

Monte Carlo Valuation of Corporate Bonds using Price Data

by

Nimesh Patel

A thesis
submitted to Victoria University of Wellington
in fulfilment of the
requirements for the degree of
Master of Commerce and Administration
in Money and Finance

Victoria University of Wellington

2010

ACKNOWLEDGEMENTS

I would like to thank Dr. Toby Daglish and Dr. John Randal for their unwaivering support and wisdom which was integral to the completion of this thesis.

Contents

1	Introduction	1
2	Dynamic Models of the Term Structure	7
2.1	Bonds and Interest Rates	7
2.2	Vasicek (1977)	10
2.3	Cox, Ingersoll and Ross (1985)	11
3	Corporate Bond Pricing Models	15
3.1	Structural Model Overview	15
3.1.1	Why Risk-Neutrality?	17
3.1.2	Merton (1974)	20
3.1.3	Black and Cox (1976)	22
3.1.4	Briys and de Varenne (1997)	26
3.2	Reduced-Form Model Overview	29
3.2.1	CIR Intensity	32

4	Review of the Empirical Literature	35
4.1	The Theory	35
4.2	The Traditional Method in Practice	38
4.3	Simulation Tests	42
4.4	The ML approach in Practice	44
4.5	The Best Method?	45
5	Approach to Estimation	47
5.1	Overview	47
5.2	Fitting the Short Rate Models	48
5.2.1	Parameter Estimation	48
5.2.2	Risk-Neutral vs Real-World	51
5.3	The ‘Portfolio of Zeroes’ Approach	54
5.4	Monte Carlo Overview	56
5.5	The Log-Likelihood Function	57
5.6	Parameter Estimation	60
5.7	Extending to a Stochastic Interest Rate	66
5.8	Survival Probabilities under Portfolio of Zeroes	71
5.9	Extension to Internal Consistency	73

5.10	Intensity Model Estimation Overview	75
5.10.1	CIR Intensity	77
5.11	Asymptotic Variance	80
5.12	One-Step Ahead Forecasting	81
5.13	Calculating Risk-Free Yields	83
6	Case Study	85
6.1	Overview	85
6.2	Summary Statistics	87
6.2.1	Bills Times Series Data	87
6.2.2	The Corporate Bond	88
7	Methodological issues	91
7.1	Overview	91
7.2	Root Finding Speed	92
7.3	Jagged Surfaces	93
7.4	Non-Existence of Roots	99
7.5	Parameter-Specific Non-Existence of Roots	100
7.6	Two Roots	107
7.7	CIR Non-Existence of Roots	110

8	Putting the Methodology into Practice	113
8.1	Models of the Term Structure	113
8.2	Structural Corporate Bond Pricing Models	117
8.2.1	Parameter Estimates	117
8.2.2	Predictive Performance	123
8.3	CIR Intensity Model	137
8.3.1	Parameter Estimates	137
8.3.2	Predictive Performance	139
9	Conclusion	143
A	Corporate Theory Appendix	149
A.1	Probability of Zero Default Events	149
A.2	Replicating Portfolio Weights	149
A.3	The Black and Cox (1976) Bond Price	150
A.4	Equity Delta under Black and Scholes	163
B	Notational Index	165
	Bibliography	165

Chapter 1

Introduction

Corporate debt securities play a large part in financial markets and hence accurate modeling of the prices of these securities is integral. Ericsson and Reneby (2005) state that the corporate bond market in the US doubled between 1995 and 2005 and is now larger than the market for US treasuries. Although the theoretical corporate bond pricing literature is vast, very little empirical research to test the effectiveness of these models has been published.

Corporate bond pricing models are split into two families of models. The first, are the structural models which endogenise default by modeling it as an event that may eventuate due to the insolvency of the underlying firm. The second family of models is the newer class of reduced-form models that exogenise default by modeling it as some random process (default intensity). The reduced-form models have been formulated largely due to the empirical failures of the structural family to accurately model prices and spreads. However as Ericsson and Reneby (2005) point out, an inadequate estimation approach may explain the poor performance of the structural

models.

Structural models are, therefore, the focus of this paper. We, however, do estimate a reduced-form model in order to make a comparison between the two types of model. There are no published papers (to my knowledge) in which both types of model are implemented.

The Merton (1974) model is the first in the structural model class and an enormous array of models have been since been formulated. This assortment of models includes the Black and Cox (1976), Geske (1977), Longstaff and Schwartz (1995), Leland and Toft (1996) and Briys and de Varenne (1997) models amongst many others. The Merton model makes the unrealistic assumption that default can only occur at the maturity of the bond. This assumption allows the use of the well-known Black and Scholes (1973) option pricing formula to price equity as an option on the firm's assets and thus the value the bond as the difference between the firm value and equity. Black and Cox attempt to rectify this restrictive assumption by allowing for safety covenants in the bond's indenture which allows bondholders to force the firm into default prior to maturity. Specifically the firm defaults if the total value of the firm falls below some default threshold. Briys and de Varenne extend this research even further by allowing for deviations from absolute priority. Their model recognises that bankruptcy costs arise with default and thus the bondholders may not receive the full value of the firm when default occurs. We choose to implement the Merton, Black and Cox, and Briys and de Varenne models to test for improvements in accuracy from adding default barriers and deviations from strict priority.

The problem with implementing structural models is that the value of the firm and its volatility are unobserved variables. There exist two main approaches to esti-

imating these inputs. The empirical studies conducted thus far (excluding Li and Wong (2008)) use the traditional approach proposed by Ronn and Verma (1986) or some variant of it. This method involves calculating firm values and their volatilities through a system of two equations based on the Black and Scholes (1973) option pricing formulae. However Duan (1994) argues that this method is misspecified and instead proposes an alternate maximum likelihood (ML) approach for estimating firm value and its parameters using equity data in the context of deposit insurance.

Ericsson and Reneby (2005) test both methods on simulated data and conclude that the ML method of Duan (1994) is far superior to the traditional method. In fact, they go on to say that the theoretical failures of the traditional method may explain the poor performance of structural models in the empirical literature. Li and Wong (2008) apply both the traditional and ML approaches to actual data and find the performance of structural models is greatly improved.

A downside of the approaches employed in the empirical literature so far is that they require equity data to be observable in order to apply the structural models and thus have no use for pricing bonds of private firms. We present an approach that employs the Duan (1994) method using bond price data and can hence be applied to firms that are private in addition to those that are publicly traded. Furthermore, we apply the models under different interest rate assumptions through the use of Monte Carlo simulation which removes the need for closed-form solutions. By doing so, we can gauge whether or not good/poor performance is due to the corporate bond pricing model assumptions or the interest rate assumption. We use the Monte Carlo approach to calibrate firm values from observed bond prices and evaluate the

log-likelihood function.

We apply the models under the two different treatments of coupon bonds. The models discussed only consider zero-coupon bonds and thus we require a means of pricing coupon bonds since these are more commonly observed in markets. We use the ‘portfolio of zeroes’ (PZ) approach outlined in Eom, Helwege and Huang (2004) where each coupon is treated as an independent zero-coupon bond. This method is simple to apply but has the weakness that the default of earlier coupons does not imply the default of later coupons. A more consistent approach is to price each coupon conditional on the survival of earlier payments. This would be difficult to implement if we wish to use closed-form solutions. However, with Monte Carlo simulation (as used in this paper), this extension is straightforward. Along a simulation path, whenever a coupon defaults, we set the value of all remaining coupons (and the principal) to 0.

In implementing these models, we encounter several problems in calibrating firm values with the Monte Carlo approach. In addition, the ML procedure, as stated in Li and Wong (2008) is very slow computationally. Therefore we present results for one bond only. This bond is that of a private firm and hence the firm does not have any observable equity data. Focusing on one bond allows us to study the estimation issues that arise. These issues are problems that must be overcome for the Monte Carlo approach to be reliable and hence discussing these issues is more important than reporting results for multiple bonds.

As with Ericsson and Reneby (2005) and Li and Wong (2008), our results suggest that the structural corporate bond pricing models vastly improve in performance under the ML approach. The mean spread errors are greatly reduced from that of

Eom, Helwege and Huang (2004). In fact, we find that the CIR intensity reduced-form model exhibits greater variability in spread errors and hence, has much less forecasting potential. These results suggest that the structural models have performed poorly in the empirical literature due to the problems associated with the traditional estimation approach of Ronn and Verma (1986) and have been dismissed without proper treatment.

The outline of the paper is as follows. Chapter 2 outlines the riskless term structure literature relevant to the paper. Chapter 3 outlines the corporate bond pricing models and Chapter 4 discusses the empirical literature to date. We then discuss the estimation approach in Chapter 5. Chapter 6 summarises the corporate bond and the treasury data used in term structure estimation, Chapter 7 is a study of the estimation issues and Chapter 8 reports our results. We then conclude the paper in Chapter 9. For convenience of reference, Appendix B contains a notational index with a list of terms commonly used throughout the paper. It is recommended that readers familiarise themselves with the notation for bond prices in the paper.

A risky bond price at t under a structural model is denoted as $\overline{B}_t(V_t, F, T, c)$ where V_t is the firm value at t , T is the maturity, F is the face value and c is the coupon rate. We use $c = 0$ to denote a zero-coupon bond.

A risky bond price at t under a reduced-form model is denoted as $\overline{B}_t(\lambda_t, F, T, c, \delta)$ where λ_t is the default intensity (hazard rate) at t and δ is the face value fractional recovery.

The price of a risk-free bond at t is denoted as $B_t(T, F, c)$.

Chapter 2

Dynamic Models of the Term Structure

2.1 Bonds and Interest Rates

In this section we study the risk-free term structure literature relevant to this paper.¹ We first need to understand what the short rate is and why it is important. We define $r_t(T)$ as the continuously compounded zero-coupon bond yield. It is the rate of interest that is used to compute the price $B_t(T, 1, 0)$, of a risk-free bond that promises \$1 at time T . i.e. the following relation holds,

$$B_t(T, 1, 0) = e^{-r_t(T)T^*} \quad T > t \quad (2.1)$$

where T^* is the time to maturity $T - t$. The short rate is the continuously compounded zero-coupon bond yield for an infinitesimally small period of time. That

¹We refer to ‘risk-free’ as meaning free of default risk.

is, it is the rate r_t where,

$$r_t = \lim_{T \rightarrow t} r_t(T)$$

We can model the riskless term structure of interest rates through a one-factor model. There is only one source of uncertainty in a one-factor model and the in models we consider, the source is the evolution of the short rate.²

The price of a risk-free zero-coupon bond paying \$1 at time T can be valued at t using the following expectation.

$$B_t(T, 1, 0) = E_{\mathbb{Q}} \left[\exp \left\{ - \int_t^T r_u du \right\} \right] \quad (2.2)$$

where r_u is the level of interest that accrues over the period between u and $u+du$ and \mathbb{Q} is the risk neutral probability measure (Harrison and Pliska (1981)). Therefore, we can use a process for the short rate to determine risk-free zero-coupon bond prices for various maturities and these prices can then be used to compute zero-coupon bond yields for these maturities. i.e. rearranging (2.1) gives,

$$r_t(T) = -\frac{1}{T-t} B_t(T, 1, 0) \quad T > t \quad (2.3)$$

There are two types of term structure model; equilibrium models and no-arbitrage models. The equilibrium family of models includes the Vasicek (1977), Dothan (1978), Brennan-Schwartz (1980) and Cox, Ingersoll and Ross (1985) (commonly known as CIR) models among others. The difference between equilibrium and no-arbitrage models is that the term structure of interest rates is an output in the equilibrium models whereas it is used as an input in the no-arbitrage models.³ Equi-

²An obvious weakness of only allowing for one factor is that all bonds have one source of uncertainty and thus changes in bond prices due to interest rate changes are perfectly correlated. In reality there is no reason for this to be true.

³Popular no-arbitrage models include Ho and Lee (1986) and Hull and White (1990).

librium models derive a process for the short rate based on behaviour observed in the market (e.g. mean reversion of interest rates) which then implies zero coupon bond prices and consequently the term structure. The problem is that the modelled term structure is unlikely to match the current term structure of interest rates. This will likely cause problems for investors who use the model. Positions in bond options are usually combined with a position in the bond itself to hedge the risk involved in the option. However this hedge will not work if the model does not correctly price the bond underlying the option. No-arbitrage models address this issue by taking the initial term structure as an input and using this data to calibrate the short rate model to match the term structure. This is enabled by allowing the coefficients of the short rate model to depend on time.

However implementing no-arbitrage models involves calibrating the coefficients in order to match the term structure and these parameters can change frequently. Choudhry (2004) argues that equilibrium models have the advantage of capturing long-term behaviour of interest rates thus dampening the effects of market distortions apparent in data. This is direct contrast to no-arbitrage models which on a given day, can have parameters that may be heavily affected by these types of distortions. For this reason, and also the fact that our main focus is modeling default risk, we restrict our attention to the Vasicek and CIR equilibrium term structure models. Our focus is not matching the initial term structure but rather simply allowing for a varying interest rate in pricing defaultable bonds.

The models we consider are of the form,

$$dr_t = a(r_t)dt + b(r_t)dW_t \tag{2.4}$$

where $a(r_t)$ and $b(r_t)$ are deterministic functions that depend on the short rate. dW_t is a Brownian motion where,

$$dW_t \sim N(0, dt)$$

The process (2.4) is a continuous time Markov process. The Markov property implies that the past of the short rate process is irrelevant in determining the future and only the current level of the short rate matters. Under the term structure models we consider, the price of a risk-free bond only depends on the current level of the short rate.⁴

Under the models we consider, closed-form solutions exist for the price of a risk-free zero-coupon bond. This price takes the form,

$$B_t(T, 1, 0) = \alpha(t, T)e^{-\beta(t, T)r_t} \quad (2.5)$$

where $\alpha(t, T)$ and $\beta(t, T)$ are functions of time and the parameters of the model. From this price, zero-coupon yields can be calculated using (2.3).

2.2 Vasicek (1977)

The Vasicek (1977) model is a commonly used model of the term structure. It assumes the short rate r_t is normally distributed and takes the following form,

$$dr_t = a(\mu_r - r_t)dt + \sigma_r dW_t \quad a, \mu_r, \sigma_r > 0 \quad (2.6)$$

⁴Some multi-factor models also allow the volatility of the short rate to be stochastic. See Fong and Vasicek (1991) and Longstaff and Schwartz (1992).

where a , μ_r and σ_r are constant. The volatility of the short rate is σ_r whilst the long run level for the short rate is μ_r and r_t is pulled to this level at rate a . This property is known as mean reversion. If r_t is below μ_r then the drift is positive and the short rate will tend to increase toward its long run level. If r_t is above μ_r then the drift is negative and the short rate will tend to decrease toward its long run level.

A closed form solution for zero-coupon bond prices exists under the Vasicek model. The zero-coupon bond price takes the form of (2.5) where,

$$\alpha(t, T) = \exp \left(\frac{(\beta(t, T) - T^*)(2a^2\mu_r - \sigma_r^2)}{2a^2} - \frac{\sigma_r^2\beta(t, T)^2}{4a} \right) \quad (2.7)$$

$$\beta(t, T) = \frac{1 - e^{-aT^*}}{a} \quad (2.8)$$

As we have said, the bond prices under the model can then be used to derive the term structure of interest rates. The Vasicek model allows the term structure to take a variety of shapes. It can be upward sloping, downward sloping or even hump-shaped.

Although the Vasicek model is quite popular it does present a problem due to the normal distribution assumption which can allow interest rates to become negative. Next, we consider the Cox, Ingersoll and Ross (1985) model which addresses the negativity issue.

2.3 Cox, Ingersoll and Ross (1985)

The Cox, Ingersoll and Ross (1985) model (CIR) is one of the most popular equilibrium models for the term structure and corrects the negative interest rate problem of the Vasicek model by assuming that the absolute volatility of innovations in

the short rate is proportional to $\sqrt{r_t}$. With the level of the short rate included in the volatility, the distribution is no longer normal but rather has a non-central chi-squared distribution. The model takes the form,

$$dr_t = a(\mu_r - r_t)dt + \sigma_r\sqrt{r_t}dW_t$$

With the $\sqrt{r_t}$ term included, the variance of the change in the short rate is proportional to the short rate itself. Therefore as the short rate approaches zero, the variance approaches zero also.

$$\begin{aligned}\text{Var}(dr_t) &= \text{Var}[a(\mu_r - r_t)dt + \sigma_r\sqrt{r_t}dW_t] \\ &= \sigma_r^2 r_t \text{Var}(dW_t) \\ &= \sigma_r^2 r_t dt\end{aligned}$$

where the second equality holds by independence. It follows that,

$$\lim_{r_t \rightarrow 0} \sigma_r^2 r_t dt = 0$$

As the short rate approaches zero, the random variation disappears and the positive drift drives the rate upwards preventing it from becoming negative. Furthermore, the upward drift is strong enough to prevent the short rate from becoming exactly zero if $2a\mu_r \geq \sigma_r^2$ (see Feller (1951)). In addition, the drift term of the CIR model is the same as that of the Vasicek model and hence the model also captures mean reversion.

As with the Vasicek model, there exists a closed form solution for zero coupon bond prices with the CIR model. The price of a risk-free zero-coupon bond under this

model takes the form (2.5) where $\eta = \sqrt{a^2 + 2\sigma_r^2}$ and,

$$\alpha(t, T) = \left(\frac{\eta e^{\frac{1}{2}(a+\eta)T^*} \beta(t, T)}{e^{\eta T^*} - 1} \right)^{\frac{2a\mu_r}{\sigma_r^2}} \quad (2.9)$$

$$\beta(t, T) = \frac{2(e^{\eta T^*} - 1)}{(a + \eta)(e^{\eta T^*} - 1) + 2\eta} \quad (2.10)$$

Empirical evidence suggests that the volatility parameter σ_r of the short rate should also be modeled using a stochastic process. Although many models have been formulated in an attempt to explain term structure volatility, there is no clear agreement on which approach is most accurate (see Chapman and Pearson (2001)). We do not consider models which assume a stochastic interest rate volatility parameter since this would require assumptions on what model of stochastic volatility is the best to use. Since our main focus is modeling default risk, any needless assumptions on interest rate volatility structure may cloud our results. Therefore we choose to use the Vasicek and CIR models since they provide a simple way to include a stochastic interest rate into our analysis.

Chapter 3

Corporate Bond Pricing Models

3.1 Structural Model Overview

We now discuss the risky bond pricing and background literature required to understand the results of this paper. Models from both the structural and reduced-form classes of corporate bond pricing are implemented. These two types of risky bond pricing models treat default in very different ways. Structural models base default on the insolvency of the underlying firm and thus are also known as firm-value models. In these models, all of the securities of a particular firm are modelled as contingent claims on the firm's assets. The first of such models is proposed by Merton (1974) in which the equity of a firm can be valued as a European call option on the assets of the firm using the Black and Scholes (1973) option pricing formula. In this model the firm has a single senior zero-coupon debt issue in addition to equity and default can only occur at its maturity. A vast array of models have since been formulated including the Black and Cox (1976), Geske (1977), Longstaff and Schwartz (1995),

Leland and Toft (1996) and Briys and de Varenne (1997) models among others.

In this paper we discuss and implement the Merton, Black and Cox, and Briys and de Varenne models from the firm-value class. The Black and Cox model extends Merton's framework by allowing for the possibility of early default by accounting for safety covenants in the bond's indenture. In this model bondholders can force the firm into bankruptcy if the total value of the firm's assets falls below a default barrier. Briys and de Varenne extend the analysis further by formulating a model that includes interest rate risk and deviations from absolute priority. Bondholders are unlikely to receive all of the assets of the firm in the event of default due to factors such as bankruptcy costs, so accounting for deviations from absolute priority is a realistic assumption to which we pay close attention. Allowing for interest rate risk is also a more realistic assumption than a constant interest rate since in reality, interest rates are anything but constant. As with Black and Cox, the Briys and de Varenne model also includes a default barrier and for this reason these models are also known as *first-passage* models since default occurs when the firm value first falls below the barrier.

Before we discuss the models individually, we need to take a small detour. We first develop the notion of risk-neutrality and its use in pricing securities. This is done through the derivation of a partial differential equation (PDE), developed in Merton (1974), which is free of preferences. The absence of preferences tells us that the price of a security that satisfies the PDE should also be free of preferences. Therefore investors with varying risk preferences should agree on this price. Since this price is consistent with any set of risk preferences, we can price securities in the risk-neutral world. This does not mean we are assuming investors are risk-neutral. All we are saying is that pricing in the risk-neutral world results in the correct price.

3.1.1 Why Risk-Neutrality?

Risk-neutral pricing is the method used in the paper. Therefore in this section we discuss why we price securities in this manner. Pricing a security in the risk-neutral world involves calculating the expected payoff of the security as if all investors are risk-neutral and then discounting back to the current time period using the risk-free rate. This method stems from a no-arbitrage argument. We develop this argument by going through the derivation of the parabolic partial differential equation derived in Merton (1974). Throughout we assume a constant interest rate r and assume the firm value V_t follows a geometric Brownian motion process under the real world probability measure \mathbb{P} ,

$$dV_t = \mu V_t dt + \sigma V_t dW_t$$

where μ is the instantaneous expected rate of return of the firm, σ is the instantaneous volatility and dW_t is a Brownian motion. We now define a contingent claim Y_t whose value depends on its maturity T and the value V_t of the underlying firm. We can write the value of this security as,

$$Y_t = f_t(V_t, T)$$

The dynamics of the security take the form,

$$dY_t = \mu_Y Y_t dt + \sigma_Y Y_t dW_t \tag{3.1}$$

Since the value of the security depends on the firm value we know that both the security and the firm value depend on the same Brownian motion. We can derive the process specifically using the process for the firm value and applying Itô's lemma.

Itô's lemma states that the process followed by the security's value is,

$$dY_t = \left(\frac{df}{dt} + \mu V_t \frac{df}{dV} + \frac{1}{2} \sigma^2 V_t^2 \frac{d^2 f}{dV^2} \right) dt + \sigma V_t \frac{df}{dV} dW_t$$

Comparing this to (3.1) we can see that,

$$\mu_Y Y_t = \frac{df}{dt} + \mu V_t \frac{df}{dV} + \frac{1}{2} \sigma^2 V_t^2 \frac{d^2 f}{dV^2} \quad (3.2)$$

$$\sigma_Y Y_t = \sigma V_t \frac{df}{dV} \quad (3.3)$$

The partial differential equation is derived by forming a portfolio π consisting of W_1 dollars in the firm's assets, W_2 dollars in the security in question and W_3 dollars in riskless bonds. We set W_3 equal to $-(W_1 + W_2)$ so that the portfolio does not require any net investment. In other words a short position is taken in riskless bonds to finance the two long positions. It is here assumptions that an investor can buy any amount of security they wish and that short sales are allowed are important. If these did not hold, then the investor may be unable to set up the portfolio. The return on the portfolio can be written as,

$$d\pi_t = W_1 \frac{dV_t}{V_t} + W_2 \frac{dY_t}{Y_t} + W_3 r dt \quad (3.4)$$

since $dB_t = rB_t dt$ where B_t is the price of a riskless bond at t . Simplifying (3.4) further gives,

$$\begin{aligned} d\pi_t &= W_1(\mu dt + \sigma dW_t) + W_2(\mu_Y dt + \sigma_Y dW_t) - W_1 r dt - W_2 r dt \\ &= (W_1(\mu - r) + W_2(\mu_Y - r))dt + (W_1\sigma + W_2\sigma_Y)dW_t \end{aligned} \quad (3.5)$$

If W_1 and W_2 are chosen so that $W_1\sigma + W_2\sigma_Y$ is zero, the stochastic term in (3.5) disappears and the return on the portfolio becomes riskless. Added to the fact that

the net investment in the portfolio is zero, the portfolio must have a return of zero to negate any possibility of arbitrage. Therefore we have,

$$\begin{aligned} W_1(\mu - r) + W_2(\mu_Y - r) &= 0 \\ W_1\sigma + W_2\sigma_Y &= 0 \end{aligned}$$

and we only obtain non-zero weights when,

$$\left(\frac{\mu - r}{\sigma} \right) = \left(\frac{\mu_Y - r}{\sigma_Y} \right) \quad (3.6)$$

A proof of (3.6) is provided in Appendix B. Taking (3.2) and (3.3) and substituting into (3.6) we obtain,

$$\begin{aligned} \left(\frac{\mu - r}{\sigma} \right) &= \frac{\frac{1}{f_t} \left(\frac{df}{dt} + \mu V_t \frac{df}{dV} + \frac{1}{2} \sigma^2 V_t^2 \frac{d^2 f}{dV^2} \right) - r}{\frac{1}{f_t} \sigma V_t \frac{df}{dV}} \\ &= \frac{\frac{df}{dt} + \mu V_t \frac{df}{dV} + \frac{1}{2} \sigma^2 V_t^2 \frac{d^2 f}{dV^2} - r f_t}{\sigma V_t \frac{df}{dV}} \end{aligned}$$

Finally rearranging and simplifying gives us,

$$\frac{df}{dt} + r V_t \frac{df}{dV} + \frac{1}{2} \sigma^2 V_t^2 \frac{d^2 f}{dV^2} - r f_t = 0 \quad (3.7)$$

We have obtained a partial differential equation that the price of a security (whose value depends on firm value and time) must satisfy. It is clear that (3.7) is not affected by the market price of risk $(\mu - r)$ and is therefore free of preferences. This tells us that we can use any set of risk preferences when pricing contingent claims. One set of preferences is risk-neutrality and therefore any security whose value depends on firm value and time can be priced in the risk-neutral world. Under this setup, any security regardless of how risky it is will grow at r since investors are

indifferent to risk. Thus the expected payoff of a security, no matter how risky, can be calculated in the risk-neutral world and then be discounted back to the current time period using the risk-free rate. This simple yet powerful result derived from a no-arbitrage argument serves as the basis for pricing corporate bonds in this entire paper. We refer to the real world probability measure as \mathbb{P} and the risk-neutral world probability measure as \mathbb{Q} . An expectation of a variable X under the real world is denoted as $E_{\mathbb{P}}(X)$ and the expectation under the real neutral world is denoted as $E_{\mathbb{Q}}(X)$.

We are now equipped with the tools we need to consider each model individually.

3.1.2 Merton (1974)

Merton (1974) proposes the first in the structural model family and his work has led to the formulation of a large number of models from many authors. The model assumes that the interest rate is constant at r thus allowing the use of the well known Black and Scholes (1973) pricing formula used to price European options. In this model it is assumed there are two claims to the firm's assets. The first is a zero-coupon bond maturing at T , with face value F and value $\bar{B}_t(V_t, T, F, 0)$ at time t . The second and residual claimant is the equity which has value $S_t(V_t)$ at time t . The firm value is assumed to follow a geometric Brownian motion process under \mathbb{Q} ,

$$dV_t = rV_t dt + \sigma V_t dW_t \tag{3.8}$$

Default in the Merton model can only occur at the maturity T of the bond. If at time T , $V_T < F$, then the firm defaults, the bondholders receive V_T and the shareholders receive nothing. If instead, $V_T \geq F$, then the bondholders receive F

and the shareholders receive $V_t - F$.

Therefore the terminal payoffs are,

$$\overline{B}_T(V_T, T, F, 0) = \min(V_T, F) \quad (3.9)$$

$$S_T(V_T) = \max(0, V_T - F) \quad (3.10)$$

From this we can see that the value of the equity at the maturity of the bond is the same as the terminal value of a European call option on the firm assets with strike price F and maturity T . Therefore the value of the equity at t is,

$$S_t(V_t) = V_t \Phi(d_1) - F e^{-rT^*} \Phi(d_2) \quad (3.11)$$

where,

$$d_1 = \frac{\ln(V_t/F) + (r + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \quad (3.12)$$

$$d_2 = d_1 - \sigma\sqrt{T^*} \quad (3.13)$$

and $\Phi(\cdot)$ is the standard normal CDF. The equations (3.11) - (3.13) used to price the equity are the well known Black and Scholes option pricing formulae and satisfy (3.7). Clearly the result is useful since it provides us with a closed-form solution to price the equity of the firm. This in turn allows us to price the zero-coupon bond. The price of the bond is given as the difference between firm value and equity.

$$\begin{aligned} \overline{B}_t(V_t, T, F, 0) &= V_t - S_t(V_t) \\ &= V_t - [V_t \Phi(d_1) - F e^{-rT^*} \Phi(d_2)] \\ &= V_t(1 - \Phi(d_1)) + F e^{-rT^*} \Phi(d_2) \end{aligned}$$

$\Phi(d_2)$ is interpreted as the probability of survival up to maturity. Although we assume that the firm does not pay dividends to shareholders, the model can be easily extended to include a continuous dividend yield.

3.1.3 Black and Cox (1976)

Black and Cox (1976) extend the analysis of Merton (1974) and provide a closed-form solution to price risky zero-coupon bonds where default can occur at any time up to the bond's maturity. The model considers the effect of including a safety covenant in the indenture of the bond which gives the bond a floor value by allowing bondholders to force the firm into bankruptcy if the firm is underperforming. Specifically a barrier is defined where bondholders force the firm into bankruptcy when the barrier is crossed. Again, the bond pays no coupons, has face value F and maturity T . The interest rate is once again assumed to be constant at r and the dynamics of the firm value in the risk-neutral world are now assumed to be,

$$dV_t = (r - \alpha)V_t dt + \sigma V_t dW_t$$

where α is the continuous dividend rate paid to shareholders.¹ We define τ as the time of default and have three possible cases scenarios concerning default.

1. $t < \tau < T$
2. $\tau = T$
3. $T < \tau$ (i.e. no default)

¹The α term can also be included in the firm value for the Merton model since the Black and Scholes option pricing formula can also account for continuous dividend yields. We do not consider dividends in this paper and hence we will assume α to be zero when pricing.

The first case refers to default before maturity and is when the firm value drops below the barrier level leading to forced bankruptcy of the firm. In this case the bondholders receive the total assets of the firm (which is equal to the level of the default barrier) and the shareholders receive nothing. The second is when the firm defaults at the maturity of the debt. Here the value of the firm is above the barrier but not high enough to pay the face value of the bond. Again the bondholders seize control of the firm and the shareholders receive nothing just as the first case. The third case denotes no default where the bondholders receive the full face value of the bond at maturity whilst the shareholders receive the remaining assets. Therefore we again have the boundary conditions (3.9) and (3.10) when default does not occur prior to maturity. When default occurs prior to maturity, the payoffs depend on the level of the default barrier since default is triggered when the firm value crosses this barrier. The barrier K_t at time t in this model is,

$$K_t = De^{-\gamma T^*} \quad D, \gamma \geq 0 \quad (3.14)$$

where D and γ are constant. Therefore in the case of early default the payoffs to the bondholders and shareholders are,

$$\begin{aligned} \overline{B}_\tau(De^{-\gamma(T-\tau)}, T, F, 0) &= De^{-\gamma(T-\tau)} \\ S_\tau(De^{-\gamma(T-\tau)}) &= 0 \end{aligned}$$

since default is triggered when $V_\tau = De^{-\gamma(T-\tau)}$. Under this setup the value of the bond is,

$$\begin{aligned}\overline{B}_t(V_t, T, F, 0) &= E_{\mathbb{Q}} \left[F e^{-rT^*} \mathbb{I}_{\{\tau \geq T, V_T \geq F\}} \right] \\ &+ E_{\mathbb{Q}} \left[V_T e^{-rT^*} \mathbb{I}_{\{\tau \geq T, V_T < F\}} \right] \\ &+ E_{\mathbb{Q}} \left[D e^{-\gamma(T-\tau)} e^{-r(\tau-t)} \mathbb{I}_{\{t < \tau < T\}} \right]\end{aligned}$$

and we hence (see Appendix B) obtain the following closed-form solution for the bond price.

$$\begin{aligned}\overline{B}_t(V_t, T, F, 0) &= F e^{-rT^*} \left(\Phi(h_1) + R_t^{2\tilde{a}} \Phi(h_2) \right) \\ &+ V_t e^{-\alpha T^*} \left(\Phi(h_3) - \Phi(h_4) + R_t^{2\tilde{a}+2} (\Phi(h_5) - \Phi(h_6)) \right) \\ &+ V_t \left(R_t^{\theta+\zeta} \Phi(h_7) + R_t^{\theta-\zeta} \Phi(h_8) \right)\end{aligned}$$

where

$$\begin{aligned}
\tilde{a} &= \frac{r - \alpha - \gamma - \frac{1}{2}\sigma^2}{\sigma^2} \\
R_t &= \frac{De^{-\gamma T^*}}{V_t} \\
\theta &= \tilde{a} + 1 \\
\zeta &= \sigma^{-2} \sqrt{(r - \alpha - \gamma - \frac{1}{2}\sigma^2)^2 + 2\sigma^2(r - \gamma)} \\
h_1 &= \frac{\ln(V_t/F) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
h_2 &= \frac{2\ln(De^{-\gamma T^*}) - \ln(V_t F) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
h_3 &= \frac{\ln(F/V_t) - (r - \alpha + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
h_4 &= \frac{\ln(D/V_t) - (r - \alpha + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
h_5 &= \frac{2\ln(De^{-\gamma T^*}) - \ln(V_t F) + (r - \alpha + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
h_6 &= \frac{2\ln(De^{-\gamma T^*}) - \ln(V_t D) + (r - \alpha + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
h_7 &= \frac{\ln(De^{-\gamma T^*}/V_t) + \zeta\sigma^2 T^*}{\sigma\sqrt{T^*}} \\
h_8 &= \frac{\ln(De^{-\gamma T^*}/V_t) - \zeta\sigma^2 T^*}{\sigma\sqrt{T^*}}
\end{aligned}$$

In this paper we consider a bond of a firm with a simple capital structure (only one bond), although this model could be extended to firms with more than one bond (see Bielecki and Rutkowski (2002)).

A problem with the Black and Cox model is that the situation where bondholders prefer default may arise. Consider $D = F$ and $\gamma < r$. In this situation, the default is barrier at any point before maturity is greater than the discounted face value. Therefore, if default occurs before maturity, bondholders will receive a payoff that can be reinvested at r to yield a higher payoff than F at time T . This payoff is

higher the earlier default occurs. The Briys and de Varenne model corrects this issue.

