

**A HIDDEN MARKOV CHAIN MODEL FOR THE TERM STRUCTURE OF
BOND CREDIT RISK SPREADS**

Lyn C. Thomas
University of Edinburgh

David E. Allen*
Edith Cowan University

Nigel Morkel-Kingsbury
Edith Cowan University

School of Finance and Business Economics Working Paper Series
March 1999
Working Paper 99.1
ISSN: 1323-9244
ISBN: 0-7298-0447-X

Corresponding author and address:

D.E. Allen
Professor of Finance
Edith Cowan University
School of Finance and Business Economics
Joondalup Campus
Joondalup WA 6027
Australia

Phone: 61-8-9400 5471
Fax: 61-8-94005271
Email: d.allen@cowan.edu.au

Abstract

This paper provides a Markov chain model for the term structure and credit risk spreads of bond process. It allows dependency between the stochastic process modeling the interest rate and the Markov chain process describing changes in the credit rating of the bonds by their mutual dependency on a hidden Markov chain. This Markov chain can be thought of as the underlying economic conditions. The model also allows a new interpretation of risk premia used in previous approaches. It also uses a linear programming approach to strip the bonds of their coupons in such a way as to guarantee there is no mis-pricing.

Acknowledgements: This paper was written while Lyn Thomas was a Visiting Professor at Edith Cowan University. We wish to acknowledge the financial support of the University that made this visit possible.

1. Introduction

Corporate bond pricing models have been in existence for twenty-five years but it is only recently that a pricing model, which incorporates a firm's credit rating as an indicator of the likelihood of default, has been developed. This is surprising since the rating of a company given by the major international credit rating agencies is the most widely available estimate of the credit risk involved in investing in the firm's bonds. The first model of bond prices to incorporate credit ratings (Jarrow, Lando, Turnbull, 1977) assumed that the stochastic process describing the rating and possible bankruptcy of the firm was independent of the stochastic process giving future interest rates and hence the default-free bond prices. This paper presents a generalization of this model in which the two processes are dependent through their relationship with the stochastic process describing the state of the underlying economy. The model also generalizes the idea of risk premia adjustments by reinterpreting them as beliefs that the future of the rating and bankruptcy process is more extreme than it has been historically. This paper also introduces a procedure based on linear programming for stripping out the zero-bond prices for risky and riskless bonds in a way that guarantees there is no mis-pricing.

Since Merton's ground breaking paper, Merton (1974), there has been a number of modeling approaches to the price of risky debt. Duffee's review, Duffee (1996), and the paper by Jarrow et al, Jarrow (1997), outline these types of models. The first model views the firm's liabilities as contingent claims against the underlying assets and assumes that bankruptcy and bond non-payment occurs when the firm's assets are exhausted. This was the model introduced by Merton (1974), but it leads to smaller credit spreads than those that actually occur. Black and Cox (1976) adjusted the model by defining bankruptcy to occur where liabilities are some fixed proportion of the assets and this leads to more realistic credit spreads. Shimko, Teejima and van Deventer (1993) generalized the model by allowing the riskless interest rates to follow a stochastic rather than a deterministic process and that the interest rate stochastic process was correlated with the firm's asset process. Kim et al (1993) allowed bankruptcy to be triggered at an exogenously specified asset value. Leland (1994) and Leland and Toft (1996) used this type of model, but with endogeneous conditions to define when bankruptcy is declared, to examine how important is the maturity of the debt as well as the amount. The difficulty with this approach is that it depends on knowledge of the firm's assets, which are not tradable and are

only partially observable. Also it has to deal with the often complex priority structure of a firm's liabilities.

The second approach assumes that on bankruptcy, the firm will pay off a pre-specified fraction of the risk-free value of the instrument where bankruptcy is again triggered when the firm's assets first reach some specified limit. This idea, first developed by Hull and White (1991) enables one to ignore the debt priority problem but still assumes knowledge of the asset value stochastic process. The Hull and White model assumed independence of the stochastic process giving the firm's asset value and the process giving the risk-free values of the instruments, which is essentially the process describing the interest rate movements. Longstaff and Schwartz (1995) generalized the model by allowing non-zero correlations between the two processes. Since in both these models bankruptcy occurs when the asset value, which is a continuous process, hits a pre-specified limit, firms never default unexpectedly. For this reason, Madan and Unal (1994) and Lando (1994), modeled the asset value as a jump process so that the firm's value can suddenly jump below the bankruptcy level.

The third approach ignores the asset value completely and again overcomes the debt liability structure by assuming that on bankruptcy a given fraction of each promised dollar is paid off. This approach assumes the bankruptcy process is specified exogenously and does not depend on the firm's underlying assets. (e.g. Jarrow and Turnbull (1995), Litterman and Iben (1991)). Lando's thesis (1994) was the first to use the evolution of the firm's credit rating as a model for the bankruptcy process. In the second essay in his thesis, he develops a continuous time Markov model in which he assumes the bankruptcy process and the process that gives the risk free bond prices are independent. This model and a discrete time equivalent model appear in the seminal paper by Jarrow et al (1997) but both assume independence between the interest rate process and the credit rating process. This assumption means their model can not take into account Duffee's point (Duffee 1996) that defaults are primarily driven by the business cycle, which derives variations in the financial variables on which most derivatives are priced. Benninga (Chapter 17, 1997) develops a similar spreadsheet model for finding the expected returns on a risky bond using the probabilities of default, the transition probabilities that the bond's credit rating will move from one level to another and the percentage recovery on the face value of the bond.

Our paper extends the model in the Jarrow, Lando, Turnbull (1997) paper in two ways. Firstly it allows dependency between the stochastic process describing interest rates and hence the risk-free bond's prices, and the stochastic process describing the movement in bond's credit

rating by linking them both to an underlying process which describes the state of the economy. The transition probabilities between states in the two processes (interest rates and credit rating) vary depending on the state of the economy process. This extension is Model 0 of section two.

Lando (1994) in the third essay of his thesis suggested a bond pricing model based on survival analysis where the probability of a bond defaulting is given by a process whose parameters can include interest rate information, although no empirical calculations are made using this model. The interest rate information affects the probability of the bond's credit rating changing but if the rating does change it does not affect the probability of which rating it will move to. Our model (Model 0) has a more indirect connection between interest rate and credit ratings but one that allows more general interactions. We also investigate how good the model's bond prices fit with empirical data.

Two different versions of this extension which allows connection between interest rates and credit rating changes are considered and one of them allows a re-interpretation of the risk premium ideas suggested by Jarrow, Lando and Turnbull (1997). In order to get their model to give prices for the risky bonds that agreed with actual values, they modified the historically derived transition matrix between ratings, P^A by making it a mixture of this and the identity matrix, namely $\pi(t) P^A + (1-\pi(t))I$. They interpret the $\pi(t)$ as risk premium but unfortunately some of the risk premium are negative and others are very large. An alternative interpretation is that this is an example of the mover-stayer Markov chain model (Frydman et al 1985) where there is a heterogeneous set of bonds, some of which will never change their ratings (the stayers) and others of which are moving around. The view that the ratings of bonds will never change is a very optimistic one because it guarantees that there will be no defaults. An equally extreme but pessimistic view is that all bonds of the given rating will default at the next period. Recently Kijima and Komoribayashi (1998) also identified that this might be a better choice of risk premium than that suggested by Jarrow et al (1997). Model 1S (S for subjective) of section two allows for the pricing of the bonds to reflect the market's relative belief between the historic movements of credit ratings and both these two extreme alternatives. The risk premium can then be interpreted as the belief the market puts on the extreme risky (or riskless) future scenarios.

