents all of the firms debt, and that in fact the bondholders as the only creditors are prone to coordination failure. This is an unrealistic assumption, but the resulting model facilitates comparisons with various earlier models. The amount recovered will be endogenously derived.

Assumption 2.B. *The bond represents a negligible proportion of aggregate debt. The firm will never be liquidated, unless it goes into bankruptcy.*

Assume that default is determined by the interaction of all creditors who together have a claim to the *aggregate debt* of the firm, and assume that the bond for which the price will be derived here constitutes a negligible amount of this aggregate debt. This means that when default is discussed, particular features of the bond such as maturity and coupon dates can be ignored. Since the exact structure of aggregate debt is not interesting (as long as it can produce coordination failure) in this instance, assume that it is uniform, pays a continuous coupon and does not mature. Assume that bondholders recover a fraction R of the principal of the bond in case of default.

The actions of agents

Suppose holders of the debt of the firm can grab assets or foreclose individually as described above, such that coordination failures can occur. To be precise, assume that in every period in which the assets of a firm are below a certain level (which might mean that the firm cannot make certain payments, or that covenants are breached), the firm is vulnerable and creditors can send out their lawyers to attempt to grab assets. This action has a certain cost associated with it. If enough creditors manage to grab assets, the firm is pushed into reorganisation (suppose there is a minimal level of assets for the firm to be able to continue operations). If the fraction of creditors attempting to grab assets is too low, the firm will not be pushed into reorganisation. Suppose that in these cases, grabbing the assets is pointless, as they are worth less than the creditor's original share of the debt of the firm, and creditors will give the assets back to the firm. Lawyers will still have to be paid.

Payoffs

Formally, let D denote the face value of aggregate debt, and let L be the critical level of assets that defines whether or not a firm is vulnerable (It would be natural to assume that $L < D$). Later, P will denote the principal/ face value of a bond (in case this is not equal to the face value of aggregate debt). If the asset value V falls below L, the firm is vulnerable. L might represent a critical covenant, or the face value of liabilities. It might, for example, be a net worth covenant, i.e. a covenant specifying that the asset value must stay above a certain value in order for the aggregate debt contract not to be violated. If the cash flow produced by the firm is proportional to the value of assets, then it could also be a cash flow covenant, specifying that the firm must have a minimum cash flow for the aggregate debt contract not to be violated. If aggregate debt pays a coupon, the firm might simply not be able to pay this coupon if the cash-flow is too low, meaning that it defaults on some payments. In all these cases, creditors would legally be in a position to attempt to secure liens on assets, or grab assets.

Note that even if aggregate debt consisted of many different individual claims, if all components of aggregate debt had cross-default clauses (as is commonly the case), then it does not matter which particular contract is violated - via the cross-provisions, default on one contract would mean a violation of all contracts. To keep the model simple here, though, aggregate debt is modelled as being homogenous.

Reorganisation will take place immediately before a time t only when the fraction of creditors who decide to grab assets l is larger than or equal to $\frac{V_t}{L}$, such that it will be impossible for the firm to be forced into bankruptcy when $V_t > L$ (the firm is not vulnerable). As V_t falls, it becomes easier for creditors to push the firm into bankruptcy. Attempting to grab assets produces an immediate cost KV_t (proportional to the value of the assets). If the firm is pushed into reorganisation, an agent that has grabbed assets receives her share of the asset value V_t, whereas agents that have not participated receive 0. If the firm is not reorganised, agents that attempted to grab assets still incur the costs but both types of agents still hold their share of the debt.

Table 2.1 illustrates the 'per unit of principal' payoffs that creditors need to take into account when making the decision whether to continue or stop lending.

	firm reorganised	firm not reorganised
grab assets	$(1-K)\,V_t$	$-KV_t$
do not grab assets	0	0

Table 2.1: Payoffs to creditors in the discrete time game

Information content of prices

A necessary ingredient for coordination failure to arise is uncertainty about the actions of other agents. Without common knowledge of the fundamentals (the asset value in our case) of an issuer, agents will not be completely sure of how other agents will act. Suppose there is private as well as public information, then provided that private information is sufficiently precise in relation to public information, i.e. there is sufficient uncertainty about the actions of others, this will create coordination failure.

In a comment on Morris and Shin's (2000) paper, Atkeson (2000) doubts that the coordination failure idea is applicable to pricing debt. He argues that if agents can see prices, there will be no coordination failure, because all information will be revealed in the prices - there is no role for private information, and hence uncertainty about the information of other agents. In the model presented here, there is no trade at the time private signals are received and agents act, hence the private information is not revealed through trading. Much hinges on the exact timing assumptions.

Suppose that agents have to make a decision as to whether or not to grab assets after they have received a signal, but before the signals are revealed to all. Subsequently, signals are revealed to all, then trading occurs and information is integrated into prices. Then there might still be coordination failure, because the private information has not been made public at the time when the agents need to act.

Timing

Time increments are of size Δ. At time t, identical agents (the creditors) know the asset value of this period, V_t. Later, the number of agents will tend to infinity, and they will subsequently be indexed by the unit interval. Relative changes in the asset value are normally distributed. The bonds trade at a price B_t which incorporates the information V_t. Let q denote a time increment that is smaller than Δ $(0 < q < \Delta)$. At $t + q$, agents receive a signal X_i about the increase in the asset value - subscript i indexes the different agents, where the time subscript is omitted to simplify notation.

They form a posterior given their information. Given their posterior, they make a decision as to whether or not to grab assets.

After it has been determined that the firm will not fail in this period, we proceed to the next period: Signals are revealed, the asset value is revealed and the price $B_{t+\Delta}$ incorporating all the information $V_{t+\Delta}$ is formed. As a consequence of these timing assumptions, only public information will be incorporated into prices. This is important as it allows pricing by standard martingale techniques.

t	$t+q$	$t+\Delta$
signals revealed	no trading	signals revealed
V_t revealed	agents receive new signals	$V_{t+\Delta}$ revealed
trading	agents form posterior	trading
B_t formed	agents act	$B_{t+\Delta}$ formed
	firm fails or does not fail	

Figure 2.1: Timing assumptions

Dynamics of asset value and signals

The relative increase in the asset value is normally distributed around a drift.

$$V_{t+\Delta} - V_t = \mu_V V_t \Delta + V_t \eta_t, \quad \eta_t \sim NID\left(0, \frac{1}{\alpha}\right)$$

At $t+q$, agents receive a signal X_i (subscript i indexes the different agents) about the impending change in V, with the distribution of the signal, conditional on the asset value V_t given by

$$X_i = V_{t+\Delta} + V_t \varepsilon_i, \quad \varepsilon_i \sim NID\left(0, \frac{1}{\beta}\right),$$

where $Cov(\eta_t, \varepsilon_i) = 0$, i.e. the noise is orthogonal to the innovations in the asset value.

From the signal X_i and the public information V_t, agents form a posterior about the value of the firm in period $t + \Delta$, $V_{t+\Delta}$ (which is also normally distributed).

2.3.2 The solution

Basic procedure

The model is solved using the procedure as in Morris and Shin (2004a). Suppose that agents follow a switching strategy around a certain posterior belief. Given the posterior belief around which agents switch, we can work out how many of them will grab assets, given the asset value in the next period (posterior beliefs will be centred around this asset value in the next period). The critical next-period asset value for which the firm will fail (given the belief in this period around which agents switch) can therefore be determined. This is the trigger point.

The discrete time trigger point

In appendix A.1, the following solution is derived (equation A.21):

$$V^*_{t+\Delta} = L\Phi\left\{\frac{\alpha}{\sqrt{\beta}}\left(\frac{V^*_{t+\Delta}}{V_t} - 1 - \mu_V\Delta\right) + \frac{\sqrt{\alpha+\beta}}{\sqrt{\beta}}\Phi^{-1}\left\{1-K\right\}\right\} \tag{2.1}$$

The trigger point V^*_t is unique if:

$$L\frac{1}{\sqrt{2\pi}}\frac{\alpha}{\sqrt{\beta}}\frac{1}{V_t} < 1$$

(cf. appendix A.1.8, proposition 1, condition A.I).

Continuous time limit

Now take the continuous time limit. If we want the asset value process to tend to a geometric Brownian motion, we need

$$\lim_{\Delta \to dt} \frac{1}{\alpha} = \sigma_V^2 dt,$$

i.e. the variance of public information about the innovation in the asset value to be proportional to time. So the variance of the innovation is $O(\Delta)$, or the precision is $O(\frac{1}{\Delta})$.

Now a sufficient condition for the uniqueness of the equilibrium described in equation (A.21) in continuous time, *regardless of the asset value, the parameter L and the face value of debt*, is that

$$\frac{1}{\beta} = o\left(\Delta^2\right),$$

i.e. that private information becomes more precise at a rate faster than Δ^2, because this ensures that condition (A.I) (s.a.) is always satisfied. This is just to say that we need the quality of private information to be sufficiently high in relation to the quality of public information in order for agents to be sufficiently uncertain about the actions of others to obtain coordination failure. As $\Delta \to dt$, $\Delta^2 \to 0$, and hence β grows at a faster rate than α. Consequently, $\frac{\alpha}{\sqrt{\beta}}$ tends to zero, so condition (A.I) will be satisfied for any permissible $V_{t+\Delta}$. Also, $\frac{\sqrt{\alpha+\beta}}{\sqrt{\beta}} \to 1$. The resulting trigger point equation then reduces to

$$V^* = L(1-K). \tag{2.2}$$

Note that the solution is constant. If we let the intermediate time period $(t+q)$

tend to the period immediately following it, the firm fails at t whenever $V(t)$ hits $L(1-K)$, i.e. when the asset value is a fraction $(1-K)$ of the vulnerability boundary (e.g. a covenant). The boundary or trigger point is a decreasing function of the cost of attempting to grab assets. Agents are reluctant to grab assets if it is costly for them to do so.

Due to the special assumptions about payoffs, this function turns out to be quite simple here - it is constant. Of course, one could allow for more general types of costs, but these would not in general produce a closed form solution.

The actions of agents in the continuous time limit

Conditional on the asset value in the next period, the probability that a creditor receives a signal which prompts it to stop lending is $\Phi\left\{\frac{1}{V_t}\sqrt{\beta}(X_t^* - V_{t+\Delta})\right\}$. As β tends to infinity, this probability tends either to 1 or to 0 for all non-marginal agents. What this means is that because all creditors essentially receive the same information (as the signal becomes infinitely precise), the agents will either all grab assets, or will all refrain from doing so. For any non-marginal creditor, the ex-ante probability of grabbing assets when the other creditors do not do so tends to zero. Also, the probability of not grabbing assets if all other creditors are grabbing assets tends to zero. This is essentially because in the limit, agents receive the same signals, and there is no uncertainty about the asset value. However, strategic uncertainty remains.

Strategic uncertainty in the continuous time limit

Strategic uncertainty remains in the sense that for the marginal agent, the fraction of creditors that forces reorganisation is still a random variable in the limit, which actually turns out to be uniformly distributed. This implies that the solution of the game is preserved in the limit. This type of result has been discussed at length elsewhere (Morris and Shin 2002). A formal proof is in appendix A.2.

Suppose that we instead start with the assumption that there is no private information, and hence no coordination failure. All creditors now have the same information, and (in the absence of asymmetries) will therefore either all grab assets lending, or all not grab assets. Suppose there is a strategy that specifies a path for the switching point. At every node of the game, it does not pay to deviate from the strategy when it specifies grabbing assets (because in this case deviating *always* loses money as the firm will be reorganised and other creditors will grab assets), and it does not pay to deviate from a strategy when it specifies not grabbing assets (deviating implies having to pay the lawyers, but *never* produces pecuniary benefits). It follows that all paths of a trigger point below L can be supported. The important difference to the coordination failure case is that there is no strategic uncertainty. Taking the limit of our discrete time coordination failure game has allowed us to eliminate all equilibria but one, even though in the continuous-time coordination failure case, agents also all have the same information.

2.3.3 Pricing

Prices are only determined in periods in which the asset value for the period is known, implying that expectations can be taken under the natural filtration of the asset value process.

Letting $t+q$ tend to the time period following it, default at time t will occur when $V(t)$ hits the trigger point V^*. Suppose τ is the stopping time at which the asset value process hits the boundary and T denotes maturity, and r is the constant riskless rate on the money market account. Assume that the asset value V is traded[2].

Pricing under assumption 2.A (bond represents entire debt)

Pricing under this assumption will facilitate comparison to models that specify a maturity for equity and aggregate debt, such as the Merton (1974) and Black and Cox (1976) models. To derive the price, it is necessary to discuss the payoffs to bondholders.