3.1.4 Briys and de Varenne (1997)

Briys and de Varenne (1997) further develop the firm value model literature by allowing for interest rate risk. The model is an extension of the Black and Cox model in which the interest rate and firm value evolve according to the following stochastic processes respectively under \mathbb{Q} ,

$$\begin{aligned} dr_t &= a(t)(\mu_r(t) - r_t)dt + \sigma_r(t)dW_{1,t} \\ dV_t &= r_t V_t dt + \sigma V_t \left[\rho dW_{1,t} + \sqrt{1 - \rho^2} dW_{2,t} \right] \end{aligned}$$

where $a(t)$, $\mu_r(t)$ and $\sigma_r(t)$ are deterministic functions of time allowing the model to be consistent with the initial riskless term structure (i.e. it is a no-arbitrage model) and ρ is the correlation between the interest rate and firm value. The bond in question is the same as that defined in the Merton, and Black and Cox models and the default barrier is now assumed to be,

$$K_t = \alpha F B_t(T, 1, 0)$$

where $0 \leq \alpha \leq 1$. The default barrier is the discounted face value of the bond scaled down by α . When α is zero, we have the Merton case of no default barrier. When α is one, the bond is riskless. Therefore we can use α to measure the strength of the safety covenant. It get stronger as α approaches one. The issue of bondholders receiving more than the value of a risk-free bond under the Black and Cox model does not exist in the Briys and de Varenne model since at most, the barrier is the

discounted value of the bond face value. Therefore, bondholders will never receive a payoff higher than that of a risk-free bond. As a result, they do not prefer default and are at best, indifferent to it.

In addition to a stochastic interest rate, the Briys and de Varenne model also accounts for deviations from the absolute priority rule. Bondholders receive a fraction f_1 ($0 \leq f_1 \leq 1$) of the firm's assets when default occurs early and receive a fraction f_2 ($0 \leq f_2 \leq 1$) of the firm's assets when default occurs at maturity. Therefore the model accounts for the fact that bankruptcy is not likely to be costless and that these costs reduce the amount of the firm's assets that the bondholders receive once default occurs. When the absolute priority rule does hold, $f_1 = f_2 = 1$ and the bondholders receive the entire firm upon default.

Again we define τ as the default time and hence we have the same scenarios as we have under the Black and Cox model.

1. $t < \tau < T$
2. $\tau = T$
3. $T < \tau$ (i.e. no default)

When default occurs prior to maturity, the bondholders receive $f_1 \alpha F B_\tau(T, 1, 0)$ at τ since the firm value has hit the barrier. When default occurs at maturity the bondholders receive $f_2 V_T$ since at T the firm value is below F but above the barrier αF . The third case is when default does not occur and the bondholders receive F at maturity. Therefore the payoff at maturity of the bond is,

$$\bar{B}_T(V_T, T, F, 0) = (f_1 \alpha F) \mathbb{I}_{\{t < \tau < T\}} + (f_2 V_T) \mathbb{I}_{\{\tau \geq T, V_T < F\}} + (F) \mathbb{I}_{\{\tau \geq T, V_T \geq F\}} \quad (3.15)$$

The first term of (3.15) is expressed in terms of payoff at time T since in the case of early default, the time τ payoff to bondholders can be invested at the risk-free rate until time T . It follows that the price of the bond at time t is,

$$\bar{B}_t(V_t, T, F, 0) = E_{\mathbb{Q}} \left[\exp \left\{ - \int_t^T r_u du \right\} \bar{B}_T(V_T, T, F, 0) \right]$$

which yields the closed-form solution,²

$$\begin{aligned} \bar{B}_t(V_t, T, F, 0) &= FB_t(T, 1, 0) \left[1 - P_E(l_t, 1) + P_E \left(q_t, \frac{l_t}{q_t} \right) \right. \\ &\quad - (1 - f_1) l_t \left(\Phi(-d_3) + \frac{\Phi(-d_4)}{q_t} \right) \\ &\quad \left. - (1 - f_2) l_t \left(\Phi(d_3) - \Phi(d_1) + \frac{\Phi(d_4) - \Phi(d_5)}{q_t} \right) \right] \end{aligned}$$

where,

$$\begin{aligned} d_1 &= \frac{\ln(l_t) + \Sigma(T^*)/2}{\sqrt{\Sigma(T^*)}} = d_2 + \sqrt{\Sigma(T^*)} \\ d_3 &= \frac{\ln(q_t) + \Sigma(T^*)/2}{\sqrt{\Sigma(T^*)}} = d_4 + \sqrt{\Sigma(T^*)} \\ d_5 &= \frac{\ln \frac{q_t^2}{l_t} + \Sigma(T^*)/2}{\sqrt{\Sigma(T^*)}} = d_6 + \sqrt{\Sigma(T^*)} \\ \Sigma(T^*) &= \int_t^T \left((\rho\sigma + \sigma_B(t, T))^2 + (1 - \rho^2)\sigma^2 \right) dt \\ l_t &= \frac{V_t}{FB_t(T)} \\ q_t &= \frac{V_t}{\alpha FB_t(T)} \\ P_E(l_t, 1) &= -l_t \Phi(-d_1) + \Phi(-d_2) \\ P_E \left(q_t, \frac{l_t}{q_t} \right) &= -q_t \Phi(-d_5) + \frac{l_t}{q_t} \Phi(-d_6) \\ \sigma_B(t, T) &= \sigma_r(t) \int_t^T \exp \left\{ - \int_t^u a(s) ds \right\} du \end{aligned}$$

²see Bielecki and Rutkowski (2002) for proof.

A difficulty of using these formulae is that firm value V_t and its volatility are not directly observable. The focus of the next chapter is on the existing empirical structural literature and how this problem has been approached.

3.2 Reduced-Form Model Overview

Firm value models model the insolvency of firms as the triggering mechanism for default. Reduced-form models, instead treat default as a ‘surprise’ event. No reason is given for default in these models, and rather than treating it as endogenous as in the structural models, default is treated exogenously. With reduced-form models, we can model the probability of survival (and hence default) as a stochastic process. Such models are called ‘intensity’ models.

Jarrow, Lando and Turnbull (1997) show that the price of a corporate zero-coupon bond can be shown as the product of the price of an indentical risk-free bond and the expected payoff ratio. In order to obtain this result, it is assumed that bondholder’s receive a fraction of the face value upon default and that this fraction is constant over time. It is also assumed that the process driving interest rates and the process driving default are independent.

The price of a risk-free zero-coupon bond $B_t(T, 1, 0)$ is equal to (2.2) and the price of an equivalent risky zero-coupon bond can be written as,

$$\begin{aligned}\overline{B}_t(\lambda_t, T, 1, 0, \delta) &= E_{\mathbb{Q}} \left[\exp \left\{ - \int_t^T r_u du \right\} (\mathbb{I}_{\{\tau^* > T\}} + \delta \mathbb{I}_{\{\tau^* \leq T\}}) \right] \\ &= E_{\mathbb{Q}} \left[\exp \left\{ - \int_t^T r_u du \right\} (\delta + (1 - \delta) \mathbb{I}_{\{\tau^* > T\}}) \right] \quad (3.16)\end{aligned}$$

where λ_t is the default intensity determining survival probabilities and $\mathbb{I}_{\{\tau^* > T\}}$ is an indicator function taking the value 1 if the modeled default time is after T . The ‘0’ represents the fact that the bond pays no coupons and δ represents the fractional recovery upon default. From the assumption of independence, we can separate this into two expectations,

$$\begin{aligned}\overline{B}_t(\lambda_t, T, 1, 0, \delta) &= E_{\mathbb{Q}} \left[\exp \left\{ - \int_t^T r_u du \right\} \right] E_{\mathbb{Q}} \left[\delta + (1 - \delta) \mathbb{I}_{\{\tau^* > T\}} \right] \\ &= B_t(T, 1, 0) \left(\delta + (1 - \delta) P_{\mathbb{Q}}(\tau^* > T) \right)\end{aligned}\tag{3.17}$$

where $P_{\mathbb{Q}}(\tau^* > T)$ is the probability, $P_t(T)$, at time t of survival up to maturity T . Therefore given a model for the term structure of interest rates, we can calculate the probability of survival to value the risky bond.

As discussed in Duffie and Singleton (2003), default is modeled as the first arrival time τ of a Poisson process with mean arrival rate (intensity) λ . Therefore the waiting time to default is exponentially distributed with pdf,

$$f(t) = \lambda e^{-\lambda t} \quad t \geq 0$$

and it follows that the probability of survival past time t is,³

$$P(t) = e^{-\lambda t}$$

However, it is not reasonable to assume that a firm’s credit quality stays constant over time. We can thus extend the setup by allowing for varying intensities. If

³See Appendix B for proof

intensity changes discretely at time 1, then the probability of survival to time 2 is,

$$\begin{aligned} P_0(2) &= P_0(1)P_1(2) \\ &= e^{-\lambda_1}e^{-\lambda_2} \\ &= e^{-(\lambda_1+\lambda_2)} \end{aligned}$$

It follows that, if λ changes at integer times, the probability of survival up to time t is,

$$P_0(t) = e^{-(\lambda_1+\lambda_2+\dots+\lambda_t)}$$

We can further extend this to allow continuously and deterministically varying intensities,

$$P_0(t) = \exp \left\{ - \int_0^t \lambda_u du \right\}$$

However, in reality, intensity does not vary deterministically. We may know the intensity for the next period but this does not mean we know the intensity for the period following the next period. At time 0 we do not know the intensity for time 2 since this intensity may depend on critical new information revealed at time 1. Therefore, the probability of survival to time 2 is actually,

$$\begin{aligned} P_0(2) &= e^{-\lambda_1} E_{\mathbb{Q}} [e^{-\lambda_2}] \\ &= E_{\mathbb{Q}} [e^{-(\lambda_1+\lambda_2)}] \end{aligned}$$

Again we can allow continuous variation in intensity to obtain,

$$P_0(t) = E_{\mathbb{Q}} \left[\exp \left\{ - \int_0^t \lambda_u du \right\} \right]$$

Finally we could consider the probability of survival up to some future time T conditional on survival to the current period t . It follows from the above results

that this is equal to,

$$P_t(T) = E_{\mathbb{Q}} \left[\exp \left\{ - \int_t^T \lambda_u du \right\} \right]$$

We can see that according to this setup, the relationship between the survival probability and the intensity is the same as the relationship (2.2) between a risk-free zero-coupon bond price and the short rate.

It should be clear that we have two sources of uncertainty in the intensity models,

1. Variation in λ .
2. Arrival time of default.

We only consider one type of intensity model covered in Duffie and Singleton (2003) since our inclusion of a reduced-form model is more for ‘completeness’ of the research. This process is where the intensity follows a Cox, Ingersoll and Ross (1985) process. Under the CIR intensity model, we can calculate probabilities of survival for the coupons and maturity payment and hence price a risky coupon-paying bond.

3.2.1 CIR Intensity

We have said that the relationship between the survival probability and the intensity is the same as the relationship between a risk-free zero-coupon bond price and the short rate. Therefore it seems natural to also model the intensity using a process commonly used for the short rate. We can assume the intensity follows a CIR process under \mathbb{Q} ,

$$d\lambda_t = a_{\lambda}(\mu_{\lambda} - \lambda_t)dt + \sigma_{\lambda}\sqrt{\lambda_t}dW_t$$

where the parameters have the same interpretation as they do for the CIR short rate model. The intensity under this setup reverts to its long-run mean of μ_λ at rate a_λ . Thus it follows that the probability of survival up to time T conditional on survival up to the current period t takes the same form (2.5) where,

$$P_t(T) = \alpha(t, T)e^{-\beta(t, T)\lambda_t} \quad (3.18)$$

and $\alpha(t, T)$ and $\beta(t, T)$ take the same form as (2.9) and (2.10). i.e.

$$\begin{aligned} \eta &= \sqrt{a_\lambda^2 + 2\sigma_\lambda^2} \\ \alpha(t, T) &= \left(\frac{\eta e^{\frac{1}{2}(a_\lambda + \eta)T^*} \beta(t, T)}{e^{\eta T^*} - 1} \right)^{\frac{2a_\lambda \mu_\lambda}{\sigma_\lambda^2}} \end{aligned} \quad (3.19)$$

$$\beta(t, T) = \frac{2(e^{\eta T^*} - 1)}{(a_\lambda + \eta)(e^{\eta T^*} - 1) + 2\eta} \quad (3.20)$$

Thus, given λ_t , a_λ , μ_λ and σ_λ we can calculate survival probabilities and price risky bonds.

As a final note, it is important to realise the consequence of the assumption of independence between the interest rate and default processes. The closed-form solution for the intensity model given above exists due to this assumption. If we assumed some sort of dependence between interest rates and default, closed-form solutions would likely be much more difficult to derive since we would not be able to go from (3.16) to (3.17). The assumption of independence means that the closed-form solutions for the probability of survival holds under any assumption of interest rate process. As a result, we can use the closed-form solution empirically regardless of any assumed interest rate process and hence simulation is not required. This makes

the reduced-form models much easier to apply than the structural models.

The main problem to overcome in estimation is obtaining V_t in the structural models and λ_t in the reduced-form models. Both are latent variables and are not directly observable. The estimation approach is discussed fully in Chapter 5.

Chapter 4

Review of the Empirical Literature

4.1 The Theory

The key issue to overcome in the implementation of the structural models is the fact that the firm asset value V_t and its volatility σ are not readily observable. In this chapter we consider methods for obtaining these values in the theoretical literature and the methods used in the empirical literature.

We firstly consider the equity-implied ‘traditional’ approach proposed by Ronn and Verma (1986) where the firm value and volatility are solved for using a system of two equations. As in Merton (1974), if a firm, which has one class of equity and a zero-coupon bond maturing at some future date T , has a value which evolves according to a geometric Brownian motion (3.8) under \mathbb{Q} , then the equity of the firm can be valued as a European call option on the assets of the firm with maturity T . Therefore equations (3.11)-(3.13) can be used to calculate the value $S_t(V_t)$ of the

equity.

Itô's rule can then be used to derive the process followed by equity under \mathbb{Q} ,

$$dS_t = \left(\frac{\partial S_t}{\partial t} + rV_t \frac{\partial S_t}{\partial V_t} + \frac{1}{2} \sigma^2 V_t^2 \frac{\partial^2 S_t}{\partial V_t^2} \right) dt + \sigma V_t \frac{\partial S_t}{\partial V_t} dW_t$$

We can then apply Itô's rule again to derive the process for the instantaneous return of equity $\ln S_t$,

$$\begin{aligned} f_t &= \ln S_t \\ df_t &= \left[\frac{\partial f_t}{\partial t} + \left(\frac{\partial S_t}{\partial t} + rV_t \frac{\partial S_t}{\partial V_t} + \frac{1}{2} \sigma^2 V_t^2 \frac{\partial^2 S_t}{\partial V_t^2} \right) \frac{\partial f_t}{\partial S_t} + \frac{1}{2} \left(\sigma V_t \frac{\partial S_t}{\partial V_t} \right)^2 \frac{\partial^2 f_t}{\partial S_t^2} \right] dt \\ &\quad + \left(\sigma V_t \frac{\partial S_t}{\partial V_t} \right) \frac{\partial f_t}{\partial S_t} dW_t \\ &= \left[\frac{\partial f_t}{\partial t} + \left(\frac{\partial S_t}{\partial t} + rV_t \frac{\partial S_t}{\partial V_t} + \frac{1}{2} \sigma^2 V_t^2 \frac{\partial^2 S_t}{\partial V_t^2} \right) \frac{1}{S_t} + \frac{1}{2} \left(\sigma V_t \frac{\partial S_t}{\partial V_t} \right)^2 \frac{-1}{S_t^2} \right] dt \\ &\quad + \left(\sigma V_t \frac{\partial S_t}{\partial V_t} \right) \frac{1}{S_t} dW_t \end{aligned}$$

where,

$$\begin{aligned} \sigma_S &= \sigma \frac{V_t}{S_t} \frac{\partial S_t}{\partial V_t} \\ &= \sigma \frac{V_t}{S_t} \Phi(d_1) \end{aligned} \tag{4.1}$$

is the volatility of the instantaneous return on equity and the second equality holds under (3.11)-(3.13).¹ Therefore we can solve the system of two equations,

$$\sigma_S S_t^{ob} = \sigma V_t \Phi(d_1) \tag{4.2}$$

$$S_t^{ob} = V_t \Phi(d_1) - F e^{-rT^*} \Phi(d_2) \tag{4.3}$$

¹See Appendix A for proof.

for V_t and σ where S_t^{ob} is the actual observation of equity at time t . This, however, does require an estimate of equity volatility σ_S . Ronn and Verma (1986) point out that equity volatility can simply be calculated using historical equity data for the relevant time period. However, as argued by Duan (1994), doing so assumes that equity volatility is constant whilst (4.2) assumes that it varies over time with the firm value. Therefore the model is misspecified. In addition, (4.2) is a redundant equation since it was derived from (4.3).

Duan (1994) proposes an alternative method to the estimation problem in the context of deposit insurance. He suggests that given a model assumes a distribution for the unobserved variable, and a relationship between the observed and unobserved variables is also assumed, we can derive the log-likelihood function for the observed data. This can then be numerically optimised to find the log-likelihood maximising parameter estimates.

In general, X and Y represent a vector of unobserved and observed variables respectively each with length n . A closed form solution exists where Y is a function of X ,

$$Y = f(X, \theta)$$

where θ is a set of parameters to be estimated. If $\ln L_X(X, \theta)$ is the log-likelihood function for the unobserved variable X , we can write as a function of our observed variables Y ,

$$\ln L(Y, \theta) = \ln L_X(f^{-1}(Y, \theta), \theta) + \ln |J(f^{-1}(Y, \theta))| \quad (4.4)$$

where $J(f^{-1}(Y, \theta))$ is the Jacobian of the variable change. In our case, since we only have one unobserved variable, that is, the firm value V_t , the Jacobian is a diagonal matrix. As Duan (1994) points out, this simplifies analysis since (4.4) can

be simplified to,

$$\begin{aligned}\ln L(Y, \theta) &= \ln L_X(f^{-1}(Y, \theta), \theta) + \sum_i \ln \left| \frac{dX}{dY} \right| \\ &= \ln L_X(f^{-1}(Y, \theta), \theta) - \sum_i \ln \left| \frac{dY}{dX} \right|\end{aligned}\tag{4.5}$$

The downside of this method is that the inverse function $f^{-1}(\cdot, \theta)$ can be difficult to derive. However, this issue can be sidestepped using a numerical procedure.

Duan presents an example of applying this method in the context of deposit insurance. A bank's value is assumed to follow a geometric Brownian motion process. This is exactly the setup of the Merton (1974) model where the equity of a firm whose value is log-normally distributed can be priced as a European call option on the assets of the firm using the Black and Scholes (1973) pricing formula. Therefore a closed-form solution of the equity value exists and a log-likelihood function can be derived in terms of equity data. Given a set of firm value parameters and the inverse of the equity price formula, firm values can thus be computed and the log-likelihood function evaluated. Using this approach requires equity price data to be available. In this paper we propose a method that uses a series of bond prices and thus can be applied to private firms. This approach is discussed in the next chapter.

4.2 The Traditional Method in Practice

The traditional approach has been widely used throughout the empirical corporate bond pricing literature. In this section we discuss these papers and their results.

The earliest empirical test of a structural model is conducted by Jones, Mason and

Rosenfeld (1984) (commonly known as JMR). In their analysis, they consider a dataset of 27 firms with simple capital structures during the period January 1975 to January 1981. The bonds in their set are mostly of investment grade and have sinking fund options. These bonds are also callable since almost all bonds issued during the considered time period are of this type. They use this data to test the effectiveness of the Merton model to price corporate bonds. The model they implement is more a variation of the Merton model since pricing their dataset requires pricing of the call and sinking fund options embedded in the bonds.

In estimating firm value and its volatility, they use a two step procedure. The first step is to form a series of firm values and calculate the volatility as the standard deviation of the returns on these firm values. This method requires the market value of equity and the market value of debt in order to calculate the firm value as the sum of these values. The second step is then to use a ML procedure based on the relationship between equity volatility and firm value volatility as implied by the traditional approach. The second step uses the volatility estimate under the first step. Using this approach requires equity and balance sheet data to estimate firm values. JMR find an average pricing error of 4.52% under the Merton type model. It is difficult, however, to appraise a model's performance through a pricing error since a small pricing error could still lead to a large spread error. JMR do not report spread errors in their paper.

Ogden (1987) consider a set of 57 bond offerings (and hence 57 prices) of firms with simple capital structures during the period 1973 through to 1985. Like JMR, the set of bonds is callable and has sinking funds, thus Ogden also applies the Merton model considered in JMR. Unlike JMR, Ogden considers the errors in predicted spreads rather than prices in order to test pricing accuracy. Their analysis involves

regressing market spreads on model spreads with the inclusion of a constant. The regression results in a constant of 1.042 and a slope coefficient of 0.925. The positive constant is an indication that the Merton model underpredicts credit spreads (overprices) as is found in the analysis of JMR. In estimating firm value and its volatility, Ogden uses an approach similar to JMR. This method, since using the relation (4.1), is misspecified since equity volatility is estimated as a constant even though the relation used assumes it is stochastic.

Lyden and Saraniti (2000) apply the Merton (1974), and Longstaff and Schwartz (1995) models to data obtained from Bridge Information Systems. Their dataset consists 56 firms with simple capital structures. Their results suggest that both models significantly overprice bonds. That is they underpredict spreads. Lyden and Saraniti estimate firm value volatility through constructing a series of firm values. This is by computing the sum of the market value of equity, book value of liabilities (other than the bond) and the market value of outstanding debt when available. In addition they compute the volatility that sets a model price equal to its corresponding observed price, i.e. a bond-implied volatility.² This is in order to check the accuracy of their historical approach. For each volatility estimate, a 95% confidence interval is formed around the estimate. Lyden and Saraniti find that the bond-implied volatility falls outside the confidence interval for 64% of the bonds under the Merton model.

Eom, Helwege and Huang (2004) consider 182 bond prices of firms with simple capital structures during for the last trading day of every December between 1986 and 1997. Their dataset contains senior bonds that contain no embedded options such as call options or sinking funds. They attempt to price these bonds under the Merton

²This is similar to the use of implied volatility in option pricing.

(1974), Geske (1977), Longstaff and Schwartz (1995), Leland and Toft (1996), and Collin-Dufresne and Goldstein (2001) models. As with JMR and Ogden, they find that the Merton model underpredicts spreads observed in the market. Although they observe only a mean pricing error of 1.69% in the Merton model, the mean error in predicted spreads is -50.42% thus showing that the percentage error in price is not sufficient enough to appraise the performance of a model. They also find that the Geske model underpredicts spreads but not as badly as Merton. The mean pricing error for the Geske model is 0.70% with a mean predicted spread error of -29.57% when assuming bondholders receive a constant fraction of the face value when default occurs.³ The results suggest that the Geske model is an improvement on the Merton model under the PZ approach and that the incorporation of the conditional probability of default aids in predicting spreads.

Eom, Helwege and Huang (2004) calculate volatility using historical equity volatility and leverage in a similar manner as JMR and Ogden. They also calculate an alternative measure of asset volatility using bond prices. This method is the same as that of Lyden and Saraniti (2000) where the volatility is chosen as the level that sets the model price under the closed-form solution equal to the actual price. This estimate of volatility results in better pricing accuracy than the above procedure, however, Eom, Helwege and Huang point out that this is not a measure of volatility as such but rather a method that takes pricing errors into account, that is, it is a “catch-all” for errors.⁴

³Eom, Helwege and Huang also apply the Geske model assuming the bondholders receive the firm when default occurs. This leads to less variation in predicted spreads, but, does not improve the mean predicted spread error. This is because assuming a constant fraction of face value incorporates the costs of financial distress which increases predicted spreads.

⁴Eom, Helwege and Huang (2004) do not report the results of using bond-implied volatility in their paper.

4.3 Simulation Tests

The current empirical literature has largely found that structural models fail to accurately measure prices and spreads. However, a failure of a model empirically does not necessarily mean the theoretical failure of the model. Estimation results can also be affected by the method of estimation. All of the empirical papers discussed above estimate firm value and its volatility consistent with the traditional approach. However, this approach may adversely affect results. Ericsson and Reneby (2005) perform a simulation experiment to test which method is more accurate. They apply both the Duan (1994) method and the traditional method outlined in Ronn and Verma (1986) to appraise the results of each and make a conclusion on which method is more efficient.

Ericsson and Reneby simulate firm values across two different dimensions of risk, the first, is risk due to the leverage level and the second, is risk due to the asset volatility level. Their method involves simulating many firm value paths and then for each path, computing an equity price path which is then used to calculate the firm value and estimate the parameters under each estimation method. This is done for the Merton (1974), Briys and de Varenne (1997), and Leland and Toft (1996) models.

They find that the ML approach outperforms the traditional approach for all models and all possible risk scenarios. Summarising across models and scenarios they find that the ML approach exhibits lower mean errors and standard deviations in estimating firm values, firm value volatility, credit spreads and prices. Table 4.1

presents their overall results.

A striking aspect of the results is the fact that the ML approach vastly outperforms the traditional approach in measuring credit spreads. As stated in Eom, Helwege and Huang (2004) the performance in the measurement of spreads is an important factor in appraising a model (or estimation method) and hence it can be said with confidence that the ML method proposed by Duan (1994) is an improvement on the Ronn and Verma (1986) method. Ericsson and Reneby even go on to say that the poor performance of the traditional approach may explain the why attempts at estimating structural bond models have largely failed. Ericsson and Reneby outline the reasons for the traditional method's failures. When equity and firm value have risen, current equity volatility is low, and thus an estimate of this volatility based on historical equity data is higher than that of the current level. This leads to a higher estimated firm value volatility and modeled bond prices lower than they should be. Conversely, if current equity volatility is high, the resulting estimate of equity volatility is too low thus leading to a low asset volatility and higher modeled bond prices.

	FV	FV Volatility	Spreads	Prices
MLE	0.1 (1.7)	-0.1 (1.9)	0.4 (13.0)	0.1 (3.1)
Traditional	4.0 (4.7)	1.9 (7.3)	23.0 (54.0)	-0.8 (9.5)

Table 4.1: Results reported in Ericsson and Reneby (2005). Mean errors and standard deviations (%) averaged across all models and risk scenarios.

4.4 The ML approach in Practice

There is only one (to my knowledge) published empirical paper on using the ML approach of Duan (1994) on actual data. This work is carried out by Li and Wong (2008). In their study, they apply the Merton (1974), Longstaff and Schwartz (1995), Leland and Toft (1996), and Briys and de Varenne (1997) models under both the traditional and ML approach to a set of bonds obtained from the Fixed Income Database. Li and Wong apply Duan's approach in which they find the parameters of the firm value model and the firm values implied by the equity data that maximises the log-likelihood function. This involves using the closed-form solution for equity prices implied by the structural model to derive root firm values. A root given by the log-likelihood maximising firm value model parameters can then be used as an input into the closed-form solution for the bond price to compute model prices. A model price is then compared to the corresponding actual price. This is precisely the method proposed by Duan (1994) in the context of deposit insurance.

The results of Li and Wong's work are consistent with that of Ericsson and Reneby (2005) in which they find that the ML approach is superior to the traditional approach. Table (4.2) summarises their results for the Merton (1974), and Briys and de Varenne (1997) models.

Li and Wong do not report percentage errors in spreads. Through price and yield errors, their results show that both the Merton, and the Briys and de Varenne models more accurately price the bond data when applied under the ML framework. They suggest the weaknesses of the traditional approach discussed above leads to serious biases in pricing.

4.5 The Best Method?

From the results of Ericsson and Reneby (2005), Li and Wong (2008), and due to the theoretical misspecifications of the Ronn and Verma (1986) method, we choose to use the ML approach of Duan (1994) in this paper. The general consensus is that the traditional method is flawed and can have an adverse effect on empirical results whereas ML is unbiased and efficient. It also allows us to estimate many parameters. However, as Li and Wong (2008) report, the downside of the ML approach is that it can be rather time-consuming computationally. This issue is discussed in Chapter 7.

In this paper, we modify the ML approach of Duan (1994). Duan proposes an approach for estimating the firm value process parameters using equity data and the closed-form solution for equity prices under the model being considered. This approach requires that equity data be observable and thus cannot be applied to private firms. Our contribution is that we present a method that applies Duan's approach using bond prices. Therefore no equity price data is required and the approach can be applied to firms that do not have readily observable equity data. In addition we choose not to use closed-form solutions for bond prices. Instead Monte Carlo simulation is used which then allows structural models to be applied different interest rate assumptions.

Therefore we have an approach which allows the pricing of corporate bonds using structural models under numerous interest rate process assumptions if a series of historical bond prices is available. The methodology is presented in the next chapter.

	Merton		Briys and de Varenne	
	Price	Yield	Price	Yield
MLE	2.6 (8.6)	-2.7 (33.7)	1.6 (6.0)	-4.8 (16.9)
Traditional	3.0 (8.3)	-3.8 (33.7)	1.9 (5.7)	-5.7 (16.2)

Table 4.2: Results reported in Li and Wong (2008). Mean errors and standard deviations (%) for the Merton, and Briys and de Varenne models.

Chapter 5

Approach to Estimation

5.1 Overview

In this chapter we discuss the approach to estimation. Our aim is to find which combination of short rate model and corporate bond pricing model best describes our data. The closed-form solution formulae presented in Chapter 3 apply to the interest rate processes assumed in the models only. It is not straightforward to obtain closed-form solutions for the corporate bond pricing models for different choices of interest rate process. However we can make use of the simple yet elegant approach of Monte Carlo simulation to implement the models under different assumptions about the term structure. With Monte Carlo we do not require closed-form solutions. We instead simulate many paths and price individual cash flows and then average across these paths.

In considering a case with stochastic interest rates, we also need the parameters of

the interest rate process. Therefore we must first estimate the parameters of the short rate models discussed in chapter 2. We do this using a ML approach. From there, we can use the corporate bond prices to obtain firm values and estimate the parameters of the firm value process by also using ML estimation.

5.2 Fitting the Short Rate Models

5.2.1 Parameter Estimation

As discussed, we use a time series approach to fit the short rate models discussed in Chapter 2. This approach involves using the yields of treasury bills as a proxy for the short rate. We use a daily series of 3 month treasury bill yields obtained from the St. Louis Federal Reserve as our proxy.¹ The short rate models are then fitted to this series using ML estimation. The parameters obtained from this estimation allow us to simulate a series of the short rate and then use bond pricing formulae to price a zero coupon bond for any maturity on any given day. This in turn allows us to calculate spot rates for any maturity for a given day and construct a yield curve.

A drawback of this method is that it is a backward looking (historical) approach and a yield curve implied by these parameter estimates need not equal the actual yield curve on any given day. This problem can be addressed by using a different approach to fitting short rate models. We could instead use a method where the yield curve is modelled directly. This is forward looking approach and is more likely to fit actual data since we would be treating the yield curve on each day separately.

¹Obviously this does not include weekends and public holidays and hence the time between observations is non-constant. However, since trading does not occur on these days, we have a rate for every date we require.

However this method is not parsimonious since it involves fitting a yield curve for every single day. This method is not difficult to apply but can lead to unstable parameter estimates. If our dataset of risk-free prices on a given day is quite thin, then fitting a yield curve to these observations may result in a good fit, but, adding an observation could potentially change the parameter estimates quite drastically meaning we have very little out-of-sample use. That is, it may over-fit the model. In addition, the estimated parameters may change from day to day. Therefore we are faced with a tradeoff. Firstly, we could fit a yield curve to every day which possibly could lead to a good fit to the sample but have no out of sample use. On the other hand, we could use the time series approach which is parsimonious and has the advantage that it captures the long term behaviour of interest rates. This advantage is the reason for our choice of the time series approach to fitting short rate models.

We make use of ML estimation to estimate the term structure models. Since we do not observe data in continuous time we must use a discrete time approach known as an Euler approximation. The following is a discrete time representation of the Vasicek model (2.6),

$$\Delta r_t = a(\mu_r - r_t)\Delta t + \sigma_r \sqrt{\Delta t} z_t$$

where z_t is a standard normal random variable. Δt represents the time between observations in our bills time series. Since we do not have an observation for every single day in our time series, it is non-constant. Δr_t is equal to $r_{t+\Delta t} - r_t$. It is straightforward to see that the differenced yield series is distributed normally,

$$\Delta r_t \sim N\left(a(\mu_r - r_t)\Delta t, \sigma_r^2 \Delta t\right)$$

and therefore has the probability density function,

$$\phi(\Delta r_t) = \frac{1}{\sigma_r \sqrt{2\pi\Delta t}} \exp\left\{-\frac{1}{2}\left(\frac{\Delta r_t - a(\mu_r - r_t)\Delta t}{\sigma_r \sqrt{\Delta t}}\right)^2\right\}$$

We can then use this to form the log-likelihood function.

$$\begin{aligned} \ln L(\Delta r_t, a, \mu_r, \sigma_r) &= \ln \prod_t \frac{1}{\sigma_r \sqrt{2\pi\Delta t}} \exp\left\{-\frac{1}{2}\left(\frac{\Delta r_t - a(\mu_r - r_t)\Delta t}{\sigma_r \sqrt{\Delta t}}\right)^2\right\} \\ &= \sum_t \ln\left(\frac{1}{\sigma_r \sqrt{2\pi\Delta t}}\right) - \frac{1}{2} \sum_t \left(\frac{\Delta r_t - a(\mu_r - r_t)\Delta t}{\sigma_r \sqrt{\Delta t}}\right)^2 \end{aligned} \quad (5.1)$$

The a , μ_r and σ_r that maximise (5.1) are our ML estimates of the true parameters.²

We present these results in chapter 6. Estimation of the parameters of the CIR model is carried out in the same way. The only difference from Vasicek is the variance term includes the short rate. In the case of CIR, the following are the respective probability density function and log-likelihood function.

$$\begin{aligned} \phi(\Delta r_t) &= \frac{1}{\sigma_r \sqrt{2\pi r_t \Delta t}} \exp\left\{-\frac{1}{2}\left(\frac{\Delta r_t - a(\mu_r - r_t)\Delta t}{\sigma_r \sqrt{r_t \Delta t}}\right)^2\right\} \\ \ln L(\Delta r_t, a, \mu_r, \sigma_r) &= \ln \prod_t \frac{1}{\sigma_r \sqrt{2\pi r_t \Delta t}} \exp\left\{-\frac{1}{2}\left(\frac{\Delta r_t - a(\mu_r - r_t)\Delta t}{\sigma_r \sqrt{r_t \Delta t}}\right)^2\right\} \\ &= \sum_t \ln\left(\frac{1}{\sigma_r \sqrt{2\pi r_t \Delta t}}\right) - \frac{1}{2} \sum_t \left(\frac{\Delta r_t - a(\mu_r - r_t)\Delta t}{\sigma_r \sqrt{r_t \Delta t}}\right)^2 \end{aligned}$$

For CIR, we are actually performing ‘Quasi’ ML since the short rate is non-central chi-squared distributed. However, it is not uncommon to use a normal approximation for small time steps.