Empirical work on bond pricing requires one to calculate the zero-coupon prices for risky bonds from the bonds in the market almost all of which have coupons. Longstaff and Schwartz (1992) take averages of bond prices and coupon rates and average maturity over a number of months and find the best regression fit. Jarrow et al (1997) takes the average price

and average coupon rate for each combination of bond rating and maturity period and then uses these values to solve a triangular system of equations to get zero coupon bond prices. However both methods can lead to mis-pricing. The zero coupon bond prices need not decrease with maturity or lower credit rating. One can overcome these difficulties and use the full details of each bond in the market rather than average values by using linear programming to find the zero-coupon bond prices which minimise the L_1 -average errors and ensure that no mis-pricing can occur.

In the next section, the bond pricing models which allow for dependency between credit rating movements and interest rate movements are introduced. Section three describes the linear programming method of ensuring no mis-pricing when stripping out the coupons from risky and riskless bonds. Section four describes the empirical data used to determine the parameters of the bond pricing models outlined in section two. The results in terms of bond prices and risk premium of using this data in the models of section two are also discussed. Section five draws some conclusion.

2. Models of bond prices

The bond prices are modeled as a discrete time trading economy both because discrete time simplifies the mathematics and because the credit rating information is summarized in discrete time formats (see Standard and Poor (1997a)). There are three interconnected processes which make up the model. Firstly the underlying economic conditions, E_t , are modeled as a discrete-time time homogenous Markov chain with two states $\{G(\text{Good}), B(\text{Bad})\}$. Let

$$g = \text{Prob}\{E_{t+1} = G \mid E_t = G\} \quad b = \text{Prob}\{E_{t+1} = B \mid E_t = B\} \quad (2.1)$$

The interest rate process over time periods $t = 0, 1, 2, \dots, T$ is a generalization of the lattice Markov chain model outlined in Pliska (Pliska, Ch.6, 1997). To be precise let I_t denote a stochastic process with initial value $I_0 = 0$ and state space $I = \{0, 1, \dots, T\}$. The transition probabilities satisfy.

$$\begin{aligned} P\{I_{t+1} = n+1 \mid I_t = n, E_t = G\} &= p_g(t, n) && \text{for } n \in I \\ P\{I_{t+1} = n \mid I_t = n, E_t = G\} &= 1 - p_g(t, n) && \text{for } n \in I \\ P\{I_{t+1} = n+1 \mid I_t = n, E_t = B\} &= p_b(t, n) && \text{for } n \in I \\ P\{I_{t+1} = n \mid I_t = n, E_t = B\} &= 1 - p_b(t, n) && \text{for } n \in I \end{aligned} \quad (2.2)$$

Although I_t is neither time homogeneous nor a Markov chain, the process (I_t, t, E_t) is both a Markov chain and time homogeneous. The process I_t gives knowledge of the spot interest rate r . If $I_t = n$, $E_t = G$, then the spot interest rate is $r_t(n, G)$ while if $I_t = n$, $E_t = B$, the interest rate is $r_t(n, B)$. Moreover it implies knowledge of future interest rates so that if $I_t = n$, $E_t = G$, spot interest rates next period will be one of $r_{t+1}(n+1, G)$, $r_{t+1}(n+1, B)$, $r_{t+1}(n, G)$ and $r_{t+1}(n, B)$. The transition probabilities defined in (2.2) are the conditional risk neutral transition probabilities for the process.

The third process R_t describes the evolution of the credit rating of the bond. Assume there are $M+1$ possible rating levels $(0, 1, 2, \dots, M)$ where 0 is the rating given to risk-free government bonds. Risky corporate bonds have a rating from 1 (most secure i.e. AAA in S and P) to $M-1$ (least secure C-grade in S and P), with M corresponding to bankruptcy. R_t is a discrete time process which is almost a Markov chain since the transition probabilities are defined by

$$\begin{aligned} P\{R_{t+1} = k \mid R_t = j, E_t = G\} &= p_{jk}^G(t) \\ P\{R_{t+1} = k \mid R_t = j, E_t = B\} &= p_{jk}^B(t) \end{aligned} \quad (2.3)$$

with $\sum p_{jk}^G(t) = \sum p_{jk}^B(t) = 1$ for all t .

Note that $p_{00}^G(t) = p_{00}^B(t) = 1$ and $p_{MM}^G(t) = p_{MM}^B(t) = 1$ for all times t . Thus (R_t, t, E_t) as well as (R_t, I_t, t, E_t) are finite state stationary Markov chains. Unlike Jarrow et al (1997) the rating process R_t and the interest rate process (I_t, t) are not now independent, but are related through their mutual dependency on the economic conditions process E_t . If we assumed E_t has only one possible state then this model reduces to the Jarrow model though no specific form of the interest process is used there. Taking there to be only one economic state for the process E_t reduces the interest rate process (I_t, t) to the lattice interest rate model detailed in Pliska (Pliska 1997). Having defined the evolution of the economic variables, it is now possible to define and calculate the bond prices in the model.

Let $Z_t^s(n, E, j)$ be the time t price of a zero-coupon bond promising to pay a dollar at time s when the bond rating is j at time t and the interest and economic conditions then are $I_t = n$ and $E_t = E$, where E is either G , B or some distribution of belief over the two possibilities. One feature of discrete time models is that several events occur in the same time period. One can choose arbitrarily what the order of these events will be. We assume that $Z_t^s(n, E, j)$ is the price of the bond at the beginning of period t , when the bond is redeemed at the end of period s . During any period, we assume all changes of state occur towards the end of the period after

the redemption date for that period, with first changes in interest rate I_t , then changes in rating R_t and finally changes in the economic condition E_t . If a company defaults, it is assumed that a fraction f of the face value of the bond will be repaid.

In order for the discounted zero coupon bond prices to be free of arbitrage opportunities, then they must be martingales, and so the price at any period must be the expected value of future bond prices under the risk neutral probabilities. Using the sequence of events within a period described above, this martingale requirement leads to the equations

$$Z_t^s(n, E, j) = \left(\frac{1}{1 + r_t(n, E)} \right) \sum_k \left[p_{jk}^E(t) \{ p_e(t, n) [e Z_{t+1}^s(n+1, E, k) + (1-e) Z_{t+1}^s(n+1, E^c, k)] \right. \\ \left. + (1 - p_e(t, n)) [e Z_{t+1}^s(n, E, k) + (1-e) Z_{t+1}^s(n, E^c, k)] \right] \\ \text{for } 0 \leq t \leq s \leq T, \quad 0 \leq n \leq t, \quad 0 \leq j \leq M-1 \quad (2.4)$$

where $E = G$ or B and if $E = G$, $e = g$ and $E^c = B$ while if $E = B$, $e = b$, $E^c = G$.