Payoffs to bondholders Let D denote the face value of aggregate debt (and here also of the face value of the bond) as before, and assume that the level of assets for which the firm becomes vulnerable L (which could be interpreted as an important covenant) is smaller than D (which will imply that $V^* < D$), then we can write the payoffs to bondholders as

1. $\Pi_1(T) = D - \max(D - V_T, 0)$, if $\tau > T$ (no default before maturity)

2. $\Pi_2(\tau) = (1 - K)V^*$, if $\tau < T$ (default before maturity)

With these payoffs, the model looks very similar to the standard Black and Cox (1976) case. The difference here is that the absorbing boundary is given by our trigger point rather than just the level of a covenant (or the level of asset for which the firm becomes vulnerable), and that the payoff upon hitting this boundary is not equal to the asset value, but to a fraction $(1 - K)$ of the asset value.

Bond price The bond price will be equal to the discounted expected value of these payoffs under the equivalent martingale measure ($\mathbb{Q}$) defined by the money market

[2]In the author's opinion the latter assumption is unpalatable but unfortunately standard in the literature.

account as a numeraire.

$$F_B(V_t, t, T) = E_t^{\mathbb{Q}}\left[e^{-r(T-t)}\Pi_1 I\left(\tau > T\right) + e^{-r(T-\tau)}\Pi_2 I\left(\tau \leq T\right)\right]$$

(Here, r denotes the constant risk free interest rate, t denotes the present and I is the indicator function)

At this stage pricing is straightforward. It was first explored by Black and Cox (1976). The exposition here follows Ericsson and Reneby (1998).

The bond can be viewed as a portfolio of barrier options. Π_1 can be interpreted as a combination of a long position in a down-and-out call with strike price 0, (which is of course just equivalent to a down-and-out position in the underlying asset value), and a short position in a down-and-out call in with a strike price of the face value of aggregate debt (equal to the face value of the bond in this case) D. This captures the fact that a bond can be viewed as a bull spread on the asset value. Π_2 can be interpreted as a long position in $(1-K)V^*$ units of a finite maturity dollar-in-boundary claim (a claim that pays one dollar in the case the boundary is hit before maturity). The price is the sum of prices of these positions:

$$\begin{aligned} B(V_t, t) = &C_{DO}(V_t, t; T, 0) - C_{DO}(V_t, t; T, D) \\ &+ (1-K)V^* G(V_t, t, |\tau < T) \end{aligned} \tag{2.3}$$

where $C_{DO}(V_t, t; T, Z)$ denotes the price of a down-and-out call with strike price Z on the underlying V at t with maturity T. Similarly, $G(V_t, t|\tau < T)$ denotes the price of a dollar-in-boundary claim with maturity T. The interested reader is referred to appendix A.3 for details of how the components are priced.

Equity price Equity can be viewed as a down-and-out call option on the asset value. Hence the price of equity is given by

$$E(V_t, t, T, V^*) = C_{DO}(V_t, t; T, D). \tag{2.4}$$

The sum of the value of equity and the value of debt equals the market value M of the firm:

$$M(V_t, t) = C_{DO}(V_t, t; T, 0) + (1 - K)V^* G(V_t, t | \tau < T) \tag{2.5}$$

This is in essence the sum of the price of the down-and-out asset value of the firm, plus the price of the down-and-in claim representing the recovery value in the case of default. Note that unlike in Merton model, the market value is not equal to the asset value of the firm unless $K = 0$, i.e. there is no cost associated with bankruptcy.

Pricing under assumption 2.B (bond is a negligible part of aggregate debt)

Pricing under this assumption will make empirical implementation of the model easier, as it easily allows for coupon bonds. Under the assumption that the bond price is a negligible proportion of aggregate debt, the firm fails if the asset value falls below below the trigger point, defined by the interaction of creditors that have each own part of the aggregate debt of the firm. The trigger point is still calculated the same way.

Payoffs to bondholders The payoffs to bondholders will be the promised payoffs at any time that default has not occurred already, or will be a fixed proportion of the face value of the bond (now different from the aggregate face value of debt) in case of default. Principal (or indeed coupon) repayments are binary options that pay of in case default occurs after their maturity.

Bond price Denote a binary option with maturity t_i as $H(V_t, t; t_i)$, then we can write the zero coupon bond price (per dollar of face value) as

$$\begin{aligned} B_{ZC}(V_t, t) = & H(V_t, t; T) \\ & + RG(V_t, t|\tau < T) \end{aligned} \tag{2.6}$$

where R is the recovery fraction. This can of course be extended easily to include coupons:

$$\begin{aligned} B(V_t, t) = & \sum_{i=1}^{M} cH(V_t, t; t_i) \\ & + H(V_t, t; T) \\ & + RG(V_t, t|\tau < T) \end{aligned} \tag{2.7}$$

Again, the interested reader is referred to appendix A.3 for details of how the components are priced.

Note that the pricing here can be seen as a special case of a very generic bond pricing approach (Ericsson and Reneby 1998) that views a bond as a portfolio of simple and barrier claims. This also goes for other bond pricing models integrating strategic interaction (Mella-Barral and Perraudin 1997). In essence, incorporating elements of strategic interaction into bond pricing often seems to produce arguments as to what the simple and barrier claims should be that make up the bond, and what determines stopping times. Strategic interaction produces different payoffs in different states of the world which determine the price.

2.4 Discussion of results

Comparisons in general are easier under assumption 2.A (the bond being priced represents the entire debt), as it facilitates the comparison to the older models that make this assumption.

2.4.1 Comparisons to Merton 1974 and Black-Cox 1976

In the Black and Cox (1976) case, the amount recovered if bondholders liquidate the firm is simply V_t. It is trivial to show that this is always more than B_t (intuitively, this is the case since B_t represents a claim to V_T in some states of the world, and a claim to a value less than V_T in others, whereas V_t represents a claim to V_T in all states of the world). Liquidation or bankruptcy will therefore always occur when the covenant is breached, as it is optimal for (coordinated or indeed uncoordinated) bondholders to liquidate.

Note that in the coordination failure model, if the cost of grabbing assets (K) is zero, the trigger point is L, which is simply the value of assets for which the firm is vulnerable (possibly the level of a critical covenant), and the amount recovered is the same as in the Black and Cox (1976) model. The prices will coincide. Essentially, if the cost of stopping lending is zero, then it becomes a dominant strategy to do so as soon as possible, if there are no bankruptcy costs.

If we let $K \longrightarrow 1$, it is always preferable to continue lending, and the barrier drops to zero. In this case, bankruptcy before maturity does not occur, and the price will coincide with the Merton (1974) price.

The effect producing the price that differs from Merton price is twofold: Firstly, the possibility of early default - i.e. receiving money before the maturity of the bond - increases the value of the bond, as in the Black-Cox case. But secondly, since there is a cost of bankruptcy $(1 - K) < 1$, this decreases the value of the bond, ceteris paribus. Because these two effects conflict, the coordination failure model does not produce a discount vis-a-vis the Merton case in all situations.

2.4.2 Comparison to the no coordination failure case

Supposing there was only one creditor, and supposing that creditors acted in a coordinated manner, we could derive an optimal boundary at which the creditor or creditors would push the firm into bankruptcy. Even if the bankruptcy cost K was the same for the single creditor case, we can deduce that the value of aggregate debt would be lower (or at most equal) to the value of debt in the coordination failure case, since the boundary derived from the coordination failure game is not optimal (it does not maximise the value of debt). If the cost K represents the cost of sending out lawyers to secure liens on assets, it is likely that the corresponding cost in the case of a single creditor would be lower (if not zero) as well, creating a further coordination failure discount.

2.4.3 Comparison to Morris-Shin 2004

Morris and Shin (2004a) refer to a version of equation (A.21), and argue that if one assumes the trigger point to be fixed, one would underprice debt, as the trigger point is actually a decreasing function of the asset value. So as the asset value decreases, the trigger point moves up. Ignoring this effect would cause overpricing. The effect mentioned by Morris and Shin (2004a) does not cause the difference in the price to the Merton (1974) model here, because the continuous time limit of the trigger point (equation 2.2) is not a function of the asset value - it is constant.

2.4.4 Thoughts on an empirical test

As argued above, a practical implementation of the model is only really feasible with the price derived above under the assumption that the bond represents a negligible proportion of overall debt (assumption 2.B). With the price under the alternative assumption, it is not clear as to how to deal with coupon bonds, and the fact that equity does not typically have a maturity.

However, any model that produces a similar default barrier will produce very similar prices (cf. e.g. Longstaff and Schwartz 1995). If the amount recovered on any particular bond is not related to the value of assets at the time of default as above, and there are no deviations of absolute priority, then the only asset that can identify the barrier will be aggregate debt. Since typically, there is no data on aggregate debt, the barrier will not be identified. The model can only identify the distance-to-default.

In conclusion, it will be difficult to distinguish between alternative models which differ according to the level of the barrier on the basis of bond and equity prices alone.

Since the default barrier is given by $V^* = (1 - K)L$, and the fraction actually recovered in the case of default is $(1 - K)V^*$, the model does make an empirically testable prediction, however. Supposing one had some data on defaulted companies, and one had a proxy variable for how costly it is to seize assets (K), and an estimate of the level of assets for which covenants are first violated (L) - possibly via the equity price at the time that the first covenants are reported to be violated, one could calculate the predicted recovered amount as $(1 - K)^2 L$, and compare this to actually recovered amounts.

Unfortunately, the author has not found data for an appropriate proxy of K, so testing this prediction of the model is an issue for future research.

2.5 Concluding Remarks

This chapter presents a model that examines the pricing of corporate bonds if coordination failures between creditors can occur. A discrete time model exhibiting coordination failures is set up, and the continuous limit is taken. A bond is priced in the context of this game under two different assumptions about the capital structure of the firm. If the assumption is made that the bond being priced represents the entire debt of the firm, the resulting price can be directly compared to the Merton (1974) and Black and Cox (1976) models, which are both limiting cases of the model presented here.

The model here will produce a discount vis-a-vis the no-coordination failure case, and could therefore present a reason as to why structural models ignoring this feature underpredict spreads.

In terms of empirical implementation, however, which will necessarily be under the second assumption about capital structure (assumption 2.B) as discussed above, it will be difficult to distinguish between other tractable models on the basis of bond and equity prices alone. As suggested above, a direct test of the model could be conducted with data on recovered amounts in bankruptcy cases.

Chapter 3

Estimating structural models of corporate bond prices

3.1 Introduction

Having examined some theoretical aspects of structural models, it is now time to turn to the econometric aspects of their implementation.

As argued in chapter 1, structural models are a natural choice in situations where both equity and credit related instruments are examined simultaneously. For example a hedge fund might be interested in hedging exposure in a credit default swap by buying and selling equity.

Alternatively, one might want to estimate hedging parameters, but not enough bond prices are available to calibrate reduced-form models. In such a case, being able to utilise equity prices would be beneficial.

Since structural models seem to be imprecise in pricing instruments, however, they are often avoided.

The typical conclusion about the problems of structural models would be that they overpredict bond prices, and underpredict spreads, as first reported by Jones, Mason, and Rosenfeld (1984). More recently, it has been argued (as discussed in more detail in chapter 1) that some models actually overpredict, but more crucially, that all models make very imprecise predictions (Eom, Helwege, and Huang 2004, Lyden and Saraniti 2001).

As argued in previous chapters, the fact that many structural models do not appear to fit the data well might be to do with the fact that some of the assumptions used to derive the prices are inappropriate. It is clear that a necessary condition for obtaining reasonable hedge parameters and predictions, however, is that estimation methods must be reasonably efficient and unbiased. Unfortunately, the standard estimation methods are neither efficient nor unbiased, as will be shown below.

The main problem that estimation methods have to confront is that although models typically assume that the asset value is traded (in order to simplify pricing), the truth is that the asset value is not even observed, except maybe at a quarterly frequency (through published balance sheets). The defining feature of different estimation methods is how they approach this problem.

Most approaches involve inverting the equity pricing function to produce equity-implied asset values. The approach presented here shows how estimating can proceed by using prices of multiple assets or any other signal of asset values (e.g. equity, bond prices, and credit default swap spreads) to obtain an implied asset value, via non-linear filtering. The approach has the advantage that all available information can be used to estimate, increasing efficiency, and that it allows assuming that assets are not perfectly priced by the model. It is shown that assuming that some assets (e.g. equity) are perfectly priced can produce biases. This is due to the fact that asset prices contain components that are not related to the modelled risk factor (asset value risk). The method presented here naturally accounts for unmodeled factors by interpreting observed asset prices as noisy signals of fundamental values.

The efficiency increase by including bond prices as well as equity prices in estimating a structural model is shown to be quite dramatic. In essence, it seems to be easier to predict bond prices if bond prices are used in estimating the model.

The rest of this chapter is organised as follows: First, the fundamental problems of estimating structural models are discussed, and some proposed estimation approaches examined. A very general description of the estimation problem is presented, followed by an exposition of how a likelihood estimation can proceed in this context. Results of a Monte Carlo experiment are discussed. Finally, an application to real data is conducted.

3.1.1 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.