²The 3 month bill yield is not a short rate exactly. Rather, it is a function of the short rate. Therefore we are actually dealing with a transformed variable type of problem discussed in Chapter 4 and the likelihood function should have a Jacobian term. However, we view the 3 monthly bill yields as being close to the short rate and hence simply assume it is a proxy for the short rate.

5.2.2 Risk-Neutral vs Real-World

We have assumed the short rate models to be defined under the risk-neutral probability \mathbb{Q} yet we apply them to real world data. Therefore we are actually estimating the process under the real world measure \mathbb{P} and consequently have to offset the estimated parameters to account for the market price of risk ν to obtain an estimate of the process followed in the risk-neutral world. For Vasicek, using the procedure discussed in 5.2.1 we model,

$$dr_t = a(\mu_r - r_t)dt + \sigma_r dW_t$$

where dW_t is a Brownian motion under \mathbb{P} . We define $d\tilde{W}_t = dW_t + \nu dt$ as a Brownian motion under \mathbb{Q} and therefore obtain the process under risk-neutral world,

$$\begin{aligned} dr_t &= a(\mu_r - r_t)dt + \sigma_r(d\tilde{W}_t - \nu dt) \\ &= \left(a(\mu_r - r_t)dt - \sigma_r \nu \right) dt + \sigma_r d\tilde{W}_t \\ &= a \left(\left[\mu_r - \frac{\sigma_r \nu}{a} \right] - r_t \right) dt + \sigma_r d\tilde{W}_t \\ &= a(\hat{\mu}_r - r_t)dt + \sigma_r d\tilde{W}_t \end{aligned}$$

where,

$$\hat{\mu}_r = \mu_r - \frac{\sigma_r \nu}{a} \tag{5.2}$$

Therefore to obtain the parameter estimates for the risk neutral world, we offset the estimate of μ_r to account for the market price of risk ν . We also need to adjust our estimates for the CIR model. We now define $d\tilde{W}_t = dW_t + \frac{\nu}{\sqrt{r_t}}dt$ as a Brownian

motion under \mathbb{Q} . We thus obtain,³

$$\begin{aligned}
dr_t &= a(\mu_r - r_t)dt + \sigma_r\sqrt{r_t}(d\tilde{W}_t - \frac{\nu}{\sqrt{r_t}}dt) \\
&= \left(a(\mu_r - r_t)dt - \sigma_r\nu\right)dt + \sigma_r\sqrt{r_t}d\tilde{W}_t \\
&= a\left(\left[\mu_r - \frac{\sigma_r\nu}{a}\right] - r_t\right)dt + \sigma_r\sqrt{r_t}d\tilde{W}_t \\
&= a(\hat{\mu}_r - r_t)dt + \sigma_r\sqrt{r_t}d\tilde{W}_t
\end{aligned}$$

as the CIR process under \mathbb{Q} where $\hat{\mu}_r$ follows from (5.2). These adjustments, though easy to apply, do however present the question, how do we find ν ? Since it is not readily observable, we must estimate it.

We estimate this value by making use of the 10-year treasury bond par-yield data available from the St. Louis Federal Reserve.⁴ 10-year maturity rates are chosen in an attempt to accurately model rates of all maturities. We obtain this data for the same date range as our 3 month bills time series used to estimate the parameters of the Vasicek and CIR models. We can use our estimated parameters for the Vasicek and CIR models to price zero-coupon bonds for many maturities and then calculate par-yields to compare to the actual par-yields.

For the Vasicek model, we fix the level of ν and offset the μ_r estimate as outlined above. For time t we then use the closed-form solutions (2.5) and (2.7) - (2.8) to price zero-coupon bonds using the original a and σ_r , the adjusted $\hat{\mu}_r$ using (5.2) and the 3 month bill yield for time t as r_t . We price zero-coupon bonds with face value \$1 for half yearly increments from 6 months up to 10 years where $B_t(t + \frac{\tau}{2}, 1, 0)$ is

³We have arbitrarily chosen the form of the market price of risk to obtain the same adjustment for the CIR model as the Vasicek model.

⁴Par-yields are coupon yields which would set the price of a bond equal to 1. The par-yield data we obtain assume semi-annual coupons.

the price of a zero-coupon bond at time t maturing at time $t + \frac{\tau}{2}$. The model par yield is $y_t(t + 10)$ where,

$$\sum_{\tau=1}^{20} \frac{1}{2} y_t(t + 10) B_t(t + \frac{\tau}{2}, 1, 0) + B_t(t + 10, 1, 0) - 1 = 0$$

We thus solve for the 10 year par yield $y_t(t + 10)$ where,

$$y_t(t + 10) = 2 \frac{1 - B_t(t + 10, 1, 0)}{\sum_{\tau=1}^{20} B_t(t + \frac{\tau}{2}, 1, 0)}$$

We then calculate the percentage error between the model par-yield and the actual par-yield for time t .

$$\epsilon_t = \frac{y_t(t + 10) - y_t^A(t + 10)}{y_t^A(t + 10)}$$

where $y_t^A(t + 10)$ is the actual 10-year par-yield for time t . We then repeat this procedure for every t so that we have a percentage error for every date in our series and calculate the sum of squared errors,

$$S = \sum_t \epsilon_t^2 \tag{5.3}$$

We then choose our market price of risk parameter as the ν that minimises (5.3). The procedure for the CIR model is exactly the same except we use (2.9)-(2.10) in (2.5) to calculate the zero-coupon bond prices that we use to calculate the par-yields. We then choose the level of ν that minimises (5.3).

5.3 The ‘Portfolio of Zeroes’ Approach

Before we can proceed to the methodology of estimating the firm value process parameters for each model we need to recognise an obvious yet difficult issue. All of the structural corporate bond pricing models we are investigating consider zero-coupon bonds. Restricting attention to bonds that pay no coupons would be very unrepresentative of bonds traded in the market. Therefore we need a means of pricing coupon bonds using the pricing models under consideration. We use the ‘portfolio of zeroes’ (PZ) type approach used in Eom, Helwege and Huang (2004), where each coupon is treated as an individual zero-coupon bond independent of all the other coupons and the principal payment. This method is easy to carry out computationally and is the method used in this paper. This approach, however simple to put into practice, is not without problems.

It is obvious that the chances of receiving a promised payment of a bond, whether it be a coupon or the face value, is conditional on the previous coupon being received. The PZ approach does not take this into account since it treats all payments of a bond as independent of each other. Under the PZ approach, each payment is treated completely in isolation and the default of an early payment does not imply the default of later payments. For example, consider the Merton model and a bond that pays a \$5 coupon at time 1 and another \$5 coupon at time 2. If our simulated firm value is less than \$5 at time 1 but greater than \$5 at time 2, then the time 1 coupon would be treated as a bond that defaults but the time 2 coupon would be treated as a bond that does not default. This is clearly undesirable since in reality a firm that defaults at time 1 ceases to exist and hence does not meet its obligations at time 2 (and all subsequent dates). This property of the PZ approach is likely to lead to overpricing of the bond along a single simulation path since the time 2

coupon would be perceived to have positive value under the model where it should actually have zero value if the time 1 coupon defaults.

By treating each coupon separately, we are assuming the distribution of the firm value is the same at each coupon date. This is not actually the case since each coupon payment is conditional on the payments of earlier coupons. Therefore the distribution of the firm value at each coupon date is actually a conditional distribution. The density used to calculate the values of later coupons is conditional on the firm value at earlier coupon dates. The PZ approach is therefore likely to lead to over pricing. To test this, in addition to the PZ approach, we also price the corporate bonds using each model in an internally consistent (IC) manner where we set the value of all post time τ payments to zero if the time τ payment defaults.

For the intensity models, Duffie and Singleton (2003) state that, assuming constant fractional recovery and independence between the default and interest rate process means that coupon bonds can be priced under the PZ approach where coupons are assumed to have zero recovery.

Under an intensity model, the price of a risky bond paying \$1 at T , with fractional face value recovery δ , and coupon rate c is $\bar{B}_t(\lambda_t, T, 1, c, \delta)$ where,⁵

$$\bar{B}_t(\lambda_t, T, 1, c, \delta) = \bar{B}_t(\lambda_t, T, 1, 0, \delta) + \frac{c}{2} \sum_{\tau} \bar{B}_t(\lambda_t, \tau, 1, 0, 0) \quad (5.4)$$

The price of the risky coupon-paying bond has been decomposed into two parts. The first is the price of a zero-coupon with the same maturity as the coupon bond and the second part is sum of the values of the coupons where each is assumed to be a zero coupon bond with zero recovery, i.e. only a fraction of the face value

⁵We are assuming that coupons are paid on a semi-annual basis.

is recovered upon default and the coupons are lost. A bond that has an assumed recovery of zero can be priced with (3.17) where δ is set to zero.

Therefore, given an intensity process driving the probability of survival, a coupon bond can easily be valued using (3.17) and (5.4)

5.4 Monte Carlo Overview

We have a total of 18 structural model cases to implement since we estimate each of the Merton (1974), Black and Cox (1976), and Briys and de Varenne (1997) models with the constant, Vasicek and CIR interest rate process assumptions under both the PZ and IC set up. In this section we summarise the Monte Carlo approach used to estimate the parameters of the firm value process for the structural models.

This approach makes use of simulation to price securities. For a given bond, we observe J prices. Each price represents the value of the bond on the day that particular transaction took place. Our approach is to value a bond on each of the transaction dates we observe in order to estimate the parameters of the firm value process for that particular bond. To do this, many series of the firm value are simulated for each transaction. Each series begins at the date of the particular transaction and ends at the maturity of the bond.

For each simulation path of a given transaction, the individual bond payments (coupons and principal) are valued and the bond itself is priced as the sum of these values. A Monte Carlo price for that transaction is then computed by averaging across the individual path bond prices. We then, for every transaction date t_j , find

a firm value that best describes the price observation on that date. That is, we use the bond prices to calibrate firm values. For each transaction, we find a firm value for the transaction date that sets the difference between the Monte Carlo model price and actual price equal to zero. These root firm values $V_{t_j}^r$ are then used along with the parameters to evaluate the log-likelihood function. ML estimation is then used to estimate the parameters required for each model.

We use a bi-weekly frequency for our simulations by necessity since the computational time taken to obtain results is quite high, particularly for the first-passage models where a high simulation frequency results in a large number of dates where the firm value is checked against the barrier. Although a high frequency here is more accurate, finding root firm values can take a long time. Finding these roots involves pricing the bond for each transaction multiple times. This involves checking for defaults and calculating payoffs repeatedly. We therefore choose to simulate at a bi-weekly frequency which then allows for more simulation paths and potentially, smoother log-likelihood surfaces.

5.5 The Log-Likelihood Function

The log-likelihood function being maximised is the same across all of the structural models since each assumes a geometric Brownian motion for the firm value. We estimate the parameters of the stochastic process under the real world probability measure \mathbb{P} since this is the measure which actually generates our data. Therefore we wish to estimate the following firm value process under \mathbb{P} ,

$$dV_t = \mu V_t dt + \sigma V_t dW_t \tag{5.5}$$

We can in fact make our lives easier by taking the log of the firm values so that we can work with the normal distribution. By an application of Itô's lemma the difference of the log firm value follows the process,

$$d \ln V_t = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW_t \quad (5.6)$$

The log of the firm value follows a normal distribution with mean and variance,

$$\begin{aligned} E(d \ln V_t) &= (\mu - \frac{1}{2}\sigma^2)dt \\ \text{Var}(d \ln V_t) &= \sigma^2 dt \end{aligned}$$

A given root $V_{t_j}^r$ is the best estimate of the firm value at t_j given the transaction that occurred at time t_j . We can easily set up the log-likelihood function for $\ln V_t$ given that it is normally distributed. However, given that we have a transformed data problem as in Duan (1994), we have to account for the Jacobian of the variable change. The log-likelihood function if we directly observe the firm values, $V_{t_j}^{ob}$, would be,

$$\ln L(\ln V_t^{ob}, \theta) = \ln \prod_{j=1}^{J-1} \frac{1}{\sigma \sqrt{2\pi \Delta t_j}} \exp \left\{ -\frac{1}{2} \left(\frac{\Delta \ln V_{t_j}^{ob} - (\mu - \frac{1}{2}\sigma^2)\Delta t_j}{\sigma \sqrt{\Delta t_j}} \right)^2 \right\} \quad (5.7)$$

where $\Delta t_j = t_{j+1} - t_j$, $\Delta \ln V_{t_j}^{ob} = \ln V_{t_{j+1}}^{ob} - \ln V_{t_j}^{ob}$ and θ is the vector of parameters we wish to estimate. However, we compute $V_{t_j}^r$ and therefore our log-likelihood function is instead a function of the observed prices \overline{B}^{ob} and the parameters θ . Therefore our

log-likelihood function is instead of the form (4.5) with,

$$\begin{aligned}
& \ln L(\ln V_t^r, \theta) \\
&= \ln \prod_{j=1}^{J-1} \frac{1}{\sigma \sqrt{2\pi \Delta t_j}} \exp \left\{ -\frac{1}{2} \left(\frac{\Delta \ln V_{t_j}^r - (\mu - \frac{1}{2}\sigma^2)\Delta t_j}{\sigma \sqrt{\Delta t_j}} \right)^2 \right\} \frac{d \ln V_{t_{j+1}}^r}{d \bar{B}_{t_{j+1}}} \\
&= \sum_{j=1}^{J-1} \ln \frac{1}{\sigma \sqrt{2\pi \Delta t_j}} \exp \left\{ -\frac{1}{2} \left(\frac{\Delta \ln V_{t_j}^r - (\mu - \frac{1}{2}\sigma^2)\Delta t_j}{\sigma \sqrt{\Delta t_j}} \right)^2 \right\} \\
&\quad + \sum_{j=1}^{J-1} \ln \left| \frac{d \ln V_{t_{j+1}}^r}{d \bar{B}_{t_{j+1}}} \right| \\
&= \sum_{j=1}^{J-1} \ln \frac{1}{\sigma \sqrt{2\pi \Delta t_j}} \exp \left\{ -\frac{1}{2} \left(\frac{\Delta \ln V_{t_j}^r - (\mu - \frac{1}{2}\sigma^2)\Delta t_j}{\sigma \sqrt{\Delta t_j}} \right)^2 \right\} \\
&\quad - \sum_{j=1}^{J-1} \ln \left| \frac{d \bar{B}_{t_{j+1}}(V_{t_{j+1}}^r, T, F, c)}{d \ln V_{t_{j+1}}^r} \right| \\
&= \sum_{j=1}^{J-1} \ln \frac{1}{\sigma \sqrt{2\pi \Delta t_j}} \exp \left\{ -\frac{1}{2} \left(\frac{\Delta \ln V_{t_j}^r - (\mu - \frac{1}{2}\sigma^2)\Delta t_j}{\sigma \sqrt{\Delta t_j}} \right)^2 \right\} \\
&\quad - \sum_{j=1}^{J-1} \ln \left| \frac{d \bar{B}_{t_{j+1}}(V_{t_{j+1}}^r, T, F, c)}{d V_{t_{j+1}}^r} \frac{d \ln V_{t_{j+1}}^r}{d V_{t_{j+1}}^r} \right| \\
&= \sum_{j=1}^{J-1} \ln \frac{1}{\sigma \sqrt{2\pi \Delta t_j}} \exp \left\{ -\frac{1}{2} \left(\frac{\Delta \ln V_{t_j}^r - (\mu - \frac{1}{2}\sigma^2)\Delta t_j}{\sigma \sqrt{\Delta t_j}} \right)^2 \right\} \\
&\quad - \sum_{j=1}^{J-1} \ln \left| \frac{d \bar{B}_{t_{j+1}}(V_{t_{j+1}}^r, T, F, c)}{d V_{t_{j+1}}^r} V_{t_{j+1}}^r \right| \tag{5.8}
\end{aligned}$$

where the extra $V_{t_{j+1}}^r$ term in the Jacobian comes from the fact that we are working with the log-likelihood function of the log-firm value rather than the firm value.⁶

The derivative term in the Jacobian can be approximated. We can offset the root $V_{t_j}^r$ for transaction j by some factor ϵ_j to obtain an offset root,

$$V_{t_j}^{r'} = V_{t_j}^r + \epsilon_j$$

⁶See the correction in Duan (2000). This term would be irrelevant if V_{t_j} were observed since it would not change between log-likelihood evaluations.

The derivative can then be approximated as,

$$\frac{d\overline{B}_{t_j}(V_{t_j}^r, T, F, c)}{dV_{t_j}^r} \approx \frac{\overline{B}_{t_j}(V_{t_j}^{r'}, T, F, c) - \overline{B}_{t_j}^{ob}(T)}{\epsilon_j} \quad (5.9)$$

since $\overline{B}_{t_j}(V_{t_j}^r, T, F, c) = \overline{B}_{t_j}^{ob}(T)$. Thus, we can find roots for all the transactions of a bond and evaluate (5.9) to approximate the derivative in the Jacobian.

The parameters values θ that maximise (5.8) are our estimates for the parameters of the firm value process we use to price that firm's bond.⁷

5.6 Parameter Estimation

In this section we outline the procedure for estimating the required parameters for the Merton, Black and Cox, and Briys and de Varenne models with the constant interest rate assumption under the PZ approach.

In applying the structural models under a constant interest rate assumption, our choice of the constant interest rate, r , is the average of a 3 month bill-yield time series obtained from the St. Louis Federal Reserve. This series is discussed in the next chapter.

In addition to the face value, a bond also pays a coupon C at fixed intervals which we must value. For pricing simplicity, we treat the maturity T coupon C and face value F as a singular payment $C + F$. A bond has M_j promised payments remaining

⁷ μ does not determine the roots since pricing occurs in the risk-neutral world. It does, however, affect the value of the log-likelihood function.

on its transaction date t_j and these payments are denoted as,⁸

$$c(m) = \begin{cases} C & \text{if } m < M_j \\ C + F & \text{if } m = M_j \end{cases}$$

We then value these payments for each simulation path. We denote the value at time t of the m^{th} remaining payment $c(m)$ as $c_t(\tau_m)$ where τ_m is the time period in which the m^{th} remaining promised payment is paid.⁹

The Black and Cox, and Briys and de Varenne models include default barriers that must be accounted for. Since each promised payment of a bond is valued separately under the PZ approach, it makes sense to assume each promised payment has its own default barrier.

For the Black and Cox model, our pricing problem is made easier by transforming the barrier so that it is constant. We hence make the same scaling change to the drift of the firm value and the barrier that is made in the calculation of the survival probability in Appendix A. Therefore we are now considering an ‘adjusted’ firm value \hat{V}_t that is assumed to follow the geometric Brownian motion process under \mathbb{Q} ,

$$d\hat{V}_t = (r - \gamma)\hat{V}_t dt + \sigma\hat{V}_t dW_t \quad (5.10)$$

We can relate this adjusted firm value to the unadjusted firm value for transaction j with,

$$\hat{V}_t = V_t e^{-\gamma(t-t_j)} \quad t \geq t_j \quad (5.11)$$

⁸ M_j differs across transactions since transactions occur at different points in time.

⁹ $c(m)$ represents a dollar payment whereas c represents the coupon rate.

Since the simulation paths for transaction j begin at t_j , \hat{V}_{t_j} and V_{t_j} coincide. We modify the Black and Cox barrier to reflect the size of the payment itself. The barrier at time t for the m^{th} remaining payment $c(m)$ for transaction j is thus set constant at,

$$K_t(\tau_m) = \alpha c(m) e^{-\gamma(\tau_m - t_j)} \quad t_j \leq t \leq \tau_m \text{ and } 0 \leq \alpha \leq 1$$

Variation in the size of the barrier is allowed with the α term. Under this setup we can be sure that a barrier accounts for the size of the payment whilst scaling with α .

We set the transaction-date firm value for observation j to some initial value V_{t_j} and simulate N series up to the maturity of the bond. The advantage of working with log-firm value is that we can simulate firm values according to the exact distribution of the firm values. From (5.6) it follows that,

$$V_{t+dt} = V_t \exp \left\{ \left(r - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t \right\}$$

under \mathbb{Q} , and therefore we can simulate path i for the Merton, and Briys and de Varenne models using,

$$V_{t+\Delta t}^i = V_t^i \exp \left\{ \left(r - \frac{1}{2} \sigma^2 \right) \Delta t + \sigma \sqrt{\Delta t} z_t^i \right\} \quad (5.12)$$

where z_t^i is a standard normal random variable and Δt represents a time step of two weeks.¹⁰ It is important that we set the seed for random number generation so that we obtain the same random numbers between likelihood evaluations. The firm

¹⁰The time steps are not always two weeks. To account for the fact that coupon dates may fall in between these time steps, we add in the coupon dates to the vector of simulation dates for a transaction.

value simulated for the Black and Cox model is an ‘adjusted’ firm value \hat{V}_t since the drift has been transformed. This firm value evolves according to the process (5.10) and is instead simulated using

$$\hat{V}_{t+\Delta t}^i = \hat{V}_t^i \exp \left\{ (r - \gamma - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}z_t^i \right\} \quad (5.13)$$

From the simulated paths we can then price the individual promised payments for a given transaction. The value at the transaction date t_j of the m^{th} remaining payment $c(m)$ for simulation path i , if no barrier default occurs, is computed as,

$$c_{t_j}^i(\tau_m) = e^{-r(\tau_m - t_j)} \left(c(m) \mathbb{I}_{\{V_{\tau_m}^i > c(m)\}} + f_2 V_{\tau_m}^i \mathbb{I}_{\{V_{\tau_m}^i < c(m)\}} \right) \quad (5.14)$$

where $f_2 = 1$ under Merton, and Black and Cox. Since ‘adjusted’ firm values are simulated for the Black and Cox model, evaluation of (5.14) requires the extra step of calculating unadjusted firm values. This is done by rearranging (5.11) for V_t and setting $t = \tau_m$. Equation (5.14) is used to price every promised payment under the Merton model since there is no default barrier.¹¹ However, under Black and Cox, and Briys and de Varenne, (5.14) only applies when default of the payment does not occur priority to maturity along its barrier. The barrier at time t for the m^{th} remaining payment $c(m)$ of transaction j can be represented as,

$$K_t(\tau_m) = \alpha c(m) \left(\xi e^{-\gamma(\tau_m - t_j)} + (1 - \xi) B_t(\tau_m, 1, 0) \right) \quad t_j \leq t \leq \tau_m \quad (5.15)$$

where $\xi = 1$ when considering Black and Cox, and $\xi = 0$ when considering Briys and de Varenne. The barrier levels for a given payment are the same across simulation paths since γ and r are constant. The price at time t of a risk-free zero-coupon bond

¹¹Much like a stock price when dividends are paid we would expect the firm’s value to drop on a coupon date. To avoid this issue we assume that the firm issues new equity to offset this drop.

maturing at time τ_m and face value \$1 is $B_t(\tau_m, 1, 0)$ where,

$$B_t(\tau_m, 1, 0) = e^{-r(\tau_m - t)} \quad (5.16)$$

The m^{th} remaining payment of transaction j and simulation path i survives the default barrier if,

$$K_t(\tau_m) < \xi \hat{V}_t^i + (1 - \xi)V_t^i \quad \forall t; \quad t_j \leq t \leq \tau_m \quad (5.17)$$

If condition (5.17) is met for a payment, then the payment can be valued using (5.14). If instead default occurs at the barrier of a payment, then the bondholders receive the value of the firm's assets (or a fraction f_1 of firm assets for the Briys and de Varenne model). Here, the value of the firm's assets is simply the level of the barrier so it follows that the value at t_j of the m^{th} remaining payment $c(m)$ of transaction j and simulation path i when default occurs at the barrier (i.e. the first time condition (5.17) is not met) is,

$$c_{t_j}^i(\tau_m) = K_{\tau_{m,b}^i}(\tau_m) e^{-r(\tau_{m,b}^i - t_j)} \left(\xi e^{\gamma(\tau_{m,b}^i - t_j)} + (1 - \xi)f_1 \right) \quad (5.18)$$

$$t_j < \tau_{m,b}^i < \tau_m \text{ and } 0 \leq f_1 \leq 1$$

where $\tau_{m,b}^i$ is the period where the firm value hit the barrier of the m^{th} remaining payment for simulation path i . The first term in the parentheses on the right-hand side of (5.18) represents the fact that we are working with a transformed barrier in the Black and Cox model. We are required to compute what the untransformed barrier level would be in this situation. We compound the payoff back up by a rate of γ to reflect what the barrier would be if we had not made the adjustment to a flat barrier. The second term in the parentheses represents deviation from absolute

priority in the Briys and de Varenne model.

Once every remaining payment of a transaction is valued, we can price the bond on the transaction date for every simulation path. For a given simulation path, we sum across the values of the payments to obtain the transaction date value of the bond for that particular path. The value of the bond for simulation path i at transaction date t_j is therefore,

$$\overline{B}_{t_j}^i(V_{t_j}, T, F, c) = \sum_{m=1}^{M_j} c_{t_j}^i(\tau_m) \quad (5.19)$$

Averaging across the simulation paths then gives us the Monte Carlo price of transaction j that we compare to the actual price for transaction j .

$$\overline{B}_{t_j}(V_{t_j}, T, F, c) = \frac{1}{N} \sum_{i=1}^N \overline{B}_{t_j}^i(V_{t_j}, T, F, c) \quad (5.20)$$

This is the Monte Carlo part of the estimation. We then use a root finder to obtain, given the parameters θ , the initial firm value V_{t_j} that sets the difference between the model price of transaction j and the actual price of transaction j to zero.¹² We find $V_{t_j} = V_{t_j}^r$ where,

$$\overline{B}_{t_j}(V_{t_j}^r, T, F, c) - \overline{B}_{t_j}^{ob}(T) = 0 \quad (5.21)$$

where $\overline{B}_{t_j}^{ob}(T)$ is the observed price for the time t_j transaction. We repeat this procedure until we have found a root for all J transactions. Each root represents the value of the firm on that particular observation's transaction date that sets the model price equal to the actual price for that transaction. The result is a time-series of firm values,

$$V_{t_1}^r, V_{t_2}^r, \dots, V_{t_J}^r$$

¹²With barriers in place, it only makes sense to search for root firm values in the range above the barrier levels at t_j for transaction j . In other words, it only makes sense to consider a range of initial firm values which represent the firm being alive (above the barriers at the transaction date).

where t_j is the time that transaction j occurred.

Once a root is found for each transaction of the bond, we can evaluate the log-likelihood function and hence use ML estimation to estimate the parameters of the firm value process and any other required parameters. The optimiser starts with some parameters, repeats the discussed procedure to price the bond on each transaction date and then evaluates the log-likelihood using the roots. This procedure is repeated until a set of parameters and roots that maximises the log-likelihood function is found.

5.7 Extending to a Stochastic Interest Rate

Extending the procedure to account for a stochastic interest rate requires only minor modifications. We are required to simulate an interest rate series for each simulation path of the transactions of a bond by using the parameters obtained in estimating the Vasicek and CIR models.¹³ For transaction j we use the prevailing 3 month bill yield at transaction date t_j as the starting point for the simulation of the interest rates. We then simulate the series for simulation path i using,

$$r_{t+\Delta t}^i = r_t^i + a(\mu_r - r_t^i)\Delta t + \sigma_r \sqrt{\Delta t} z_t^{i,r} \quad (5.22)$$

for the Vasicek model, or,

$$r_{t+\Delta t}^i = r_t^i + a(\mu_r - r_t^i)\Delta t + \sigma_r \sqrt{r_t^i \Delta t} z_t^{i,r} \quad (5.23)$$

¹³Simulating an interest rate series for each firm value path enables us to allow for correlation between the interest rate and firm value if required.

for the CIR model where $z_t^{i,r}$ is a standard normal variable different from z_t^i in (5.12). Like the firm value simulations, the seed for random number generation is fixed so that we obtain the same interest rate simulations between log-likelihood evaluations. We can then simulate the firm values for path i using,

$$V_{t+\Delta t}^i = V_t^i \exp \left\{ (r_t^i - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}z_t^i \right\} \quad (5.24)$$

for the Merton, and Briys and de Varenne models and,

$$\hat{V}_{t+\Delta t}^i = \hat{V}_t^i \exp \left\{ (r_t^i - \gamma - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}z_t^i \right\}$$

for the Black and Cox model. The only change from firm value simulation in the constant interest rate set up is that the simulated interest rate at time t is now used in the drift term when simulating $V_{t+\Delta t}$ rather than the constant interest rate r . This requires simulating the interest rates at the same frequency as the firm values so as to be sure on which rate to use in the drift of the firm value simulation.

The shift to stochastic interest rates has to be accounted for in valuing individual payments. The value at the transaction date t_j of the m^{th} payment $c(m)$ for simulation path i when barrier default does not occur would be computed as,

$$c_{t_j}^i(\tau_m) = \exp \left\{ - \int_{t_j}^{\tau_m} r_t^i dt \right\} \left(c(m) \mathbb{I}_{\{V_{\tau_m}^i > c(m)\}} + f_2 V_{\tau_m}^i \mathbb{I}_{\{V_{\tau_m}^i < c(m)\}} \right)$$

if we could simulate a continuous interest rate series. Since our series is discrete, we approximate the integral with a summation. Therefore the value of the payment is computed as,

$$c_{t_j}^i(\tau_m) = \exp \left\{ - \sum_{t=t_j}^{\tau_m - \Delta\tau_m} r_t^i \Delta t \right\} \left(c(m) \mathbb{I}_{\{V_{\tau_m}^i > c(m)\}} + f_2 V_{\tau_m}^i \mathbb{I}_{\{V_{\tau_m}^i < c(m)\}} \right) \quad (5.25)$$

Again, this formula is used to price every payment under the Merton model since there is no default barrier. For the Black and Cox, and Briys and de Varenne models, however, we must account for the possibility of early default along the payment's barrier. We also have to account for the stochastic interest rate when calculating the barrier level for the Briys and de Varenne model. The prices of risk-free zero-coupon bonds cannot be calculated using (5.16) any more since the interest is no longer assumed to be constant. Instead the prices depend on which interest rate model we are assuming. For the Vasicek model, the price at time t of a risk-free zero-coupon bond maturing at time τ_m takes the same form (2.5) discussed in Chapter 2,

$$B_t^i(\tau_m, 1, 0) = \alpha(t, \tau_m) e^{-\beta(t, \tau_m) r_t^i} \quad (5.26)$$

where r_t is obtained from simulation and,

$$\begin{aligned} \alpha(t, \tau_m) &= \exp\left(\frac{(\beta(t, \tau_m) - (\tau_m - t))(2a^2\mu - \sigma^2)}{2a^2} - \frac{\sigma^2\beta(t, \tau_m)^2}{4a}\right) \\ \beta(t, \tau_m) &= \frac{1 - e^{-a(\tau_m - t)}}{a} \end{aligned}$$

Under the CIR model the price at time t of a risk-free zero-coupon bond maturing at time τ_m also takes the form (5.26) with,

$$\begin{aligned} \alpha(t, \tau_m) &= \left(\frac{\eta e^{\frac{1}{2}(a+\eta)(\tau_m - t)} \beta(t, \tau_m)}{e^{\eta(\tau_m - t)} - 1} \right)^{\frac{2a\mu}{\sigma^2}} \\ \beta(t, \tau_m) &= \frac{2(e^{\eta(\tau_m - t)} - 1)}{(a + \eta)(e^{\eta(\tau_m - t)} - 1) + 2\eta} \\ \eta &= \sqrt{a^2 + 2\sigma^2} \end{aligned}$$

The barrier at time t for the m^{th} remaining payment $c(m)$ of transaction j and simulation path i is now represented as,

$$K_t^i(\tau_m) = \alpha c(m) (\xi e^{-\gamma(\tau_m - t_j)} + (1 - \xi) B_t^i(\tau_m, 1, 0)) \quad t_j \leq t \leq \tau_m \quad (5.27)$$

and therefore, condition (5.17) becomes,

$$K_t^i(\tau_m) < \xi \hat{V}_t^i + (1 - \xi)V_t^i \quad \forall t; \quad t_j \leq t \leq \tau_m \quad (5.28)$$

We have accounted for the fact that the barrier levels are different across simulation paths under the Briys and de Varenne models with a stochastic interest rate. The risk-free bond price $B_t^i(\tau_m, 1, 0)$ is the time τ_m risk-free zero-coupon bond price at t for simulation path i under either the Vasicek or CIR model. The difference in simulated interest rates across paths leads to the difference in risk-free zero-coupon bond prices and hence different levels of the barrier across simulation paths. However, the barrier for the m^{th} remaining payment of transaction j will be the same across simulation paths on the transaction date t_j since we simulate all our interest rates paths for transaction j from the 3 month bill yield at date t_j as a starting point.