Also following the ordering of events within a period, defined above, one gets

$$Z_t^t(n, E, j) = \frac{1}{1 + r_t(n, E)} \text{ if } j \neq M \text{ and } Z_t^s(n, E, M) = f \quad \forall t, s, n, E. \quad (2.5)$$

At $t = 0$, assume $I_0 = 0$, and E_0 is either G , B or a distribution $(p, 1-p)$ over (G, B) . The prices of bonds at time $t = 0$ can be used to identify the prices of zero-coupon bonds $Z_0^s(0, E_0, j)$ by using methods discussed more fully in section three of this paper. Thus the model appears to have $3 + 2T(T+1) + 2T(M-1)(M-2)$ parameters $-g, b, f, 2T(T+1)$ parameters of the form $p_g(t, n), p_b(t, n), r_t(n, G), r_t(n, B)$, and $2T(M-1)(M-2)$ of the form $p_{jk}^G(t), p_{jk}^B(t)$ given $\sum_k p_{jk}^G(t) = \sum_k p_{jk}^B(t) = 1$.

Ideally the model will satisfy $TM+1$ constraints in that it should closely fit the zero coupon bond prices $Z_0^s(0, E_0, j)$, for $s = 1, \dots, T$, and $j = 0, M-1$ and at time $t = 0$ satisfies (2.5). Since there are more parameters than constraints one could expect to impose other conditions on the parameters. However there is less freedom than seems the case. If for example the transition matrices $p_{jk}^G(t)$ and $p_{jk}^B(t)$ are assumed to be stationary and given by past history, there should appear to be more than enough other parameters - $2T(T+1) + 3$ - to satisfy $(M+1)(T+1)$ conditions. However, the number of parameters $p_g(t, n), r_t(n, G)$ etc increases linearly with t , so there are only 3 parameters at $t=0$ to set the time-1 prices and only 8 parameters at $t = 1$ to set the time-2 prices. This is not enough to define the $M+1$ bond prices given for each t -time for the early t -times.

parameters $p_g(t,n)$, $r_t(n,G)$ etc. are more than enough to define the risk-free bond prices $Z_0^s(0, E_0, 0)$, and can help define good approximations to the risky bond prices $Z_0^s(0, E_0, j)$ $j \neq 0$. This reflects the fact that there are an infinite number of future stochastic evolutions of interest rates which give the current yield curve for riskless government bonds, but there is information in the yield structure of the risky bonds which helps define which evolution is being assumed by the market.

Model 1: Simple Model

One obvious way of simplifying the number of parameters in Model 0 is to assume the rating transitions are stationary and the interest rate transitions are state independent. Also, one can assume the underlying economic state of the system only affects the probability of changes in the interest rates and not the interest rate levels. This corresponds to keeping g , b , f as in model 0 and defining

$$\begin{aligned} p_g(t,n) &= p_g(t); \quad p_b(t,n) = p_b(t) \quad \forall n,t \\ (1 + r_t(n,G)) &= (1 + r_t(n,B)) = (1 + r_t(0))/c(t)^n \quad \forall n,t, j, k, t \end{aligned} \quad (2.6)$$

The $c(t)$ can be interpreted as measures of the volatility of the time t spot interest rates. One advantage of this model is that one can obtain the basic interest rate levels $r_t(0)$ as an analytic expression of the other parameters, g , b , $c(t)$, $p_g(t)$ and $p_b(t)$. Before proving this result note that the definition of $Z_t^s(n,E,0)$ and the assumption in (2.6) mean that we can define $z_t(n)$ by

$$z_t(n) = Z_t^t(n,E,0) = \frac{1}{1 + r_t(n,E)} = \frac{c(t)^n}{1 + r_t(0)} = c(t)^n z_t(0) \quad (2.7)$$

Also define the vector

$$\mathbf{Z}_t^s(\mathbf{n}, \mathbf{0}) = \begin{pmatrix} Z_t^s(n, G, 0) \\ Z_t^s(n, B, 0) \end{pmatrix} \quad (2.8)$$

Lemma

i): Define the following 2x2 matrices:

$$\begin{aligned} \mathbf{P}_t^s &= \begin{pmatrix} g(1-p_g(t) + p_g(t)c(t+1) \dots c(s)), & (1-g)(1-p_g(t) + p_g(t)c(t+1) \dots c(s)) \\ (1-b)(1-p_b(t) + p_b(t)c(t+1) \dots c(s)), & b(1-p_b(t) + p_b(t)c(t+1) \dots c(s)) \end{pmatrix} \end{aligned} \quad (2.9)$$

then

$$\mathbf{Z}_t^s(\mathbf{n}, \mathbf{0}) = \Pi_{j=t}^{s-1} \mathbf{P}_j^s \begin{pmatrix} 1 \\ 1 \end{pmatrix} z_j(\mathbf{n}) z_s(\mathbf{n}) \quad (2.10)$$

$$\text{ii)} \quad z_s(0) = \frac{1}{1 + r_s(0)} = \frac{Z_0^s(0, E_0, 0)}{Z_0^{s-1}(0, E_0, 0)} \frac{\mathbf{e}_0 \Pi_{j=0}^{s-2} \mathbf{P}_j^{s-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}}{\mathbf{e}_0 \Pi_{j=0}^{s-1} \mathbf{P}_j^s \begin{pmatrix} 1 \\ 1 \end{pmatrix}} \quad (2.11)$$

where $\mathbf{e}_0 = (1, 0)$ if $E_0 = G$ and $\mathbf{e}_0 = (0, 1)$ if $E_0 = B$.

Proof:

i) can be proved using backward induction on t in $Z_t^s(\mathbf{n}, E, 0)$ of (2.10) starting with $t = s$. We will concentrate on $Z_t^s(\mathbf{n}, G, 0)$ as the proof for the other component is the same.

For $t = s$, $Z_s^s(\mathbf{n}, G, 0) = \frac{1}{1 + r_s(\mathbf{n}, G)} = Z_s(\mathbf{n})$

Assume (2.10) is true for $Z_k^s(\mathbf{n}, 0)$ for $k \geq t+1$. By (2.4)

$$\begin{aligned} Z_t^s(\mathbf{n}, G, 0) &= z_t(\mathbf{n}) (p_g(t) (g Z_{t+1}^s(\mathbf{n}+1, G, 0) + (1-g) Z_{t+1}^s(\mathbf{n}+1, B, 0)) + (1-p_g(t)) (g Z_{t+1}^s(\mathbf{n}, G, 0) \\ &\quad + (1-g) Z_{t+1}^s(\mathbf{n}, B, 0))) \\ &= z_t(\mathbf{n}) \{ p_g(t) [g (1-0) (\prod_{j=t+1}^{s-1} \mathbf{P}_j^s \begin{pmatrix} 1 \\ 1 \end{pmatrix} z_j(\mathbf{n}+1)) z_s(\mathbf{n}+1) + (1-g) (0-1) (\prod_{j=t+1}^{s-1} \mathbf{P}_j^s \begin{pmatrix} 1 \\ 1 \end{pmatrix} z_j(\mathbf{n}+1)) z_s(\mathbf{n}+1)] \\ &\quad + (1-p_g(t)) [g (1-0) (\prod_{j=t+1}^{s-1} \mathbf{P}_j^s \begin{pmatrix} 1 \\ 1 \end{pmatrix} z_j(\mathbf{n})) z_s(\mathbf{n}) + (1-g) (0-1) (\prod_{j=t+1}^{s-1} \mathbf{P}_j^s \begin{pmatrix} 1 \\ 1 \end{pmatrix} z_j(\mathbf{n})) z_s(\mathbf{n})] \} \end{aligned} \quad (2.12)$$