3.1.2 Other issues for estimation

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

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 obtained from any one party 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. As a relatively recent development (July 2002), the National Association of Security Dealers has started distributing high quality bond pricing data, so hopefully this will be less of an issue in the future.

In addition to the problem of data frequency, 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.

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 2004b) 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.

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.

3.1.3 Estimation approaches

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):

$$\sigma_E = \sigma_V \frac{V}{E} \frac{\partial F_E}{\partial V} \tag{3.1}$$

Note that here, F_E depends on the particular structural model to be estimated, of course. 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 (2001) 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.

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.

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 used 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.

$$\begin{aligned} E &= F_E(V, \sigma_V) && (3.2) \\ \sigma_E &= \sigma_V \frac{V}{E} \frac{\partial F_E}{\partial V}. && (3.3) \end{aligned}$$

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 3.3. 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).

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.

3.2 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.

3.2.1 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_t of the firm (which determines the value of equity and debt) follows a geometric Brownian motion:

$$dV_t = \mu V_t dt + \sigma V_t dW_t. \tag{3.4}$$

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

$$\int_t^T d\log V_s = (\mu - \frac{1}{2}\sigma^2) \int_t^T ds + \sigma \int_t^T dW_s \sim N\left\{\mu (T-t), \sigma^2 (T-t)\right\}. \tag{3.5}$$

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

$$\alpha_t = d + \alpha_{t-1} + \eta_t, \quad \eta_t \sim NID\{0, \sigma_\eta\}, \tag{3.6}$$

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.

3.2.2 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.

$$\xi_{it} = F_{it}(\alpha_t; \psi_i) e^{\varepsilon_t}, \quad \varepsilon_t \sim NID\{0, \Sigma_\varepsilon\}. \tag{3.7}$$

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 or components that are outside the structural model (for an example, see chapter 4).

3.2.3 The system

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

$$\alpha_t = \quad d + \alpha_{t-1} + \eta_t, \quad \eta_t \sim NID\{0, \sigma_\eta\}, \tag{3.8}$$

$$y_t = \quad Z_t(\alpha_t; \psi) + \varepsilon_t, \quad \varepsilon_t \sim NID\{0, \Sigma_\varepsilon\} \tag{3.9}$$

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.

3.2.4 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 $\sigma_i = \infty, \quad \forall i \neq E$:

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.

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

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

$$p(E, A_1, A_2, \cdots A_n) = p(A_1, A_2, \cdots A_n|E)p(E) \tag{3.10}$$

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 in chapter 4.

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. The procedure described below can of course also be used to estimate the systems described under Case 1 and 2, and is more general than ERMLE.

3.3 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

$$L(\psi) = p(y|\psi) = \int p(\alpha, y|\psi)d\alpha. \tag{3.11}$$

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 3.3.1)[2]. We can now write the likelihood as

$$\begin{aligned} L(\psi) =& g(y) \int \frac{p(a, y|\psi)}{g(a, y|\psi)} g(\alpha|y, \psi)d\alpha & (3.12) \\ =& L_g(\psi)E_g\left[w(a, y)\right], & (3.13) \end{aligned}$$

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

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

$$w\left(a,\ y\right)=\frac{p(a,y|\psi)}{g(a,y|\psi)}. \tag{3.14}$$

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 $E_g\left[w(a,y)\right]$ 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.
3. To obtain a correction factor, simulate α from the importance density, and numerically calculate the expectation term.

We can write

$$\hat{L}(\psi)=L_g(\psi)\bar{w} \tag{3.15}$$

where

$$\bar{w} = \frac{1}{M}\sum_{i=1}^{M} w_i, \quad w_i = \frac{p\left(\alpha^i, y|\psi\right)}{g\left(a^i,\; y \mid \psi\right)}, \tag{3.16}$$

and α^i is drawn from the importance density (for details, see appendix B). 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):

$$\log \hat{L}(\psi) = \log L_g(\psi) + \log \bar{w} + \frac{s_w^2}{2M\bar{w}^2}, \tag{3.17}$$

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 more computationally intensive than ERMLE, but more general. The ERMLE likelihood function is nested in the likelihood function estimated here. Estimating the ERMLE likelihood function with the method proposed here is simply an issue of imposing some restrictions.

3.3.1 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:

$$Z_t(\alpha_t) \cong Z_t(\tilde{\alpha}_t) + \dot{Z}_t(\tilde{\alpha}_t)(\alpha_t - \tilde{\alpha}_t), \tag{3.18}$$

where

$$\dot{Z}_t(\tilde{\alpha}_t) = \left.\frac{\partial Z_t(\alpha_t)}{\partial \alpha_t{}'}\right|_{\alpha_t=\tilde{\alpha}_t}. \tag{3.19}$$

Defining

$$\tilde{y}_t = y_t - Z_t(\tilde{\alpha}_t) + \dot{Z}_t(\tilde{\alpha}_t)\tilde{\alpha}_t, \tag{3.20}$$

we can approximate the observation equation by

$$\tilde{y}_t = \dot{Z}_t\left(\tilde{\alpha}_t\right)\alpha_t + \varepsilon_t. \tag{3.21}$$

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:

$$\alpha_t = d + \alpha_{t-1} + \eta_t \tag{3.22}$$

$$\tilde{y}_t = \dot{Z}_t(\tilde{\alpha}_t)\alpha_t + \varepsilon_t \tag{3.23}$$

where

$$\eta_t \sim NID\{0, \sigma_\eta\} \tag{3.24}$$

$$\varepsilon_t \sim NID\{0, \Sigma_\varepsilon\} \tag{3.25}$$

and

$$\eta_t \perp \varepsilon_t. \tag{3.26}$$

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).

3.3.2 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.

3.3.3 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.

3.3.4 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.

3.4 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[3].

3.4.1 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

[3]The estimation technique was implemented using Ox Version 3.32 (Doornik 2002) and SsfPack Version 3.0 beta 2 (Koopman, Shephard, and Doornik 1999)

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, as discussed in chapter 1 and 2. This has the advantage that 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 (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[4].

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

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

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).

3.4.2 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.

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 3.1 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 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.

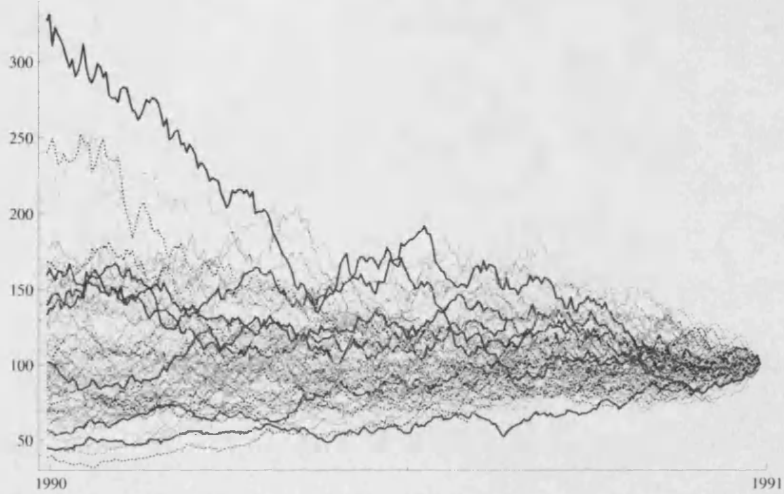

Figure 3.1: Alternative asset value paths that could have led to an asset value of 100 in 1991

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 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 more on average), it is possible to anticipate that the performance of the calibration technique should be worse in this case.

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 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 3.1 and Table 3.2 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 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

	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 3.1: Monte Carlo results (high quality issuer)

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 3.2).

Both SMLE and ERMLE allow the calculation of posterior densities for spread or yield predictions. Figures 3.2 and 3.3 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

	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 3.2: Result of Monte Carlo (low quality issuer)

for the actual yield is calculated. The quantiles according to the posterior density can then be plotted against the quantiles of 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.

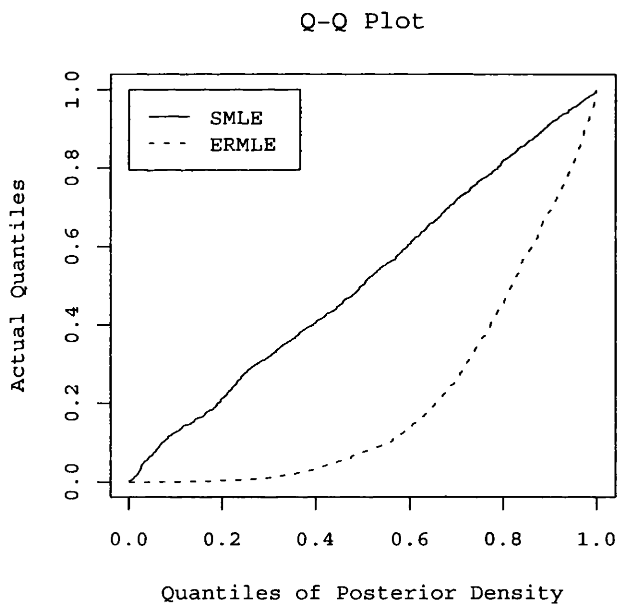

Figure 3.2: SMLE and ERMLE actual versus predicted quantiles (high quality issuer)

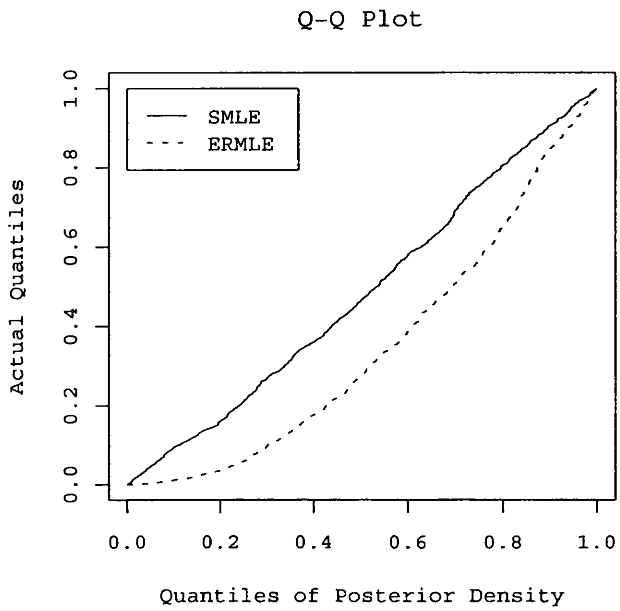

Figure 3.3: SMLE and ERMLE actual versus predicted quantiles (high quality issuer)

3.4.3 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.

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 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 model has 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 (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, 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 39 issuers.

Note that despite the cleaning of the data, there are issues with its quality. It is not clear, for example, whether the reported dates for the transactions are transaction or settlement dates. Close to coupon dates, the exact date can make a large difference to cash prices. This is especially important as the prices in the database are flat prices, and the provided numbers from which accrued interest can be calculated look very unreliable (a large proportion is negative, for example). Also, some of the observations appear to be affected by data entry errors - for example, there are bonds which trade frequently in one particular month with prices between 101 and 105, followed by a price in the same month of 135, followed by prices in the previous price range (still in the same month). This is of course a problem when trying to determine the accuracy of pricing methods, especially in trying to compare the results to those reached in other papers with alternative datasets.

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 unambiguously matched up were ignored. This left 33 issuers, detailed in table B.2 (see appendix B for details). Bond price data for which no equity price data was obtainable was ignored.

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.

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[5].

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

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.

As a simple workaround, it would be possible to e.g. fix the standard deviation of the observation errors at 1% and not estimate them. SMLE was run here once without restrictions on estimation errors, and once with this restriction imposed.

The results are reported in Table 3.3. As can be seen, the SMLE estimates are substantially less biased (a mean error of -12 bp for SMLE (restricted) 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 (restricted) and 125 and 128 bp for ERMLE and Calibration respectively).

The performance of SMLE is good, considering the fact that apart from the asset value, only a single parameter was estimated, and that the bond pricing data is of low quality. SMLE can be seen to be substantially less biased and more precise than either of the ERMLE and calibration techniques.

Several things can be noted:

- Including bond prices in the estimation reduces the out-of-sample bond price

	Mean error	Mean abs. error	RMSE
Spread or Yield error (bp)			
SMLE[1]	−27	36	49
SMLE[2]	−12	52	88
ERMLE	−77	95	125
Calibration	−78	95	128
Yield percentage error			
SMLE[1]	−3.84%	5.45%	6.71%
SMLE[2]	−0.50%	10.20%	17.55%
ERMLE	−7.25%	18.44%	25.84%
Calibration	−7.50%	18.41%	25.92%
Bond price error ($ per 100$ face value)			
SMLE[1]	0.79	3.88	7.47
SMLE[2]	−0.22	3.31	6.71
ERMLE	4.08	6.39	9.71
Calibration	4.26	6.50	10.13
Bond price percentage error			
SMLE[1]	1.04%	3.53%	6.20%
SMLE[2]	0.04%	3.03%	5.57%
ERMLE	4.43%	6.26%	9.39%
Calibration	4.68%	6.44%	10.12%

[1]: unrestricted
[2]: s.d. of obs. errors restricted to 1%

Table 3.3: Prediction errors on real data

prediction bias.