As with maturity payoffs, the change to a stochastic interest rate must also be accounted for when valuing barrier payoffs. If condition (5.28) is not met for the m^{th} remaining payment of transaction j and simulation path i then the value at time t_j is,

$$c_{t_j}^i(\tau_m) = K_{\tau_{m,b}^i}^i(\tau_m) \exp \left\{ - \sum_{t=t_j}^{\tau_{m,b}^i - \Delta\tau_{m,b}^i} r_t^i \Delta t \right\} \left(\xi e^{\gamma(\tau_{m,b}^i - t_j)} + (1 - \xi)f_1 \right)$$

$$t_j < \tau_{m,b}^i < \tau_m \text{ and } 0 \leq f_1 \leq 1$$

From here we can proceed in the usual way of calculating the bond's price along each simulation path and then averaging across paths to obtain the Monte Carlo price of transaction j . This procedure is repeated for each transaction and a root

	Merton	Black and Cox	Briys and de Varenne
μ	Y	Y	Y
σ	Y	Y	Y
α	N	Y	Y
γ	N	Y	N
f_1	N	N	Y
f_2	N	N	Y

Table 5.1: Parameters to be estimated under each model.

$V_{t_j}^r$ is found for each transaction j to evaluate the log-likelihood function.

The ML approach has the advantage that it can be used to estimate multiple parameters. Table (5.1) summarises which parameters are estimated under each structural model.

The Merton model does not include a default barrier hence only μ and σ are required to be estimated. The Black and Cox model requires γ and α in addition to μ and σ to account for the default barrier. The Briys and de Varenne model does not require γ but does account for deviations from absolute priority with f_1 and f_2 .

Before moving on, we need to acknowledge a difference between the constant interest rate and stochastic interest rate applications. When considering a stochastic interest rate, for a given transaction, we simulate rates for an observation using the prevailing transaction date 3 month bill yield as a starting point. That is, we attempt to simulate interest rates consistent with actual rates at the time of the transaction. In applying a model under the constant interest rate setup, we simply use the average of the 3 month bill series. Using the prevailing 3 month bill yield at the time of a transaction as the constant interest rate, however, would be more consistent with the stochastic interest setup. However, since we have many transactions for a given bond, this would presumably mean assuming many different constant interest rates

for a single bond. This contradicts the assumption that interest rates are constant and therefore, we choose to use the same constant interest rate for every transaction of a bond.

5.8 Survival Probabilities under Portfolio of Zeros

Along with estimating the parameters required for each model, we can also calculate probabilities of survival for each promised payment. These are risk-neutral survival probabilities since they are calculated from the paths simulated in accordance with the risk-neutral process for V_t . Calculating the probability of survival of a payment is simple given we can use the procedure in place. Given the ML estimates, we can simulate many paths of the firm value for each transaction using their respective roots as starting points.

Under Merton, the m^{th} remaining payment of transaction j survives along path i if at time τ_m , the firm value V_{τ_m} is greater than the promised level of the payment $c(m)$. We can estimate the probability of survival of this payment under the Merton model using an indicator function.

$$P_{t_j}(\tau_m) = \frac{1}{N} \sum_{i=1}^N \mathbb{I}_{\{V_{\tau_m}^i > c(m)\}} \quad (5.29)$$

We can simply calculate the probability at t_j of survival of the m^{th} remaining payment $c(m)$ as the proportion of simulated time τ_m firm values that are greater than $c(m)$ since default only occurs at maturity in the Merton model. The formula (5.29) applies to all payments of the bond since the default of a payment is not affected

by the default of a different time period payment under the PZ approach.

When we consider the probability of survival of a payment under the Black and Cox model, we need to account for the probability of early default through the payment's barrier. The probability of survival of the m^{th} remaining payment is estimated as the proportion of simulation paths that have a firm value greater than $c(m)$ at τ_m and that have not hit the barrier for the time τ_m payment. A payment survives the default barrier if condition (5.17) (or (5.28) for a stochastic interest rate case) is met. The probability at t_j that the m^{th} remaining payment $c(m)$ of transaction j survives is estimated as,

$$P_{t_j}(\tau_m) = \frac{1}{N} \sum_{i=1}^N \mathbb{I}_{\{\hat{V}_t^i > K_t^i(\tau_m); \quad \forall t; t_j \leq t \leq \tau_m\}} \cdot \mathbb{I}_{\{V_{\tau_m}^i > c(m)\}} \quad (5.30)$$

where the unadjusted payment date firm value V_{τ_m} is calculated by a simple rearranging of (5.11) and setting $t = \tau_m$. By using the sum of the product of the indicator functions, we are assured that we obtain the number of paths that survived both at maturity and the barrier at all dates up to maturity.

The calculation of survival probabilities under the Briys and de Varenne model is almost the same as that of the Black and Cox model for the PZ case. The prescence of deviations from absolute priority does not affect the probability of survival since f_1 and f_2 do not determine the firm values or the level of the barrier. Rather, they simply determine the payoff once defaults occurs. For this reason, we only make one adjustment to the calculation of survival probabilities under the Briys and de Varenne model. We modify (5.30) to account for the fact we are now concerned

with unadjusted firm values.

$$P_{t_j}(\tau_m) = \frac{1}{N} \sum_{i=1}^N \mathbb{I}_{\{V_t^i > K_t^i(\tau_m); \forall t; t_j \leq t \leq \tau_m\}} \cdot \mathbb{I}_{\{V_{\tau_m}^i > c(m)\}} \quad (5.31)$$

5.9 Extension to Internal Consistency

Estimation under an IC set up would be rather difficult if we were to require closed-form solutions. However, with Monte Carlo simulation, it is easy. Along a simulation path, we can price in an IC manner by setting the values of subsequent payments to zero when a payment defaults. This accounts for the fact that once default occurs and the bondholders take control of the firm's assets, the firm ceases to exist.

Under the IC approach, the default time is defined as the time the first default occurs. For the Black and Cox, and Briys and de Varenne models, this could be at either the maturity of a payment or along the default barrier of a payment. With IC, we essentially have one default barrier for all payments of the bond. At any one time, the default barrier is the maximum level of all the individual default barriers. As soon as this barrier is hit by the firm value, all remaining payments default. The log-likelihood function is exactly as discussed and we proceed in the usual manner of finding root firm values to evaluate the log-likelihood function.

We have to account for this change when calculating survival probabilities. Under Merton and an IC approach, the survival of the m^{th} remaining payment also requires the survival of all payments that occur before τ_m . We account for this through a modification of (5.29). The probability of survival at time t_j of the m^{th} remaining

payment $c(m)$ of transaction j is estimated as,

$$P_{t_j}(\tau_m) = \frac{1}{N} \sum_{i=1}^N \left\{ \prod_{k=1}^m \mathbb{I}_{\{V_{\tau_k}^i > c(k)\}} \right\} \quad (5.32)$$

This treatment of the probability of survival conditions on survival of earlier payments for each simulation path. A path i is only counted (assigned 1) if every payment for path i survives up to and including the time τ_m payment.

With the Black and Cox model, we also have to account for the default of *any* other payment through its respective default barrier prior to τ_m .¹⁴ Therefore, we modify (5.30) to account for the fact we are now calculating a conditional probability. Therefore, the survival probability of the m^{th} remaining payment of transaction j is,

$$P_{t_j}(\tau_m) = \frac{1}{N} \sum_{i=1}^N \left\{ \prod_{k=1}^m \mathbb{I}_{\{V_{\tau_k}^i > c(k)\}} \right\} \left\{ \prod_{h=1}^{M_j} \mathbb{I}_{\{\hat{V}_t^i > K_t^i(\tau_h); \quad \forall t; t_j \leq t \leq \min(\tau_m, \tau_h)\}} \right\} \quad (5.33)$$

The same argument applies to the Briys and de Varenne model. In modifying (5.31) we make the same adjustment as we do for the Black and Cox IC case.

$$P_{t_j}(\tau_m) = \frac{1}{N} \sum_{i=1}^N \left\{ \prod_{k=1}^m \mathbb{I}_{\{V_{\tau_k}^i > c(k)\}} \right\} \left\{ \prod_{h=1}^{M_j} \mathbb{I}_{\{V_t^i > K_t^i(\tau_h); \quad \forall t; t_j \leq t \leq \min(\tau_m, \tau_h)\}} \right\} \quad (5.34)$$

It follows from these formulae that the risk-neutral probability of default at time t_j of the m^{th} remaining payment of transaction j for any model can be estimated as $1 - P_{t_j}(\tau_m)$.

¹⁴Specifically, the only later payment's barrier a coupon can default at is the barrier of the maturity payment since the payment is larger than a coupon.

5.10 Intensity Model Estimation Overview

We now turn to the estimation of the reduced-form model. Estimation of this models is much easier than that of the structural models since simulation is not required.

As stated in Chapter 3, through the assumption of independence between the interest rate and default processes, the closed-form solution for the CIR default intensity models can be used regardless of interest rate process. Therefore we do not need to perform simulation to estimate the parameters of the process for λ_t . This makes estimation for the reduced-form model much easier than that of the structural models.

We again use ML estimation to estimate the parameters of the default intensity process under the real world probability measure \mathbb{P} . Thus we are again faced with the problem of observability we had with the structural models. We do not observe the default intensity λ_t . If we did, we could simply use these values to evaluate the log-likelihood function and find which parameters maximise the likelihood function. Since we only observe prices, the transformed variable ML method developed in Duan (1994) is again used. The bond prices are used to derive root intensities which can then be used to evaluate the log-likelihood along with the parameters.

Risky coupon-paying bonds can be priced using (3.17) and (5.4). Specifically, for a given transaction, a survival probability is calculated for each remaining payment of the bond in order to compute the model price. Probabilities are calculated using the closed-form solution presented in Chapter 3.

Since the pricing of a risky coupon bond in the reduced-form model assumes that

the coupon payments have zero recovery, we must separately value the maturity coupon and face value unlike in the structural models where the maturity payoff is valued as a single payment. The value at the transaction j date t_j of \$1 in promised coupons at time τ_m with zero recovery is,

$$\bar{B}_{t_j}(\lambda_{t_j}, \tau_m, 1, 0, 0) = B_{t_j}(\tau_m, 1, 0)P_{t_j}(\tau_m)$$

This formula is simply (3.17) with $\delta = 0$ and maturity $T = \tau_m$. The value of \$1 of promised principal is also valued in the same way except a constant recovery of δ is assumed. Therefore, we can price transaction j of the risky semi-annual coupon-paying bond as,

$$\begin{aligned} \bar{B}_{t_j}(\lambda_{t_j}, T, F, c, \delta) &= F\bar{B}_{t_j}(\lambda_{t_j}, T, 1, c, \delta) \\ &= F\bar{B}_{t_j}(\lambda_{t_j}, T, 1, 0, \delta) + \frac{c}{2}F \sum_{m=1}^{M_j} \bar{B}_{t_j}(\lambda_{t_j}, \tau_m, 1, 0, 0) \\ &= F\bar{B}_{t_j}(\lambda_{t_j}, T, 1, 0, \delta) + C \sum_{m=1}^{M_j} \bar{B}_{t_j}(\lambda_{t_j}, \tau_m, 1, 0, 0) \quad (5.35) \end{aligned}$$

where c is the coupon rate, $C \equiv \frac{c}{2}F$ is the coupon payment in dollars and the second equality is given by (5.4).¹⁵ The input required to use (5.35) is the recovery parameter δ . As with Duffee (1999), we use $\delta = 0.44$ as the fractional face value recovered upon default. Duffee justifies this choice by pointing out that Moodys reports senior bondholders receive around 44% of face value when default occurs.¹⁶

Discount factors (risk-free zero-coupon bond prices) for each payment date are given

¹⁵This is for semi-annual coupons. For annual coupons we have $C \equiv cF$.

¹⁶Jorion (2007) reports that the fraction is 53% for the period 1982-2002. Numerous factors such as the business cycle and the debt structure of a firm affect the recovery rate.

by the assumed interest rate process and are independent of the default process. Therefore, for a given set of parameters for the default intensity process, we can calculate survival probabilities and price a bond. We can then use a root finder to obtain the root that sets the model price equal to the actual price. i.e. find $\lambda_{t_j} = \lambda_{t_j}^r$ such that,

$$\overline{B}_{t_j}(\lambda_{t_j}^r, T, F, c, \delta) - \overline{B}_{t_j}^{ob}(T) = 0$$

We find the root intensity for each transaction j which are then used to evaluate the likelihood function along with the default intensity parameters.

5.10.1 CIR Intensity

We wish to estimate the parameters for the process under the real world probability measure \mathbb{P} , since this is the measure that actually generates our data. Therefore, as with the structural models, we price in the risk-neutral world in order to estimate the parameters under the real world. For consistency, we assume the same form for the market price of default risk ν as that of the CIR short rate model discussed in section 5.2.2 and estimate the market price of default risk as a part of the ML procedure.

The process followed by the default intensity under \mathbb{P} is,

$$d\lambda_t = a_\lambda(\mu_\lambda^P - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dW_t \quad (5.36)$$

where dW_t is a Brownian motion under \mathbb{P} . From our earlier definition of market price of risk, it then follows that we must have,

$$\mu_\lambda^P = \mu_\lambda + \frac{\sigma_\lambda \nu}{a_\lambda} \quad (5.37)$$

where μ_λ is the mean intensity level under the risk-neutral probability measure \mathbb{Q} . Given a set of parameters a_λ , μ_λ^P , σ_λ and ν , we calculate the mean intensity level μ_λ for the risk neutral world and calculate survival probabilities. Under the \mathbb{Q} measure CIR model for default intensity, the probability at transaction date t_j of survival to τ_m is of the form,

$$P_{t_j}(\tau_m) = \alpha(t_j, \tau_m) e^{-\beta(t_j, \tau_m) \lambda_{t_j}} \quad (5.38)$$

where $\alpha(t, \tau_m)$ and $\beta(t, \tau_m)$ are computed using (3.20) and (3.19) with $T = \tau_m$. We can calculate these probabilities for each payment date and use (3.17) and (5.35) to price the bond given an interest rate process. Under a constant interest rate setup, the price at t_j of a risk-free zero coupon bond paying \$1 at τ_m is,

$$B_{t_j}(\tau_m, 1, 0) = e^{-r(\tau_m - t_j)}$$

where, again, r is the average of the 3 monthly bill yield time-series. Under a Vasicek or CIR setup, risk-free bond prices can be calculated using the closed-form solutions discussed in Chapter 2.

Once a root intensity is obtained for each transaction, we can evaluate the likelihood function for the real world process (5.36). As with the short rate under the CIR model, the intensity is non-central chi-squared distributed. However, we apply the same normality approximation when estimating the intensity model parameters as we do for the CIR short rate model. Therefore the log-likelihood if λ_t were observed

would be,

$$\begin{aligned}
\ln L(\lambda_t^{ob}, \theta) &= \ln \prod_{j=1}^{J-1} \frac{1}{\sigma_\lambda \sqrt{2\pi \lambda_{t_j}^{ob} \Delta t_j}} \exp \left\{ -\frac{1}{2} \left(\frac{\Delta \lambda_{t_j}^{ob} - a_\lambda(\mu_\lambda^P - \lambda_{t_j}^{ob}) \Delta t_j}{\sigma_\lambda \sqrt{\lambda_{t_j}^{ob} \Delta t_j}} \right)^2 \right\} \\
&= \sum_t \ln \left(\frac{1}{\sigma_\lambda \sqrt{2\pi \lambda_{t_j}^{ob} \Delta t_j}} \right) - \frac{1}{2} \sum_t \left(\frac{\Delta \lambda_{t_j}^{ob} - a_\lambda(\mu_\lambda^P - \lambda_{t_j}^{ob}) \Delta t_j}{\sigma_\lambda \sqrt{\lambda_{t_j}^{ob} \Delta t_j}} \right)^2
\end{aligned}$$

where $\Delta \lambda_t^{ob} = \lambda_{t_{j+1}}^{ob} - \lambda_{t_j}^{ob}$ and θ is a vector of the default intensity model parameters. However, as with the firm value in the structural models, we do not observe λ_{t_j} directly but rather as a function of the bond price $\bar{B}_{t_j}^{ob}$. Therefore we need to account for the Jacobian term. The log-likelihood function is hence,

$$\begin{aligned}
&\ln L(\lambda_t^r, \theta) \\
&= \ln \prod_{j=1}^{J-1} \frac{1}{\sigma_\lambda \sqrt{2\pi \lambda_{t_j}^r \Delta t_j}} \exp \left\{ -\frac{1}{2} \left(\frac{\Delta \lambda_{t_j}^r - a_\lambda(\mu_\lambda^P - \lambda_{t_j}^r) \Delta t_j}{\sigma_\lambda \sqrt{\lambda_{t_j}^r \Delta t_j}} \right)^2 \right\} \frac{d\lambda_{t_{j+1}}}{d\bar{B}_{t_{j+1}}} \\
&= \sum_{j=1}^{J-1} \ln \frac{1}{\sigma_\lambda \sqrt{2\pi \lambda_{t_j}^r \Delta t_j}} \exp \left\{ -\frac{1}{2} \left(\frac{\Delta \lambda_{t_j}^r - a_\lambda(\mu_\lambda^P - \lambda_{t_j}^r) \Delta t_j}{\sigma_\lambda \sqrt{\lambda_{t_j}^r \Delta t_j}} \right)^2 \right\} \\
&\quad + \sum_{j=1}^{J-1} \ln \left| \frac{d\lambda_{t_{j+1}}}{d\bar{B}_{t_{j+1}}} \right| \\
&= \sum_{j=1}^{J-1} \ln \frac{1}{\sigma_\lambda \sqrt{2\pi \lambda_{t_j}^r \Delta t_j}} \exp \left\{ -\frac{1}{2} \left(\frac{\Delta \lambda_{t_j}^r - a_\lambda(\mu_\lambda^P - \lambda_{t_j}^r) \Delta t_j}{\sigma_\lambda \sqrt{\lambda_{t_j}^r \Delta t_j}} \right)^2 \right\} \\
&\quad - \sum_{j=1}^{J-1} \ln \left| \frac{d\bar{B}_{t_{j+1}}(\lambda_{t_{j+1}}^r, T, F, c, \delta)}{d\lambda_{t_{j+1}}^r} \right| \tag{5.39}
\end{aligned}$$

The Jacobian can be approximated as in the structural case. We can offset the root $\lambda_{t_j}^r$ for transaction j by some factor ϵ_j to obtain an offset root,

$$\lambda_{t_j}^{r'} = \lambda_{t_j}^r + \epsilon_j$$

The derivative is then approximated as,

$$\frac{d\overline{B}_{t_j}(\lambda_{t_j}, T, F, c, \delta)}{d\lambda_{t_j}} \approx \frac{\overline{B}_{t_j}(\lambda_{t_j}^{r'}, T, F, c, \delta) - \overline{B}_{t_j}^{ob}(T)}{\epsilon_j} \quad (5.40)$$

since $\overline{B}_{t_j}(\lambda_{t_j}^r, T, F, c, \delta) = \overline{B}_{t_j}^{ob}(T)$. We can thus find roots for all the transactions of a bond and evaluate the log-likelihood function using (5.40) to approximate the derivative in the Jacobian. The set of parameters a_λ , μ_λ^P , σ_λ and ν (along with the roots they imply) that maximise (5.39) are the ML estimates for the CIR intensity model.

5.11 Asymptotic Variance

We can also estimate the asymptotic variance of our ML estimates for all of the discussed models in order to determine their accuracy. Greene (2000) outlines a procedure to estimate the variance-covariance matrix of the ML estimates based on the first derivatives of the log-likelihood. The variance-covariance matrix is defined as,

$$[I(\theta)]^{-1} = \left\{ -E \left[\frac{\partial^2 \ln L(\theta)}{\partial \theta \partial \theta'} \right] \right\}^{-1} \quad (5.41)$$

where θ is a vector of the actual parameters and $I(\theta)$ is the inverse variance-covariance matrix. Greene (2000) states that (5.41) can be easily approximated by using the fact that the variance-covariance matrix of the first derivative is equal to the expected second derivatives matrix. (5.41) can therefore be approximated using,

$$[\hat{I}(\theta)]^{-1} = \left[\sum_{p=1}^n \hat{g}_p \hat{g}_p' \right]^{-1} \quad (5.42)$$

where

$$\hat{g}_p = \frac{\partial \log f(V, \hat{\theta})}{\partial \hat{\theta}} \quad (5.43)$$

is a vector of length J with V being the root firm values or intensities, $\hat{\theta}$ being a vector of the estimated parameters and $f(\dots)$ returning the individual log-likelihood of a transaction. We calculate \hat{g}_p with a discrete approximation by offsetting parameter p , calculating new roots and then evaluating the individual log-likelihoods. For each parameter estimated under a given model, we then calculate the difference between the log-likelihoods due to the offset and divide by the change in the parameter. Our choice of offset for parameter θ_p is to add $0.0001\theta_p$ to θ_p . So we have

$$\hat{g}_p = \frac{\log f(V_\epsilon, \hat{\theta} + \epsilon) - \log f(V, \hat{\theta})}{\epsilon} \quad (5.44)$$

where V_ϵ is the root firm values when parameter p has been offset by ϵ .

5.12 One-Step Ahead Forecasting

Once we have obtained parameter estimates for a given model, we can perform one-step ahead forecasting to test the accuracy of the estimates. We can apply drift to the root of transaction j and use this as the initial value in pricing transaction $j + 1$. In the structural model case, this method is to take the root $V_{t_j}^r$ for transaction j and apply drift using parameter estimates to obtain an estimate of the firm value $V_{t_{j+1}}$ for transaction $j + 1$. This firm value can then be used as the starting point for simulation and along with the log-likelihood maximising parameters, can be used to compute the model price $\bar{B}_{t_{j+1}}(V_{t_{j+1}}, T, F, c)$ for transaction $j + 1$.¹⁷ This price can

¹⁷It is important here to use the same seed that is used in estimation.

then be compared to the actual price $\overline{B}_{t_{j+1}}^{ob}$.

Since the firm value follows a geometric Brownian motion, it follows that,¹⁸

$$E_{\mathbb{P}}[V_{t+\Delta t}] = V_t e^{\mu \Delta t}$$

Therefore we can compute an estimate of $V_{t_{j+1}}$ using $V_{t_j}^r$ through,

$$V_{t_{j+1}}' = V_{t_j}^r e^{\mu \Delta t_j} \quad (5.45)$$

Therefore, the application of structural bond pricing models discussed above does not require the use of equity price data or balance sheet data. All that is required to price a bond under the given models is information about the bond issue and a series of observed prices for the bond. Using these prices we can estimate the required parameters and use the root values given by the log-likelihood maximising parameters to test our results.

We can also compute an estimate of the one-step ahead intensity under the CIR model. Under CIR,¹⁹

$$E_{\mathbb{P}}[\lambda_{t+\Delta t}] = \lambda_t e^{-a\lambda \Delta t} + \mu_{\lambda}^P (1 - e^{-a\lambda \Delta t})$$

The one-step forecast of $\lambda_{t_{j+1}}$ is then,

$$\lambda_{t_{j+1}}' = \lambda_{t_j}^r e^{-a\lambda \Delta t_j} + \mu_{\lambda}^P (1 - e^{-a\lambda \Delta t_j})$$

¹⁸See Hull (2006).

¹⁹See Cox, Ingersoll and Ross (1985).

5.13 Calculating Risk-Free Yields

As stated in Eom, Helwege and Huang (2004), the percentage error in price is not adequate to gauge the performance of a model. This is because a small error in price may still lead to a large error in spread. Therefore we wish to compare the spreads predicted by the models to the actual spreads.

To calculate a spread for a bond transaction, we need to compute the yield of equivalent risk-free bond. This would be straightforward if we observe risk-free zero coupon yields for each payment date. Since here, calculating the risk-free yield would simply require discounting all the promised payments at their respective yields and computing a risk-free price. This price would then be used to calculate the risk-free yield.

However, we only observe par-yield data and require a means of estimating zero-coupon yields. We can do so using cubic interpolation and the minimisation of errors.²⁰

For a given transaction j and i^{th} par-yield maturity of T^i years, we can interpolate risk-free zero-coupon yields (and hence risk-free zero-coupon bond prices) for six month increments up to T^i years. This is done using a vector of initial zero-coupon yields for all of the par-yield maturities we observe up to T^i years.²¹ We can then price a risk-free semi-annual coupon bond at t_j maturing in T^i years as,

$$B_{t_j}^i(t_j + T^i, 1, y_{t_j}(t_j + T^i)) = \sum_{\tau=1}^{2T^i} \frac{y_{t_j}(t_j + T^i)}{2} B_{t_j}^i(t_j + \frac{\tau}{2}, 1, 0) + B_{t_j}^i(t_j + T^i, 1, 0) \quad (5.46)$$

²⁰Cubic interpolation is discussed in Martellini, Priaulet and Priaulet (2003).

²¹We observe par yields for maturities of 6 months, 1 year, 2 years, 3 years, 5 years, 7 years, 10 years, 20 years and 30 years.

where $B_{t_j}(t_j + \frac{\tau}{2}, 1, 0)$ is the price of a risk-free zero-coupon bond paying \$1 at time $t_j + \frac{\tau}{2}$ and $y_{t_j}(t_j + T^i)$ is the par-yield at time t_j for maturity at time $t_j + T^i$.

The coupon paying bond price (5.46) should equal one since $y_{t_j}(t_j + T^i)$ is the par-yield at t_j for a term of T^i years. We can thus calculate an error for the i^{th} par-yield.

$$\epsilon_i = \left(B_{t_j}^i(t_j + T^i, 1, y_{t_j}(t_j + T^i)) - 1 \right)^2$$

We repeat this procedure for all N par yield maturities we observe and compute the total error as,

$$S = \sum_{i=1}^N \epsilon_i \quad (5.47)$$

We then vary the initial vector of risk-free zero-coupon yields until (5.47) is minimised. This procedure is repeated for all J transactions so that we have a set of risk-free zero-coupon yields for every par-yield maturity for every transaction. The zero-coupon yields for the par-yield maturities for transaction j are then used to interpolate the risk-free zero-coupon yields on every remaining payment date of the corporate bond. These yields are used to discount the individual promised payments and calculate the price of a risk-free equivalent. This in turn gives us the yield of a risk-free equivalent. Spreads can then be calculated as the difference between the risky coupon bond's yield and the risk-free equivalent's yield.

Chapter 6

Case Study

6.1 Overview

In applying the estimation procedure discussed in Chapter 5, we run into many issues. These issues range from no roots, multiple roots and non-convergence. We only estimate all of the models discussed for one bond as a case study. This is so that we can delve deeper into the results of the estimation and the issues that arise during estimation. These issues, discussed in the next chapter, are problems we encounter using the Monte Carlo ML estimation approach and are issues that must be overcome in order to obtain results. Our Monte Carlo method can be applied to bonds consistent with the model assumptions if price data is available. The downside is that the root finding discussed in chapter 5 can be rather unreliable. In addition, there are issues with obtaining successful convergence when optimising. Analysis of these issues is more important in the context of practitioners applying

this approach rather than the reporting of results for multiple bonds.

Since our approach uses price data and thus does not require equity prices to be observable, it can be applied to firms that are private. Therefore it seems to be natural to test our approach on a bond belonging to a firm that is private. Information about this bond issue is obtained from the Fixed Investment Securities Database (FISD). Price observations for this bond are obtained from the National Association of Insurance Companies (NAIC). We obtain 3 month Treasury bill yields from the St. Louis Federal Reserve as a proxy for the short rate. We use this yield series since it is the shortest maturity yield series available that covers our entire date range. Longer maturity Treasury yields are also obtained from the St. Louis Federal Reserve in order to test the results of the term structure models.

We are required to choose a bond from the database. Our focus is purely on default risk and this is complicated by the fact that the database contains bonds with many different provisions and the fact that these properties affect the price of a bond cannot be ignored. We wish to avoid these types of bonds since including them would involve trying to price the implicit options embedded in a bond's price due to these provisions.

We choose a bond issued by URC Holdings Corp as the bond we price using the structural and reduced-form models. The firm is private and the bond has no implicit options we are required to price. That is, it is not callable, puttable or convertible. It has no sinking fund options and is not part of a unit offering. The bond also pays a fixed coupon and is denominated in US dollars.

In addition the bond is senior and the firm in question has no other bonds in the

database. This is important because the properties of the underlying firm should closely match the properties of the hypothetical simple capital structure firm in the models. It is not reasonable to assume that the model can price a bond that is not first in line to the assets when the bond really is not. In this situation it is not clear whether a large discrepancy between an actual price and a model price is due to the model performing poorly or whether it is due to a dataset that is inconsistent with the assumptions of the firm in the models.¹

6.2 Summary Statistics

In this section we analyse the bond we price. We also examine the 3 month bill-yield time series used to estimate the parameters of the Vasicek and CIR models.

6.2.1 Bills Times Series Data

In reality we do not observe short rates. Therefore we need to find some proxy to serve as a short rate series. Figure 6.1 plots the three-month bill yield data from the St. Louis Federal Reserve that we use as the proxy for the short rate in order to estimate the parameters of the Vasicek and CIR term structure models. We view the three-month bills yields as relatively short maturity securities appropriate to use to estimate the term structure models. The data covers the period from the start of 1994 through to the end of 2006 in order to cover the date range of our corporate price observations. Table 6.1 gives summary statistics for the bills data we use.

¹This method of filtering is quite artificial however, since the appearance of only one bond for a firm in the database does not necessarily mean that the firm has issued just one bond. However this type of filtering is consistent with the filtering in existing bond pricing literature and therefore

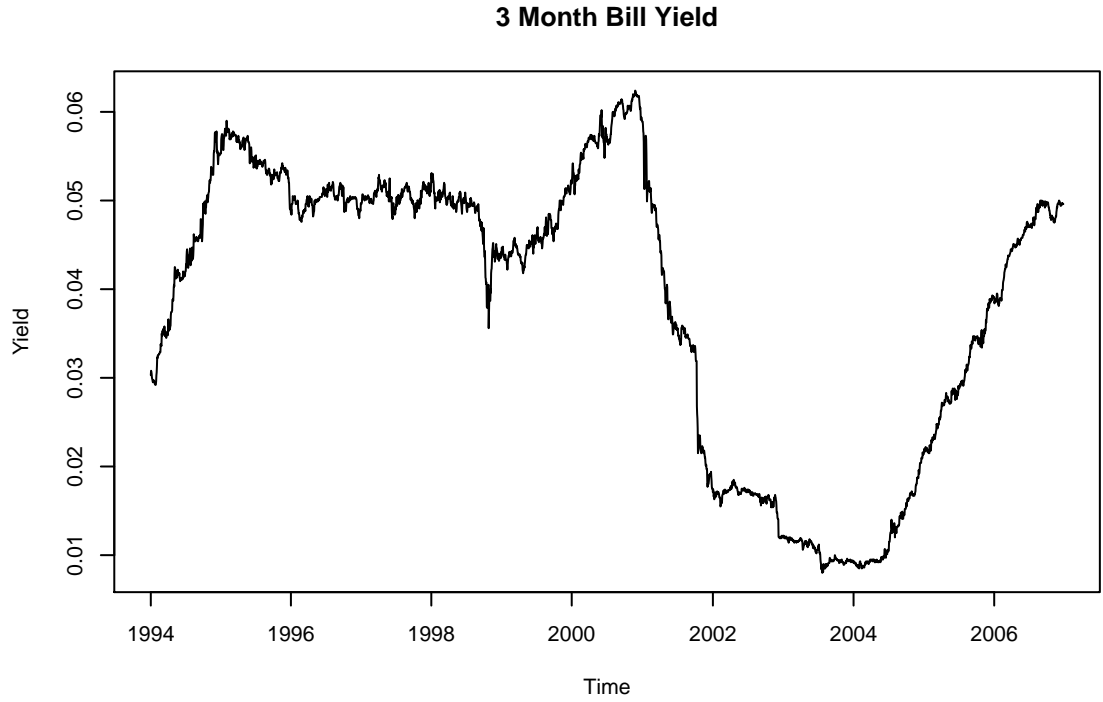


Figure 6.1: 3 Month Bill Yields.

Mean	0.0386
Standard Deviation	0.0163
Skewness	-0.6353
Kurtosis	-1.0655
Max	0.0624
Min	0.0080

Table 6.1: 3 Month Bill Yield Time-Series Statistics.

6.2.2 The Corporate Bond

We observe 71 price observations for the corporate bond we study in this paper. However, many of these transactions are for the same day with the same price. We

we do the same.

Firm	URC HOLDINGS CORP
Industry	Insurance
Issued	20/06/1996
Maturity	30/06/2006
Coupon Rate (%)	7.875
Coupon Frequency	Semi-Annual
Mean Yield (%)	7.01

Table 6.2: Summary information on the corporate bond priced in this paper.

observe 34 unique transactions for the bond and these are transactions we use in estimation. All of the prices are based on a \$100 par value basis. Table (6.2) contains information for the bond issue. From 26/07/2000 through to 4/10/2000 the bond is classed as speculative grade whilst investment grade at all other times. Of the 34 observations, all of the transactions occur when the bond is classed as investment grade except for one transaction that occurred on 31/08/2000. The yield of this transaction is 7.67% whilst the mean yield of the investment grade transactions is 6.99%.

Chapter 7

Methodological issues

7.1 Overview

In this chapter, we discuss issues that arise when implementing the bond pricing models. The Monte Carlo ML approach is slow and optimising can be difficult due to a jagged estimated log-likelihood surface. In addition, we encounter situations where roots do not exist. Evaluation of the log-likelihood function requires estimates of firm values. For each transaction and under given parameters, we derive a root firm value that best describes the actual bond price observed on the transaction date. However, situations arise where these roots do not actually exist. All of the examples given in this chapter are taken from our work with the corporate bond applied in this paper.

7.2 Root Finding Speed

We use bisection to find root firm values. This method evaluates the price difference at the midpoint of a user-specified interval for the initial firm value. If the price differential is positive then this midpoint becomes the upper limit in the next price difference evaluation.¹ If instead the price differential is negative, it sets the midpoint as the lower limit in the next evaluation. The result is that the root finder zeroes in on the correct root. The downside of this method is that it is very slow since it continually halves the interval until it finds the root.