Since $z_j(\mathbf{n}+1) = c(j) z_j(\mathbf{n})$ for $j = t+1, \dots, s-1$, (2.12) becomes

$$\begin{aligned} &(g(1-p_g(t) + p_g(t)c(t+1) \dots c(s)) - (1-g)(1-p_g(t) + p_g(t)c(t+1) \dots c(s))) z_t(\mathbf{n}) \Pi_{j=t+1}^{s-1} \mathbf{P}_j^s \begin{pmatrix} 1 \\ 1 \end{pmatrix} z_j(\mathbf{n}) z_s(\mathbf{n}) \\ &= (1-0) \Pi_{j=t+1}^{s-1} \mathbf{P}_j^s \begin{pmatrix} 1 \\ 1 \end{pmatrix} z_j(\mathbf{n}) z_s(\mathbf{n}) \end{aligned}$$

ii) From (2.10) we have that the ratio of $Z_0^s(0, E_0, 0)$ and $Z_0^{s-1}(0, E_0, 0)$ is

$$\begin{aligned} \frac{Z_0^s(0, E_0, 0)}{Z_0^{s-1}(0, E_0, 0)} &= \frac{\mathbf{e}_0 \prod_{j=0}^{s-1} \mathbf{P}_j^s \begin{pmatrix} 1 \\ 1 \end{pmatrix} z_j(0) z_s(0)}{\mathbf{e}_0 \prod_{j=0}^{s-2} \mathbf{P}_j^{s-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} z_j(0) z_{s-1}(0)} \\ &= z_s(0) \frac{e_0 \prod_{j=0}^{s-1} P_j^{s-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}}{e_0 \prod_{j=0}^{s-2} P_j^{s-2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}} \end{aligned}$$

Hence (2.11) follows.

The lemma implies that under model 1 once the 3T parameters $p_g(t)$, $p_b(t)$, $c(t)$, $t=0, \dots, T-1$ are given the parameters $r_i(0)$ can be chosen to ensure the zero-risk bond prices are met. We will assume the economy transition probabilities g and b are given by estimating from historic data. Thus all that remains to be fixed in the model are the rating transition matrices $p_{jk}^G(t)$ and $p_{jk}^B(t)$.

This paper investigates two different ways of defining the rating transition matrices.

Model 1H: Historical Data

In this the transitions are estimated from the actual transitions in ratings in the past and the transition matrices are assumed to be time-independent. Thus one defines the transition matrices $P^{AG}, (P^{AB})$ -actual good (actual bad) - from actual historic data so that

$$P_{jk}^G(t) = p_{jk}^{AG}, \quad p_{jk}^B(t) = p_{jk}^{AB}, \quad \forall t, 0 \leq j, k \leq M \quad (2.13)$$

It is clear that this model cannot hope to obtain all the risky bond prices completely accurately, as there are $T(M-1)$ risky bond prices and now only $3T + 1$ parameters available, namely $p_g(t)$, $p_b(t)$, $c(t)$, and f .

An alternative approach is to assume that the market does not accept that the historic movements in ratings are the ones that will occur in the future. The markets view is a mixture of beliefs, some based on historic movements, some on more extreme views of the movements. We consider two extreme positions: catastrophe and no change.

The catastrophe view (C) is that in the coming year all risky bonds of all rating will default. This corresponds to a transition matrix

$$P^C = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

The no change position (NC) is the lazy view that all bonds will keep the same rating in the coming year and corresponds to as transition matrix $\underline{P}^{NC} = \underline{I}$.

Model 1S – Subjective Ratings

Assume that in the good'time periods, the market takes as the rating transition matrix, a mixture of the historic rating changes in good times and the extreme view that there will be no changes. In the bad'time periods, the market's view is a mix of the historic rating changes in bad periods and the extreme view that all bonds will default. We assume that the ratio of the mixture can differ for different bond ratings and for different time periods. This leads to the definition:

$$\begin{aligned} P_{jk}^G(t) &= \pi^G(j,t) p_{jk}^{NC} + (1-\pi^G(j,t)) p_{jk}^{AG} \\ P_{jk}^B(t) &= \pi^B(j,t) p_{jk}^C + (1-\pi^B(j,t)) p_{jk}^{AB} . \end{aligned} \quad (2.14)$$

The values $\pi^G(j,t)$, $\pi^B(j,t)$ could then be considered a type of risk premium measure. In this case, it is how much weight the market puts on the extreme view of the future. In fact, one can reinterpret the risk premium which Jarrow et al (1997) introduced into their paper as $1-\pi^G(j,t)$ in this formulation if one assumes the only underlying state is G. This may explain why they end up with negative risk premium in their calculation. They have only allowed for the market to have a more optimistic view than the historic one of the future. This reinterpretation however only makes sense if $\pi^G(j,t) \leq 1$ whereas Jarrow et al (1997) allow values greater than 1. Model 1S seeks to allow the market to have both a more optimistic and a more pessimistic view of the future than was the historic average and for simplicity restricts optimistic views to good years and pessimistic views to bad years..

3. Using Linear Programming to strip out coupons

Models of bond prices take zero-coupon bonds as their basic entity, whereas most bonds have coupons which involve part payments during the life of the bond, as well as the redemption value to be paid on maturity. Thus there is a need to strip out the coupons and calculate what the market price of the bond implies about the value of a bond that will just pay 1 unit at time t. Some authors (Longstaff, Schwartz (1995)) take the average bond price, coupon rate and maturity each month for over a given time period and fit a regression line. The data however will include the changes over time in market sentiment and so does not reflect the position at a given time. Jarrow, Lando and Turnbull (1997) split bonds into classes

depending on their credit rating and their maturity. For each class the average market price and average yield were taken to be the values for bonds of that rating and maturity. Solving a triangular system of equations gave the zero-coupon bond prices. However, there was some mis-pricing of their bonds with their calculated zero-coupon bond prices not necessarily increasing as the credit rating improved nor decreasing as maturity increased. Alderson and Zivney (1994) report similar examples of mis-pricing in junk bonds and they show that reported bond yields depend on which investment strategies are assumed.

One can set up the problem of stripping out the coupons to get zero-coupon bond prices from bonds with coupons as a linear program, in the following way. Assume bonds are given one of M credit ratings $0, 1, 2, \dots, M-1$ (credit rating M corresponds to default). Suppose there are N bonds and all have maturity and coupon payments within the next T periods. Assume bond i $1 \leq i \leq N$ has a current market price of p_i , a credit rating of $d(i)$ and the coupons and redemption payments involve a payment of $c_i(t)$ in period t , $t = 1, \dots, T$. Let $R_j = \{i \mid d(i) = j\}$ $j = 0, \dots, M-1$ be the set of bonds with rating j . Let the present value of a zero-coupon, j -rated bond which pays one unit at time t , be $v_j(t)$, $j = 0, \dots, M-1$, $t = 1, \dots, T$.