- The largest difference between techniques appears to depend on whether or not bond prices were used in the estimation. ERMLE does not seem to significantly outperform the calibration technique for this dataset, possibly due to relatively constant leverage over the sample period. This is in spite of the fact that only a relatively small number of bond price observations were used in estimation (the bond price data is not daily data).

- All methods appear to underpredict spreads on average. It is likely that this is due to the theoretical model producing spreads that are too low to match empirically observed spreads.

- It is clear that for this theoretical model, the underprediction of spreads it produces outweighs the overprediction effect from ERMLE.

- Imposing the restriction that the standard deviation of observation errors is 1% for both equity and bonds appears to produce a less biased out-of-sample prediction. It does appear to raise the variability of the prediction on some measures. A likelihood ratio test conducted firm-by-firm to test the null that the restriction is valid leads to no rejections at 5% for any of the 33 firms.

Eom, Helwege, and Huang (2004) conclude that there are problems with the accuracy of structural models, testing several models. For any metric that they present, it is always the case that at least one of the two SMLE results present here is better than their best model. The results here also compare favourably to those presented by Lyden and Saraniti (2001), who report mean absolute yield or spread prediction errors in terms of 80 to 90 basis points (the mean absolute yield error for ERMLE, Calibration and unrestricted SMLE are 95 bp, 95 bp and 36 bp respectively).

This is *in spite* of the fact that the data utilised here is unreliable. In chapter 4, another dataset is presented: the Lehman Brothers Fixed Income Securities Database (the same data used by Eom, Helwege, and Huang (2004)). Estimating the same theoretical model with the ERMLE technique, produces RMSEs in the range of 91 bp, as opposed to 125 bp. This indicates that the NAIC pricing data is of poor quality, as discussed above. Unfortunately, it was not possible to use the Lehman FISD data here for licensing reasons.

It is possible that the conclusion of Eom, Helwege, and Huang (2004) that "accuracy is a problem", which is broadly supported by other studies (e.g. Lyden and Saraniti 2001) is partly to do with their estimation methods, and it would be interesting to repeat the analysis presented here with more reliable bond data, in particular, data at a higher frequency and for different classes of assets and derivatives (e.g. CDS spread data).

3.5 Concluding Remarks

This chapter 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).

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.

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.

Chapter 4

The non-default components of corporate yield spreads [1]

4.1 Introduction

Many structural models appear to underpredict bond spreads. So far, there have been two distinct responses to this problem. As detailed in the introduction, many have argued that in order to improve the fit of structural models, the calculated default probabilities have to be increased and/or payoffs to bondholders have to be decreased such that predicted credit spreads rise.

This is of course only one possible answer to the problem; the other answer is that the default component of the spread is accurately predicted by structural models, but that there are components in spread which are not related to default. Possible proposed causes for this components include taxes and liquidity. There seems to be a growing awareness in academia that default risk explains only part of observed

[1]This chapter is based on joint work with Joel Reneby.

corporate bond spreads.

Clearly, it would be interesting to see whether it is possible to

- estimate a structural model explicitly assuming that there is a non-default component to corporate bond spreads, and
- to see whether the estimated model is better at out-of-sample prediction of bond spreads.

Ideally, one would like to have a theoretical model that jointly explains the default and non-default components of spreads, but as of yet no empirically tractable models have been suggested. A tentative explanation is that including illiquidity or transaction costs would destroy the frictionless setting in which the models are derived.

In this chapter, a more pragmatic approach is taken instead. A structural credit risk model is used to determine the default risk premium, and it is hypothesised that the non-default premium of a bond is a linear function of market-wide liquidity proxies. The main finding is that even with a very simple model of the non-default component, out-of-sample prediction ability of the combined model is substantially improved.

Several recent papers have investigated the components of corporate spreads. E.g. Elton, Gruber, Agrawal, and Mann (2001) back out default spreads from historical default rates from transition matrices, and find that compensation for expected losses only account for a small part of the spread. Huang and Huang (2003) do a similar exercise but match historical equity premium as well. They, too, find that the default spread can only constitute a fraction of total spreads. Longstaff, Mithal, and Neis (2004) extract a non-default component from corporate bond spreads using the premium on corresponding credit default swaps as a proxy for the default premium. They find that although the majority of the corporate bond spread is default-related there is a large non-default component. Averaging across time allows them to regress the non-default component of different firms on proxy variables for non-default factors. They find that coupons, the average bid-ask spread of a bond, amount outstanding, maturity, rating and a financial sector dummy are important. Averaging across firms allows them to regress their non-default component on time series. They find that money market mutual fund flows as well as the spreads between off-the run and on-the-run treasury bonds are significant.

In contrast to the studies cited above, Collin-Dufresne and Goldstein (2001a) investigate causes of *changes* in corporate spreads. They show that when performing simple linear regressions of credit spread changes on several explanatory variables that one would find as inputs to typical structural models, only about a quarter of the variation in credit spread changes can be explained. When examining the residuals, they find that they are correlated across firms, indicating an unidentified market-wide factor. They also attempt a regression where they include liquidity proxies, such as for example the difference in yields of 30 year on-the-run versus off-the-run treasury bonds. This increases the percentage of explained variation, but still does not seem to eliminate the market-wide factor, which they believe is related to liquidity but cannot be captured with their liquidity proxies.

Delianedis and Geske (2001) attempt to identify the default component with a structural model, and then run regressions to explain the non-default component, and find that equity volume as a proxy for liquidity appears to be a relevant explanatory variable.

In sum, the results above indicate that non-default components make up a substantial portion of yield spreads. Existing bond pricing models, in contrast, are typically designed to price default risk only. To better understand and predict bond prices therefore, it seems necessary to include non-default components in bond pricing models.

As argued above, one would ideally like to have a theoretical model that jointly explains the default and non-default components of spreads, but as of yet no empirically tractable models have been suggested.

Here, this problem is avoided by using an ad-hoc model: the total bond spread is modelled as the sum of the default spread predicted by the structural model and a non-default spread component. The non-default spread is, in turn, modelled as a linear function of a set of market-wide variables. Inspired by the findings of Longstaff, Mithal, and Neis (2004), the following variables are chosen: the liquidity on the treasury market as proxied by the spread between on- and off-the-run treasuries, as well as the flow in and out of money market mutual funds as reported by the Federal Reserve. The structural model that is used is the modified version of the Leland (1994) model as presented in previous chapters.

Using stock prices and bond spreads, balance sheet data and the chosen market factors, the combined model of default and non-default risk is first estimated with maximum likelihood. Next, a prediction of the corporate bond spread one month out-of-sample is obtained from the stock price and the value of the market factors at this point. If the market factors indeed can explain the non-default component of bond spreads, this estimate should be superior to the corresponding estimate from a model without the non-default component. Comparing the two models, the findings are that while the default-risk-only model underpredicts spreads by about 60 basis points, the prediction combining the default component with the non-default component is virtually unbiased, and reduces the variability of out-of-sample forecasts.

The rest of this chapter is organised as follows: first the model for the default and non-default components is discussed, followed by the estimation methodology. Then the data set used to implement the procedure is presented. Finally, results are produced and conclusions drawn.

4.2 The model

4.2.1 The default component

The structural model utilised here is the same as in chapter 3. This is a modified version of the Leland (1994) model, extended to allow for violations of the absolute priority rule and future debt issues. Aggregate debt and aggregate equity are viewed as perpetuities, with default occuring when the asset value hits the level at which shareholders are no longer willing to contribute funds to stave off financial distress (i.e. the default boundary is endogenous). As argued above, different assumptions about the barrier are feasible, but probably will not lead to large differences in predicted credit spreads.

As above, an individual bond is presumed to be a negligible proportion of total debt, and priced as a portfolio of binary options (one for each coupon, and one for the principal), and an option that pays of a recovery fraction of the principal at the time of default (if default occurs before maturity). Again, this allows combining a realistic model of a bond (i.e. the bond is coupon-paying and has a finite maturity) with analytical tractability. The model has been shown by Ericsson and Reneby (2002b) to work well in an empirical setting.

The yield predicted by this structural model is denoted Y^{CRED}.

4.2.2 The non-default component

The non-default component of the yield Y^{ND} is assumed to be a linear function of some selected proxy variables X_t:

$$Y_t^{ND} = X_t'\gamma \tag{4.1}$$

such that the total yield becomes

$$Y_t = Y_t^{CRED} + Y_t^{ND} \tag{4.2}$$

Later, an observation error will be added to this total yield.

The proxy variables utilised here will be differences between yields of the on-the-run and off-the-run government bonds at various maturities, and money market fund flows. Other proxy variables could be found, but the number of proxy variables that can be used in estimation will be constrained by the amount of data available.

4.3 Estimation

4.3.1 Setup

The estimation method will be based on maximum likelihood estimation. In order to proceed, we first need to write down how observed bond and equity prices and predicted prices or yields relate, and make distributional assumptions about errors.

Let α_{it} denote the log asset value of a firm i at time t. Let $Z_t^E(\alpha_{it}; \theta)$ denote the (log) equity pricing function, where θ_i is a vector of parameters of the credit model which are to be estimated. Suppose we have data on bond prices and equity prices, and would like to use both to estimate the model.

We can write the system in state space as

$$\alpha_{i,t+1} = \alpha_{it} + \eta_{i,t+1} \tag{4.3}$$

$$\log(E_{it}) = Z_t^E(\alpha_{it}; \theta_i) \tag{4.4}$$

$$Y_t = Y_t^{cred}(\alpha_{it}; \theta_i) + X_t'\gamma_i + \varepsilon_{it} \tag{4.5}$$

assuming that

$$\eta_{i,t+1} \sim NID(0, \sigma^2_{\eta,i}) \tag{4.6}$$

$$\varepsilon_{it} \sim NID(0, \sigma^2_{Y,i}). \tag{4.7}$$

Alternatively, this could be specified similarly to the system in chapter 3:

$$\alpha_{i,t+1} = \alpha_{it} + \eta_{i,t+1} \tag{4.8}$$

$$\log(E_{it}) = Z_t^E(\alpha_{it}; \theta_i) \tag{4.9}$$

$$\log(B_{it}) = Z_t^B(\alpha_{it}; \theta_i) + X_t'\gamma_i + \varepsilon_{it} \tag{4.10}$$

$$\tag{4.11}$$

Note that if the bond was a zero coupon bond, its spread would be $-(1/\tau)\log B$, where τ is the time to maturity. We can see that the above setup, in choosing the log bond price, is almost equivalent to choosing the spread as the observed variable (at least for zero coupon bonds). As it turns out, the resulting non-default component parameters will depend on time to maturity (over and above bond age) if estimated in terms of prices, so the alternative specification in terms of yields was adopted here.

For this specification, the resulting parameters appear to be invariant with respect to maturity.

This setup allows for the simultaneous (maximum likelihood) estimation of the credit parameters θ, as well as the liquidity parameters γ.

4.3.2 Evaluating the likelihood function

Although it would be possible to estimate the model using the SMLE procedure presented in chapter 3, this method was not adopted here, as the method can become slow for larger datasets. As argued in that chapter, the assumption that equity is priced perfectly introduces a bias, but the size of the bias appears to be reasonably small for the given structural model. We can also see that if the number of bond observations is relatively small compared to the number of equity pricing observations, the SMLE procedure will estimate a very low equity observation error in any case, unless restrictions are imposed.

It was concluded in that chapter, though, that including information on bond prices or yields appears to be important. This is of course necessary in any case if the objective is to estimate a non-default component of yields. Consequently, a modified version of the ERMLE/Duan, to also take into account yields, was utilised here.

As described in chapter 3, the method first proposed by Duan (1994) utilises a change of variables from asset values to equity prices to arrive at a density (and hence likelihood) for the observed equity prices: Let E^{obs} denote the observed equity price. Given that certain conditions on Z_t^E hold, $p(E^{obs})$ can be calculated easily from the density of the log asset value α as

$$p(E^{obs}) = p(\alpha) \left[\frac{\partial Z_t^E(\alpha;\theta)}{\partial \alpha} \right]^{-1} \Bigg|_{\alpha=\alpha(E^{obs};\theta)}$$

where $\alpha(E^{obs};\theta)$ is the asset value implied by the observed equity price, given a parameter vector θ.

As argued in chapter 3 section 3.2.4, we note that the joint density of equity and yields Y (or bond prices, or other contingent claims) can be written as

$$p(E, Y) = p(Y|E)p(E) \tag{4.12}$$

$$\log p(E, Y) = \log p(Y|E) + \log p(E). \tag{4.13}$$

Now given a parameter and an observed equity price, we can work out the equity-implied asset value. Since the event that $E = \bar{E}$ is equivalent to $\alpha(E;\theta) = \alpha(\bar{E};\theta)$ for a given theta, we can write the conditional $p(Y|E^{obs})$ as $p(Y|\alpha(E^{obs};\theta))$, where $\alpha(E^{obs};\theta)$ is the same equity-implied log asset value.