We could instead use a method that relies on the slope of the pricing function (e.g. Newton-Raphson, Secant). These methods have the advantage that they converge to the root much quicker than bisection but also have the disadvantage of being rather unreliable. The slope of the pricing function is flat at either extreme. As the initial firm value is increased, the bond price goes toward being risk free and the price function thus becomes flat. As the initial firm value is decreased, the bond's price hits a floor (zero for Merton, the minimum barrier payoffs for Black and Cox, and Briys and de Varenne). As a result, it is not uncommon for a method relying on the slope of the function to become 'stuck' at either extreme and give an incorrect root.

One solution to this is to restrict the range in which the root is searched for. This would involve using bisection with a low precision to determine a bounded range in which the root exists. A root finding method relying on slope is then used to find the root. However this 'hybrid' method is also unreliable if the slope of the pricing function changes sharply. Thus, convergence at one of the boundaries is still likely

¹This assumes that the price is always increasing in initial firm value. However, as we shall discuss, this is not always true.

to occur. The bisection precision could be increased to restrict ourselves to a smaller range but in this case, bisection may as well be used to find the root itself.

Therefore we are faced with a tradeoff. We can use a root finding method that is fast but has performance issues, or a root finding method that is slow but is guaranteed to find the actual root. We use bisection as the root finding method because of its reliability. The downside, is that a maximum likelihood approach, including this method, becomes very slow, often taking more than couple minutes to perform one log-likelihood evaluation (with 600 simulation paths) for a 10-year bond under the Black and Cox, and the Briys and de Varenne models.²

7.3 Jagged Surfaces

Another significant problem in the estimation approach is that of non-convergence. A rather robust optimiser is required to estimate the parameters for the models. This is due to the estimated log-likelihood surface being quite jagged. Presumably, increasing the number of simulation paths would aid in smoothing out the surface. However, this would require more computation and therefore, more time.

²Li and Wong (2008) also encounter speed issues with the maximum likelihood approach.

Merton – PZ: Negative Log-Likelihood

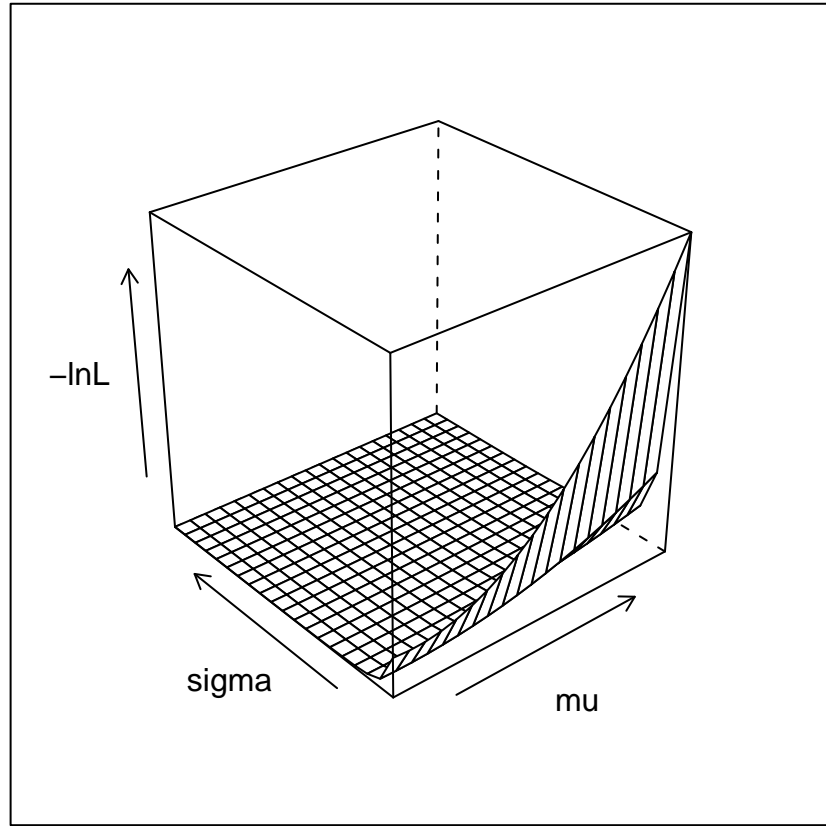


Figure 7.1: Negative log-likelihood surface for the Model model under a constant interest rate and the PZ approach. The surface is plotted for μ and σ between 0 and 1.

Figure 7.1 plots the negative of the log-likelihood surface of the Merton model under the PZ approach and a constant interest rate assumption for the risky bond we price.³ The function seems rather flat and smooth. However, if we zoom into the range of convergence (figure 7.2), we find the surface is actually rather rugged across differing levels of σ .

³Most optimisers by default, find a minimum thus we change the sign of the log-likelihood function.

Merton – PZ: $-\ln L$ around Convergence Point

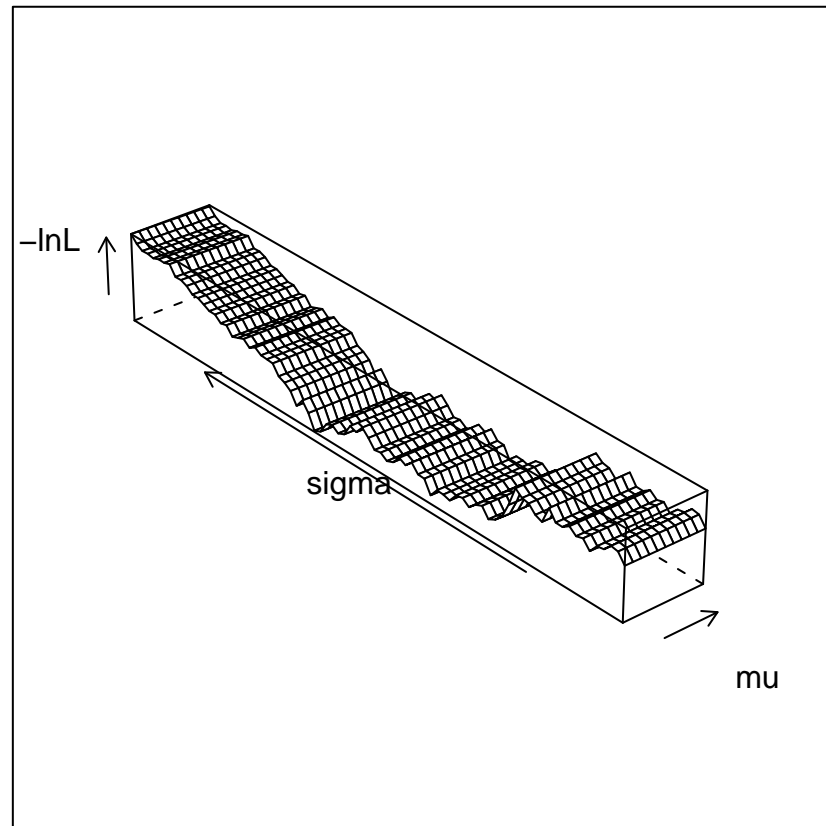


Figure 7.2: Negative log-likelihood surface for Merton - constant rate - PZ for $\mu \pm 0.05$ and $\sigma \pm 0.4$.

This issue is even more apparent for the internally consistent (IC) case of the Merton model under a constant interest rate (figures 7.3 and (7.4)).

Merton – IC: Negative Log–Likelihood

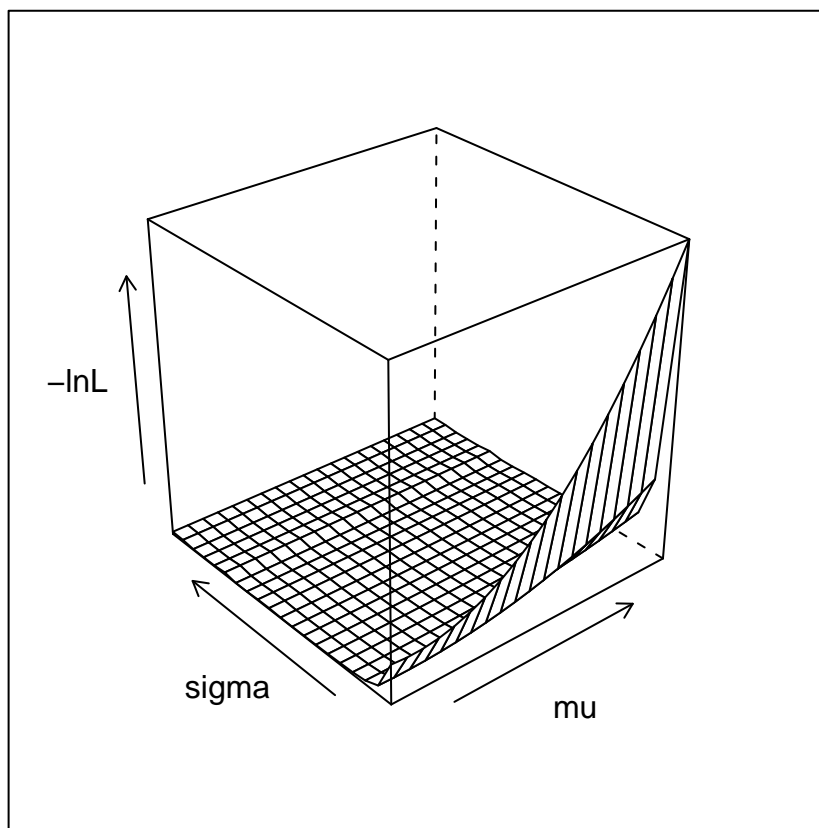


Figure 7.3: Negative log-likelihood surface for Merton - constant rate - IC. The surface is plotted for μ and σ between 0 and 1.

Merton – IC: $-\ln L$ around Convergence Point

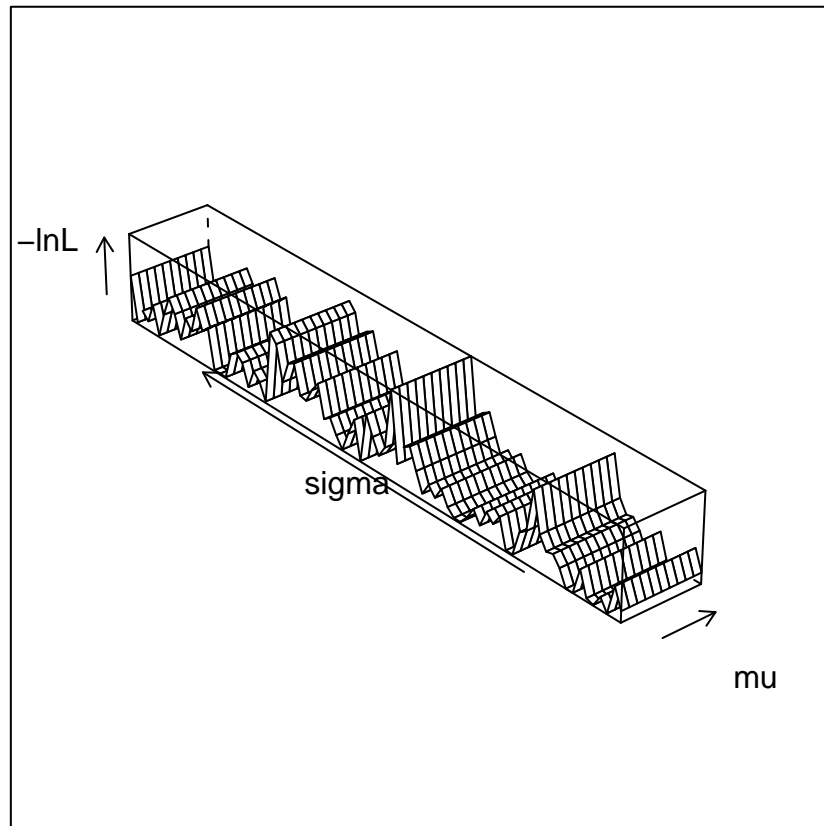


Figure 7.4: Negative log-likelihood surface for Merton - constant rate - IC for $\mu \pm 0.05$ and $\sigma \pm 0.4$.

Here the convergence region is even more jagged. This follows from the fact that the consequences of default are more severe for an IC case than a PZ case. Under an IC scenario, given a set of initial firm values, a small change in volatility can lead to large price changes. For a given transaction, a volatility increase may lead to the early default of paths that were previously surviving many payments. Thus the price of the bond along these paths would drop significantly and as a result, the initial firm value for the transaction would need to rise by a potentially large amount to set the Monte Carlo price equal to the actual price. Conversely, a small decrease in

volatility may lead to the survival of previously defaulting paths and thus large price increases. Here the initial firm value would need to drop significantly.⁴ This large change in root would hence lead to large movements in the log-likelihood function.

Under the PZ approach, default of payments do not imply default of other payments so the consequences of default are much less severe. Therefore small changes in volatility do not lead to overly large changes in prices and hence do not lead to large shifts in roots. As a result the log-likelihood function does not change significantly.

The consequence of having a jagged likelihood surface is that the optimiser has a difficult time searching for the log-likelihood-maximising parameters. Coupled with the fact that root finding is slow, estimation is made quite difficult. Parameters often do not shift significantly from initial values. This is because the initial value may be at a local minimum. This problem is more apparent for optimisers that rely on gradients to search for the optimum point. Therefore, it is better to use an optimiser that only uses function values and does not calculate gradients. These methods, however, typically do not allow bounds to be set on the parameters. This can be problematic since we have parameters that we wish to restrict (α , f_1 and f_2).

A way around this issue may be to perform a grid search to find appropriate starting values for the parameters. However, this is expensive in terms of computational time when considering the Black and Cox, and Briys and de Varenne models since these models have more parameters that need to be estimated than the Merton model. Calculating roots for these models is also slower than that of the Merton model since further code-optimisation is not possible. For a given transaction, simulation

⁴The effect on model prices of a change in σ , which is discussed in section 7.4, depends on the overall level of σ itself.

path and promised payment, checking for barrier default requires checking a vector of firm values to the barrier values. Therefore it is not possible to check multiple payments for barrier default at the same time. It also seems necessary to run the optimiser multiple times for a single model. Each consecutive run of the optimiser uses a smaller minimum parameter step than the last in an attempt to differentiate between a local optimum and a global optimum. Again, this is expensive computationally.

There is a clear trade-off in the implementation of these models. We can either have a large number of simulation paths and (possibly) a smoother likelihood surface but slow root finding, or we can have a smaller number of paths which leads to faster evaluations of the likelihood surface but potentially less reliable optimisation.

7.4 Non-Existence of Roots

Non-parameter-specific non-existence of roots is a problem associated with the term structure models. Summing the discounted face value of all remaining payments of a bond results in the price of an equivalent risk-free bond. If the actual price of the transaction exceeds this risk-free price, then no root exists for the transaction. This issue arises due to the remaining payments being discounted at interest rates that are too high. In this case, the only solution is to omit these transactions since the risk-free price sets a ceiling on the price of the risky bond.

This scenario is more likely to occur for the constant interest rate set up, since the constant interest rate we use does not account for the overall level of rates at the time of the transaction. For a stochastic interest rate case, we simulate rates consistent

with interest rates at the time of the transaction and therefore non-existence of roots is less likely. In order to calculate the price of a risk-free equivalent under a stochastic interest rate, each promised payment is valued as a zero-coupon bond using the Vasicek or CIR closed-form solutions presented in Chapter 2. These prices are the limit of the prices under the interest rate simulations and hence, give a reliable indicator of which transactions are likely to be problematic for the assumed data generating process.

7.5 Parameter-Specific Non-Existence of Roots

We can run into problems with non-existence of roots for specific combinations of parameter values in a model. This situation, which arises for the Black and Cox, and Briys and de Varenne models, is when the default barriers create a floor for the Monte Carlo price that exceeds the actual price. One solution is to restrict what values a parameter can take. However, this is easier said than done since the problem is not due to individual parameter levels but rather a combination of parameter levels.

Consider a single transaction under the Black and Cox model where γ exceeds r . Here, the Monte Carlo price is an increasing function of initial firm value since a payment's barrier at any point is less than the discounted face value of the payment. Therefore bondholders do not prefer default over survival and the Monte Carlo bond price is higher the greater the probability of survival. If the default barriers are quite high, then the Monte Carlo price will be high even for the lowest possible initial firm value. A root will exist if the difference in the Monte Carlo price and the actual price is negative for the lowest possible firm value. However, a small change in one

of the parameters may lead to the situation where the price differential is positive at the lowest initial firm value. In this case, no root exists for the transaction.

There are two effects of an increase in σ . The first is increased variability in the firm value simulations. This increased variability is in both directions, hence, can cause both earlier default and later survival. The second effect is in the drift of the log-firm value. A high level of σ corresponds to stronger downward drift in the firm value and therefore, more defaults.

Given the initial firm value V_{t_j} , the effect of increasing the volatility σ depends on the level of σ itself. If it is quite low, then an increase in σ may lead to an increase in price, the reason being increased variation will create higher firm values along a path and potentially, later survival of payments along that path. An increase in σ may also lead to lower firm values along a path and hence earlier default, though for low σ , we observe an increase in the Monte Carlo price. As σ gets very high, a strong downward drift drives down simulated firm values leading to more defaults and a lower Monte Carlo price.

The price increase due to an increase in σ may create the situation where, even at the lowest possible initial firm value, the price differential is positive. Here no root can be found for the transaction. One solution to this problem is to exclude transactions where roots do not exist as they arise. That is, for each evaluation, only compute the log-likelihood function for the transactions where roots exist. However, this is likely to cause problems when roots cannot be found for a large proportion of the transactions. We may find that the optimiser reaches its optimum point at these parameter levels. This is clearly not desirable since this set of parameter values cannot explain a large number of our actual prices and it is unreasonable to assume

that this is the optimum point.

Another solution to this non-existence problem is to set the log-likelihood to $-\infty$ when no root can be found. This forces the optimiser to only converge at a set of parameter values where roots exists for all transactions. However this leads to the situation where we have a discontinuous log-likelihood surface and thus makes optimisation difficult. This type of solution seems too extreme since the failure to find a root for one transaction would lead to a viable combination of parameter values being disregarded due to a potential ‘outlier’ observation.

Figure 7.5 plots the negative of the log-likelihood for the Black and Cox model for different levels of σ and given levels of μ , α and γ . The discontinuity is caused by the non-existence of a root for one (transaction two) of the 34 transactions and can make optimisation difficult.

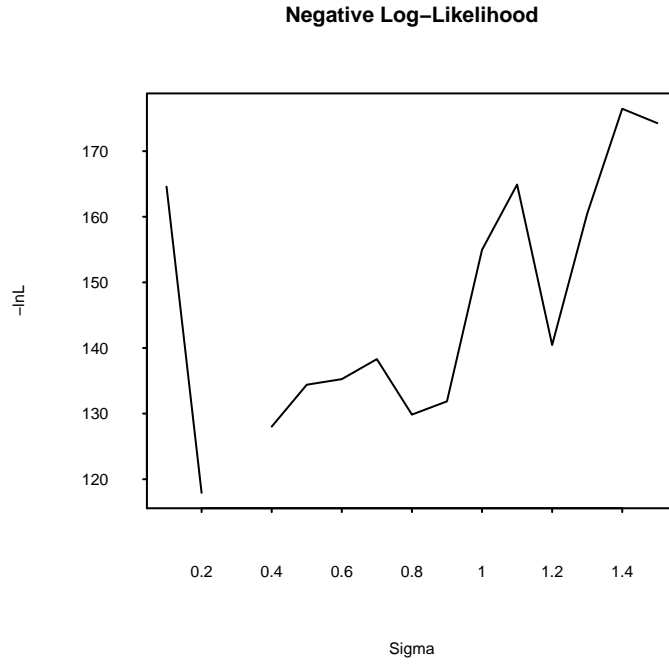


Figure 7.5: Negative log-likelihood for Black and Cox - constant rate - PZ for different σ and $\mu = 0.4$, $\alpha = 0.9$ and $\gamma = 0.09$. The discontinuity at $\sigma = 0.3$ is caused by the non-existence of a root for one transaction.

Figure 7.6 plots the lowest price differential of the second transaction for different σ . The increase in σ from 0.2 to 0.3 causes an increase in the Monte Carlo price and hence pushes the lowest price differential above 0.



Figure 7.6: Lowest possible price differential of transaction two for Black and Cox - constant rate - PZ for differing σ . At $\sigma = 0.3$, no root exists for the transaction.

For both the Black and Cox, and Briys and de Varenne models, whenever we encounter a transaction where a root does not exist for a specific combination of parameter values, we could set the the root to a level just above the highest of the initial barrier levels. This would be done in order to keep the log-likelihood surface continuous. This solution can, however, be problematic when the optimiser finds an optimal set of parameter values where many transactions have no roots. These points are high barrier points where roots do not exist. Figure 7.7 plots this type of situation for the transaction although it occurs for many transactions.

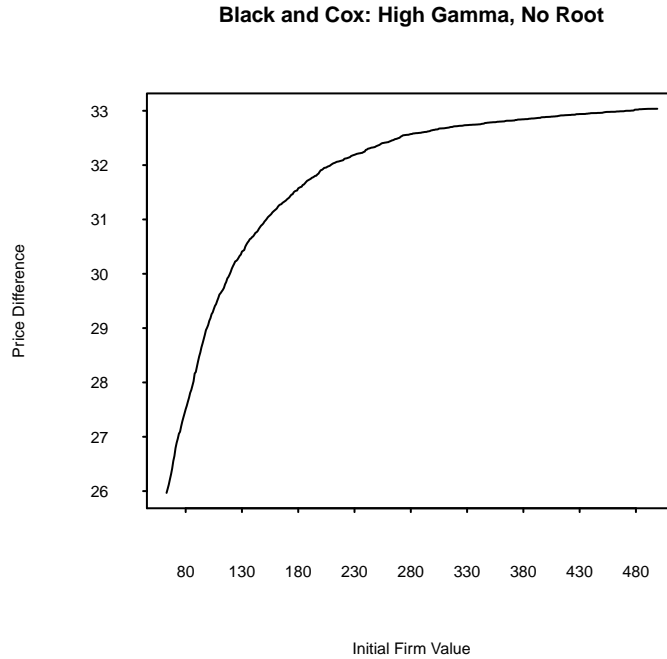


Figure 7.7: Price differential for differing initial firm values for transaction two under Black and Cox - constant rate - PZ. The initial firm value is bounded below by the highest initial barrier and no root exists since the payment barriers are too high. γ ($=0.05$) is higher than r ($=0.0386$) and hence the price function is increasing in initial firm value.

Since setting the log-likelihood function to $-\infty$ is too extreme and setting the roots of many transactions to some level above the initial barrier is also too extreme, we set a limit on how many transactions we are prepared to manually set the root. For a given combination of parameter levels, if 90% of transactions have roots that exist, then the remaining 10% of transactions' roots are set to be just above the initial barrier level. If instead more than 10% of the transactions have a non-existent root, then the log-likelihood is set to $-\infty$. The result is that extreme parameter values are avoided whilst the log-likelihood function is kept continuous. Figure 7.8 plots

the result of doing so.

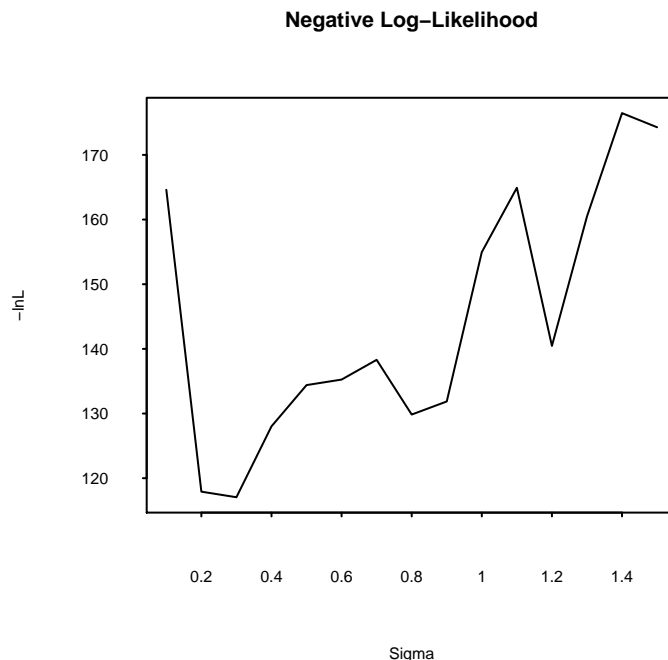


Figure 7.8: Negative log-likelihood for Black and Cox - constant rate - PZ for different σ and $\mu = 0.4$, $\alpha = 0.9$ and $\gamma = 0.09$. When roots are found for at least 90% of the transactions, the roots for the remaining transactions are set to the lowest possible firm values. This allows us to work with a continuous likelihood surface.

The downside of this solution is that it contributes to the problem of a jagged surface thus making optimisation difficult. It should be noted that the non-existence issue is more likely to occur for the PZ approach than the IC approach. This is again because of the consequences of default under the IC set up being more severe.

When the root of a transaction is set to just above the highest initial barrier level, we require a different treatment of the Jacobian approximation (5.9) since we have manually set the root to a value that does not actually explain the actual price. i.e.

$\bar{B}_{t_j}(V_{t_j}^r, T, F, c)$ does not equal the actual price $\bar{B}_{t_j}^{ob}(T)$ since $V_{t_j}^r$ is not an actual root for transaction j . Therefore we need to calculate what the model price is at the manually set root. The Jacobian for these transactions is thus computed as,

$$\frac{d\bar{B}_{t_j}(V_{t_j}^r, T, F)}{dV_{t_j}^r} \approx \frac{\bar{B}_{t_j}(V_{t_j}^{r'}, T, F) - \bar{B}_{t_j}(V_{t_j}^r, T, F)}{\epsilon_j} \quad (7.1)$$

7.6 Two Roots

We also encounter situations where two roots exist for a given transaction. This issue is unique to the Black and Cox model and can occur when γ is less than r . Here, the barrier level for the m^{th} remaining payment $c(m)$ of transaction j before the payment's maturity is $\alpha c(m)$ discounted back at a rate lower than r . In this situation, the bondholder prefers early barrier default to late barrier default since the present value of the payoff is higher the earlier default occurs. When the initial firm value is low, barrier default is highly likely and hence the Monte Carlo price is high. As the initial firm value is increased, barrier defaults begin to occur later and the value of the bond drops. However, as the initial firm value is increased even higher, the bond becomes risk free and the value increases. We are not saying that the bondholder prefers default over survival, since, we must account for α scaling the barriers down. The low level of γ simply leads to the preference of early default over later default.

We have two potential problematic scenarios that depend on the level of α . The first is when α is sufficiently high so that the Monte Carlo price always exceeds the actual price. In this situation no root exists. Figure 7.9 plots this situation for transaction two.

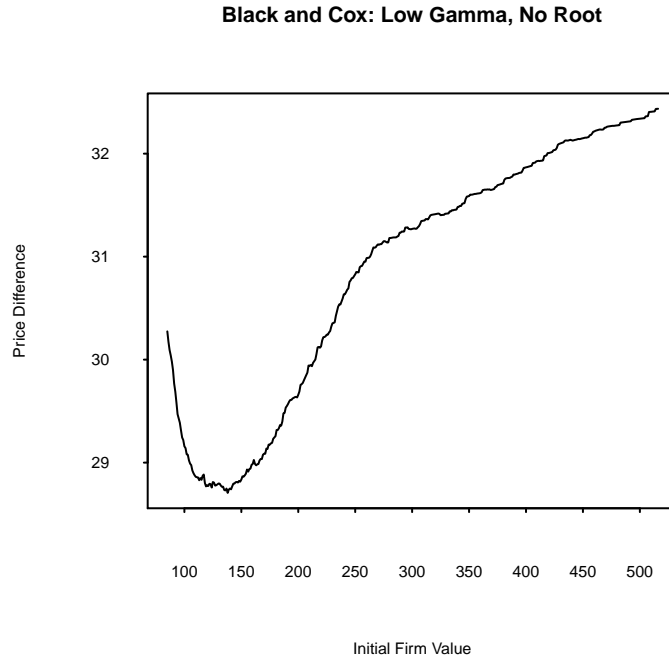


Figure 7.9: Price differential for differing initial firm values for transaction two under Black and Cox - constant rate - PZ. Here no root exists since the barriers are too high. Here, $\alpha=1$ and the quadratic type shape arises from the fact that γ ($=0.02$) is less than r ($=0.0386$).

The second scenario is when α is sufficiently low such that the price differential is negative for intermediate initial firm values levels but the barrier is still high such that the price differential is positive for low and high initial firm values. This is the situation in which two roots exist. The first root is when the price is decreasing in the initial firm value (later defaults which lower price) and the second is where the price is increasing in initial firm value (less risk). Figure 7.10 plots an example of this situation for transaction 33.

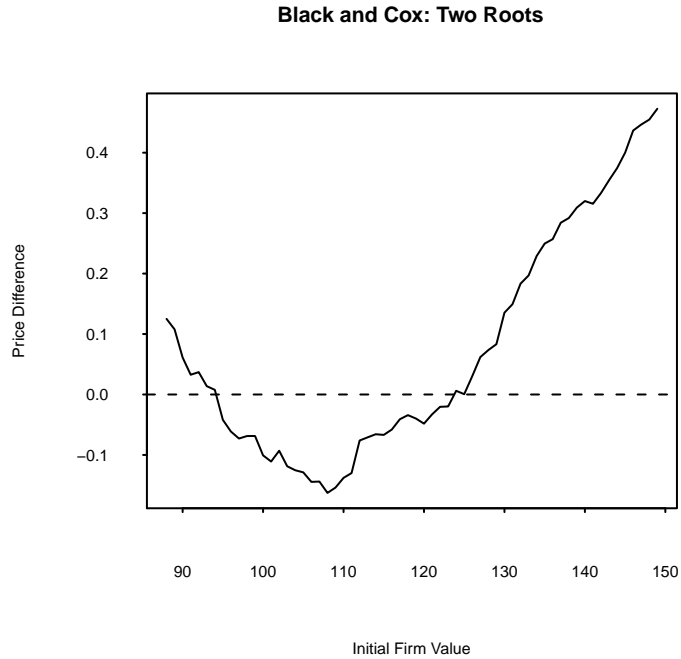


Figure 7.10: Price differential for differing initial firm values for transaction 33 under Black and Cox - constant rate - PZ. Here two roots exist since γ ($=0.025$) is less than r ($=0.0386$) and α ($=0.93$) is sufficiently low.

Our solution to this issue is to restrict attention to the higher root. This is to ensure we are dealing with the part of the pricing function where the bond's price is increasing in firm value. This ensures we focus on a more realistic region where survival is strictly preferred.

We can avoid the issue of early default preference by restricting γ . In the constant interest rate case, this is simple. We can impose the condition that $\gamma > r$. However, under a stochastic interest rate case, it is not clear how to restrict the values that γ can take since interest rates are not fixed. Therefore to stay consistent with the stochastic interest rate cases, we do not restrict γ to be always greater than r in the

constant interest rate cases of Black and Cox.

7.7 CIR Non-Existence of Roots

We also encounter a problem of non-existent roots in the CIR intensity model. A high level of σ_λ results in low survival probabilities and hence low corporate bond prices. Figure 7.11 plots an example of this type of situation.

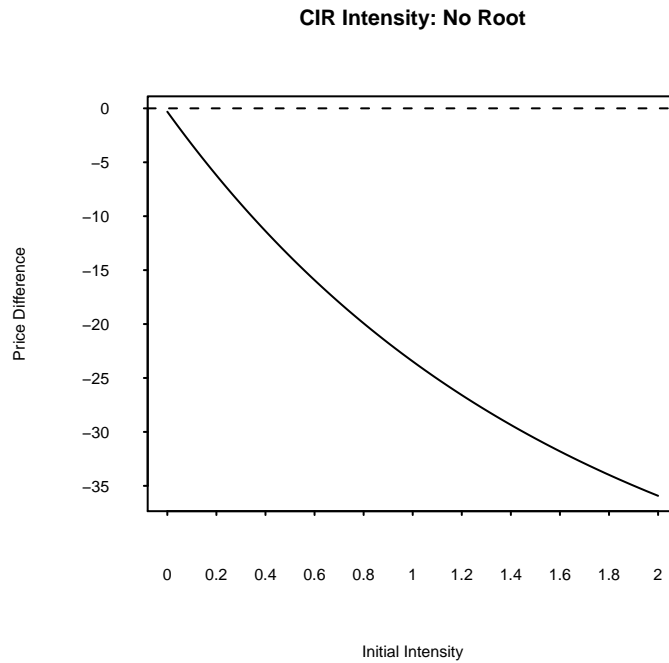


Figure 7.11: Price differential for different levels of initial intensity for transaction 34. No root exists since even at the lowest intensity level of 0, the differential is negative.

Even at the lowest possible level of $\lambda = 0$, a bond's price may be less than the

actual price. When this situation arises for a transaction, no root exists and our best solution is to manually set the root intensity for the transaction to 0.

However, as with the structural models, to prevent convergence to a point where many roots cannot be found, we restrict the number of transactions for which this occurs to 10%. Transactions where the root does not exist for specific combinations of parameters again require a different treatment of the Jacobian (5.40). Since we have manually set the root, the Jacobian is instead approximated as,

$$\frac{d\overline{B}_{t_j}(\lambda_{t_j}, T, F, c, \delta)}{d\lambda_{t_j}} \approx \frac{\overline{B}_{t_j}(\lambda_{t_j}^{r'}, T, F, c, \delta) - \overline{B}_{t_j}(\lambda_{t_j}^r, T, F, c, \delta)}{\epsilon_j}$$

since $\overline{B}_{t_j}(\lambda_{t_j}^r, T, F, c, \delta)$ does not equal $\overline{B}_{t_j}^{ob}(T)$.

Chapter 8

Putting the Methodology into Practice

In this chapter we present the results of our estimation work. We first consider the estimation results for the term structure models. From there, we proceed to the corporate bond model estimation results.