Ideally one has

$$p_i = \sum_{t=1}^T c_i(t) v_{d(i)}(t) \quad \forall i = 1 \dots N \quad (3.1)$$

However, one cannot guarantee this will occur, so instead one requires

$$p_i + a_i = \sum_{t=1}^T c_i(t) v_{d(i)}(t) + b_i \quad (3.2)$$

where $a_i \geq 0$, $b_i \geq 0$ are the 'above' or 'below' errors in the market price.

Thus one can find $v_j(t)$ $j = 0, \dots, M-1$, $t = 1, \dots, T$ by solving the following linear program, LP1.

$$\text{Minimise } \sum_{i=1}^N (a_i + b_i)$$

$$p_i + a_i = \sum_{t=1}^T c_i(t) v_{d(i)}(t) + b_i \quad i = 1, \dots, N$$

$$v_j(t) \geq (1 + m(t)) v_j(t+1) \quad j = 0, \dots, M-1, t = 1, \dots, T-1 \quad (3.3)$$

$$v_j(t) \geq v_{j+1}(t) \quad j = 0, \dots, M-2, t = 1, \dots, T \quad (3.4)$$

$$a_i \geq 0 \quad b_i \geq 0 \quad v_j(t) \geq 0 \quad i = 1, \dots, N, j = 0, \dots, M-1, t = 1, \dots, T.$$

where $m(t)$ is the minimum possible interest rate in period t .

(3.3) guarantees that the zero-bond prices satisfy the obvious financial maturity requirements. If $m(t) = 0$, (3.3) reduces to the requirement that the price of bonds decreases with increasing maturity. Conditions (3.4) ensure that there is no mis-pricing on credit ratings so the bonds with the best (lowest) credit rating have the highest prices.

The linear program has $MT+2N$ variables and $N+M(T-1) + (M-1)T$ constraints. Such programs can be solved by dedicated linear programming solvers or by solvers in spreadsheet packages such as EXCEL. The solver in EXCEL 97 can only deal with 200 variables, so only less than 100 bonds can be dealt with at a time.. However, one can solve the problem as a set of nested linear programs if one has more than 100 bonds, by solving first the prices of the 0-rated bonds only, dropping constraint (3.4) and applying (3.2) only to $i \in R_0$. Then solve for the $v_1(t)$ only replacing (3.4) by $v_1(t) \leq v_0(t)$, $t=1, \dots, T$ where $v_0(t)$ was obtained from the previous linear program and (3.2) only holds for $i \in R_1$. Repeating this procedure for all the ratings in turn corresponds to solving LP1 but with an objective function

$$\text{Minimise } \sum_{j=0}^{M-1} \sum_{i \in R_j} L_j (a_i + b_i)$$

where L_i is an order of magnitude greater than L_{i+1} for $i=0, \dots, M-2$.

The linear program LP1 gives the prices that best fit the actual bond data in the sense of minimising the average absolute error, while also ensuring there is no mis-pricing.

4. Data and results of an example using US bond prices

The models of section two and three were applied to data on US bond prices and credit ratings obtained from DATASTREAM and Standard and Poor (Standard and Poor 1997a, 1997b) respectively. 64 of the US Treasury Bonds which make up the DATASTREAM US yield curve data set in 1995 and 1996 were taken as the riskless bonds. Their market price on 3rd July 1996 was taken -the data being chosen as an example of a mid-week, mid-year, pre-holiday period. The set of risky bonds satisfied three criteria. They were in the DATASTREAM database of US industrial and US financial bonds; their market prices and S&P rating for 3rd July 1996 were available; there were no callable dates. The extra option that being callable gives a bond is more difficult to strip out of the price than the coupons. There were 178 such bonds in total (7 rated AAA, 24 rated AA, 61 rated A, 68 rated BBB, 12 rated BB, 6 rated B). DATASTREAM does not usually record the prices of C-rated speculative

bonds but there were 8 bonds in the set that moved from C to investment grade or vice versa during the year (Standard and Poor 1997b) and hence we were able to obtain the 1996 market price when they were C-rated.

TABLE 1 Zero-coupon bond prices

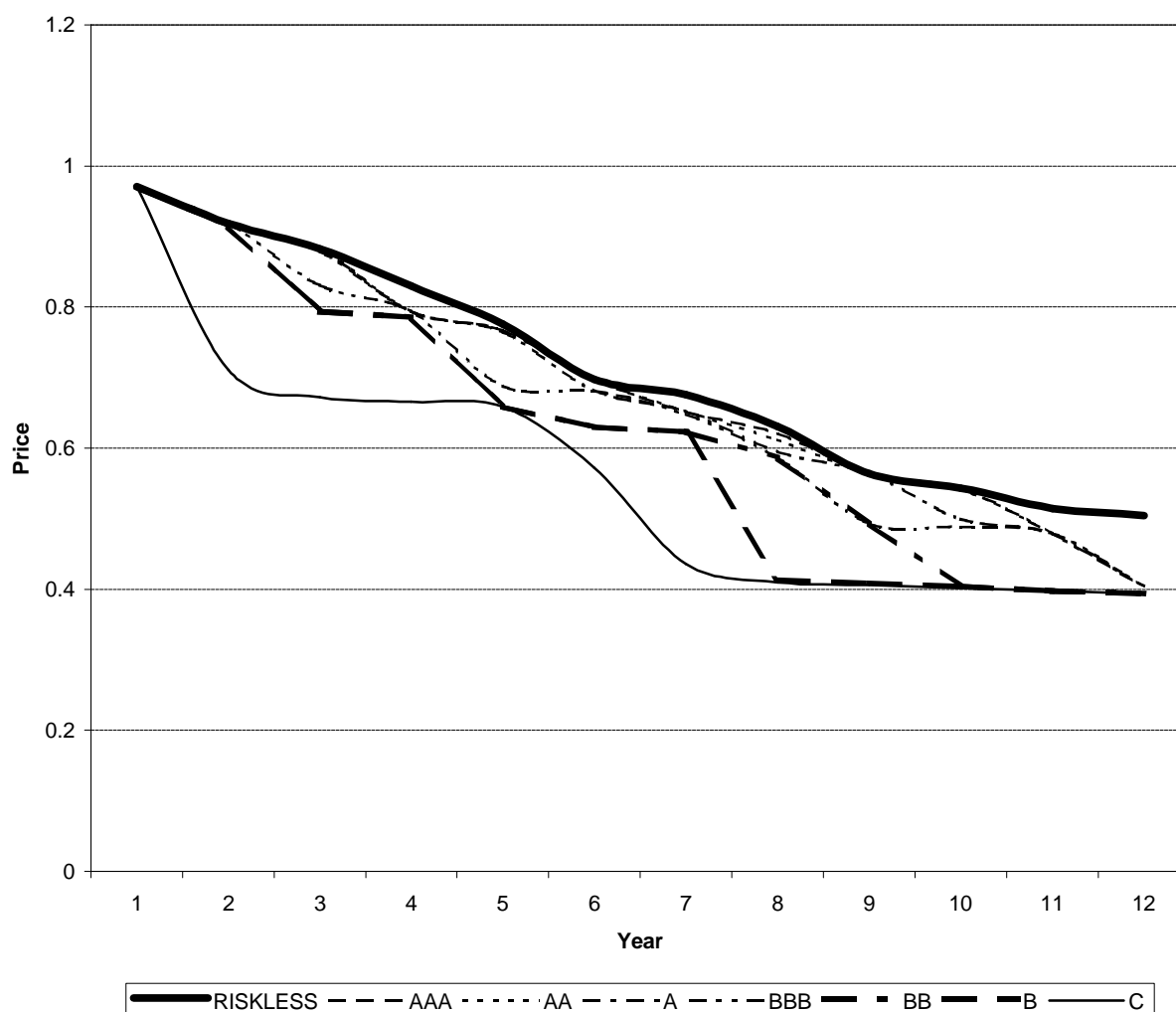
Year	1996	1997	1998	1999	2000	2001	2002	2003	2004
RISKLESS	0.9713	0.9187	0.8827	0.8300	0.7760	0.6979	0.6754	0.6305	0.5638
AAA	0.9713	0.9187	0.8821	0.7939	0.7667	0.6979	0.6521	0.6208	0.5638
AA	0.9713	0.9187	0.8821	0.7939	0.7667	0.6979	0.6521	0.6117	0.5638
A	0.9713	0.9187	0.8786	0.7939	0.7645	0.6814	0.6521	0.5949	0.5638
BBB	0.9713	0.9187	0.8309	0.7939	0.6879	0.6810	0.6476	0.5865	0.4931
BB	0.9713	0.9147	0.7936	0.7857	0.6590	0.6296	0.6233	0.5865	0.4931
B	0.9713	0.9147	0.7936	0.7857	0.6590	0.6296	0.6233	0.4126	0.4085
C	0.9713	0.7110	0.6723	0.6656	0.6590	0.5722	0.4362	0.4098	0.4058
Year	2005	2006	2007	2008	2009	2010	2011to20	2021+	
RISKLESS	0.5435	0.5147	0.5046	0.4947	0.4850	0.4755	0.2317	0.1402	
AAA	0.5419	0.4789	0.4039	0.3999	0.3959	0.3920	0.2317	0.1402	
AA	0.5419	0.4789	0.4039	0.3999	0.3959	0.3920	0.2317	0.1259	
A	0.4994	0.4789	0.4039	0.3999	0.3467	0.3433	0.2087	0.1259	
BBB	0.4882	0.4789	0.4039	0.3825	0.3467	0.3387	0.2087	0.1259	
BB	0.4045	0.3978	0.3938	0.3825	0.3467	0.3387	0.2087	0.1259	
B	0.4045	0.3978	0.3938	0.3825	0.3421	0.3387	0.2087	NA	
C	0.4017	0.3978	0.3938	NA	NA	NA	NA	NA	