The algorithm for evaluating the likelihood function is as follows:

1. For the given parameter vector, compute the equity-implied asset value.

2. Calculate the Jacobian of the equity pricing function at this asset value.

3. Calculate the density of this asset value.

4. Calculate the yield implied by this asset value, and calculate its density.

5. Combine all components to calculate the likelihood at this parameter vector.

This allows numerical maximisation of the likelihood function.

4.4 Data

4.4.1 Bond price data

The corporate bond data we use was the Lehman Brothers Fixed Income Securities Database (FISD). This database contains data for bonds that were part of the Lehman Brothers indices. It contains monthly prices for a large array of fixed income securities. Here, the focus is exclusively on industrial bonds for firms based in the US during the coverage of the database, i.e. between January 1994 and February 1998.

The FISD database distinguishes between trader quotes and matrix quotes. As matrix quotes represent interpolation mechanisms rather than actual market prices, they are eliminated. Also, bonds with embedded optionality, sinking fund clauses etc are removed, such that the remaining bonds were plain vanilla unsecured bullet bonds only. Since a time series of reasonable length is a prerequisite for estimating the relationship between the non-default component and time-varying market-wide liquidity proxies, all bonds with less than 24 consecutive months of observations were eliminated. All bonds for which prices were clearly stale were also eliminated, where data entry errors were apparent or for which corresponding equity or balance sheet data was simply not available. Seven firms were dropped because of convergence problems in the estimation procedure.

The resulting dataset comprises 108 bond issues by as many firms; the number of monthly price observations ranges from 24 to 50.

The FISD database also provides all the characteristics of the individual bonds such as coupon, maturity, issue size and ratings.

4.4.2 Equity prices and accounting data

The empirical measure of the total debt of the firm is the total liabilities as reported in firms' EDGAR filings compiled and provided by Mergent. In addition to this, daily market values for the outstanding equity as well as information about the risk free interest term structures over time were obtained from DATASTREAM.

4.4.3 Non-default proxies

Differences between yields of on-the-run and off-the-run treasuries at different maturities (2, 5, 10 and 30 years) are used as proxy variables for market-wide liquidity (these were calculated from treasury market data). They were augmented by mutual money market fund flows (as suggested by Longstaff, Mithal, and Neis (2004)). The fund flows are provided by Federal Reserve.

4.4.4 Assumed and estimated parameter values

Due to the limited amount of data, the total number of estimated parameters needs to be kept to a minimum. Only the asset value volatility σ_η, the yield error σ_Y and the coefficients of the liquidity proxies (these include a constant) γ_i are estimated. This means that assumptions about the other credit model parameters need to be made. It is assumed that the cash payout rate (β) is 2%, the recovery rate to equity/deviation from absolute priority (ϵ) is 5%, the bankruptcy cost (κ) is 75%, and the corporate tax rate (ζ) is 20%. The recovery fraction of individual bonds (ψ) is assumed to match those reported by Altman and Kishore (1996) (matched by industry).

4.5 Results

For each bond, the model is estimated on all available data (at least 24 months) prior to January 31, 1998. Using the equity price observation for January 31, 1998, and the estimated model parameters, the spread is predicted for that date. This is not an out-of-sample forecast in the sense that the contemporaneous equity price observation is used in the forecast, but reflects possible uses of structural models such as hedging exposure in (illiquid) corporate bonds with equity - the assumption being that equity prices are always readily available. In the following, this will be referred to as an out-of-sample prediction, as it is out-of-sample with respect to the bond price/ spread being predicted (which is not part of the estimation sample). This is similar to the notion used in chapter 3.

4.5.1 Out-of-sample prediction performance

In order to compare, we first run the model without specifying a non-default component. The mean out-of-sample yield or spread prediction error for the model without non-default component is -61 bp (i.e. the model underpredicts by 61 bp on average). This is the now familiar result that most structural models on their own underpredict spreads or overpredict bond prices, and parallels the conclusions first drawn by Jones, Mason, and Rosenfeld (1985). The model without a non-default component also produces a fairly high root mean square error of 91 bp.

We then introduce the non-default component. As a first specification, X will only consist of a constant. In a second specification, proxies X_t include a (firm-specific) constant, and the spread between off-the-run and on-the-run treasury bonds at maturities of 2 and 30 years, and institutional money market mutual fund flows[2]. As discussed before, the coefficients on the market-wide variables are estimated as firm specific coefficients.

Introducing a constant firm-specific non-default component already produces a dramatic performance increase. The mean error drops to -11 bp, and the RMSE drops to 25 bp. The mean non-default component is 51 bp. It appears that the model underprediction is constant, and improved predictions can simply be achieved by incorporating a constant in the estimated model. There is of course a question as to whether the constant simply corrects for the misspecification of the structural model. It is interesting in this context that the constant varies across firms - and as a form of testing whether it picks up credit model misspecification how the constant varies in the cross-section will be examined below.

Introducing the on-the-run off-the-run spreads (ONOFF2 and ONOFF30) and the money market fund flows (MMMF), the mean yield error drops to -9 bp. More importantly, the RMSE drops another 5 bp to 20 bp. The average in-sample predicted non-default component is now 52 bp. This appears to indicate that the variables can explain some of the variability of the non-default component.

This can be cast interpreted in terms of an R^2 in the following way: The RMSE of the constant-only model can be interpreted as the square root of the Total Sum of Squares (TSS) divided by the number of observations, the RMSE of the full model can be interpreted as the square root of the Residual Sum of Squares (RSS) divided by the number of observations. A quick calculation shows that the R^2, i.e. the percentage of variation in the difference between the true yield and the yield of the default model

[2]Off-the-run on-the-run spreads at maturity of 5 and 10 years, as well as the retail money market mutual fund flows turned out to be insignificant for our dataset, so we omit presenting the results pertaining to them here. The difference in results when including them is very small.

explained by the proxy variables is 0.35, or 35%. This appears to be rather higher than the R^2s reported e.g. in Delianedis and Geske (2001), which are all in the single digits.

4.5.2 Estimated parameters

The estimated parameters (the σ_η and the γ vector) are presented in table 4.1.

As can be seen, it is clear that the constant is positive and significant - on average, there is a positive non-default component.

The coefficient for ONOFF2 also seems to be positive and significant for most firms, indicating that the liquidity premium in the treasury market tends to move in the same direction as the non-default component (possibly a corporate bond liquidity premium) in corporate bonds.

The evidence for ONOFF30 seems to be more mixed. The median maturity of the corporate bonds examined here is around 7 years, so possibly the liquidity premium in the long end of the treasury market matters less for (shorter maturity) corporate bonds.

Selecting only those bonds which have a maturity larger than 20 years, the mean of the the t-statistics is 3.72, suggesting that the longer-maturity treasury liquidity premium is indeed positively related (on average) to the non-default component in the longer-maturity corporate bonds.

The negative coefficient on the MMMF variable indicates that as money flows into money market funds, the non-default component decreases. Since these mutual funds can and do invest in corporate funds, these inflows represent inflows of cash into the entire bond market. One could interpret the negative coefficient as evidence

	parameter value	t-statistic
σ_η		
1st Quart.	0.1185713	
2nd Quart.	0.1359962	
3rd Quart.	0.1544704	
4th Quart.	0.1860867	
Constant		
1st Quart.	0.003541493	7.137775
2nd Quart.	0.234743061	8.852635
3rd Quart.	0.370158930	10.479838
4th Quart.	0.582756919	11.597603
ONOFF2		
1st Quart.	−1.3366848	−1.4486486
2nd Quart.	−0.2060235	0.5376329
3rd Quart.	0.5477506	2.1353796
4th Quart.	1.7370544	2.9645993
ONOFF30		
1st Quart.	−1.30905615	−1.9023132
2nd Quart.	−0.50104812	−0.4057316
3rd Quart.	0.09100273	0.9478363
4th Quart.	0.96930335	1.9913497
MMMF		
1st Quart.	−0.028574870	−5.915437
2nd Quart.	−0.017750038	−4.989686
3rd Quart.	−0.012228793	−3.663055
4th Quart.	−0.006617398	−1.806942

Table 4.1: Mean parameter estimates (and t-stats) for the different quartiles

that as cash flows into the market, liquidity increases, and the non-default component decreases.

The sign of the coefficient on MMMF is opposite to the sign of the coefficient on a similar variable as reported by Longstaff, Mithal, and Neis (2004). They argue that their variable captures inflows into money market mutual funds and therefore defensive money market positions, and that inflows should therefore be associated with increases in the non-default component. It would be interesting to compare the exact make up of the variables used.

Overall, the results seems to suggest that liquidity premia in corporate bond and treasury markets move together, rather than in opposite directions.

4.5.3 Cross-section properties of estimated parameters

In order to examine how the estimated non-default parameters vary across different types of firms and bonds, the coefficients of the constant, ONOFF2 and MMMF where regressed on firm/bond-specific variables including rating (with 1 being the best rating, and 18 being the worst rating), a dummy that indicates whether a bond is considered non-investment grade, time to maturity, amount outstanding for the bond issue, coupon (as a possible proxy for taxes), leverage, as well as stock market capitalisation (as a measure of the size of the firm).

The results of the regressions are as reported in tables C.1 - C.4 (see appendix C). As can be seen from table C.1, the constant seems to be affected by whether or not the bond is an investment grade bond or not: A positive coefficient with a p-value of 0.0573 indicates that non-investment grade bonds have a large non-default component on average (in absolute terms, this is not relative to the total spread). This could be explained with reference to the institutional restrictions placed on holding non-investment grade bonds, which are likely to reduce the liquidity of such bonds, and hence increase spreads. It is interesting to note that once controlling for whether or not a bond is an investment grade bond, rating has no significant effect. This seems to indicate that the relationship between rating and non-default component is non-linear. Note that this result is similar to the one reported by Longstaff, Mithal, and Neis (2004) - they find that an AAA/AA dummy is significant and negative.

The fact that rating per se and leverage are not significant is reassuring, as it indicates that the model is not misspecified.

One would, however, expect that the amount outstanding and the size of the firm is relevant for the liquidity of the bond being traded and hence the size of the non-default component - there does not appear to be any conclusive evidence that this is the case here.

Longstaff, Mithal, and Neis (2004) also report that there is a positive relationship between maturity and the non-default component, arguing that longer maturity bonds are less liquid. This effect does not appear in the data here.

For the sensitivity of the non-default component to the off-the-run on-the-run spread (ONOFF2), the investment grade status also seems to play a role, the coefficient is positive and significant at the 5% level. It appears that for non-investment grade bonds, the non-default component is more likely to be positively related to the liquidity premium in the treasury market - a rising liquidity premium in the treasury market is more likely to be associated with a rising non-default component for non-investment grade bonds.

None of the other variables appear to be significant.

As can be seen in table C.3, the only variable which appears to have a significant influence on the effect of the 30-year off-the-run on-the-run spread (ONOFF30) is the coupon. The effect of coupons is positive and highly significant. This seems to suggest that the non-default spread component of corporate bonds with higher coupons is more likely to be positively affected by longer maturity treasury market liquidity premium for corporate bonds with high coupons. It is difficult to find an explanation for this result. None of the other cross-sectional variables appear to be relevant for explaining the effect of ONOFF30 on the non-default component.

Table C.4 shows that the effect of money market fund flows again changes depending on whether one is looking at investment grade or non-investment grade bonds, with the coefficient here being negative and significant at the 5% level. For non-investment grade bonds, it appears that money market fund flows are more important, and an inflow of the same size is associated with a drop in the non-default component of a larger size. This is surprising, as one would expect money market funds to be restricted to investing in investment grade securities, one would therefore expect non-investment grade securities to be less affected by fund flows.

It also appears that bonds with higher coupons have a non-default component that is affected to a greater extent by money market fund flows (again, the coefficient is negative and significant at the 5% level). One could conjecture that this might be due to the fact that money market funds are affected less by the tax treatment of coupons than other investors. As can be seen in comparison with table C.1, the coupon does not appear to have an effect on the overall level of the non-default component. This is in contrast to the result presented in Longstaff, Mithal, and Neis (2004).

The regression also indicates that maturity and leverage are relevant - these effects are rather more difficult to explain.

4.6 Concluding remarks

Many structural models appear to underpredict bond spreads. Possible explanations include either that the model predicted default spreads are too low (because default probabilities are underestimated, and payoffs are overestimated), or that there are components in spreads which are not related to default.