8.1 Models of the Term Structure

We firstly discuss the results of the estimation of the Vasicek and CIR term structure models. Table 8.1 presents the results of the estimation of the real world processes for both Vasicek and CIR. The most striking aspect is that the volatility parameter is much higher for the CIR model. However, this is because the CIR model has the $\sqrt{r_t}$ term included which, once accounted for, gives volatility levels consistent with

	a_r	μ_r	σ_r
Vasicek	0.064040772	0.046586808	0.007591073
CIR	0.04168732	0.05915043	0.03899186

Table 8.1: Vasicek and CIR under \mathbb{P}

	ν
Vasicek	-0.4631468
CIR	-0.07837819

Table 8.2: Market price of risk estimates for the Vasicek and CIR interest rate models.

the Vasicek model. Table 8.2 gives the estimates of the market price of risk ν under each short rate model. The market risk premium parameter is lower in absolute terms for the CIR model than that of the Vasicek model. However, it should be remembered that we assumed the market risk premium under the CIR model is divided by the square root of the short rate. Once adjusting for the market price of risk, we obtain the parameter estimates contained in Table 8.3 for the processes under the risk neutral probability \mathbb{Q} . Since our market price of risk estimates under both models are negative, the adjustment of the mean interest rate level results in a higher μ_r under the risk neutral world than that of the real world. The result is that our interest rate simulations will result in yield curves that are largely upward sloping. This is consistent with what is commonly observed in the real world.

Under the parameter estimates of Table 8.3, we calculate the par-yields for treasury bonds for maturities of 6 months up to 30 years and then compare to actual par-yields obtained from the St. Louis Federal Reserve. Table 8.4 displays the mean percentage errors whilst Table 8.5 displays the mean absolute percentage errors. The

	a_r	μ_r	σ_r
Vasicek	0.064040772	0.101485915	0.007591073
CIR	0.04168732	0.1324608	0.03899186

Table 8.3: Vasicek and CIR under \mathbb{Q}

	6 months	1 year	2 years	3 years	5 years	7 years	10 years	20 years	30 years
Vasicek	-2.01	-3.73	-6.80	-7.07	-6.17	-4.70	0.04	4.62	10.07
CIR	-2.30	-4.17	-7.36	-7.61	-6.52	-4.79	0.38	6.09	12.12

Table 8.4: Vasicek and CIR par-yield mean errors (%) by maturity.

	6 months	1 year	2 years	3 years	5 years	7 years	10 years	20 years	30 years
Vasicek	7.01	9.02	11.94	12.45	13.51	13.72	13.77	12.84	13.94
CIR	6.89	9.08	12.21	12.97	14.30	14.60	14.92	14.18	15.80

Table 8.5: Vasicek and CIR par-yield mean absolute errors (%) by maturity.

Vasicek and CIR models both underpredict the par-yields for short maturity treasury bonds whilst overpredicting the par-yields for longer maturity treasury bonds. That the errors are quite small for the 10-year bonds is simply a result of our choice to calculate the market risk premium by minimising the sum of squared percentage errors of the 10-year par-yields.

The maximum mean absolute percentage error is around 13.9% for Vasicek and 15.8% for CIR whilst the minimum absolute percentage error is around 7% for Vasicek and 6.9% for CIR. Therefore, both models are producing more accurate estimates for shorter maturities rather than longer maturities. As a result of calibrating the market price of risk to the 10-year par-yields, the models are underpredicting the yields on maturities shorter than 10 years whilst overpredicting the yields on maturities longer than 10 years.

We find that the models cannot explain large variations between yields of different maturities on a given day. Therefore we observe clustering in the percentage errors with estimated yields that are too high when the actual yield curve is sloping downward and estimated yields that are too low when the actual yield curve is sloping upward thus averaging out over a large set of dates. This is due to the estimated yield curves for effectively all our dates being upward sloping because of the large estimates of μ_r for both models. This high level coupled with the evidence of mean reversion in the data (positive estimates of a_r) results in largely upward sloping yield curves. This as, we will see in the next section, creates over-discounting problems in pricing corporate bonds.

These results suggest that the two models do not have enough flexibility to accurately measure the term structure. The clustering of errors suggests the short rate

alone is not enough to explain the variations in the yield curve and that we need to take stochastic interest rate volatility into account. What the volatility structure should be, however, is not so clear.

8.2 Structural Corporate Bond Pricing Models

8.2.1 Parameter Estimates

We now turn to the results of our estimation of the structural corporate bond pricing models. Table 8.6 presents the parameter estimates for the structural models and Table 8.7 presents the standard deviation estimates for these parameters using the method of Greene (2000). We can be reasonably confident in the standard deviation estimates for μ in all of the models since it plays no part in simulation and hence there is no jagged estimated log-likelihood surface problem, as discussed in Chapter 7, in the μ direction. However for the other parameters, the standard deviation estimates need to be interpreted with caution. The jaggedness of the estimated log-likelihood surface leads to very inaccurate approximations of (5.43). Therefore, it seems necessary to increase the number of simulation paths to increase the reliability of this ML approach.

We encounter difficulties in estimating the standard deviation for f_1 in the Briys and de Varenne model. This is because the α estimates in these cases are quite small. A small α represents a very low barrier and hence a barrier that is not hit by the firm value very often (or at all). Therefore the fraction of firm value received upon barrier default is insignificant. This leads to the approximated derivative (5.44)

being very small (or even zero) for all transactions and hence a large or infinite estimated standard deviation. In some cases it is necessary to use a larger offset for f_1 in (5.44) to obtain significant changes in log-likelihoods. Similar problems also arise in estimation of the standard deviation of γ in the Black and Cox model.

We find that μ does not shift by a large amount from the initial value used in the optimiser even when running the optimiser multiple times with different levels of minimum parameter step sizes. Most of the estimates of μ are close to the initial value of 0.3. This suggests that the estimated log-likelihood is quite invariant to changes in μ . The large, overall, standard deviation estimates for μ are evidence of this.

The estimates of σ are, overall, higher for the Merton model than that of the Black and Cox, and Briys and de Varenne models. This is because the absence of a default barrier makes default less likely in the Merton model and hence leads to high model prices. Since μ plays no part in simulation, σ must do all of the work to explain actual prices. In order to drive a model price down, the level of σ is increased so that the probability of default increases. The default of more payments over simulation paths (or a lower firm value payoff when default occurs at maturity) lowers the monte carlo model price. This may explain why the Merton model has underpredicted spreads (overpriced) in the existing empirical literature. The estimated firm value volatilities using the traditional method in these papers may be too low.

We find that the estimated σ in the Merton model is lower in the constant rate cases than the stochastic interest rate cases. The mean short rate parameter μ_r is quite high for both the Vasicek and CIR models (Table 8.3) after the market price of risk adjustment, particularly for the CIR model, which then leads to largely

Model	Interest Rate Assumption	Coupon Bond Treatment	μ	σ	α	γ	f_1	f_2
Merton	CR	PZ	0.2001	0.5884	-	-	-	-
		IC	0.2500	0.7001	-	-	-	-
	V	PZ	0.0995	1.4903	-	-	-	-
		IC	0.0954	1.4953	-	-	-	-
	CIR	PZ	-0.2241	2.4434	-	-	-	-
		IC	0.0981	2.3955	-	-	-	-
Black and Cox	CR	PZ	0.3000	0.7092	0.2345	0.6568	-	-
		IC	0.2994	0.7362	0.0252	1.0913	-	-
	V	PZ	0.2992	0.4036	0.2545	0.7096	-	-
		IC	0.2953	0.6969	0.0000	1.1178	-	-
	CIR	PZ	0.1849	1.0221	0.0000	1.5626	-	-
		IC	-0.0050	1.2915	0.0288	1.0159	-	-
Briys and de Varenne	CR	PZ	0.2980	0.3574	0.0026	-	0.3365	0.9115
		IC	0.2991	0.4518	0.3478	-	0.4258	0.4219
	V	PZ	0.0940	0.8299	0.0061	-	0.7432	0.9243
		IC	0.2953	0.6708	0.0099	-	0.5542	0.8099
	CIR	PZ	0.0025	0.1280	0.9923	-	0.0026	0.8180
		IC	0.2999	0.3860	0.4687	-	0.3180	1.0000

Table 8.6: Structural model parameter estimates.

Model	Interest Rate Assumption	Coupon Bond Treatment	μ	σ	α	γ	f_1	f_2
Merton	CR	PZ	0.2037	0.0478	-	-	-	-
		IC	0.2257	0.0313	-	-	-	-
	V	PZ	0.3125	0.1031	-	-	-	-
		IC	0.3168	0.1141	-	-	-	-
	CIR	PZ	0.4690	0.3309	-	-	-	-
		IC	0.4650	0.2966	-	-	-	-
Black and Cox	CR	PZ	0.2199	0.0002	0.0001	3.67e-05	-	-
		IC	0.2360	0.1442	0.0008	74.8369	-	-
	V	PZ	0.0854	0.0681	0.0445	0.0001	-	-
		IC	0.1404	0.0004	0.0006	0.4142	-	-
	CIR	PZ	0.2073	0.0692	2.5268	1473.0	-	-
		IC	0.2461	0.0655	0.0851	40.770	-	-
Briys and de Varenne	CR	PZ	0.0967	1.89e-05	0.0051	-	-	0.0260
		IC	0.1405	0.0037	0.0004	-	0.0008	0.0006
	V	PZ	0.1737	0.0019	0.0348	-	148.32	0.0024
		IC	0.1346	1.23e-05	0.0018	-	-	0.0004
	CIR	PZ	0.0007	8.60e-06	0.0176	-	1.32e-07	0.0126
		IC	0.0776	0.0012	0.0013	-	1.63e-06	0.0031

Table 8.7: Structural model parameter standard deviation estimates.

upward sloping interest rate simulation paths. As a result, long-term payments are discounted at high rates and the transaction date values of these payments are quite low. Therefore, for a given transaction, we may encounter the situation where a simulation path has simulated rates that are sufficiently high such that the price (5.19) of the bond along this individual path is less than the actual value of the bond for any level of initial firm value, i.e. a path that cannot explain the actual bond price. If this occurs for many simulation paths for a given transaction, the model price is heavily affected by these paths and the initial firm value need to be very high to explain the actual price. For example, the root for the first transaction under the Merton model with a constant rate and the PZ approach is 193.69 whereas for the CIR case it is $1.194737e+17$! Along a path where the interest rate simulation path is quite low in comparison to the high interest rate paths, a high initial firm value leads to a price for that path that exceeds the actual price of the bond. Therefore a large σ is required to drive the firm value down and cause more defaults.¹

The parameter estimates for the Merton model are very similar across the PZ and IC set ups. This is because there are no default barriers and the coupon size is low. Therefore there is a very small probability that default will occur at a coupon date and the only payment with a significant probability of default is the maturity payment. Of course, since the maturity payment is the last payment of the bond, the consequence of default of this payment is exactly the same for the PZ and IC set up.

In the Black and Cox, and Briys and de Varenne models, the σ estimates are on average, lower than that of the Merton model. With the Black and Cox model, we obtain high estimates of γ . This causes strong downward drift in the simulation

¹Since the individual coupons are small, the large volatility may not lead to many defaults of coupons, rather, a small payoff for default at maturity.

of ‘adjusted’ firm values and hence more defaults even with a very low level of α . When α is zero, γ is essentially doing the work of σ in creating strong downward drift since compounding the barrier back up in this case still results in a payoff of zero. Our estimates of σ are still quite high for the Black and Cox model, but not as high as the Merton case due to the presence of default barriers.

Our estimation results suggest significant deviations from absolute priority when default occurs along a barrier under the Briys and de Varenne model. However, as discussed, we find the effect of the barriers is small due to a low α (except for one case) and hence f_1 is insignificant. The results also suggest deviations from absolute priority when default occurs at the maturity of a payment. The result is, given an initial firm value, a lower price of the bond along a simulation path when default occurs on a payment date in comparison to the Merton case. Therefore, σ does not need to be as high as in the Merton model to explain actual prices and hence our estimates of σ are lower for the Briys and de Varenne model.

We find estimates of α to be quite low across both the Black and Cox, and Briys and de Varenne model except for the Briys and de Varenne case of CIR interest rate and PZ approach. In this case, we find a α of around 1 and a lower level of σ . Therefore the high barrier accounts for more defaults rather than an increased diffusion due to σ . Here, barrier default is likely and bondholders are estimated to receive approximately 82% of the firm value when default occurs. This ‘outlier’ case in terms of parameter values suggests that individual parameter estimates are not overly significant but rather the combination of parameter estimates.

When analysing the results for one bond only, it is difficult to determine the effects of single parameters and rather, we are required to interpret the combinations of

parameter estimates that eventuate. In order to discover the true effects of each parameter, the estimation would have to be repeated for multiple bonds. With different bonds we would be able to compare parameter estimates for bonds differing by factors such as rating, time-to-maturity and coupon size. Here, we would be able to see how each parameter adjusts in each case and gain a better understanding of what features each parameter captures and how these models perform overall.

8.2.2 Predictive Performance

In terms of modeling our bond, the structural models exhibit superior performance under the ML approach than that of the structural models under the traditional approach in the existing empirical literature. Table 8.8 and 8.9 presents our summary results for the structural models. We find that the mean spread error for the constant rate version of the Merton model under the PZ approach is 1.06% in contrast to the -50.42% error in Eom, Helwege and Huang (2004). Although these comparisons are for different bonds, the difference is quite large suggesting that we have removed the bias in pricing that was apparent under the traditional approach.

While we do obtain results consistent with the consensus that the Merton model underpredicts spreads, we find the mean spread error in our results to be quite small. However, we can see that the mean of the absolute spread prediction errors is quite large for all of the models. There is a large variation in the predictions of spreads. Since we find very small pricing errors, our results are consistent with the view of Eom, Helwege and Huang (2004) who state that price errors are not an adequate measure of model performance. Our results are an improvement on Eom, Helwege

Model	Interest Rate Assumption	Coupon Bond Treatment	Spread		
			Mean	SD	Mean Absolute
Merton	CR	PZ	1.062	43.796	31.951
		IC	-1.456	42.384	31.545
	V	PZ	-1.038	47.151	36.242
		IC	-1.051	46.942	36.191
	CIR	PZ	-1.443	46.107	35.963
		IC	-1.453	47.459	36.870
Black and Cox	CR	PZ	-2.400	42.809	32.038
		IC	-3.053	41.859	31.485
	V	PZ	-8.675	47.114	35.367
		IC	-8.801	46.980	35.552
	CIR	PZ	-8.798	47.840	36.426
		IC	-9.004	47.353	36.077
Briys and de Varenne	CR	PZ	-3.424	41.249	31.177
		IC	-6.195	41.028	31.340
	V	PZ	-6.022	47.402	35.228
		IC	-8.714	47.092	35.753
	CIR	PZ	10.351	72.653	43.751
		IC	-11.656	46.998	36.257

Table 8.8: Structural model spread prediction errors (%). Errors are obtained by subtracting actual from predicted and then dividing by actual.

Model	Interest Rate Assumption	Coupon Bond Treatment	Price		
			Mean	SD	Mean Absolute
Merton	CR	PZ	0.224	3.199	2.458
		IC	0.411	3.154	2.458
	V	PZ	0.848	3.353	2.696
		IC	0.853	3.351	2.697
	CIR	PZ	1.130	3.286	2.671
		IC	1.110	3.363	2.725
Black and Cox	CR	PZ	0.487	3.199	2.515
		IC	0.531	3.161	2.479
	V	PZ	0.776	3.350	2.638
		IC	0.770	3.357	2.661
	CIR	PZ	0.748	3.395	2.709
		IC	0.750	3.374	2.690
Briys and de Varenne	CR	PZ	0.559	3.116	2.451
		IC	0.758	3.168	2.524
	V	PZ	0.587	3.377	2.650
		IC	0.765	3.357	2.670
	CIR	PZ	-0.279	3.991	2.891
		IC	0.946	3.353	2.682

Table 8.9: Structural model price prediction errors (%). Errors are obtained by subtracting actual from predicted and then dividing by actual.

and Huang (2004) who observed a standard deviation in spread prediction errors of 71.84% for the Merton model, however, our standard deviations are still quite large. Although we have been successful at removing a large bias in structural model pricing, we have not been able to remove the large variation in predictions provided by these models.

Surprisingly, it is the simplest case that performs the best in modeling the corporate bond. Although verification would be required in terms of testing the approach on multiple bonds, it appears that nothing is gained from considering default barriers and deviations from absolute priority. In fact, these features have reduced accuracy. However, the presence of default barriers may exacerbate the issue of jagged log-likelihood surfaces. The inclusion of default barriers results in extra parameters that must be estimated. Since the estimated log-likelihood function using the Monte Carlo approach is not smooth, estimation of many parameters is difficult. Hence, it is possible that the parameter estimates under Black and Cox, and Briys and de Varenne are less precise than that of the Merton model.

We also find that there is little difference between the performance of the PZ approach and the IC approach. Eom, Helwege and Huang (2004) find that the Geske (1977) model is a significant improvement on the Merton model. We, however, find very little difference between the PZ and IC cases for the Merton model with a constant interest rate.

For example, Table 8.10 gives the estimated risk-neutral default probabilities of the first transaction for the Merton model under a constant interest rate and both the PZ and IC approaches. There is very little difference in the default probabilities between the two approaches and hence the PZ approach does not seem to be much

of a compromise.² As another example, we also reach the same conclusion when comparing the estimated probabilities for the Black and Cox model (Table 8.11) under the Vasicek interest rate assumption. Again, there is no overly large difference in the probabilities. We find that the default barriers in most cases are quite small, thus, early default is unlikely. Since the sizes of the individual coupons are not large, it appears that the probabilities of default of the coupons are negligible. We find this is the case across all of the models.

We also find that the constant interest rate case, overall, performs better than the Vasicek and CIR short rate models. The spread errors seem to be closely related to the assumed interest rate process rather than the bond pricing model. This is because of the large variation in roots under the stochastic interest rate cases caused by the high discounting of long-term payments. With highly variable roots, forecasting firm values accurately is more difficult and large spread errors are more likely. We are not concluding in favour of a constant interest rate over a stochastic interest rate, merely over the Vasicek and CIR models. It is apparent that these models have difficulties in accurately measuring the riskless term structure and this leads to problems in corporate bond valuation.

We observe large variations in the spread prediction errors. Therefore, we now attempt to explain why these large variations arise. Figures 8.1, 8.2 and 8.3 give box plots of the spread errors for Merton, Black and Cox, and Briys and de Varenne respectively.

²Although it can be argued that the two sets of probabilities are not directly comparable since they have different parameter estimates and different roots, we view these as comparable since they are the probabilities at the optimum in each case

Coupon Bond Treatment	1	2	3	4	5	6	7	8	9	10
PZ	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.67	1.00
IC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.83	1.33
Coupon Bond Treatment	11	12	13	14	15	16	17	18	19	20
PZ	1.17	1.00	1.67	2.67	3.17	3.00	4.50	4.83	7.33	67.33
IC	2.00	2.83	3.83	4.50	5.67	7.00	8.83	10.17	12.00	62.50

Table 8.10: Individual payment estimated default probabilities (%) for the 20 remaining payments of the first observation with Merton under a constant interest rate and both the PZ and IC set up.

Coupon Bond Treatment	1	2	3	4	5	6	7	8	9	10
PZ	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.17
Coupon Bond Treatment	11	12	13	14	15	16	17	18	19	20
PZ	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	41.33
IC	0.33	0.33	0.50	0.50	0.50	0.50	1.33	1.67	2.67	34.00

Table 8.11: Individual payment estimated default probabilities (%) for the 20 remaining payments of the first observation with Black and Cox under a Vasicek interest rate assumption and both the PZ and IC set up.

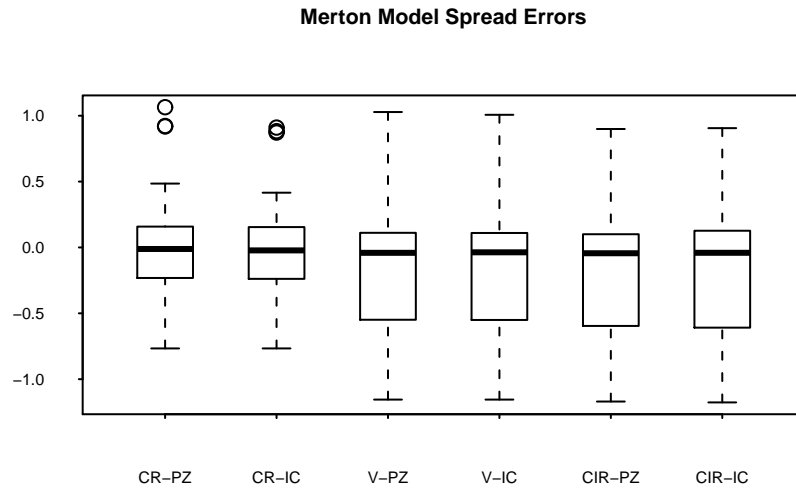


Figure 8.1: Box plots of the spread errors for the Merton cases. Circles represent outlier transactions which have large spread prediction errors. The constant, Vasicek and CIR interest rates are represented with ‘CR’, ‘V’ and ‘CIR’ respectively. ‘PZ’ refers to ‘portfolio of zeroes’ and ‘IC’ refers to internal consistency.

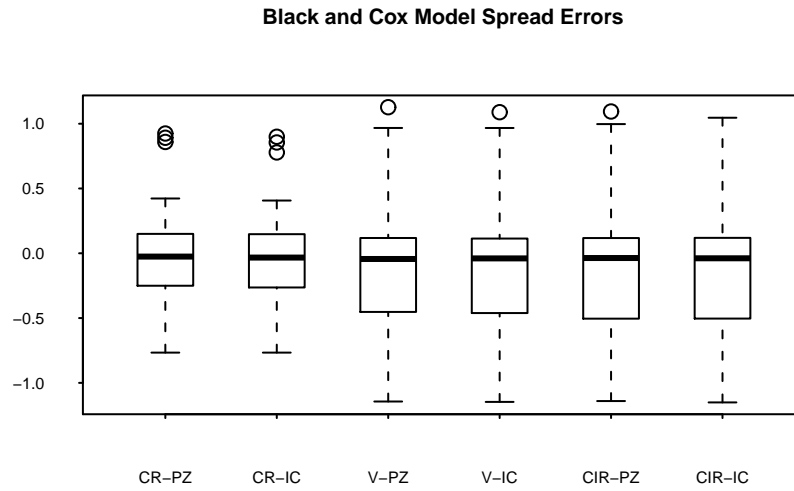


Figure 8.2: Box plots of the spread errors for the Black and Cox cases. Circles represent outlier transactions which have large spread prediction errors. The constant, Vasicek and CIR interest rates are represented with ‘CR’, ‘V’ and ‘CIR’ respectively. ‘PZ’ refers to ‘portfolio of zeroes’ and ‘IC’ refers to internal consistency.

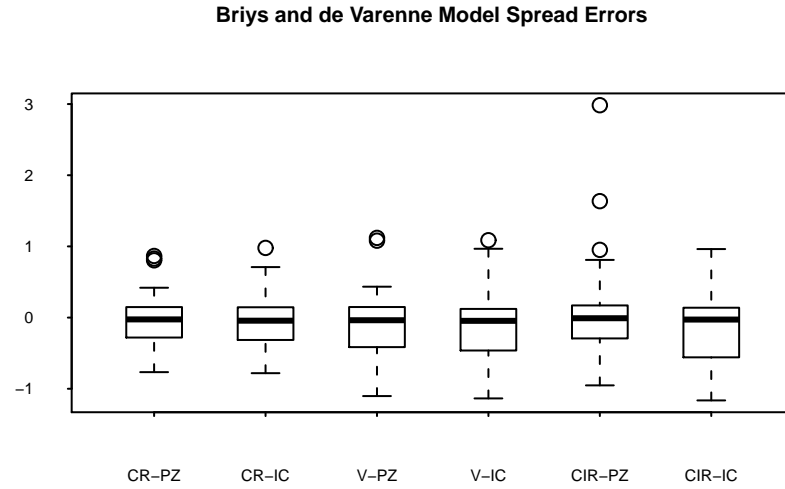


Figure 8.3: Box plots of the spread errors for the Briys and de Varenne cases. Circles represent outlier transactions which have large spread prediction errors. The constant, Vasicek and CIR interest rates are represented with ‘CR’, ‘V’ and ‘CIR’ respectively. ‘PZ’ refers to ‘portfolio of zeroes’ and ‘IC’ refers to internal consistency.

For each model, we have 32 spread predictions of the 34 transactions. This is because we cannot forecast a price for the first transaction since there is no root to forecast from. We also encounter a transaction where there is a non-existent root problem as discussed in section 7.4. However, for all of the structural models, convergence is achieved at an optimum where there are no parameter-specific root-non-existence problems. It is clear from the box plots that the constant interest rate cases outperform the stochastic interest rate cases. This is due to the large nature of the roots in the stochastic interest rate cases discussed in 8.2.1. There is more variability in the roots for the stochastic interest rate cases and hence, more difficulty in forecasting. Performance in a stochastic interest rate case may be improved using a no-arbitrage

model since there would be no issues of over or under-discounting payments on a given day.

An obvious feature of these plots is that there seem to exist transactions that none of the models can explain. These large errors occur when there is sharp short term movement in actual prices. If transaction j has a low price, then it has a low root initial firm value which is used to estimate the initial firm value for transaction $j+1$. If transaction $j+1$ has a high price, then its root is also high. Thus the forecasted firm value for transaction $j+1$ is too low, the subsequent forecasted price is too low and the spread overpredicted.

We also expect large pricing errors to occur when the time between transactions is quite large. Here, a root will have drift applied to it for a longer amount of time. i.e. a large Δt_j is used in (5.45). In this situation there is a higher chance that the predicted firm value will ‘overshoot’ to a high level and hence create a large spread error.

There is also an issue in explaining drops in initial firm value across transactions. Unless μ is negative, the geometric Brownian motion cannot explain drops in firm value over time in forecasting since the expectation is always an increase. An example of this is the forecasted spread of the last transaction. This transaction occurs over 2 years after the penultimate transaction and is close to maturity with only 3 payments remaining and hence a much lower price than the rest of the transactions we observe. The penultimate transaction has a high price, a high root and hence a high predicted firm value for the final transaction. This leads to the final transaction being overpriced. The spread error is only prevented from being very large by the fact that the model price is bounded by the price of a risk-free equivalent. Figure

8.4 plots the actual price against the predicted price for the Merton model under a constant interest rate and PZ. The prediction for the last transaction is only pulled down because it is bounded by the risk-free equivalent price. Here, the root is so high that the model price is equal to the risk free equivalent price.

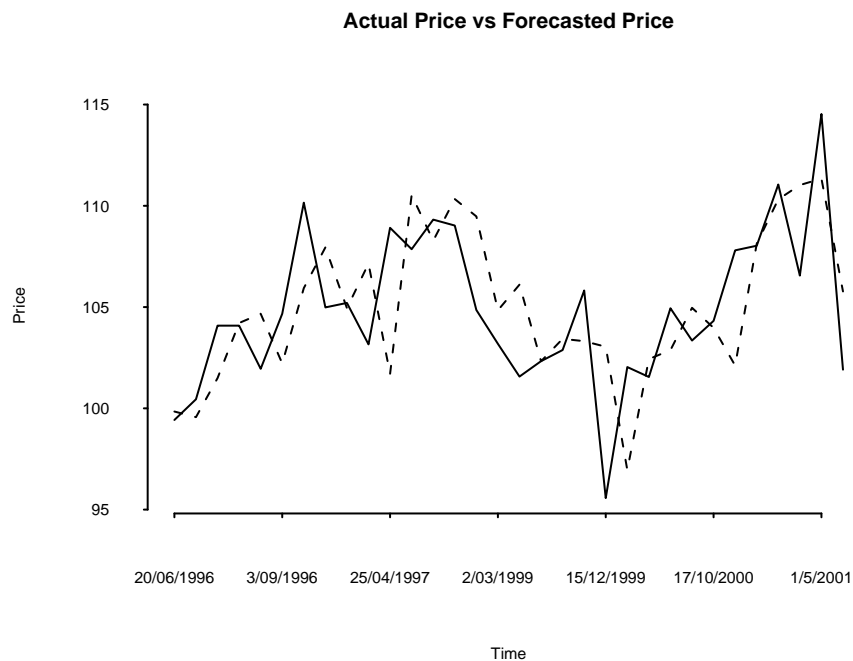


Figure 8.4: Actual (solid) vs predicted (dashed) prices for the Merton model under a constant interest rate and PZ. The price predictions closely follow the actual prices.

Overall, 8.4 indicates that the model has performed very well in modeling the actual prices. Although it is true that the pricing errors are quite small, investigation of the plot of predicted spreads against actual spreads in figure 8.5 for the same case of the Merton model shows that the predictions contain large errors. However,

on average, the predictions are somewhat accurate as compared to attempts in the existing literature. The spread error of the last transaction is quite large, however, due to the large time difference between the last and second-to-last transaction and subsequent large root difference. Nonetheless, our results suggest that the structural bond pricing models do have the ability to be reliable in corporate bond pricing. It follows from our results that the approach is likely to perform better in modeling the price of a bond issued by a low-risk firm since there would likely be less variation in the prices. This lower level of variation would lead to less variable root estimates and more accurate forecasts. In addition, the approach is likely to exhibit superior performance for a liquid bond since there would be less time between observations and less ‘overshooting’ problems in forecasting. These points also suggest that forecasting future prices is likely to be more reliable if not forecasting too far into the future.

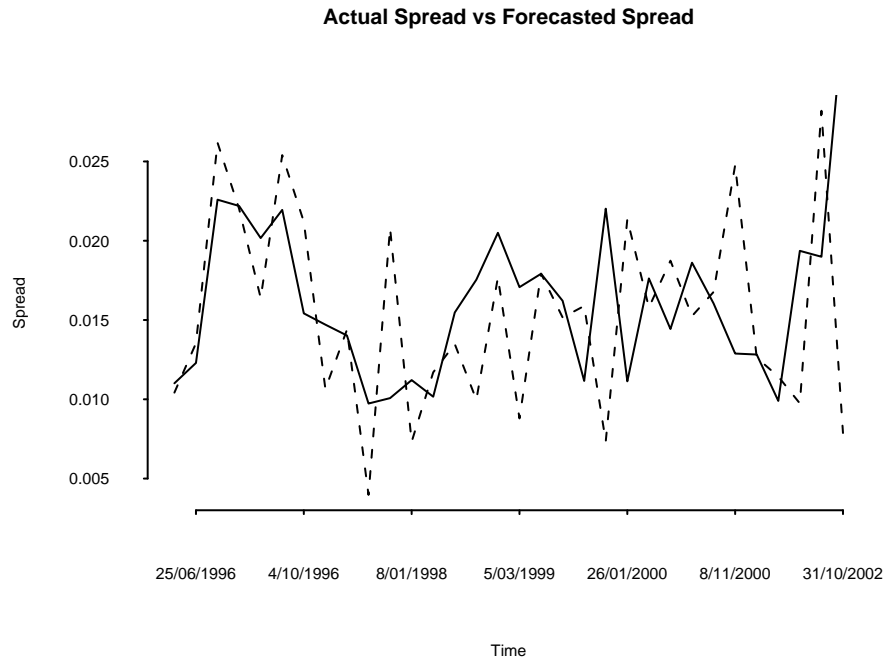


Figure 8.5: Actual (solid) vs predicted (dashed) spreads for the Merton model under a constant interest rate and PZ.

Performing a grid search to find appropriate starting values for the parameters may aid in obtaining better results. We find that since μ is not used in simulation, its estimate does not shift much from the initial value. Yet it is an important parameter in the sense that it is used in predicted one-step ahead forecasts of firm values.

Model	Interest Rate	a_λ	μ_λ^P	σ_λ	ν
CIR Intensity	CR	0.3846	0.0002	1.220	-1.32812
		(0.1326)	(5.51e-05)	(0.0405)	(1.66e-05)
	V	1.3056	1.1779	1.1111	-0.1222
		(0.5736)	(0.5512)	(0.3952)	(1.3540)
	CIR	1.3266	1.1841	1.1147	-0.1086
		(0.5783)	(0.5621)	(0.4009)	(1.3891)

Table 8.12: CIR intensity model parameter estimates. Numbers in parentheses are the standard deviation estimates of the parameter estimates.

8.3 CIR Intensity Model

8.3.1 Parameter Estimates

To complete our analysis of the estimation and forecasting results, we now turn our attention to the CIR intensity model also estimated in this paper. Table 8.12 presents the parameter estimates for this model.

The most striking feature of these parameter estimates is the fact that the estimated mean level of the default intensity under \mathbb{P} is much lower for the constant interest rate case. This suggests lower default intensities and lower probabilities of default in the constant interest rate case. However, upon inspection of the estimated market price of default risk ν , we find the opposite is true. Since pricing occurs in the risk neutral world, the mean level of the default intensity that is actually relevant to pricing is the estimate μ_λ under \mathbb{Q} . This is calculated by rearranging (5.37) for μ_λ . This is much higher under the constant interest rate case and hence calculated survival probabilities will be lower than that of the Vasicek and CIR interest rate cases. This makes sense since our estimated mean interest rate level μ_r in both the Vasicek and CIR interest rate models is quite high. The result is lower calculated risk-free bond prices meaning default probabilities do not need to be as low to explain actual

prices.

We can place more confidence in the standard deviation estimates of the CIR intensity model parameter estimates than those of the structural models since convergence is easy to obtain in the CIR intensity model. We find no issues associated with a jagged estimated log-likelihood surface as we did with the structural models. This is because we are using the closed-form solution for bond prices rather than Monte Carlo simulation. Figure 8.6 plots the estimated negative log-likelihood surface for the CIR intensity model under a Vasicek interest rate assumption for fixed a_λ and ν and different μ_λ and σ_λ .