The linear program developed in section three was applied to the 250 bonds in total, and the zero-coupon prices for each individual year 1996 to 2010, one common price for years 2011 to 2020, and one common price for all years beyond 2020 were calculated. It was assumed that there would be no bankruptcies in the rest of 1996 so $v_j(1996)$ was assumed constant for all ratings j . The results are given in table 1 where there was no price available for C-rated zero-coupon prices beyond 2007 because there were no such bonds with maturity beyond this date.

The yield curves for the various rated bonds are given in Figure 1.

Figure 2 graphs the spreads for risky bonds for their yields compared with those riskless bonds. The curves are not smooth and the spreads of differently rated bonds converge and then separate at several points but the general shape seems reasonable. The highest spread is for the C-rated bonds of early maturity and in general the spreads for these bonds decrease with time. The lowest spreads are for the AAA-rated bonds of early maturity and these spreads slowly increase with time. Note that the linear programme derives the same values for the price of C-rated bonds with a greater than twelve year maturity as for B-rated bonds of the same maturity because there were no examples of the former for the linear programme to use.

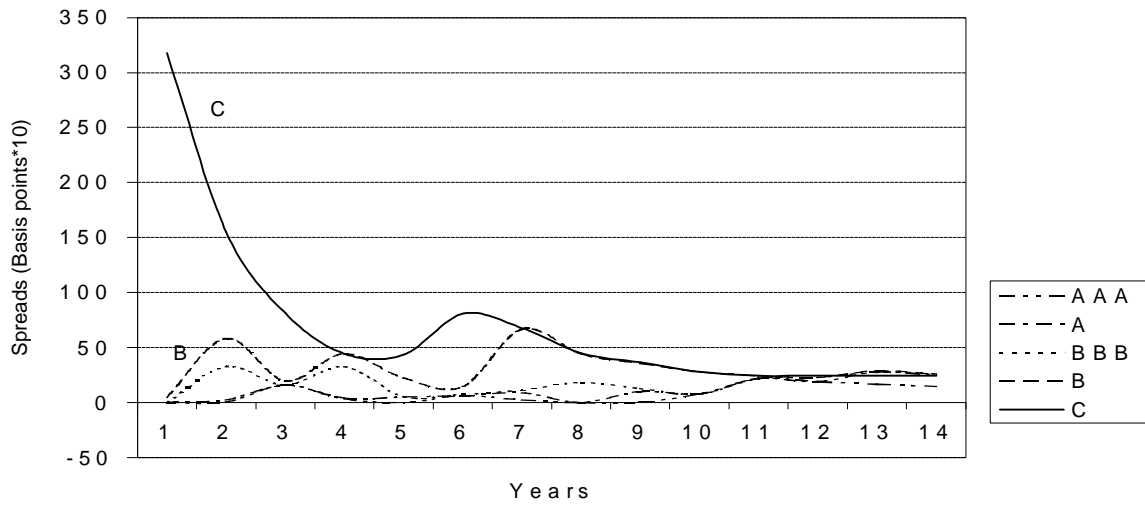
Figure 1: Bond Prices from Linear Programming



The data for the rating process was obtained from the Standard and Poor Rating Performance (S&P 1997a) which gives the number of bonds making each possible annual ratings transition for the years 1981-1996. The decision on which were the 'good' and 'bad'

years in the underlying process used two sets of data. Firstly the annual rating transitions were investigated and % downgradings + % defaults - % upgradings taken as a measure of the ratings changes in that year. The years were then ordered according to this measure and the highest rated were taken as good. Secondly, after examining Dow Jones Index long-term and short-term interest rates, US unemployment data, US CPS Industry Production, leading indicators and yield spreads, we classified the years as good or bad. With one change, the two

Fig2:Spreads from LP prices as of 1996



sequences agreed with one another with the ratings sequence lagging one year behind the subjective economic sequence. This lagged economic sequence was thus used and 81, 84, 87, 92, 93, 94, 96 were classified as good years and 82, 83, 85, 86, 88, 89, 90, 91, 95 classified as bad. Totalling the annual bond ratings changes for these two sequences separately and then translating into percentages led to the annual transition matrices for good and bad years respectively given in table 2. In this table the rows represent the bond rating at the beginning of a year and the columns represent the bond rating at the beginning of the next year, while the values are the probability of such a transition in good and bad years.