This chapter looks at the the second possible explanation, and examines whether the predictive performance can be improved if a non-default component is explicitly accounted for. With a very simple setup, it is shown that

- The predictive performance of the combined model is substantially better than the model that does not account for a non-default component.
- Although a large part of the level of the non-default component seems to be captured here by a firm-specific constant, some of the time variation in the non-default component can be explained by market-wide variables that are proxies for e.g. liquidity in treasury markets, and furthermore, the coefficients do vary across the cross-section in a meaningful way (e.g. whether or not the bonds are non-investment grade does seem to matter).
- Many of the results here are either in line with results of previous research, or in line with expectations.

A point of particular interest is that a part of the non-default component (possibly a liquidity premium) in corporate bond spreads appears to be *positively* related to the treasury market liquidity premium. This lends support to "flight to quality" explanations of liquidity premia in bond markets, and suggests that to some degree off-the-run treasury bonds and corporate bonds behave as complements, not substitutes, and “flights to quality” are really flights towards on-the-run treasury bonds.

Overall, it appears that the non-default component of credit spreads is an interesting area of research, as much of the time variation in the non-default component is still not explained. In future, it may be possible to improve the estimate of the credit component by including more information on assets (such as bond prices at a daily frequency, CDS spreads) in estimation, for example via the procedure described in chapter 3. Also, it would be interesting to find more variables that appear to be related to this non-default component - both time-varying firm specific variables as well as market-wide variables. Possibly, this can inform the theoretical discussion of why and how spreads can contain non-default components.

Chapter 5

Conclusion

Structural models would be a natural choice for pricing/ hedging in situations where an explicit link to equity prices is desirable, e.g. when a hedge is to be executed by buying and selling equity, or when models need to be estimated and bond pricing data is relatively scarce.

Unfortunately, structural models have often been avoided in these kind of applications because of a perceived inability to price assets with a reasonable degree of precision. In particular, it appears that a lot of structural models underpredict spreads. Models that impose less structure (so-called reduced-form models) are often used instead.

There are several facets to understanding the problems of structural models in explaining the data.

As was argued in chapter 3, part of the problem up until now seems to have been insufficiently precise estimation methods. A method, based on non-linear filtering, allowing for the inclusion of prices of more than one asset class was presented. The method has the advantage that prices of more than one asset can be used in estimation, and that it is not necessary to assume that any one asset is priced perfectly by the model. It was shown that when estimating structural models, including even relatively small amounts of bad quality bond price data in addition to equity data can produce large increases in performance in the sense that out-of-sample spreads are predicted more accurately. Possibly, part of the perceived problem with structural models is to do with inaccurate estimation methods.

Since the traditional Merton (1974) model appears to underpredict spreads, an important response to attempting to improve the performance of structural models is to provide models that produce a higher credit spread. In chapter 2, an argument was put forward that a model incorporating coordination failures between creditors can increase spreads. It was argued that in a lot of cases, creditors are unlikely to act in a coordinated manner, and that since they often have incentives to attempt to enforce claims on the firm at the expense of other creditors, races to do so are possible - even in legal regimes that have relatively strict bankruptcy arrangements such as the US. It is obvious that this has an effect on the price of debt. It was argued that the effect is such that the resulting continuous-time bond pricing model can be seen to have the original Merton (1974) and Black and Cox (1976) models as limiting cases, depending on the parameter that governs the incentives to attempt to grab assets: The legal costs of such actions. It was also argued that this effect can produce a discount in bond prices (and hence higher spreads). Although this is a convincing story as to how bankruptcy happens, and how strategic interaction has an effect on the price of debt, simply estimating the model on bond prices is unlikely to differentiate it from other models that produce similar barriers and payoffs. It was instead argued that an explicit empirical test can be based on the level of the barrier and the fraction of assets recovered in case of bankruptcy, rather than on the bond price.

When implemented on bond prices, models need to be tractable. As was argued throughout this thesis, models must be able to price coupon bonds, and must treat equity as a perpetuity. An alternative to modelling the typically very complicated debt structure of a firm, including differences in seniority of different layers of debt, as well as differences in maturities, and covenants, embedded optionality and complicated repayment conditions, is to choose a simple model for aggregate debt that treats it as a perpetuity, and to price the debt claim in question as a negligible proportion of aggregate debt. For most corporate bond issues, this is a reasonable assumption. Naturally, this needs to be combined with the properties of a first passage model, allowing for default before maturity of the infinite maturity aggregate debt. Although different arguments can be made as to what the chosen barrier should be, the prices of any such model are likely to be very similar in practice.

Lastly, chapter 4 argued that possibly, the underprediction of spreads by structural models is not necessarily a sign of producing default probabilities that are too low, or payoffs that are too high, but that there are components in the spread which are not related to default, as suggested by recent empirical research (Longstaff, Mithal, and Neis 2004, Delianedis and Geske 2001). Instead of attempting to produce models that exhibit a higher default component of spreads, by e.g. including jumps, reducing information sets, lowering barriers and raising bankruptcy costs, or introducing coordination failures, it might be necessary to expand models by explicitly including a non-default component of spreads. Chapter 4 presented a very simple model of a non-default component, and estimated the combined model on data. From the results, it appears to be the case that this improves out-of-sample forecasting performance, and produces parameter estimates with reasonable, and indeed interesting signs and magnitudes.

Structural models have been avoided in pricing and hedging applications, but it is clear that after some reconsideration, they might still prove valuable and indeed necessary in these kind of applications.

Appendix A

Pricing coordination failure

A.1 Solution of the discrete time model

A.1.1 Basic procedure

We follow the same procedure as Morris and Shin (2004a) to solve the model. Suppose that agents follow a switching strategy around a certain posterior belief. Given the posterior belief around which agents switch, it is possible to derive the fraction of them that will foreclose, given the asset value in the next period (posterior beliefs will be centred around this asset value in the next period). We can therefore work out what the critical next-period asset value is for which the firm will fail, given the belief in this period around which agents switch.

Also, we can use the fact that agents will switch if they believe that they will obtain a higher utility from doing so. Once we have defined utilities, this allows us to derive the critical posterior belief, given a critical next period asset value for which the firm fails.

So we have two equations in two unknowns, which can then be solved for the critical asset value for which the firm fails - the trigger point.

A.1.2 Information

For convenience, the assumptions about information are restated here. The relative increase in the asset value is normally distributed around a drift.

$$V_{t+\Delta} - V_t = \mu_V V_t \Delta + V_t \eta_t, \quad \eta_t \sim NID\left(0, \frac{1}{\alpha}\right) \tag{A.1}$$

Agents receive a signal X_i (subscript i indexes the different agents) about this increase with a distribution conditional on the asset value V_t given by

$$X_i = V_{t+\Delta} + V_t \varepsilon_i, \quad \varepsilon_i \sim NID\left(0, \frac{1}{\beta}\right), \tag{A.2}$$

with $Cov(\eta_t, \varepsilon_i) = 0$, i.e. the noise is orthogonal to the innovations in the fundamental.

A.1.3 Posteriors

From the signal X_i and the public information V_t, agents form a posterior about the value of the firm in period $t + \Delta$, $V_{t+\Delta}$ which is normal with mean and variance given by

$$\rho_i = E[V_{t+\Delta}|X_i] = \frac{\alpha(1 + \mu_V \Delta)V_t + \beta X_i}{\alpha + \beta} \tag{A.3}$$

and

$$Var(V_{t+\Delta}|X_i) = \frac{(V_t)^2}{\alpha + \beta}. \tag{A.4}$$

A.1.4 Critical value of $V_{t+\Delta}$ for which the firm fails

Given the posterior belief around which agents switch, we work out how many of them will foreclose, given the asset value in the next period (posterior beliefs will be centred around this asset value in the next period). We then work out what the critical next-period asset value is for which the firm fails, given the belief in this period around which agents switch.

Suppose agents follow a switching strategy around ρ^*, i.e. agents foreclose when their posterior is below ρ^*. Then an agent will not foreclose if and only if the private signal is bigger than

$$X^* = \frac{\alpha+\beta}{\beta}\rho^* - \frac{\alpha}{\beta}(1+\mu_V\Delta)V_t. \tag{A.5}$$

Conditional on state $V_{t+\Delta}$, the distribution of X_i is normal with mean $V_{t+\Delta}$ and precision $\frac{\beta}{V_t^2}$. So the ex-ante probability for any agent of foreclosing is equal to

$$\Phi\left\{\frac{1}{V_t}\sqrt{\beta}(X^* - V_{t+\Delta})\right\}. \tag{A.6}$$

where Φ is the cumulative standard normal density function.

As the number of agents tends to infinity, the fraction of agents that foreclose will be equal to this ex ante probability for any individual agent by the law of large numbers.

Since the firm fails if the fraction that forecloses is

$$l \geq \frac{V_{t+\Delta}}{L}, \tag{A.7}$$

the critical value of $V_{t+\Delta}$ (denoted by $V^*_{t+\Delta}$) for which the firm fails at t is given by

$$V^*_{t+\Delta} = L\Phi\left\{\frac{1}{V_t}\sqrt{\beta}\left(X^* - V^*_{t+\Delta}\right)\right\} \tag{A.8}$$

or

$$V^*_{t+\Delta} = L\Phi\left\{\frac{1}{V_t}\left(\frac{\alpha}{\sqrt{\beta}}\left(\rho^* - (1+\mu_V\Delta)V_t\right) + \sqrt{\beta}\left(\rho^* - V^*_{t+\Delta}\right)\right)\right\}. \tag{A.9}$$

A.1.5 Utility

Consumption in any intermediate period $t+q$ per unit of aggregate debt is as described in the payoff matrix in the main text, reproduced here for convenience:

	firm reorganised (R)	firm not reorganised ($\neg R$)
grab assets (G)	$(1-K)V_t$	$-KV_t$
do not grab assets ($\neg G$)	0	0

Let ξ denote the fraction of aggregate debt that a creditor holds, then consumption in the case the agent decides to grab assets and the firm is reorganised is

$$c_{t+q}(G, R) = (1-K)V_t\xi, \tag{A.10}$$

consumption in the case the agent decides to grab assets, but the firm is not reorganised is

$$c_{t+q}(G, \neg R) = -KV_t\xi, \tag{A.11}$$

and consumption in the case the agent decides not to grab assets, whether or not the firm is reorganised is

$$c_{t+q}(\neg G, R) = c_{t+q}(\neg G, \neg R) = c_{t+q}(\neg G) = 0. \qquad \text{(A.12)}$$

The utility function is additively separable across time. Letting $u\left(c_{t+q}(V_{t+\Delta})\right)$ denote Bernoulli utility at time $t+q$ (this only stresses the dependences on $V_{t+\Delta}$, but of course the utility will depend on the actions of the agents as well as on whether or not the firm fails), which will depend on the actions taken at that time, and letting $U_{t+\Delta}$ denote future utility and letting δ denote a discount factor, we can write the agents utility function as

$$EU_{t+q} = E\left[u\left(c_{t+q}(V_{t+\Delta})\right) + \delta_{t+q} U_{t+q}\left(V_{t+\Delta}\right)\right] \qquad \text{(A.13)}$$

The expectation is taken under the agent's belief.

A.1.6 Critical value of ρ

We now use the fact that agents will switch if they believe that they will obtain a higher utility from doing so. We then derive the critical posterior belief, given a critical next period asset value for which the firm is reorganised.

Now the marginal agent (one that is indifferent between forcing reorganisation or not) has a posterior over the asset value which has its mean just at the switching point (i.e. ρ for this agent is equal to ρ^*). For her the expected utility of not grabbing assets should just equal the expected utility of grabbing assets. This *defines* the switching point. Use F to denote the posterior cumulative distribution (given the belief) over the asset value $V_{t+\Delta}$. We can write:

$$\begin{aligned} \int_{-\infty}^{V^*_{t+\Delta}} u\left(c_{t+q}(G,R)\right) dF + & \int_{V^*_{t+\Delta}}^{\infty} u\left(c_{t+q}(G,\neg R)\right) dF \\ + & \delta_{t+q} \int_{-\infty}^{\infty} U_{t+\Delta} dF \\ = \qquad 0 \; + & \delta_{t+q} \int_{-\infty}^{\infty} U_{t+\Delta} dF \qquad \text{(A.14)} \end{aligned}$$

In effect, the stage games are similar to the prisoners' dilemma. Since the game ends whenever cooperation is not achieved, any equilibria based on punishing strategies are infeasible. This implies that the sequential game is actually turned into a sequence of one-shot prisoners' dilemma games. Future utility will not depend on today's actions, and agents cannot use strategies that specify future actions other than the solutions to the one-stage prisoners' dilemma games. Future utility will not be related to today's actions, and will therefore be the same, whether or not the agent decides to grab assets. Also note that the payoff is in terms of V_t, which does not depend on $V_{t+\Delta}$. We can therefore write:

$$u\left(c_{t+q}(G,R)\right) \Pr\left(V_{t+\Delta} < V^*_{t+\Delta}\right) + u\left(c_{t+q}(G,\neg R)\right) \Pr\left(V_{t+\Delta} > V^*_{t+\Delta}\right) = 0 \qquad \text{(A.15)}$$

We can write this probability as:

$$\Pr\left(V_{t+\Delta} > V^*_{t+\Delta}\right) = \Phi\left\{\frac{\sqrt{\alpha+\beta}}{V_t}\left(\rho^* - V^*_{t+\Delta}\right)\right\} \qquad \text{(A.16)}$$
$$\left(= 1 - \Pr\left(V_{t+\Delta} < V^*_{t+\Delta}\right)\right).$$

where Φ denotes the cumulative normal density function.