CIR Intensity: $-\ln L$ around Convergence Point

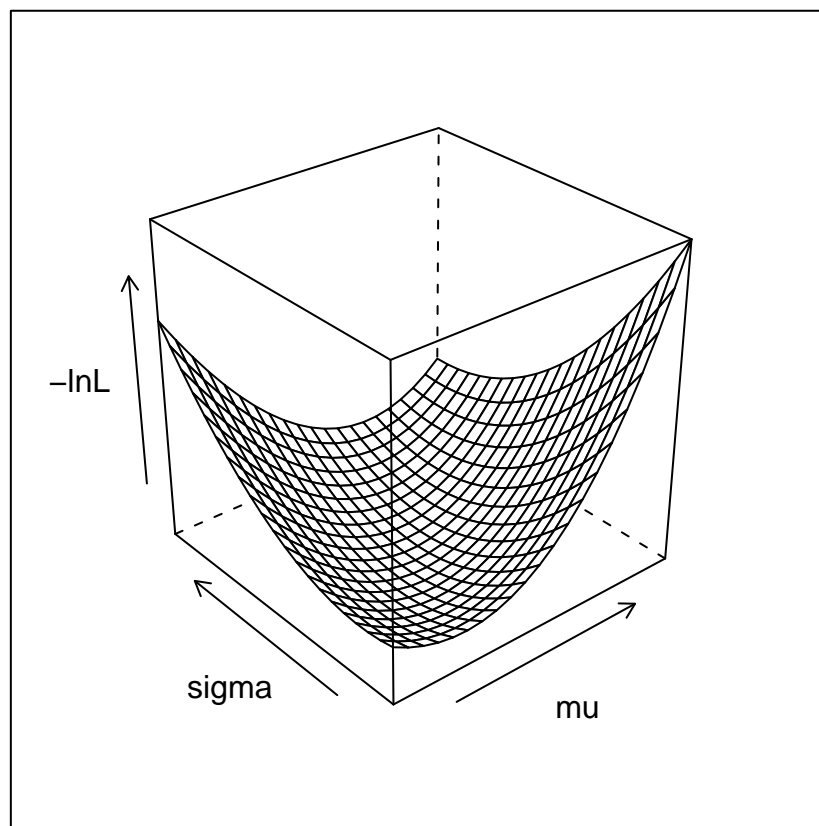


Figure 8.6: Negative log-likelihood surface with the CIR intensity model under a Vasicek interest rate for $\mu_\lambda \pm 0.3$ and $\sigma_\lambda \pm 0.3$

8.3.2 Predictive Performance

The surface is much smoother than that of the structural models and thus makes convergence easier to obtain. However, ease of application does not imply better results. Table 8.13 plots our summary results for the CIR intensity model. We obtain small mean spread errors consistent with the smaller mean spread errors under the structural models. However, the variation in the spread errors is much

larger than that of the structural models. The reduced-form model does not exhibit any bias in spread prediction but has large spread errors in both directions.

The standard deviations of the spread and pricing errors are very large for all three cases and much larger than that of the structural models. Figure 8.7 plots the predicted spread against the actual spread for the CIR intensity model with a Vasicek interest rate assumption. The forecasts of the default intensity appear to ‘overshoot’, causing large spread errors. A root does not exist for the last transaction in any of the CIR intensity cases since the model price is less than the actual price even at a root intensity of zero. These results suggest that the CIR model is a poor choice of process for the default intensity.

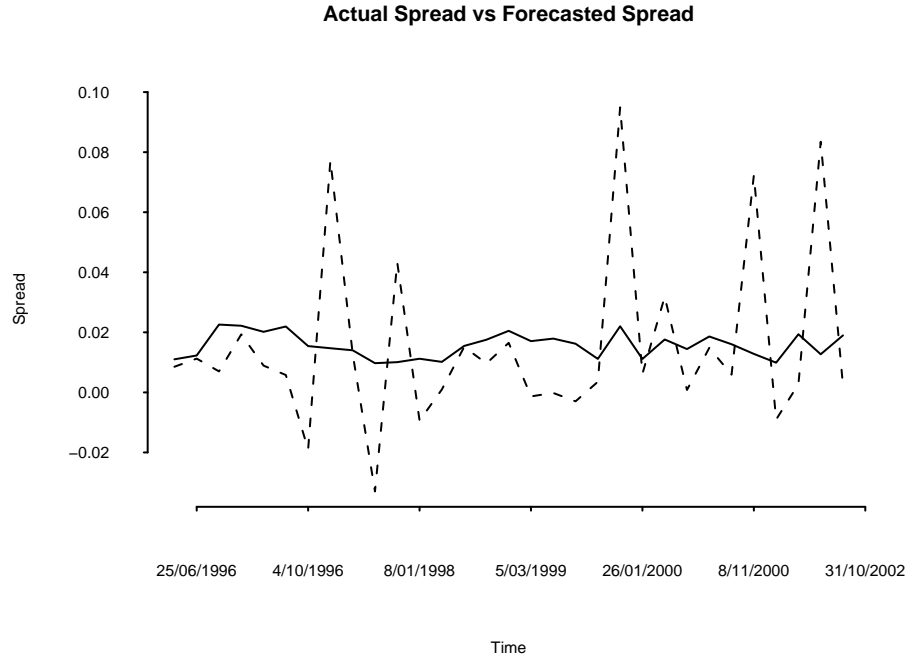


Figure 8.7: Actual (solid) spread vs predicted (dashed) spread for the CIR intensity model with a Vasicek interest rate. The spreads exhibit a large level of variability.

Model	Interest Rate	Spread			Price		
		Mean	SD	Mean Absolute	Mean	SD	Mean Absolute
CIR Intensity	CR	-3.371	229.1	149.6	1.925	157.8	111.6
	V	-1.599	211.3	142.7	0.734	160.8	105.9
	CIR	-1.718	210.7	142.4	0.737	160.6	105.8

Table 8.13: CIR intensity model spread and price prediction errors (%). Errors are obtained by subtracting actual from predicted and then dividing by actual.

Nevertheless, these results give even more strength to the argument that the failure of structural models to accurately price risky bonds in the existing literature is due to the inadequacies of the traditional estimation approach of Ronn and Verma (1986). We find that the structural models, in fact, outperform the reduced-form model. However, we cannot conclude that structural models dominate reduced-form models since we have only estimated one reduced-form model. Making such a conclusion would require testing the approach on many more different processes for the default intensity.

The low mean spread errors present some hope for structural bond pricing models. Due to their poor empirical performance under the traditional estimation method, structural models have been largely written-off by practitioners who now favour the reduced-form models. We however, find superior performance under the structural models. Although there is large variation in our spread prediction errors for all of the structural models we implement, there is even bigger spread error variation under the CIR intensity model. Our results are consistent with the view of Ericsson and Reneby (2005) and Li and Wong (2008) in that the performance of structural models of corporate bond pricing are greatly improved under a ML approach.

Chapter 9

Conclusion

Empirical evidence suggests that structural corporate bond pricing models cannot adequately price risky bonds and that reduced-form models should instead be used. However, these failures seem to be more based on an inadequate estimation approach rather than theoretical failures of the models. The main problem in applying structural models is that the firm value V_t and its volatility σ are unobservable. All of the existing empirical papers of structural bond pricing (except for Li and Wong (2008)) use some form of the traditional estimation method of Ronn and Verma (1986) to estimate firm value and its volatility. However, the traditional estimation approach, as argued by Duan (1994) is misspecified and uses a redundant equation in estimation. Duan instead proposes a maximum likelihood (ML) approach that derives firm values from equity prices. Ericsson and Reneby (2005) test both the traditional and ML estimation approaches and find that the performance of the structural models is greatly improved under the ML approach. Li and Wong (2008) test both approaches on actual bond data and obtain results consistent with Eric-

son and Reneby's conclusions.

We modify the ML approach of Duan (1994) by proposing a Monte Carlo method that can be used if a time-series of historical bond prices is available. Therefore, the approach can be applied to firms that do not have observable equity prices, i.e. firms that are private. In addition, by the use of Monte Carlo, we can easily test model assumptions using a variety of interest rate processes since closed-form solutions are not required.

We apply the Merton (1974), Black and Cox (1976) and Briys and de Varenne (1997) structural models under the constant, Vasicek and CIR interest rate assumptions to one bond belonging to a private firm. We also test each model assuming two different treatments of coupon bonds. The first is the 'portfolio of zeroes' (PZ) approach used by Eom, Helwege and Huang (2004) where each individual payment of a bond is treated as a zero-coupon bond. However, this method does not account for the fact that once default occurs, a firm ceases to exist. Therefore, we also apply the models under an internally consistent (IC) assumption. For completeness, we also apply the CIR intensity reduced-form model outlined by Duffie and Singleton (2003).

Our results are consistent with that of Ericsson and Reneby (2005) and Li and Wong (2008) in that the performance of the structural models is greatly improved under an ML approach. This suggests that the inadequacies of the traditional estimation approach limit the performance of structural models. We, in fact, find that the structural models perform better than the CIR intensity model. However, although an ML approach removes the bias that is apparent when using the traditional method, there is still a large variability in spread prediction errors. These errors are likely to

be larger when the approach is applied to a bond that is illiquid and is issued by a high-risk firm since estimated roots would be quite variable and forecasting would likely result in more ‘overshooting’.

We find that the Merton model outperforms the Black and Cox, and Briys and de Varenne models in accurately predicting spreads. We gain nothing and rather lose accuracy from including default barriers and deviations from absolute priority. However, this may be the result of difficulties in estimating many parameters with an estimated log-likelihood function that is quite jagged. We also find that at the optimum points in each of the models, there is very little difference between the default probabilities under the PZ and IC approaches. This suggests that assuming a coupon bond is a portfolio of zero-coupon bonds is not an overly inaccurate assumption since coupons are small and unlikely to default. However, further research should test this theory on bonds that pay large coupons since large coupons are more likely to trigger default.

Although we find that the Vasicek and CIR interest rate models perform poorly in comparison to a constant interest rate assumption, we cannot conclude in favour of a constant interest rate. Rather, it is likely that a more complex model of the short rate is required. The use of a no-arbitrage model would remove the issues of under and over-discounting we observe with the Vasicek and CIR interest rate models and possibly greatly improve the performance of the structural models. Therefore further research should explore this avenue.

The results presented in this paper should be taken with caution. Since we have only estimated the models for one bond, we cannot be sure how the Monte Carlo ML approach will perform on bonds for differing risk profiles and coupon sizes. As

we have stated, this would require a large amount of computational time. This time required is an increasing function of the time-to-maturity and coupon frequency of a bond. Li and Wong (2008) argue that it may be necessary to perform estimation across multiple computers.

In addition, we obtain estimated log-likelihood surfaces that are quite jagged due to the use of Monte Carlo simulation. Therefore it seems we require many simulation paths (higher N) to smooth out the estimated functions and improve accuracy. This increases the computational time required to perform estimation. As a test on the effect of the number of simulation paths, it may be worthwhile repeating estimation on a single bond for many different levels of N to determine a reasonable number of paths required.

The numerous problems associated with calibrating firm values outlined in Chapter 7 make the Monte Carlo ML approach difficult to apply. However, this may not represent a weakness of the approach but rather, misspecified models. Through the use of Monte Carlo simulation instead of closed-form solutions, we gain a better insight to what features of a model are problematic in application. The Black and Cox barrier definition seems inadequate since the situation of default preference can arise. This may explain why we find no previous applications of this model. In addition, the main advantages of assuming a geometric Brownian motion for firm value are that it does not allow negative values and permits the use of the normal distribution for the log-firm value. Thus the geometric Brownian motion is used more out of convenience.

The results of this paper suggest many avenues for further research. We have dismissed the use of the Vasicek and CIR interest rate models in this paper as a result of

finding better model performance under a constant interest rate assumption. Further research should focus on using no-arbitrage models of the riskless term structure in pricing corporate bonds. This is because no-arbitrage models match the initial term structure and there are no problems with under or over discounting apparent in equilibrium term structure models. Secondly, the Monte Carlo ML approach should be applied on different assumed processes for the firm value. With Monte Carlo, we can easily test which processes perform better than others since we do not require closed-form solutions. Thirdly, it seems necessary to use ML estimation to estimate more reduced-form models. To be sure that the performance of structural models dominate the performance of reduced-form models under ML estimation, we would have to test the approach on multiple reduced-form models. Fourthly, we have only estimated the discussed models for one bond. Implementing the models on more bonds would allow us to make more concrete conclusions on model performance. Furthermore, applying bonds of differing risk profiles and coupon sizes would aid in interpretation of individual parameter estimates. Finally, since we have encountered a problem with jagged estimated log-likelihood surfaces, it is also necessary to repeat estimation for multiple levels of N to investigate whether parameter estimates change significantly when N is increased. This would aid in determining the number of simulation paths necessary to have confidence in the Monte Carlo approach. However, as we have said, this is expensive computationally.

Appendix A

Corporate Theory Appendix

A.1 Probability of Zero Default Events

If X is a Poisson distributed random variable with rate λ then,

$$P(X = x) = \frac{\lambda t^x e^{-\lambda t}}{x!} \quad (\text{A.1})$$

for an interval of length t . Survival past t implies zero Poisson events between 0 and t so it follows from (A.1) that,

$$\begin{aligned} P(X = 0) &= \frac{\lambda t^0 e^{-\lambda t}}{0!} \\ &= e^{-\lambda t} \end{aligned}$$

as required.

A.2 Replicating Portfolio Weights

We have a system of equations,

$$\begin{aligned} W_1(\mu - r) + W_2(\mu_Y - r) &= 0 \\ W_1\sigma + W_2\sigma_Y &= 0 \end{aligned}$$

We can denote this matrix form $\mathbf{A}w = \mathbf{b}$ where,

$$\mathbf{A} = \begin{pmatrix} \mu - r & \mu_Y - r \\ \sigma & \sigma_Y \end{pmatrix}$$

and

$$\mathbf{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

We first construct the augmented matrix.

$$\begin{pmatrix} \mathbf{A} & : & \mathbf{b} \end{pmatrix} \\ = \begin{pmatrix} \mu - r & \mu_Y - r & 0 \\ \sigma & \sigma_Y & 0 \end{pmatrix}$$

and then perform row operations on this matrix.

$$\xrightarrow{r_1 \frac{1}{\mu - r}} \begin{pmatrix} 1 & \frac{\mu_Y - r}{\mu - r} & 0 \\ \sigma & \sigma_Y & 0 \end{pmatrix} \\ \xrightarrow{r_2 - \sigma r_1} \begin{pmatrix} 1 & \frac{\mu_Y - r}{\mu - r} & 0 \\ 0 & \sigma_Y - \sigma \frac{\mu_Y - r}{\mu - r} & 0 \end{pmatrix}$$

Therefore we have the condition,

$$\left(\sigma_Y - \sigma \frac{\mu_Y - r}{\mu - r} \right) W_2 = 0$$

This holds for non-zero W_2 only if,

$$\left(\sigma_Y - \sigma \frac{\mu_Y - r}{\mu - r} \right) = 0$$

which rearranged gives,

$$\left(\frac{\mu - r}{\sigma} \right) = \left(\frac{\mu_Y - r}{\sigma_Y} \right)$$

as required.

A.3 The Black and Cox (1976) Bond Price

The price of a risky zero-coupon bond using the Black and Cox (1976) model is,

$$\begin{aligned} \overline{B}_t(V_t, T, F, 0) &= F e^{-rT^*} \left(\Phi(h_1) + R_t^{2\tilde{a}} \Phi(h_2) \right) \\ &+ V_t e^{-\alpha T^*} \left(\Phi(h_3) - \Phi(h_4) + R_t^{2\tilde{a}+2} (\Phi(h_5) - \Phi(h_6)) \right) \\ &+ V_t \left(R_t^{\theta+\zeta} \Phi(h_7) + R_t^{\theta-\zeta} \Phi(h_8) \right) \end{aligned} \tag{A.2}$$

where

$$\begin{aligned}
\tilde{a} &= \frac{r - \alpha - \gamma - \frac{1}{2}\sigma^2}{\sigma^2} \\
R_t &= \frac{De^{-\gamma T^*}}{V_t} \\
\theta &= \tilde{a} + 1 \\
\zeta &= \sigma^{-2} \sqrt{(r - \alpha - \gamma - \frac{1}{2}\sigma^2)^2 + 2\sigma^2(r - \gamma)} \\
h_1 &= \frac{\ln(V_t/F) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
h_2 &= \frac{2\ln(De^{-\gamma T^*}) - \ln(V_t F) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
h_3 &= \frac{\ln(F/V_t) - (r - \alpha + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
h_4 &= \frac{\ln(D/V_t) - (r - \alpha + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
h_5 &= \frac{2\ln(De^{-\gamma T^*}) - \ln(V_t D) + (r - \alpha + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
h_6 &= \frac{2\ln(De^{-\gamma T^*}) - \ln(V_t D) + (r - \alpha + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
h_7 &= \frac{\ln(De^{-\gamma T^*}/V_t) + \zeta\sigma^2 T^*}{\sigma\sqrt{T^*}} \\
h_8 &= \frac{\ln(De^{-\gamma T^*}/V_t) - \zeta\sigma^2 T^*}{\sigma\sqrt{T^*}}
\end{aligned}$$

Bielecki and Rutkowski (2002) provide a brief proof to (A.2). We will now attempt to work through the proof step-by-step to ensure we understand where each term comes from. There are three possible outcomes for a holder of a risky zero coupon bond in the Black and Cox model. The default time is defined as τ .

1. $t < \tau < T$
2. $\tau = T$
3. $T < \tau$ (i.e. no default)

To price the bond, we need to account for all three outcomes. Therefore the price of the bond can said to be,

$$\overline{B}_t(V_t, T, F, 0) = E_{\mathbb{Q}}\left(Fe^{-rT^*}\mathbb{I}_{\{\tau \geq T, V_T \geq F\}}\right) \quad (\text{A.3})$$

$$+ E_{\mathbb{Q}}\left(V_T e^{-rT^*}\mathbb{I}_{\{\tau \geq T, V_T < F\}}\right) \quad (\text{A.4})$$

$$+ E_{\mathbb{Q}}\left(De^{-\gamma(T-\tau)}e^{-r(\tau-t)}\mathbb{I}_{\{t < \tau < T\}}\right) \quad (\text{A.5})$$

Where \mathbb{I} is the indicator function and \mathbb{Q} is the risk neutral measure. First let's consider (A.3) which is the case of no default.

$$E_{\mathbb{Q}}\left(Fe^{-rT^*}\mathbb{I}_{\{\tau \geq T, V_T \geq F\}}\right) = Fe^{-rT^*}P_{\mathbb{Q}}\left(\tau \geq T, V_T \geq F\right)$$

We can assume the barrier is constant at $De^{-\gamma T^*}$ if we adjust the drift of the log firm value appropriately. Hence the probability we are trying to evaluate is,

$$P_{\mathbb{Q}}\left(\ln(V_t) > \ln(F) - \gamma T^*, m_T > \ln(De^{-\gamma T^*})\right)$$

where m_T is the running minimum log-firm value up to time T . The probability is equal to (see Bielecki and Rutkowski (2002)),

$$\begin{aligned} & \Phi\left[\frac{\ln(V_t/F) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right] \\ & - \exp\left\{-2\frac{(r - \alpha - \gamma - \frac{1}{2}\sigma^2)}{\sigma^2}(\ln(V_t) - \ln(De^{-\gamma T^*}))\right\} \\ & \Phi\left[\frac{2\ln(De^{-\gamma T^*}) - \ln(V_tF) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right] \end{aligned} \quad (\text{A.6})$$

where

$$\begin{aligned} \exp\left\{-2\frac{(r - \alpha - \gamma - \frac{1}{2}\sigma^2)}{\sigma^2}(\ln(V_t) - \ln(De^{-\gamma T^*}))\right\} &= \left[\frac{De^{-\gamma T^*}}{V_t}\right]^{\frac{2(r - \alpha - \gamma - \frac{1}{2}\sigma^2)}{\sigma^2}} \\ &= R_t^{2\tilde{a}} \end{aligned}$$

Therefore (A.3) turns out to be,

$$Fe^{-rT^*}\left(\Phi(h_1) + R_t^{2\tilde{a}}\Phi(h_2)\right) \quad (\text{A.7})$$

as required. The second and third parts are more complicated. Let's now consider (A.4). This refers to the case of default at maturity. Here, the firm value is above the barrier but below the face value of the bond.

$$E_{\mathbb{Q}}\left(V_Te^{-rT^*}\mathbb{I}_{\{\tau \geq T, V_T < F\}}\right) = e^{-rT^*}E_{\mathbb{Q}}\left(V_T\mathbb{I}_{\{\tau \geq T, V_T < F\}}\right)$$

This involves evaluating the following integral,

$$e^{-rT^*} \int_D^F x dP_{\mathbb{Q}}\left(\tau \geq T, V_T < x\right)$$

We need to derive $dP_{\mathbb{Q}}\left(\tau \geq T, V_T < x\right)$. It is straightforward to see that,

$$P_{\mathbb{Q}}\left(\tau \geq T, V_T < x\right) = P_{\mathbb{Q}}\left(\tau \geq T\right) - P_{\mathbb{Q}}\left(\tau \geq T, V_T \geq x\right) \quad (\text{A.8})$$

We already know the second part of (A.8) since we can simply use (A.6) by replacing F with x . Therefore we have,

$$\begin{aligned} P_{\mathbb{Q}}\left(\tau \geq T, V_T \geq x\right) &= \Phi\left[\frac{\ln(V_t/x) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right] \\ &- R_t^{2\tilde{a}}\Phi\left[\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t x) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right] \end{aligned}$$

It turns out we can also use (A.6) to solve the first part of (A.8) by replacing F with D . This is the probability of the firm value being above the barrier until maturity.

$$\begin{aligned} P_{\mathbb{Q}}\left(\tau \geq T\right) &= P_{\mathbb{Q}}\left(\tau \geq T, V_T \geq D\right) \\ &= \Phi\left[\frac{\ln(V_t/D) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right] \\ &- R_t^{2\tilde{a}}\Phi\left[\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t D) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right] \end{aligned}$$

Therefore (A.8) is equal to,

$$\begin{aligned} P_{\mathbb{Q}}\left(\tau \geq T, V_T < x\right) &= \Phi\left[\frac{\ln(V_t/D) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right] \\ &- R_t^{2\tilde{a}}\Phi\left[\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t D) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right] \\ &- \Phi\left[\frac{\ln(V_t/x) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right] \\ &+ R_t^{2\tilde{a}}\Phi\left[\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t x) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right] \end{aligned}$$

It is obvious that $P_{\mathbb{Q}}\left(\tau \geq T\right)$ does not depend on x so it follows that,

$$\begin{aligned} dP_{\mathbb{Q}}\left(\tau \geq T, V_T < x\right) &= -d\Phi(A) + R_t^{2\tilde{a}}d\Phi(B) \\ &= \left(-\Phi'(A)\frac{\partial A}{\partial x} + R_t^{2\tilde{a}}\Phi'(B)\frac{\partial B}{\partial x}\right)dx \\ &= \left(-\Phi'(-A)\frac{\partial A}{\partial x} + R_t^{2\tilde{a}}\Phi'(B)\frac{\partial B}{\partial x}\right)dx \\ &= \left(\Phi'(-A)\frac{\partial(-A)}{\partial x} + R_t^{2\tilde{a}}\Phi'(B)\frac{\partial B}{\partial x}\right)dx \\ &= d\Phi(-A) + R_t^{2\tilde{a}}d\Phi(B) \end{aligned}$$

where

$$\begin{aligned} A &= \frac{\ln(V_t/x) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\ B &= \frac{2\ln(De^{-\gamma T^*}) - \ln(V_t x) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \end{aligned}$$

Thus we obtain,

$$\begin{aligned} dP_{\mathbb{Q}}\left(\tau \geq T, V_T < x\right) &= d\Phi\left(\frac{\ln(x/V_t) - (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) \\ &+ R_t^{2\tilde{a}} d\Phi\left(\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t x) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) \end{aligned}$$

So our integral now becomes,

$$\begin{aligned} &e^{-rT^*} \int_D^F x dP_{\mathbb{Q}}\left(\tau \geq T, V_T < x\right) \\ &= e^{-rT^*} \int_D^F x \left[d\Phi\left(\frac{\ln(x/V_t) - (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) \right. \\ &\quad \left. + R_t^{2\tilde{a}} d\Phi\left(\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t x) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) \right] \end{aligned} \quad (\text{A.9})$$

We can expand the integral into two parts and integrate each separately. The first part is,

$$\begin{aligned} &\int_D^F x d\Phi\left(\frac{\ln(x/V_t) - (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) \\ &= \int_D^F x \phi\left(\frac{\ln(x/V_t) - (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) \frac{1}{x\sigma\sqrt{T^*}} dx \\ &= \frac{1}{\sigma\sqrt{T^*}} \int_D^F \phi\left(\frac{\ln(\frac{x}{V_t}) - (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) dx \end{aligned}$$

Our integral can be made much simpler if we perform a change of variable. We define a new variable z where,

$$\begin{aligned} z &= \frac{\ln(x/V_t) - (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\ \frac{dz}{dx} &= \frac{1}{x\sigma\sqrt{T^*}} \end{aligned}$$

and

$$\begin{aligned} x &= \exp\left\{z\sigma\sqrt{T^*} + \ln(V_t) + (r - \alpha - \frac{1}{2}\sigma^2)T^*\right\} \\ dx &= x\sigma\sqrt{T^*} dz \\ &= \exp\left\{z\sigma\sqrt{T^*} + \ln(V_t) + (r - \alpha - \frac{1}{2}\sigma^2)T^*\right\} \sigma\sqrt{T^*} dz \end{aligned}$$

After the change of variable the integral becomes,

$$\begin{aligned}
& \int_Q^R \phi(z) e^{z\sigma\sqrt{T^*} + \ln(V_t) + (r - \alpha - \frac{1}{2}\sigma^2)T^*} dz \\
&= \int_Q^R \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} e^{z\sigma\sqrt{T^*} + \ln(V_t) + (r - \alpha - \frac{1}{2}\sigma^2)T^*} dz \\
&= e^{\frac{1}{2}\sigma^2 T^* + \ln(V_t) + (r - \alpha - \frac{1}{2}\sigma^2)T^*} \int_Q^R \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z - \sigma\sqrt{T^*})^2} dz \\
&= V_t e^{(r - \alpha)T^*} \int_Q^R \phi(z - \sigma\sqrt{T^*}) dz \\
&= V_t e^{(r - \alpha)T^*} \int_Q^R d\Phi(z - \sigma\sqrt{T^*})
\end{aligned}$$

where

$$\begin{aligned}
Q &= \frac{\ln(F/V_t) - (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
R &= \frac{\ln(D/V_t) - (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}
\end{aligned}$$

Evaluating the integral give us,

$$\begin{aligned}
& V_t e^{(r - \alpha)T^*} \left[\Phi\left(\frac{\ln(F/V_t) - (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} - \sigma\sqrt{T^*}\right) \right. \\
& \quad \left. - \Phi\left(\frac{\ln(D/V_t) - (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} - \sigma\sqrt{T^*}\right) \right] \\
&= V_t e^{(r - \alpha)T^*} \left[\Phi\left(\frac{\ln(F/V_t) - (r - \alpha + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) - \Phi\left(\frac{\ln(D/V_t) - (r - \alpha + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) \right] \\
&= V_t e^{(r - \alpha)T^*} \left(\Phi(h_3) - \Phi(h_4) \right)
\end{aligned}$$

We now turn to the second part of the integral,

$$\begin{aligned}
& \int_D^F x d\Phi\left(\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t x) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) \\
&= \int_D^F x \phi\left(\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t x) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) \left(-\frac{1}{x\sigma\sqrt{T^*}}\right) dx \\
&= -\frac{1}{\sigma\sqrt{T^*}} \int_D^F \phi\left(\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t x) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) dx
\end{aligned}$$

We again perform a change of variable to simplify the integral.

$$\begin{aligned}
z &= \frac{2\ln(De^{-\gamma T^*}) - \ln(V_t x) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
\frac{dz}{dx} &= -\frac{1}{x\sigma\sqrt{T^*}}
\end{aligned}$$

and

$$\begin{aligned}
x &= \exp\left\{2\ln(De^{-\gamma T^*}) - \ln(V_t) + (r - \alpha - \frac{1}{2}\sigma^2)T^* - z\sigma\sqrt{T^*}\right\} \\
dx &= -x\sigma\sqrt{T^*}dz \\
&= -\exp\left\{2\ln(De^{-\gamma T^*}) - \ln(V_t) + (r - \alpha - \frac{1}{2}\sigma^2)T^* - z\sigma\sqrt{T^*}\right\}\sigma\sqrt{T^*}dz
\end{aligned}$$

After the change of variable the integral becomes,

$$\begin{aligned}
&\int_Q^R \phi(z) e^{2\ln(De^{-\gamma T^*}) - \ln(V_t) + (r - \alpha - \frac{1}{2}\sigma^2)T^* - z\sigma\sqrt{T^*}} dz \\
&= \int_Q^R \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} e^{2\ln(De^{-\gamma T^*}) - \ln(V_t) + (r - \alpha - \frac{1}{2}\sigma^2)T^* - z\sigma\sqrt{T^*}} dz \\
&= e^{\frac{1}{2}\sigma^2 T^* + 2\ln(De^{-\gamma T^*}) - \ln(V_t) + (r - \alpha - \frac{1}{2}\sigma^2)T^*} \int_Q^R \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z + \sigma\sqrt{T^*})^2} dz \\
&= V_t R_t^2 e^{(r - \alpha)T^*} \int_Q^R \phi(z + \sigma\sqrt{T^*}) dz \\
&= V_t R_t^2 e^{(r - \alpha)T^*} \int_Q^R d\Phi(z + \sigma\sqrt{T^*})
\end{aligned}$$

where

$$\begin{aligned}
Q &= \frac{2\ln(De^{-\gamma T^*}) - \ln(V_t F) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} \\
R &= \frac{2\ln(De^{-\gamma T^*}) - \ln(V_t D) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}
\end{aligned}$$

Evaluating the integral give us,

$$\begin{aligned}
&V_t R_t^2 e^{(r - \alpha)T^*} \left[\Phi\left(\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t F) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} + \sigma\sqrt{T^*}\right) \right. \\
&- \left. \Phi\left(\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t D) + (r - \alpha - \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} + \sigma\sqrt{T^*}\right) \right] \\
&= V_t R_t^2 e^{(r - \alpha)T^*} \left[\Phi\left(\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t F) + (r - \alpha + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) \right. \\
&- \left. \Phi\left(\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t D) + (r - \alpha + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right) \right] \\
&= V_t R_t^2 e^{(r - \alpha)T^*} \left(\Phi(h_5) - \Phi(h_6) \right)
\end{aligned}$$

Therefore (A.9) is equal to,

$$\begin{aligned}
&e^{-rT^*} \left[V_t e^{(r - \alpha)T^*} \left(\Phi(h_3) - \Phi(h_4) \right) + R_t^{2\tilde{a}} \left(V_t R_t^2 e^{(r - \alpha)T^*} \left(\Phi(h_5) - \Phi(h_6) \right) \right) \right] \\
&= V_t e^{-\alpha T^*} \left(\Phi(h_3) - \Phi(h_4) \right) + V_t R_t^{2\tilde{a}+2} e^{-\alpha T^*} \left(\Phi(h_5) - \Phi(h_6) \right) \\
&= V_t e^{-\alpha T^*} \left(\Phi(h_3) - \Phi(h_4) + R_t^{2\tilde{a}+2} (\Phi(h_5) - \Phi(h_6)) \right)
\end{aligned}$$

as required. Finally we turn to (A.5). We have,

$$\begin{aligned}
& E_{\mathbb{Q}} \left(De^{-\gamma(T-\tau)} e^{-r(\tau-t)} \mathbb{I}_{t < \tau < T} \mid \Omega_t \right) \\
&= De^{-\gamma T} e^{rt} E_{\mathbb{Q}} \left(e^{(\gamma-r)\tau} \mathbb{I}_{t < \tau < T} \mid \Omega_t \right) \\
&= De^{-\gamma T} e^{rt} \int_t^T e^{(\gamma-r)s} dP_{\mathbb{Q}}(\tau < s)
\end{aligned} \tag{A.10}$$

The probability of default before time s is,

$$\begin{aligned}
P_{\mathbb{Q}}(\tau < s) &= 1 - P_{\mathbb{Q}}(\tau \geq s) \\
&= 1 - \Phi \left[\frac{\ln(V_t) - \ln(De^{-\gamma(T-s)}) + (r - \alpha - \frac{1}{2}\sigma^2)(s-t)}{\sigma\sqrt{s-t}} \right] \\
&+ R_t^{2\tilde{a}} \Phi \left[\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t) - \ln(De^{-\gamma(T-s)}) + (r - \alpha - \frac{1}{2}\sigma^2)(s-t)}{\sigma\sqrt{s-t}} \right] \\
&= \Phi \left[\frac{\ln(De^{-\gamma(T-s)}) - \ln(V_t) - (r - \alpha - \frac{1}{2}\sigma^2)(s-t)}{\sigma\sqrt{s-t}} \right] \\
&+ R_t^{2\tilde{a}} \Phi \left[\frac{2\ln(De^{-\gamma T^*}) - \ln(V_t) - \ln(De^{-\gamma(T-s)}) + (r - \alpha - \frac{1}{2}\sigma^2)(s-t)}{\sigma\sqrt{s-t}} \right] \\
&= \Phi \left[\frac{\ln(De^{-\gamma T^*}/V_t) - (r - \alpha - \gamma - \frac{1}{2}\sigma^2)(s-t)}{\sigma\sqrt{s-t}} \right] \\
&+ R_t^{2\tilde{a}} \Phi \left[\frac{\ln(De^{-\gamma T^*}/V_t) + (r - \alpha - \gamma - \frac{1}{2}\sigma^2)(s-t)}{\sigma\sqrt{s-t}} \right]
\end{aligned}$$