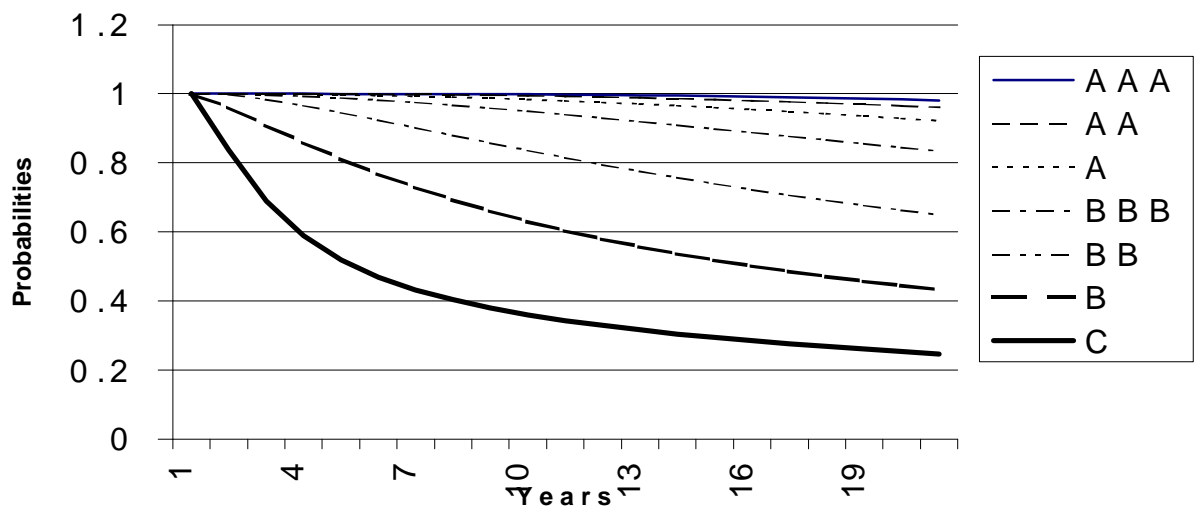
TABLE 2: Credit ratings transition matrices for good and bad years

Good	AAA	AA	A	BBB	BB	B	C	DEFAULT
AAA	0.898588	0.09371	0.006418	0.001284	0	0	0	0
AA	0.007007	0.923331	0.065952	0.00371	0	0	0	0
A	0.000448	0.021281	0.933692	0.039651	0.002912	0.001568	0.000224	0.000224
BBB	0.000351	0.001405	0.053057	0.896697	0.040408	0.007027	0.000351	0.000703
BB	0.000491	0.001473	0.004912	0.078094	0.828094	0.079077	0.005403	0.002456
B	0	0	0.004162	0.005945	0.087396	0.837099	0.026754	0.038644
C	0	0	0.004651	0.004651	0.027907	0.15814	0.64186	0.162791
Bad	AAA	AA	A	BBB	BB	B	C	DEFAULT
AAA	0.915984	0.07377	0.008197	0	0.002049	0	0	0
AA	0.006141	0.896622	0.085977	0.007506	0.001024	0.002388	0.000341	0
A	0.00119	0.026775	0.896271	0.06188	0.010115	0.002975	0	0.000793
BBB	0.000323	0.004523	0.063974	0.853958	0.058158	0.014216	0.001939	0.002908
BB	0	0.000936	0.007491	0.072566	0.802903	0.084738	0.013577	0.01779
B	0	0.001665	0.001249	0.003331	0.054538	0.822648	0.04746	0.069109
C	0.003401	0	0.006803	0.017007	0.027211	0.078231	0.598639	0.268707

One can check using χ^2 tests that the transitions for a given row are significantly different in the two matrices.

Looking at the pattern of good and bad years in the sequence which gives these transition matrices and adding in years 1980 and 1997 both of which are classified as good years enables us to estimate g and b . There are 5 times a bad year is followed by a bad year and 4 times it is followed by a good year. Good years were followed by good years 4 times and by bad years 4 times. This data leads to the estimates $g=4/8$ and $b=5/9$. We assume that in mid-1996 when the bond prices are taken it is not yet clear if the economic conditions in 1996 are

Fig 3: Survival Probabilities for 2-state model starting in Good State



good or bad and so we assume that since 1995 was a bad year the chance 1996 is bad is $5/9$.

Hence we assume $E_0=(4/9, 5/9)$. and the price of j -rated t -maturity zero-coupon bonds in 1996 is taken as $4/9Z_0^t(0,G,j)+ 5/9Z_0^t(0,B,j)$.

The effect on the rating transitions of a hidden underlying 2-state model of the economy can be shown by looking at the survival probabilities (i.e. the probabilities of not defaulting) after t periods for bonds rated j now. Figure 3 shows the results of this for the 2-state model starting in a good year .One can do similar calculations for both the 2-state model starting in a bad year and the 1-state model where one calculates the transition matrix from all years put together The one-state survival probabilities all lie below the two-sate survival probabilities when the current year is assumed good and are all above the two-state survival probabilities

when the current year is bad, The largest differences in the three graphs occur in the early years of B and C rated bonds. Of course eventually the survival probabilities will become 0 for all ratings in both models since in both cases the ratings process is an absorbing Markov chain.

All the parameters in the ratings and the underling economic processes have been defined above from historical data. The only parameters left to be defined in Model 1H are those describing the interest rate process $-c(t)$, $p_g(t)$, $p_b(t)$ for all t up to T - and f , the fraction of the face value repaid if a bond is defaulted. We will concentrate on the time interval 1996 2006, so seek to build a model that matches the July 3rd 1996 bond prices for zero-coupon bonds maturity at the end of each of these 11 years. Our ordering of the events during a year means that no bond will default during the rest of 1996 and hence the price of bonds maturing in this period is the same for all credit ratings. This follows from (2.5) and the discussion preceding it.

One possible approach is to choose reasonable values for the interest rate parameters and find the value of f that gives the best match with the prices obtained in Table 2. As an example we chose $c(t)=1$, $p_g(t)=0.6$, $p_b(t)=0.4$ for all t and then find the value of f which minimises the mean square error (MSE) over the 8 different ratings (T-bill, AAA, AA, A, BBB, BB, B, C) and the 11-years of the models zero-coupon bond prices compared with the actual bond prices. The best value is $f=0.3631$ with error 0.001200 and the implied risk free interest rates and bond prices are given in Table 3.

TABLE 3: Parameter values for Model 1H with optimal f and chosen c, p .

	96	97	98	99	2000	2001	2002	2003	2004	2005	2006
$c(t)$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$p_g(t)$	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60
$p_b(t)$	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40
$r(0)\%$	2.95	5.73	4.07	6.35	6.96	11.19	3.33	7.11	11.84	3.73	5.60
$g=$	0.50	$b=$	0.56	$f=$	0.36	MSE= 0.001200					

One can seek a better fit by optimizing over the interest rate parameters as well as f . There are limits on the parameters -the $p_g(t)$, $p_b(t)$ must be probabilities and $c(t)$ must be less than or equal to 1 in order that the actual interest rates $r_t(n, E)$ are monotonically increasing in n , which ensures that the interest rates $r_t(n, E)$ reflect the underlying ordering in the interest rate space I . To avoid the modeling partially collapsing to a deterministic one we will in fact impose $0.05 \leq p_g(t), p_b(t) \leq 0.95$. and to ensure volatility is not too great we require $c(t) \geq 0.5$.