We insert this and rearrange to obtain

$$\rho^* - V^*_{t+\Delta} = \frac{V_t}{\sqrt{\alpha - \beta}} \Phi^{-1} \left\{ \frac{u\left(c_{t+q}(G,R)\right)}{u\left(c_{t+q}(G,R)\right) - u\left(c_{t+q}(G,\neg R)\right)} \right\} \tag{A.17}$$

Insert consumption as specified above, and take limits as the number of agents goes to infinity. This will imply that the fraction of the loan held by any individual agent goes to zero, $\xi_t \to 0$. Note that we have a fraction of functions of ξ_t, and can apply l'Hopital's rule. In the limit, all c_{t+q} are equal, so

$$\lim_{\xi_t \to 0} \frac{u\left(c_{t+q}(G,R)\right)}{u\left(c_{t+q}(G,R)\right) - u\left(c_{t+q}(G,\neg R)\right)} = \tag{A.18}$$

$$\frac{u'(c_{t+q})(1-K)V_t}{u'(c_{t+q})(1-K)V_t - u'(c_{t+q})(-K)V_t} = 1 - K. \tag{A.19}$$

So the limit of our equation is

$$\rho^* - V^*_{t+\Delta} = \frac{V_t}{\sqrt{\alpha + \beta}} \Phi^{-1} \left\{1 - K\right\}. \tag{A.20}$$

This equation together with (A.9) pins down the critical value of beliefs and the asset value.

A.1.7 Equilibrium forced reorganisation

Combining equations (A.20) and (A.9) we can solve for the failure point at which the asset value in the next period causes failure in this period:

$$V_{t+\Delta}^{*} = L\Phi\left\{\frac{\alpha}{\sqrt{\beta}}\left(\frac{V_{t+\Delta}^{*}}{V_t} - 1 - \mu_V\Delta\right) + \frac{\sqrt{\alpha+\beta}}{\sqrt{\beta}}\Phi^{-1}\{1-K\}\right\} \tag{A.21}$$

Reorganisation at time $t+q$ will occur when V hits V^* at $t+\Delta$.

A.1.8 Uniqueness

To simplify notation, define

$$Z = \frac{\alpha}{\sqrt{\beta}}\left(\frac{V_{t+\Delta}^{*}}{V_t} - 1 - \mu_V\Delta\right) + \frac{\sqrt{\alpha+\beta}}{\sqrt{\beta}}\Phi^{-1}\{1-K\} \tag{A.22}$$

and

Condition A.I. $L\frac{1}{\sqrt{2\pi}}\frac{\alpha}{\sqrt{\beta}}\frac{1}{V_t} < 1$

Proposition 1. *The trigger point $V_{t+\Delta}^{*}$ is unique if condition (A.I) is satisfied.*

Proof. This is a version of the proof in Morris and Shin (2004a). A sufficient condition for a unique solution is that the slope of

$$L\Phi\{Z\} \tag{A.23}$$

is less than one everywhere. This slope is equal to

$$L\varphi\{Z\}\frac{\alpha}{\sqrt{\beta}}\frac{1}{V_t}. \tag{A.24}$$

It reaches a maximum where the argument of the normal density is 0, the maximum there will be $\frac{1}{\sqrt{2\pi}}$. Hence a sufficient condition for a unique solution is that

$$L\frac{1}{\sqrt{2\pi}}\frac{\alpha}{\sqrt{\beta}}\frac{1}{V_t} < 1. \tag{A.25}$$

■

A.2 Uncertainty in the limit

It can be shown that the marginal or pivotal agent views the fraction of banks that attempt to foreclose as a random variable that is uniformly distributed in the continuous-time limit, and hence that strategic uncertainty remains. Note that these kind of results have been discussed at length elsewhere (Morris and Shin 2002).

Proposition 2. *The distribution of l given the belief ρ^* of the marginal agent is uniform in the limit.*

Proof. The proportion of people who receive a lower signal X^* is

$$l = \Phi\left\{\frac{\sqrt{\beta}}{V_t}\left(X^* - V_{t+\Delta}\right)\right\}. \tag{A.26}$$

The question to ask is: What is the probability that a fraction less than z of the other bondholders receive a signal higher than that of the marginal agent, conditional on the marginal agent's belief, or what is

$$\Pr\left((1-l) < z \mid \rho^*\right)? \tag{A.27}$$

Now the event

$$1 - l < z \tag{A.28}$$

is equivalent to

$$1-\Phi\left\{\frac{\sqrt{\beta}}{V_t}\left(X^*-V_{t+\Delta}\right)\right\}<z \tag{A.29}$$

or (rearranging)

$$V_{t+\Delta}<X^*+\frac{V_t}{\sqrt{\beta}}\Phi^{-1}\{1-z\}. \tag{A.30}$$

So the probability we are looking for is

$$\Pr\left(V_{t+\Delta}<X^*+\frac{V_t}{\sqrt{\beta}}\Phi^{-1}\{1-z\}\mid\rho^*\right). \tag{A.31}$$

The posterior of the marginal agent over $V_{t+\Delta}$ has mean ρ^* and variance $\frac{V_t^2}{\alpha+\beta}$, hence this probability is

$$\Pr\left((1-l)<z\mid\rho^*\right)=\Phi\left\{\frac{\sqrt{\alpha+\beta}}{V_t}\left(X^*+\frac{V_t}{\sqrt{\beta}}\Phi^{-1}\{1-z\}-\rho^*\right)\right\}. \tag{A.32}$$

Now as we take limits, $\rho^*\to X^*$, since private information becomes infinitely more precise than public information (the agent attaches all weight to the signal and none to the mean of the prior), and $\frac{\sqrt{\alpha+\beta}}{\sqrt{\beta}}\to 1$. It follows that

$$\Pr\left((1-l)<z\mid\rho^*\right)=1-z, \tag{A.33}$$

or

$$\Pr\left(l<1-z\mid\rho^*\right)=1-z. \tag{A.34}$$

So the cumulative distribution of l is the identity function, which implies that the density will be uniform. ■

A.3 Pricing the bond

To be more comparable to the existing literature, assume that the drift of the asset value is defined as $\mu_V = r + \lambda\sigma - \gamma$, i.e. it is equal to the risk free rate, plus a market price of risk component, minus a cash payout rate.

A.3.1 The default process

Define the default process

$$Y_t \equiv \frac{1}{\sigma} \ln \left(\frac{V_t}{V^*} \right). \tag{A.35}$$

Default occurs at time τ, when $Y_\tau = 0$. For an asset value process of

$$dV_t = (r + \lambda\sigma - \gamma)V dt + \sigma V dW, \tag{A.36}$$

the dynamics under the objective measure $\mathbb{P}$ of Y_t are (by Itô's lemma):

$$dY_t = \left(\frac{r + \lambda\sigma - \gamma - \frac{1}{2}\sigma^2}{\sigma} \right) dt + dW \tag{A.37}$$

The dynamics under the measure $\mathbb{Q}$, defined by the money market account as the numeraire are

$$dY_t = \left(\frac{r - \gamma - \frac{1}{2}\sigma^2}{\sigma} \right) dt + d\hat{W} \tag{A.38}$$

The dynamics under the measure $\mathbb{V}$, defined by the asset value as the numeraire are

$$dY_t = \left(\frac{r - \gamma + \frac{1}{2}\sigma^2}{\sigma} \right) dt + d\hat{\hat{W}} \tag{A.39}$$

Denote the drifts of Y_t under these measure as $\mu_Y^{\mathbb{P}}$, $\mu_Y^{\mathbb{Q}}$, $\mu_Y^{\mathbb{V}}$.

A.3.2 The first passage density

The first passage density for Y_t under a measure $\mathbb{M}$ is

$$f^{\mathbb{M}}\left(Y_t; s\right) = \frac{Y_t}{\sqrt{2\pi(s-t)^3}} exp\left\{-\frac{1}{2}\left(\frac{Y_t + \mu^{\mathbb{M}}(s-t)}{\sqrt{s-t}}\right)^2\right\} ds \tag{A.40}$$

By completing the square it can be shown that

$$\int_t^{\infty} e^{-\rho(s-t)} f^{\mathbb{M}}\left(Y_t; s\right) ds = exp\left\{-Y_t\left(\sqrt{\left(\mu^{\mathbb{M}}\right)^2 + 2\rho} + \mu^{\mathbb{M}}\right)\right\} \tag{A.41}$$

as long as $\left(\mu^{\mathbb{M}}\right)^2 + 2\rho \geq 0$.

A.3.3 Perpetual dollar-in-default claim

The price of a perpetual dollar-in-default claim (a claim that pays 1 in the case of default) $G(V_t, t)$ is given by

$$G(V_t, t) = E^{\mathbb{Q}}\left[e^{-r(\tau - t)}\right] \tag{A.42}$$

Using (A.41) and (A.38) it is easy to show that

$$G(V_t, t) = \left(\frac{V_t}{V^*}\right)^{-\theta} \tag{A.43}$$

where

$$\theta = \frac{\sqrt{\left(\frac{r-\gamma-\frac{1}{2}\sigma^2}{\sigma}\right)^2 + 2r} + \frac{r-\gamma-\frac{1}{2}\sigma^2}{\sigma}}{\sigma} \tag{A.44}$$

Define the measure $\mathbb{G}$ by the numeraire G. It can be shown that the drift of Y_t under $\mathbb{G}$ is

$$\mu_Y^{\mathbb{G}} = \mu_Y^{\mathbb{Q}} - \theta\sigma \tag{A.45}$$

A.3.4 Survival probabilities

The survival probabilities $Q_t^{\mathbb{M}}(\tau > T)$ (probabilities of the event $\tau > T$ at t under the measure $\mathbb{M}$, $\mathbb{M} = \{\mathbb{Q}, \mathbb{G}\}$ are given by

$$Q_t^{\mathbb{M}}(\tau > T) = \Phi\left(k^{\mathbb{M}}\left(\frac{V_t}{V^*}\right)\right) - \left(\frac{V_t}{V^*}\right)^{-\frac{2}{\sigma}\mu_Y^{\mathbb{M}}} \Phi\left(k^{\mathbb{M}}\left(\frac{V^*}{V_t}\right)\right) \tag{A.46}$$

where

$$k^{\mathbb{M}} = \frac{\ln(x) + \sigma\mu_Y^{\mathbb{M}}(T-t)}{\sigma\sqrt{T-t}} \tag{A.47}$$

and Φ denotes the cumulative standard normal distribution function.