So it follows that (A.10) becomes,

$$\begin{aligned}
& De^{-\gamma T} e^{rt} \int_t^T e^{(\gamma-r)s} dP_{\mathbb{Q}}(\tau < s) \\
&= De^{-\gamma T} e^{rt} \int_t^T e^{(\gamma-r)s} \left[d\Phi \left(\frac{\ln(De^{-\gamma T^*}/V_t) - (r - \alpha - \gamma - \frac{1}{2}\sigma^2)(s-t)}{\sigma\sqrt{s-t}} \right) \right. \\
&+ \left. R_t^{2\tilde{a}} d\Phi \left(\frac{\ln(De^{-\gamma T^*}/V_t) + (r - \alpha - \gamma - \frac{1}{2}\sigma^2)(s-t)}{\sigma\sqrt{s-t}} \right) \right]
\end{aligned} \tag{A.11}$$

We can evaluate (A.11) by separating the integral into two parts and evaluating each part. The first part we need to evaluate is (ignoring the term outside the integral),

$$\int_t^T e^{(\gamma-r)s} d\Phi \left(\frac{\ln(De^{-\gamma T^*}/V_t) - (r - \alpha - \gamma - \frac{1}{2}\sigma^2)(s-t)}{\sigma\sqrt{s-t}} \right)$$

For notational simplicity let's denote this as,

$$\int_t^T e^{as} d\Phi \left(\frac{b - c(s-t)}{\sqrt{s-t}} \right)$$

where

$$\begin{aligned} a &= \gamma - r \\ b &= \sigma^{-1} \ln \left(\frac{De^{-\gamma T^*}}{V_t} \right) \\ c &= \sigma^{-1} (r - \alpha - \gamma - \frac{1}{2} \sigma^2) \end{aligned}$$

We can express the integral as,

$$\int_t^T e^{as} \phi \left(\frac{b - c(s-t)}{\sqrt{s-t}} \right) \left(-\frac{b}{2(s-t)^{\frac{3}{2}}} - \frac{c}{2\sqrt{s-t}} \right) ds \quad (\text{A.12})$$

Let's define the following,

$$\begin{aligned} g(s) &= e^{b(c-d)} \Phi \left(\frac{b-d(s-t)}{\sqrt{s-t}} \right) \\ g'(s) &= e^{b(c-d)} \phi \left(\frac{b-d(s-t)}{\sqrt{s-t}} \right) \left(-\frac{b}{2(s-t)^{\frac{3}{2}}} - \frac{d}{2\sqrt{s-t}} \right) \end{aligned}$$

It follows that,

$$\begin{aligned} \phi \left(\frac{b-d(s-t)}{\sqrt{s-t}} \right) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{b-d(s-t)}{\sqrt{s-t}} \right)^2} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{b^2 - 2bd(s-t) + d^2(s-t)^2}{s-t} \right)} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \frac{b^2}{s-t}} e^{bd} e^{-\frac{1}{2} (c^2 - 2a)(s-t)} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \frac{b^2}{s-t}} e^{bd} e^{-\frac{1}{2} c^2 (s-t)} e^{a(s-t)} \end{aligned}$$

Thus

$$\begin{aligned} e^{b(c-d)} \phi \left(\frac{b-d(s-t)}{\sqrt{s-t}} \right) &= e^{bc} e^{-bd} e^{bd} e^{a(s-t)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \frac{b^2}{s-t} - \frac{1}{2} c^2 (s-t)} \\ &= e^{a(s-t)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{b^2 + c^2 (s-t)^2 - 2bc(s-t)}{s-t} \right)} \\ &= e^{a(s-t)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{b-c(s-t)}{\sqrt{s-t}} \right)^2} \\ &= e^{a(s-t)} \phi \left(\frac{b-c(s-t)}{\sqrt{s-t}} \right) \end{aligned}$$

Therefore

$$g'(s) = e^{a(s-t)} \phi \left(\frac{b-c(s-t)}{\sqrt{s-t}} \right) \left(-\frac{b}{2(s-t)^{\frac{3}{2}}} - \frac{d}{2\sqrt{s-t}} \right)$$

We also define,

$$\begin{aligned} h(s) &= e^{b(c+d)} \Phi \left(\frac{b+d(s-t)}{\sqrt{s-t}} \right) \\ h'(s) &= e^{b(c+d)} \phi \left(\frac{b+d(s-t)}{\sqrt{s-t}} \right) \left(-\frac{b}{2(s-t)^{\frac{3}{2}}} + \frac{d}{2\sqrt{s-t}} \right) \end{aligned}$$

It follows that,

$$\begin{aligned}
\phi\left(\frac{b+d(s-t)}{\sqrt{s-t}}\right) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b+d(s-t)}{\sqrt{s-t}}\right)^2} \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b^2+2bd(s-t)+d^2(s-t)^2}{s-t}\right)} \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{b^2}{s-t}} e^{-bd} e^{-\frac{1}{2}(c^2-2a)(s-t)} \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{b^2}{s-t}} e^{-bd} e^{-\frac{1}{2}c^2(s-t)} e^{a(s-t)}
\end{aligned}$$

Thus

$$\begin{aligned}
e^{b(c+d)}\phi\left(\frac{b+d(s-t)}{\sqrt{s-t}}\right) &= e^{bc}e^{bd}e^{-bd}e^{a(s-t)}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\frac{b^2}{s-t}-\frac{1}{2}c^2(s-t)} \\
&= e^{a(s-t)}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{b^2+c^2(s-t)^2-2bc(s-t)}{s-t}\right)} \\
&= e^{a(s-t)}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{b-c(s-t)}{\sqrt{s-t}}\right)^2} \\
&= e^{a(s-t)}\phi\left(\frac{b-c(s-t)}{\sqrt{s-t}}\right)
\end{aligned}$$

Therefore

$$h'(s) = e^{a(s-t)}\phi\left(\frac{b-c(s-t)}{\sqrt{s-t}}\right)\left(-\frac{b}{2(s-t)^{\frac{3}{2}}} + \frac{d}{2\sqrt{s-t}}\right)$$

So we have,

$$\begin{aligned}
g'(s) + h'(s) &= e^{a(s-t)}\phi\left(\frac{b-c(s-t)}{\sqrt{s-t}}\right)\left(-\frac{b}{2(s-t)^{\frac{3}{2}}} - \frac{d}{2\sqrt{s-t}}\right) \\
&+ e^{a(s-t)}\phi\left(\frac{b-c(s-t)}{\sqrt{s-t}}\right)\left(-\frac{b}{2(s-t)^{\frac{3}{2}}} + \frac{d}{2\sqrt{s-t}}\right) \\
&= e^{a(s-t)}\phi\left(\frac{b-c(s-t)}{\sqrt{s-t}}\right)\left(-\frac{b}{2(s-t)^{\frac{3}{2}}} - \frac{d}{2\sqrt{s-t}} - \frac{b}{2(s-t)^{\frac{3}{2}}} + \frac{d}{2\sqrt{s-t}}\right) \\
&= e^{a(s-t)}\phi\left(\frac{b-c(s-t)}{\sqrt{s-t}}\right)\left(-\frac{2b}{2(s-t)^{\frac{3}{2}}}\right) \\
&= -e^{a(s-t)}\phi\left(\frac{b-c(s-t)}{\sqrt{s-t}}\right)\frac{b}{(s-t)^{\frac{3}{2}}}
\end{aligned}$$

and

$$\begin{aligned}
g'(s) - h'(s) &= e^{a(s-t)}\phi\left(\frac{b-c(s-t)}{\sqrt{s-t}}\right)\left(-\frac{b}{2(s-t)^{\frac{3}{2}}} - \frac{d}{2\sqrt{s-t}}\right) \\
&- e^{a(s-t)}\phi\left(\frac{b-c(s-t)}{\sqrt{s-t}}\right)\left(-\frac{b}{2(s-t)^{\frac{3}{2}}} + \frac{d}{2\sqrt{s-t}}\right) \\
&= e^{a(s-t)}\phi\left(\frac{b-c(s-t)}{\sqrt{s-t}}\right)\left(-\frac{b}{2(s-t)^{\frac{3}{2}}} - \frac{d}{2\sqrt{s-t}} + \frac{b}{2(s-t)^{\frac{3}{2}}} - \frac{d}{2\sqrt{s-t}}\right) \\
&= e^{a(s-t)}\phi\left(\frac{b-c(s-t)}{\sqrt{s-t}}\right)\left(-\frac{2d}{2\sqrt{s-t}}\right) \\
&= -e^{a(s-t)}\phi\left(\frac{b-c(s-t)}{\sqrt{s-t}}\right)\frac{d}{\sqrt{s-t}}
\end{aligned}$$

We can now write (A.12) as,

$$\frac{1}{2}e^{at} \int_t^T \left[g'(s) + h'(s) + \frac{c}{d}(g'(s) - h'(s)) \right] ds$$

Since

$$\begin{aligned} & \int_t^T \left[\frac{1}{2}e^{at} \left(-e^{a(s-t)} \phi \left(\frac{b-c(s-t)}{\sqrt{s-t}} \right) \frac{b}{(s-t)^{\frac{3}{2}}} \right) + \frac{c}{2d}e^{at} \left(-e^{a(s-t)} \phi \left(\frac{b-c(s-t)}{\sqrt{s-t}} \right) \frac{d}{\sqrt{s-t}} \right) \right] ds \\ &= \int_t^T \left[-e^{as} \phi \left(\frac{b-c(s-t)}{\sqrt{s-t}} \right) \frac{b}{2(s-t)^{\frac{3}{2}}} - e^{as} \phi \left(\frac{b-c(s-t)}{\sqrt{s-t}} \right) \frac{c}{2\sqrt{s-t}} \right] ds \\ &= \int_t^T e^{as} \phi \left(\frac{b-c(s-t)}{\sqrt{s-t}} \right) \left(-\frac{b}{2(s-t)^{\frac{3}{2}}} - \frac{c}{2\sqrt{s-t}} \right) ds \\ &= (A.12) \end{aligned}$$

So we obtain the following,

$$\frac{1}{2}e^{at} \left[g(T) + h(T) + \frac{c}{d}(g(T) - h(T)) - g(t) - h(t) - \frac{c}{d}(g(t) - h(t)) \right]$$

However note that,

$$\begin{aligned} g(t) &= e^{b(c-d)} \Phi \left(\frac{b-dT^*}{\sqrt{T^*}} \right) \\ &= e^{b(c-d)} \Phi \left(\frac{b}{0} \right) \\ &= e^{b(c-d)} \Phi(-\infty) \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} h(t) &= e^{b(c+d)} \Phi \left(\frac{b+dT^*}{\sqrt{T^*}} \right) \\ &= e^{b(c+d)} \Phi \left(\frac{b}{0} \right) \\ &= e^{b(c+d)} \Phi(-\infty) \\ &= 0 \end{aligned}$$

Since $b = \sigma^{-1} \ln \left(\frac{De^{-\gamma T^*}}{V_t} \right) < 0$ because we assume the firm is currently above the default barrier. Therefore we are left with,

$$\begin{aligned} & \frac{1}{2}e^{at} \left(g(T) + h(T) + \frac{c}{d}(g(T) - h(T)) \right) \\ &= e^{at} \left(\frac{c+d}{2d}g(T) - \frac{c-d}{2d}h(T) \right) \end{aligned}$$

If we bring the term from outside the integral back in then we have,

$$\begin{aligned}
& De^{-\gamma T} e^{rt} e^{at} \left(\frac{c+d}{2d} g(T) - \frac{c-d}{2d} h(T) \right) \\
= & De^{-\gamma T} e^{rt} e^{(\gamma-r)t} \left(\frac{c+d}{2d} g(T) - \frac{c-d}{2d} h(T) \right) \\
= & De^{-\gamma T^*} \left(\frac{c+d}{2d} g(T) - \frac{c-d}{2d} h(T) \right) \\
= & De^{-\gamma T^*} \frac{c+d}{2d} g(T) - De^{-\gamma T^*} \frac{c-d}{2d} h(T) \\
= & De^{-\gamma T^*} \frac{c+d}{2d} e^{b(c-d)} \Phi \left(\frac{b-dT^*}{\sqrt{T^*}} \right) - De^{-\gamma T^*} \frac{c-d}{2d} e^{b(c+d)} \Phi \left(\frac{b+dT^*}{\sqrt{T^*}} \right)
\end{aligned}$$

Let's consider the first term,

$$\begin{aligned}
& De^{-\gamma T^*} \frac{c+d}{2d} e^{b(c-d)} \Phi \left(\frac{b-dT^*}{\sqrt{T^*}} \right) \\
= & De^{-\gamma T^*} \frac{\sigma \tilde{a} + \sigma \zeta}{2\sigma \zeta} e^{\ln(\frac{De^{-\gamma T^*}}{V_t})(\sigma^{-1}c - \sigma^{-1}d)} \Phi \left(\frac{b-dT^*}{\sqrt{T^*}} \right) \\
= & De^{-\gamma T^*} \frac{\tilde{a} + \zeta}{\zeta} R_t^{\tilde{a}-\zeta} \Phi \left(\frac{b-dT^*}{\sqrt{T^*}} \right) \\
= & \frac{V_t(\tilde{a} + \zeta)}{2\zeta} R_t^{\tilde{a}-\zeta+1} \Phi \left(\frac{b-dT^*}{\sqrt{T^*}} \right) \\
= & \frac{V_t(\tilde{a} + \zeta)}{2\zeta} R_t^{\theta-\zeta} \Phi \left(\frac{b-dT^*}{\sqrt{T^*}} \right) \\
= & \frac{V_t(\tilde{a} + \zeta)}{2\zeta} R_t^{\theta-\zeta} \Phi \left(\frac{\sigma b - \sigma dT^*}{\sigma \sqrt{T^*}} \right) \\
= & \frac{V_t(\tilde{a} + \zeta)}{2\zeta} R_t^{\theta-\zeta} \Phi \left(\frac{\ln(De^{-\gamma T^*}/V_t) - \zeta \sigma^2 T^*}{\sigma \sqrt{T^*}} \right) \\
= & \frac{V_t(\tilde{a} + \zeta)}{2\zeta} R_t^{\theta-\zeta} \Phi(h_8)
\end{aligned}$$

The second term is equal to,

$$\begin{aligned}
& De^{-\gamma T^*} \frac{c-d}{2d} e^{b(c+d)} \Phi \left(\frac{b+dT^*}{\sqrt{T^*}} \right) \\
= & De^{-\gamma T^*} \frac{\sigma \tilde{a} - \sigma \zeta}{2\sigma \zeta} e^{\ln(\frac{De^{-\gamma T^*}}{V_t})(\sigma^{-1}c + \sigma^{-1}d)} \Phi \left(\frac{b+dT^*}{\sqrt{T^*}} \right) \\
= & De^{-\gamma T^*} \frac{\tilde{a} - \zeta}{2\zeta} R_t^{\tilde{a}+\zeta} \Phi \left(\frac{b+dT^*}{\sqrt{T^*}} \right) \\
= & \frac{V_t(\tilde{a} - \zeta)}{2\zeta} R_t^{\tilde{a}+\zeta+1} \Phi \left(\frac{b+dT^*}{\sqrt{T^*}} \right) \\
= & \frac{V_t(\tilde{a} - \zeta)}{2\zeta} R_t^{\theta+\zeta} \Phi \left(\frac{b+dT^*}{\sqrt{T^*}} \right) \\
= & \frac{V_t(\tilde{a} - \zeta)}{2\zeta} R_t^{\theta+\zeta} \Phi \left(\frac{\sigma b + \sigma dT^*}{\sigma \sqrt{T^*}} \right) \\
= & \frac{V_t(\tilde{a} - \zeta)}{2\zeta} R_t^{\theta+\zeta} \Phi \left(\frac{\ln(De^{-\gamma T^*}/V_t) + \zeta \sigma^2 T^*}{\sigma \sqrt{T^*}} \right) \\
= & \frac{V_t(\tilde{a} - \zeta)}{2\zeta} R_t^{\theta+\zeta} \Phi(h_7)
\end{aligned}$$

So the first part of (A.11) is

$$\frac{V_t(\tilde{a} + \zeta)}{2\zeta} R_t^{\theta - \zeta} \Phi(h_8) - \frac{V_t(\tilde{a} - \zeta)}{2\zeta} R_t^{\theta + \zeta} \Phi(h_7) \quad (\text{A.13})$$

The final part of this proof is to evaluate the second part of (A.11). The integral is the following (again ignoring the term outside the integral),

$$\int_t^T e^{(\gamma - r)s} d\Phi \left(\frac{\ln\left(\frac{De^{-\gamma T^*}}{V_t}\right) + (r - \alpha - \gamma - \frac{1}{2}\sigma^2)(s - t)}{\sigma\sqrt{s - t}} \right)$$

This is very similar to the first part of (A.11). The only difference is that the '-' in the numerator is now a '+'. So again we define this integral as

$$\int_t^T e^{as} d\Phi \left(\frac{b - c(s - t)}{\sqrt{s - t}} \right)$$

where a and b are the same as before and c is now $-\sigma^{-1}(r - \alpha - \gamma - \frac{1}{2}\sigma^2)$. So again we obtain,

$$e^{at} \left(\frac{c + d}{2d} g(T) - \frac{c - d}{2d} h(T) \right)$$

Bringing all of the terms from outside the integral back in gives us,

$$\begin{aligned} & De^{-\gamma T} e^{rt} R_t^{2\tilde{a}} e^{at} \left(\frac{c + d}{2d} g(T) - \frac{c - d}{2d} h(T) \right) \\ &= De^{-\gamma T} e^{rt} R_t^{2\tilde{a}} e^{(\gamma - r)t} \left(\frac{c + d}{2d} g(T) - \frac{c - d}{2d} h(T) \right) \\ &= De^{-\gamma T^*} R_t^{2\tilde{a}} \left(\frac{c + d}{2d} g(T) - \frac{c - d}{2d} h(T) \right) \end{aligned}$$

Let's consider the first term,

$$\begin{aligned} & De^{-\gamma T^*} R_t^{2\tilde{a}} \frac{c + d}{2d} g(T) \\ &= De^{-\gamma T^*} R_t^{2\tilde{a}} \frac{-\sigma\tilde{a} + \sigma\zeta}{2\sigma\zeta} e^{b(c-d)} \Phi\left(\frac{b - dT^*}{\sqrt{T^*}}\right) \\ &= De^{-\gamma T^*} R_t^{2\tilde{a}} \frac{-\tilde{a} + \zeta}{2\zeta} e^{b(c-d)} \Phi\left(\frac{b - dT^*}{\sqrt{T^*}}\right) \\ &= V_t R_t^{2\tilde{a}+1} \left(-\frac{\tilde{a} - \zeta}{2\zeta} \right) e^{\ln\left(\frac{De^{-\gamma T^*}}{V_t}\right)(\sigma^{-1}c - \sigma^{-1}d)} \Phi\left(\frac{\sigma b - \sigma dT^*}{\sigma\sqrt{T^*}}\right) \\ &= -\frac{V_t(\tilde{a} - \zeta)}{2\zeta} R_t^{2\tilde{a}+1} R_t^{-\tilde{a} - \zeta} \Phi\left(\frac{\sigma b - \sigma dT^*}{\sigma\sqrt{T^*}}\right) \\ &= -\frac{V_t(\tilde{a} - \zeta)}{2\zeta} R_t^{\tilde{a}+1-\zeta} \Phi\left(\frac{\sigma b - \sigma dT^*}{\sigma\sqrt{T^*}}\right) \\ &= -\frac{V_t(\tilde{a} - \zeta)}{2\zeta} R_t^{\theta - \zeta} \Phi\left(\frac{\ln(De^{-\gamma T^*}/V_t) - \zeta\sigma^2 T^*}{\sigma\sqrt{T^*}}\right) \\ &= -\frac{V_t(\tilde{a} - \zeta)}{2\zeta} R_t^{\theta - \zeta} \Phi(h_8) \end{aligned}$$

Then consider the second term,

$$\begin{aligned}
& De^{-\gamma T^*} R_t^{2\tilde{a}} \frac{c-d}{2d} h(T) \\
&= De^{-\gamma T^*} R_t^{2\tilde{a}} \frac{-\sigma\tilde{a} - \sigma\zeta}{2\sigma\zeta} e^{b(c+d)} \Phi\left(\frac{b+dT^*}{\sqrt{T^*}}\right) \\
&= De^{-\gamma T^*} R_t^{2\tilde{a}} \frac{-\tilde{a} - \zeta}{2\zeta} e^{b(c+d)} \Phi\left(\frac{b+dT^*}{\sqrt{T^*}}\right) \\
&= V_t R_t^{2\tilde{a}+1} \left(-\frac{\tilde{a} + \zeta}{2\zeta}\right) e^{\ln\left(\frac{De^{-\gamma T^*}}{V_t}\right)(\sigma^{-1}c + \sigma^{-1}d)} \Phi\left(\frac{\sigma b + \sigma dT^*}{\sigma\sqrt{T^*}}\right) \\
&= -\frac{V_t(\tilde{a} + \zeta)}{2\zeta} R_t^{2\tilde{a}+1} R_t^{-\tilde{a} + \zeta} \Phi\left(\frac{\sigma b + \sigma dT^*}{\sigma\sqrt{T^*}}\right) \\
&= -\frac{V_t(\tilde{a} + \zeta)}{2\zeta} R_t^{\tilde{a}+1+\zeta} \Phi\left(\frac{\sigma b + \sigma dT^*}{\sigma\sqrt{T^*}}\right) \\
&= -\frac{V_t(\tilde{a} + \zeta)}{2\zeta} R_t^{\theta+\zeta} \Phi\left(\frac{\ln(De^{-\gamma T^*}/V_t) + \zeta\sigma^2 T^*}{\sigma\sqrt{T^*}}\right) \\
&= -\frac{V_t(\tilde{a} + \zeta)}{2\zeta} R_t^{\theta+\zeta} \Phi(h_7)
\end{aligned}$$

Therefore the second part of (A.11) is equal to,

$$\frac{V_t(\tilde{a} + \zeta)}{2\zeta} R_t^{\theta+\zeta} \Phi(h_7) - \frac{V_t(\tilde{a} - \zeta)}{2\zeta} R_t^{\theta-\zeta} \Phi(h_8) \quad (\text{A.14})$$

Therefore (A.5) is equal to,

$$\begin{aligned}
(\text{A.13}) + (\text{A.14}) &= \frac{V_t(\tilde{a} + \zeta)}{2\zeta} R_t^{\theta-\zeta} \Phi(h_8) - \frac{V_t(\tilde{a} - \zeta)}{2\zeta} R_t^{\theta+\zeta} \Phi(h_7) \\
&+ \frac{V_t(\tilde{a} + \zeta)}{2\zeta} R_t^{\theta+\zeta} \Phi(h_7) - \frac{V_t(\tilde{a} - \zeta)}{2\zeta} R_t^{\theta-\zeta} \Phi(h_8) \\
&= R_t^{\theta+\zeta} \Phi(h_7) \left(\frac{V_t(\tilde{a} + \zeta)}{2\zeta} - \frac{V_t(\tilde{a} - \zeta)}{2\zeta}\right) \\
&+ R_t^{\theta-\zeta} \Phi(h_8) \left(\frac{V_t(\tilde{a} + \zeta)}{2\zeta} - \frac{V_t(\tilde{a} - \zeta)}{2\zeta}\right) \\
&= V_t \left(R_t^{\theta+\zeta} \Phi(h_7) + R_t^{\theta-\zeta} \Phi(h_8)\right)
\end{aligned}$$

as required.

A.4 Equity Delta under Black and Scholes

We have the Black and Scholes pricing formula (3.11)-(3.13) to price equity as a European call option on the firm assets maturing at time T . Differentiating (3.11) with respect to firm value V_t gives us,

$$\frac{\partial S_t}{\partial V_t} = \Phi(d_1) + V_t \Phi'(d_1) \frac{1}{V_t \sigma \sqrt{T^*}} - F e^{-rT^*} \Phi'(d_2) \frac{1}{V_t \sigma \sqrt{T^*}} \quad (\text{A.15})$$

$\Phi'(x)$ is simply $\phi(x)$ where

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\}$$

is the standard normal density of x . We can take the ratio of the standard normal density of d_1 and the standard normal density of d_2 ,

$$\frac{\Phi'(d_1)}{\Phi'(d_2)} = \exp\left\{-\frac{1}{2}d_1^2 + \frac{1}{2}d_2^2\right\}$$

and then substitute in (3.12) and (3.13) for d_1 and d_2 .

$$\begin{aligned} \frac{\Phi'(d_1)}{\Phi'(d_2)} &= \exp\left\{-\frac{1}{2}\left(\frac{\ln(\frac{V_t}{F}) + (r + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}}\right)^2 + \frac{1}{2}\left(\frac{\ln(\frac{V_t}{F}) + (r + \frac{1}{2}\sigma^2)T^*}{\sigma\sqrt{T^*}} - \sigma\sqrt{T^*}\right)^2\right\} \\ &= \exp\left\{-\frac{1}{2\sigma^2 T^*}\left(\ln\left(\frac{V_t}{F}\right) + \left(r + \frac{1}{2}\sigma^2\right)T^*\right)^2 - \left(\ln\left(\frac{V_t}{F}\right) + \left(r - \frac{1}{2}\sigma^2\right)T^*\right)^2\right\} \\ &= \exp\left\{-\frac{1}{2\sigma^2 T^*}\left(\ln\left(\frac{V_t}{F}\right)^2 + \left(r + \frac{1}{2}\sigma^2\right)^2(T^*)^2 + 2\ln\left(\frac{V_t}{F}\right)\left(r + \frac{1}{2}\sigma^2\right)T^* \right. \right. \\ &\quad \left. \left. - \ln\left(\frac{V_t}{F}\right)^2 - \left(r - \frac{1}{2}\sigma^2\right)^2(T^*)^2 - 2\ln\left(\frac{V_t}{F}\right)\left(r - \frac{1}{2}\sigma^2\right)T^*\right)\right\} \\ &= \exp\left\{-\frac{1}{2\sigma^2 T^*}\left(2\ln\left(\frac{V_t}{F}\right)\left(\left(r + \frac{1}{2}\sigma^2\right)T^* - \left(r - \frac{1}{2}\sigma^2\right)T^*\right) + 2r\sigma^2(T^*)^2\right)\right\} \\ &= \exp\left\{-\frac{1}{2\sigma^2 T^*}\left(2\ln\left(\frac{V_t}{F}\right)(\sigma^2 T^*) + 2r\sigma^2(T^*)^2\right)\right\} \\ &= \exp\left\{-\ln\left(\frac{V_t}{F}\right) - rT^*\right\} \\ &= \exp\left\{-\ln\left(\frac{V_t}{F}\right)\right\} \exp\left\{-rT^*\right\} \\ &= \frac{F}{V_t} e^{-rT^*} \end{aligned}$$

Rearranging gives,

$$V_t \Phi'(d_1) = F e^{-rT^*} \Phi'(d_2)$$

which can then be substituted into (A.15) to obtain,

$$\begin{aligned} \frac{\partial S_t}{\partial V_t} &= \Phi(d_1) + V_t \Phi'(d_1) \frac{1}{V_t \sigma \sqrt{T^*}} - V_t \Phi'(d_1) \frac{1}{V_t \sigma \sqrt{T^*}} \\ &\quad \Phi(d_1) \end{aligned}$$

as required.

Appendix B

Notational Index

The following is a list of commonly encountered terms used throughout the paper.

- T = Bond maturity where $T > t$.
- T^* = Time to maturity $T - t$.
- V_t = Firm value at time t .
- λ_t = Default intensity at time t .
- $S_t(V_t)$ = Structural model equity price at time t .
- $B_t(T, F, c)$ = Risk-free zero coupon bond price at t with maturity T , face value F and coupon rate c .
- $\bar{B}_t(V_t, T, F, c)$ = Structural model risky bond price at t with maturity T , face value F and coupon rate c .
- $\bar{B}_t(\lambda_t, T, F, c, \delta)$ = Intensity model risky bond price at t with maturity T , face value F , coupon rate c and recovery δ .
- $\bar{B}_t^{ob}(T)$ = Observed risky bond price at time t with maturity T .
- S_t^{ob} = Observed equity price at time t .
- W_t = Brownian motion at time t .
- $P_t(T)$ = Probability at time t of survival up to time T .
- r_t = Instantaneous risk-free interest rate (short rate) at t .
- $r_t(T)$ = Spot rate (risk-free zero-coupon yield) from t to T .
- t_j = Date of transaction j .
- c = Coupon rate.
- $c(m)$ = Dollar amount of the m^{th} remaining payment of a bond.
- $y_t(T)$ = Par-yield at time t for maturity at time T .
- δ = Fractional face value recovery in reduced-form model.

Bibliography

- Bielecki, T. R. & Rutkowski, M. (2002). *“Credit Risk: Modeling, Valuation and Hedging”*, Springer-Verlag, Berlin.
- Black, F. & Cox, J. C. (1976). “Valuing Corporate Securities: Some Effects of Bond Indenture Provisions”, *Journal of Finance* **31**(2): 351–367.
- Black, F. & Scholes, M. (1973). “The Pricing of Options and Corporate Liabilities”, *Journal of Political Economy* **81**: 637–659.
- Brennan, M. J. & Schwartz, E. S. (1980). “Analyzing Convertible Bonds”, *Journal of Financial and Quantitative Analysis* **15**: 907–929.
- Briys, E. & de Varenne, F. (1997). “Valuing Risky Fixed Rate Debt: An Extension”, *Journal of Financial and Quantitative Analysis* **32**(2): 239–248.
- Chapman, D. A. & Pearson, N. D. (2001). “Recent Advances in Estimating Term Structure Models”, *Financial Analysts Journal* **57**(4): 77–95.
- Choudhry, M. (2004). *“Advanced Fixed Income Analysis”*, Elsevier Butterworth-Heinemann, Oxford.
- Collin-Dufresne, P., Goldstein, R. & Martin, S. (2001). “The Determinants of Credit Spread Changes”, *Journal of Finance* **56**: 2177–2207.
- Cox, J. C., Ingersoll, J. E. & Ross, S. A. (1985). “A Theory of the Term Structure of Interest Rates”, *Econometrica* **53**(2): 385–407.
- Dothan, U. L. (1978). “On the Term Structure of Interest Rates”, *Journal of Financial Economics* **6**: 59–69.
- Duan, J.-C. (1994). “Maximum Likelihood Estimation Using Price Data of the Derivative Contract”, *Mathematical Finance* **4**(2): 155–167.
- Duan, J.-C. (2000). “Correction: Maximum Likelihood Estimation Using Price Data of the Derivative Contract”, *Mathematical Finance* **10**(4): 461–462.
- Duffee, G. R. (1999). “Estimating the Price of Default Risk”, *Review of Financial Studies* **12**(1): 197–226.
- Duffie, D. & Singleton, K. J. (2003). *“Credit Risk: Pricing, Measurement, and Management”*, Princeton University Press, New Jersey.
- Eom, Y., Helwege, J. & Huang, J.-Z. (2004). “Structural Models of Corporate Bond Pricing: An Empirical Analysis”, *Review of Financial Studies* **17**(2): 499–544.
- Ericsson, J. & Reneby, J. (2005). “Estimating Structural Bond Pricing Models”, *Journal of Business* **78**(2): 707–735.
- Feller, W. (1951). “Two Singular Diffusion Problems”, *The Annals of Mathematics, Second Series* **54**: 173–182.

- Fong, G. & Vasicek, O. (1991). "Fixed-Income Volatility Management", *Journal of Portfolio Management* **17**(4): 41–46.
- Geske, R. (1977). "The Valuation of Corporate Liabilities as Compound Options", *Journal of Financial and Quantitative Analysis* **12**(4): 541–552.
- Greene, W. H. (2000). *"Econometric Analysis"*, Prentice-Hall, New Jersey.
- Harrison, J. M. & Pliska, S. R. (1981). "Martingales and Stochastic Integrals in the Theory of Continuous Trading", *Stochastic Processes and their Applications* **11**(3): 215–260.
- Ho, T. S. Y. & Lee, S.-B. (1986). "Term Structure Movements and Pricing Interest Rate Contingent Claims", *Journal of Finance* **41**(5): 1011–1029.
- Hull, J. C. (2006). *"Options, Futures, and Other Derivatives"*, Pearson Prentice Hall, New Jersey.
- Hull, J. & White, A. (1990). "Pricing Interest-Rate-Derivative Securities", *Review of Financial Studies* **3**(4): 573–592.
- Jarrow, R., Lando, D. & Turnbull, S. (1997). "A Markov Model for the Term Structure of Credit Risk Spreads", *Review of Financial Studies* **10**: 481–523.
- Jones, E., Mason, S. & Rosenfeld, E. (1984). "Contingent Claims Analysis of Corporate Capital Structures: An Empirical Investigation", *Journal of Finance* **39**(3): 611–625.
- Jorion, P. (2007). *"Financial Risk: Manager Handbook, 4th edition"*, Wiley, New York.
- Leland, H. E. & Toft, K. B. (1996). "Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads", *Journal of Finance* **51**(3): 987–1019.
- Li, K. L. & Wong, H. Y. (2008). "Structural Models of Corporate Bond Pricing with Maximum Likelihood Estimation", *Journal of Empirical Finance* **15**: 751–777.
- Longstaff, F. A. & Schwartz, E. S. (1995). "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt", *Journal of Finance* **50**: 789–820.
- Longstaff, F. & Schwartz, E. (1992). "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model", *Journal of Finance* **47**(4): 1259–1282.
- Lyden, S. & Saraniti, D. (2000). "An Empirical Examination of the Classical Theory of Corporate Security Valuation", *Barclays Global Investors* .
- Martellini, L., Priaulet, P. & Priaulet, S. (2003). *"Fixed-Income Securities"*, John Wiley and Sons, West Sussex.
- Merton, R. C. (1974). "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", *Journal of Finance* **29**(2): 449–470.
- Ogden, J. P. (1987). "Determinants of the Ratings and Yields on Corporate Bonds: Tests of the Contingent Claims Model", *Journal of Financial Research* **10**(4): 329–339.
- Ronn, E. I. & Verma, A. K. (1986). "Pricing Risk-Adjusted Deposit Insurance: An Option-Based Model", *Journal of Finance* **41**(4): 871–895.
- Vasicek, O. (1977). "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics* **5**(2): 177–188.