A.3.5 The finite-maturity dollar-in-default claim

The price of this claim can be written as

$$G\left(V_t, t | \tau < T\right) = G(V_t, t) - e^{-r(T-t)} E^{\mathbb{Q}}\left[G(V_t, T) I_{\{\tau > T\}}\right] \tag{A.48}$$

We can rewrite this expression as

$$\begin{aligned} G\left(V_t, t | \tau < T\right) = & \quad G(V_t, t) - e^{-r(T-t)} E^{\mathbb{Q}}\left[G(V_t, T)\right] E^{\mathbb{G}}\left[I_{\{\tau > T\}}\right] & \text{(A.49)} \\ = & \quad G(V_t, t)\left(1 - Q^{\mathbb{G}}(\tau > T)\right) & \text{(A.50)} \end{aligned}$$

A.3.6 The down-and-out binary option

The price of this claim is simply given by

$$H(V_t, t; T) = e^{-r(T-t)} Q_t^{\mathbb{Q}}(\tau > T) \tag{A.51}$$

A.3.7 Probability of survival and being in-the-money

The probabilities of the event

$$A = \{\tau > T, V_T > Z\} \tag{A.52}$$

under measures $\mathbb{M} = \{\mathbb{Q}, \mathbb{V}, \mathbb{G}\}$ are given by

$$Q^{\mathbb{M}}(A) = \Phi\left(k^{\mathbb{M}}\left(\frac{V_t}{Z}\right)\right) - \left(\frac{V_t}{V^*}\right)^{-\frac{2}{\sigma}\mu_Y^{\mathbb{M}}} \Phi\left(k^{\mathbb{M}}\left(\frac{(V^*)^2}{ZV_t}\right)\right) \tag{A.53}$$

A.3.8 The price of a down-and-out call

Let $C_{DO}(V_t, t, ; T, Z)$ denotes the price of a (European) down-and-out call with strike price Z, maturity T, with a barrier of V^*. A formula for this was first derived by Merton (1973). Using (A.53), this can be written as

$$C_{DO}(V_t, t; T, Z) = V_t e^{-\gamma(T-t)} Q^{\mathbb{V}}(A) - e^{-r(T-t)} Q^{\mathbb{Q}}(A) \tag{A.54}$$

A.3.9 Pricing under assumption 2.A: (Bond represents entire debt)

As mentioned in the main text, under assumption 2.A, the bond can be seen as a portfolio of two down-and-out calls and the dollar-in-boundary claim:

$$\begin{aligned} B(V_t, t) =& C_{DO}(V_t, t; 0) - C_{DO}(V_t, t; D) \\ &+ (1 - K)V^* G(V_t, t, |\tau < T) \end{aligned} \tag{A.55}$$

A.3.10 Pricing under assumption 2.B: (Bond represents negligible proportion of debt)

Under assumption 2.B a plain vanilla bullet bond with coupon c paid at times $\{t_i : i = 1, \cdots M\}$, for which a fraction R of the principal is recovered in default can be priced (in terms of dollars per dollar face value) as follows:

$$\begin{aligned} B(V_t, t) =& \sum_{i=1}^{M} cH(V_t, t; t_i) \\ &+ H(V_t, t; T) \\ &+ RG(V_t, t, |\tau < T) \end{aligned} \tag{A.56}$$

Appendix B

Simulated maximum likelihood estimation of structural models

B.1 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

$$\begin{aligned} \alpha_t &= d + \alpha_{t-1} + \eta_t \\ \tilde{y}_t &= \dot{Z}_t(\tilde{\alpha}_t)\alpha_t + \varepsilon_t \end{aligned} \tag{B.1}$$

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

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

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

ing 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.2 Description of selected issuers and bonds

	Notional ($m)	Coupon	Time to maturity (years)
Min.	11.0	5.250	0.620
1st Qu.	250.0	6.450	3.475
Median	475.0	7.362	5.525
Mean	601.5	7.349	8.593
3rd Qu.	750.0	8.188	8.418
Max.	3250.0	9.375	27.250

Table B.1: Summary of chosen bonds

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 B.2: Issuers chosen for empirical analysis

issue_id†	Issuer name	Notional ($m)	Coupon	Maturity*
2565	Boeing Co	300	6.350	1.48
3027	CSX Corp	300	9.000	4.64
6019	Dow Chem Co	300	9.000	19.27
6023	Dow Chem Co	150	7.375	21.19
10507	GTE Corp	500	9.375	0.67
18941	Ralston Purina Co	200	9.250	8.05
23782	Weyerhaeuser Co	150	8.375	5.13
25750	Enron Corp	150	7.125	5.66
25799	Viacom Inc	1000	7.750	3.45
34274	Union Pac Corp	250	6.400	4.10
46294	Coca Cola Enterprises Inc	250	6.950	24.89
51170	Archer Daniels Midland Co	350	7.500	25.22
52337	Penney J C Inc	425	7.600	5.28
54434	Norfolk Southn Corp	750	7.350	5.39
57203	Philip Morris Cos Inc	1000	7.000	3.55
58864	Raytheon Co	1000	6.450	0.62
71458	Adelphia Communications Corp	325	9.250	1.04
72800	General Mtrs Corp	600	6.750	26.35
72845	GTE Corp	450	6.360	6.05
81220	Procter & Gamble Co	750	5.250	1.71
81287	International Business Machs Corp	185	5.370	1.73
83731	Dayton Hudson Corp	200	5.875	6.87
87654	Ford Mtr Co Del	1500	6.375	27.11
89620	Lucent Technologies Inc	1360	6.450	27.25
98345	Wal Mart Stores Inc	3250	6.875	7.62
100417	Charter Communications Hldgs Llc	600	8.250	5.26
101679	Adelphia Communications Corp	500	9.375	8.16
101895	Lockheed Martin Corp	1000	8.200	7.98
102919	Lowes Cos Inc	11	8.150	2.45
104014	Disney Walt Co	750	7.300	3.11
104722	Honeywell Intl Inc	1000	7.500	8.17
106726	Deere & Co	500	7.850	8.38
107814	Lowes Cos Inc	500	8.250	8.43
114386	Consolidated Edison Co NY Inc	350	6.625	3.97
114499	Worldcom Inc	1000	7.375	4.08
117184	Hertz Corp	500	7.400	9.98
123033	Union Pac Corp	200	5.840	2.41
131630	Emerson Elec Co	250	5.750	9.84

*: Time to maturity (years) on date of last bond price observation

†: issue_id is the unique identifier in Mergent's FISD

Table B.3: Bonds chosen for empirical analysis

Appendix C

The non-default component of corporate yield spreads

C.1 Cross-sectional regressions

	Estimate	Std. Error	t value	$Pr(>\|t\|)$	signif.
(Intercept)	$-2.741e-02$	$4.645e-01$	-0.059	0.9531	
Rating	$4.331e-01$	$2.252e-01$	1.923	0.0573	.
D_NI	$8.291e-03$	$4.164e-02$	0.199	0.8426	
Maturity	$1.172e-02$	$9.777e-03$	1.199	0.2333	
Outstanding	$-7.059e-08$	$6.162e-07$	-0.115	0.9090	
Coupon	$4.755e-02$	$5.411e-02$	0.879	0.3816	
Leverage	$-1.160e-01$	$5.090e-01$	-0.228	0.8202	
Size	$4.018e-07$	$2.152e-06$	0.187	0.8522	

Residual standard error: 0.6637 on 100 degrees of freedom
Multiple R-Squared: 0.0921, Adjusted R-squared: 0.02854
F-statistic: 1.449 on 7 and 100 DF, p-value: 0.1943
Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Table C.1: Regression of Constant on cross-sectional variables

	Estimate	Std. Error	t value	Pr(> $\|t\|$)	signif.
(Intercept)	$-1.229e+00$	$2.036e+00$	-0.604	0.5474	
Rating	$2.019e+00$	$9.873e-01$	2.045	0.0434	*
D_NI	$3.175e-01$	$1.825e-01$	1.639	0.1012	
Maturity	$4.516e-02$	$4.286e-02$	1.054	0.2946	
Outstanding	$1.639e-07$	$2.702e-06$	0.061	0.9517	
Coupon	$6.466e-02$	$2.372e-01$	0.273	0.7857	
Leverage	$-3.041e+00$	$2.232e+00$	-1.363	0.1760	
Size	$7.037e-06$	$9.434e-06$	0.746	0.4574	

Residual standard error: 2.91 on 100 degrees of freedom
Multiple R-Squared: 0.1635, Adjusted R-squared: 0.105
F-statistic: 2.792 on 7 and 100 DF, p-value: 0.01073
Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Table C.2: Regression of ONOFF2 on cross-sectional variables

	Estimate	Std. Error	t value	Pr(> $\|t\|$)	signif.
(Intercept)	-3.702$e+00$	$1.405e+00$	-2.634	0.009774	**
Rating	-3.622$e-02$	$6.813e-01$	-0.053	0.957707	
D_NI	$3.993e-02$	$1.260e-01$	0.317	0.751936	
Maturity	-1.251$e-02$	$2.958e-02$	-0.423	0.673254	
Outstanding	-1.122$e-06$	$1.864e-06$	-0.602	0.548746	
Coupon	$6.108e-01$	$1.637e-01$	3.731	0.000317	***
Leverage	-2.683$e-01$	$1.540e+00$	-0.174	0.862071	
Size	-3.729$e-06$	$6.510e-06$	-0.573	0.568071	

Residual standard error: 2.008 on 100 degrees of freedom
Multiple R-Squared: 0.1561, Adjusted R-squared: 0.09706
F-statistic: 2.643 on 7 and 100 DF, p-value: 0.01504
Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Table C.3: Regression of ONOFF30 on cross-sectional variables

	Estimate	Std. Error	t value	Pr(> $\|t\|$)	signif.
(Intercept)	$1.014e-02$	$1.572e-02$	0.645	0.5204	
Rating	$1.538e-03$	$1.406e-03$	1.094	0.2765	
D_NI	$-1.527e-02$	$7.166e-03$	-2.130	0.0356	*
Maturity	$5.390e-04$	$3.111e-04$	1.732	0.0863	.
Outstanding	$2.317e-08$	$1.961e-08$	1.182	0.2402	
Coupon	$-3.813e-03$	$1.722e-03$	-2.214	0.0291	*
Leverage	$-2.866e-02$	$1.670e-02$	-1.716	0.0893	.
Size	$-2.335e-08$	$6.977e-08$	-0.335	0.7386	

Residual standard error: 0.02112 on 99 degrees of freedom
Multiple R-Squared: 0.1955, Adjusted R-squared: 0.1304
F-statistic: 3.006 on 8 and 99 DF, p-value: 0.004619
Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Table C.4: Regression of MMMF on cross-sectional variables

Bibliography

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, 1996, "Design and Valuation of Debt Contracts," *Review of Financial Studies*, 9, 37–68.

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.

Atkeson, A., 2000, "Comment on: Rethinking Multiple Equilibria in Macroeconomic Modelling (Morris, S., Shin, H.)," *NBER Macroeconomics Annual.*

Baird, D., and T. Jackson, 1990, *Cases, Problems and Materials on Bankruptcy*, Little Brown and Company, Boston MA USA.

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

Black, F., and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–654.

Briys, E., and F. de Varenne, 1997, "Valuing Risky Fixed Rate Debt: An Extension," *Journal of Financial and Quantitative Analysis*, 31, 239–248.

Collin-Dufresne, P., and R. Goldstein, 2001a, "The Determinants of Credit Spread Changes," *Journal of Finance*, 56, 2177–2208.

Collin-Dufresne, P., and R. Goldstein, 2001b, "Do Credit Spreads Reflect Stationary Leverage Ratios?," *Journal of Finance*, 56, 1929–1957.

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.

Duffie, D., and D. Lando, 2001, "Term Structures of Credit Spreads with Incomplete Accounting Information," *Econometrica*, 69, 633–664.

Duffie, D., and K. Singleton, 1999, "Modelling Term Structures of Defaultable Bonds," *Review of Financial Studies*, 12, 687–720.

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.

Elton, E., M. Gruber, D. Agrawal, and C. Mann, 2001, "Explaining the Rate Spread on Corporate Bonds," *Journal of Finance*, 56, 247–277.

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.

Gatev, E., and P. Strahan, 2003, "Banks' advantage in hedging liquidity risk: theory and evidence from the commercial paper market," NBER working paper 9956.

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.

Geske, R., 1979, "The Valuation of Compound Options," *Journal of Financial Economics*, 7, 63–81.

Giesecke, K., 2004, "Correlated Default with Incomplete Information," *Journal of Banking and Finance*, 28, 1521–1545.

Giesecke, K., and L. Goldberg, 2004, "Forecasting Default in the Face of Uncertainty," *Journal of Derivatives*, 12, 14–25.

Hilberink, B., and L. Rogers, 2002, "Optimal Capital Structure and Endogenous Default," *Finance and Stochastics*, 6, 237 – 263.

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.

Jackson, T., 1986, *The Logic and Limits of Bankruptcy*, Harvard University Press, Cambridge MA 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.

Jones, P., S. Mason, and E. Rosenfeld, 1985, "Contingent Claims Analysis of Corporate Capital Structures: Theory and Empirical Tests," in B. Friedman (ed.), *Corporate Capital Structures in the United States* . pp. 239–261, University of Chicago Press, Chicago, USA.

Kim, J., K. Ramaswamy, and S. Sundaresan, 1993, "Does Default Risk in Coupons Affect the Valuation of Corporate Bonds?: A Contingent Claims Model," *Financial Management*, 22, 117–131.

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, 2004, "Corporate Yield Spreads: Default Risk or Liquidity. New Evidence from the Credit-Default Swap Market," UCLA Working Paper.

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.

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

Mella-Barral, P., and W. Perraudin, 1997, "Strategic Debt Service," *Journal of Finance*, 52, 531–556.

Merton, R., 1973, "The theory of rational option pricing," *Bell Journal of Economics and Management Science*, 4.

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, 2000, "Rethinking Multiple Equilibria in Macroeconomic Modelling," *NBER Macroeconomics Annual*, pp. 139–161.

Morris, S., and H. Shin, 2002, "Global Games: Theory and Applications," in M. Dewatripont, L. Hansen, and S. Turnovsky (ed.), *Advances in Economics and Econometrics, the Eighth World Congress*, Cambridge University Press.

Morris, S., and H. Shin, 2004a, "Co-ordination Risk and the Price of of Debt," *European Economic Review*, 48, 113–153.

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

Neustadter, G., 2002, "U.C.C. Article 9," Electronic Casebook: http://www.scu.edu/law/FacWebPage/Neustadter/article9/.

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.

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.

Zhou, C., 2001, "The Term Structure of Credit Spreads with Jump Risk," *Journal of Banking and Finance*, 25, 2015–2040.