Finally to ensure consistency with the assumptions underlying the derivation of the real zero-coupon prices we will assume the risk-free spot interest rates in all years are at least 1%. Using the non-linear solver in Excel gives the parameter values in Table 4 as the ones that minimise MSE with a MSE value of 0.000856.

TABLE 4: Parameter values for model 1H with optimal f , c and p .

	96	97	98	99	2000	2001	2002	2003	2004	2005	2006
$c(t)$	1.00	0.97	0.97	1.00	1.00	1.00	0.98	0.96	0.98	0.99	1.00
$p_g(t)$	0.14	0.95	0.05	0.05	0.05	0.05	0.05	0.95	0.21	0.60	0.61
$p_b(t)$	0.95	0.92	0.05	0.05	0.05	0.05	0.05	0.28	0.05	0.35	0.46
$r(0)\%$	2.95	5.31	1.00	6.35	6.96	11.19	1.00	1.00	7.12	1.00	5.60
$g=$	0.50	$b=$	0.56	$f=$	0.4387	MSE		0.000856			

The results of table 4 suggest a model where the interest rates are expected to rise considerably in the years 1999-2001 and in all years except 96 the chance of interest rates rising is higher if it is good year than if it is a bad year.

In section two it was pointed out that one could not match all the bond prices with the historical model 1H but a nearer fit may be possible if one allowed subjective views of the future ratings transitions as suggested in Model 1S. Initially one would expect that with the extra flexibility that the risk premiums parameters, $\pi^G(j,t)$, $\pi^B(j,t)$ give one could match the real prices exactly. The time 1-risk premiums $\pi^G(j,1)$, $\pi^B(j,1)$ can be used to match the time-1 bond prices $Z_0^1(0, E_0, j)$, then the time-2 risk premiums can be defined to get the time-2 maturity bonds and so on. In each case the bond price is a linear function of the corresponding risk premium and so the solution can be obtained by solving linear equations or by linear programming. However since (2.5) implies the 0-maturity bond prices are the same for all ratings, the risk premium $\pi^G(j,t)$ in the good state has no effect if it is defined in this way unless there is a chance there is an immediate transition to the default state M. From Table 2 one can see that there is no such chance of default for AAA and AA bonds. A second problem is the stability of such a calculation. The time-1 risk premiums are set by the time-1 bond prices but they in turn are a factor in all the longer maturity bonds. Any error in the time-1 bond prices is then reflected in the time-1 premium and the time-2 premium have to correct for these if they want to match the time 2 bond prices. Thus any errors grow as the risk premiums seek to compensate for errors in earlier risk premiums.

TABLE 5: risk premium using interest rate date of table 3

π^G	96	97	98	99	0	1	2	3	4	5
AAA	0	0	0	0.0006	0.0237	0.0270	0.0195	0.0159	0	0
AA	0.0185	0	0	0	0	0	0	0.2503	0	0
A	0.1026	0	0	0	0	0	0	0.2513	0	0
BBB	0	0	0	0	0	0	0	0	0	0
BB	0.2616	0.2600	0.2645	0.2714	0.3008	0.0353	0	0	0.0940	0
B	0	0	0	0	0	0	0	0.5166	0.9802	1
C	0.1647	0.5050	0.7201	0.4875	0.6492	0.4439	0.3853	0.3836	1.0000	1
π^B										
AAA	0.0023	0.0015	0.0673	0	0	0	0	0	0	0.148228
AA	0.0091	0	0.0593	0	0	0	0	0	0	0.239304
A	0.0167	0	0.0784	0	0	0.0142	0.0767	0	0	0
BBB	0	0.1578	0	0.0214	0	0	0	0.0638	0	0
BB	0	0.2378	0	0	0	0	0.3763	0.2638	0	0
B	0	0	0	0	0	0	0.7824	0.0000	0	0
C	0.1407	0	0	0	0	0	0	0	0	0
c(t)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1
.p _g (t).	0.6000	0.6000	0.6000	0.6000	0.6000	0.6000	0.6000	0.6000	0.6000	0.6
.p _b (t).	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4
r(0)%	2.9548	5.7256	4.0726	6.3550	6.9631	11.1916	3.3336	7.1093	11.8366	3.731883
g=	0.5000	B=	0.5556	f=	0.3631	MSE=	0.000565			

An alternative is to give up the advantage of linearity and use all the risk premia to match all the bond prices in one go. Since bond prices of maturity t depend on products of the risk premium for all times up to t this is a non-linear problem. One can solve the problem using non-linear algorithms including the ones in Excel. This is the approach we adopt here. Table 5 shows the risk premia that arise if one uses the data for the interest rate process given by Table 3. Using risk premia reduces the Mean Square Error over the 88 prices to 0.000565.

Table 6 shows the results when the data of table 4 which was the interest rate parameters that minimized mean square error the most were used. In this case the adding of risk premium brings the MSE down from 0.0008564 to 0.000533.

TABLE 6: Risk premium using risk data of Table 4

π^G	96	97	98	99	0	1	2	3	4	5
AAA	0.0002	0.0002	0.0003	0.0571	0.0767	0.0715	0.0622	0.1823	0	0

AA	0.3685	0.0418	0.0001	0.0001	0	0	0.0001	0.2501	0	0
A	0.1730	0	0	0	0	0	0	0.2513	0	0
BBB	0	0	0	0	0.0335	0	0	0	0.0043	0
BB	0.2623	0.2623	0.2667	0.2764	0.1682	0	0	0	0.1526	0.0050
B	0	0	0	0	0	0	0	0.5940	1.0000	1.0000
C	0.1675	0.5462	0.7365	0.4559	0.6651	0.4490	0.7031	0.3953	1.0000	1.0000
π^B										
AAA	0	0	0	0	0	0	0	0	0	0.0045
AA	0	0	0	0	0	0	0	0	0	0
A	0	0	0	0	0	0	0	0	0	0
BBB	0	0.0883	0	0	0	0	0	0	0	0
BB	0	0.1799	0	0.0582	0	0	0.1545	0.0392	0	0
B	0	0	0	0	0	0	0.7839	0	0	0
C	0.2724	0	0	0	0	0	0	0	0	0
c(t)	1.0000	0.9725	0.9725	1.0000	1.0000	1.0000	0.9822	0.9561	0.9774	0.9867
$\cdot p_g(t)$	0.1446	0.9500	0.0500	0.0500	0.0500	0.0500	0.0500	0.9500	0.2125	0.6000
$\cdot p_b(t)$	0.9500	0.9234	0.0500	0.0500	0.0500	0.0500	0.0500	0.2802	0.0500	0.3512
r(0)%	2.9548	5.3055	1.0000	6.3550	6.9631	11.1916	1.0000	1.0000	7.1243	1.0000
g=	0.5000	B=	0.5556	f=	0.4387	MSE= 0.000533				

Instead of first finding the interest rate data that best fits the bond price structure and then finding the best risk premium for this interest rate data, one could seek to optimize over interest rate data and the risk premium at the same time to try and find a good fit to the bond price structure. Table 7 gives the results of doing exactly that and leads to a fit where the Mean Square Error is 0.000198.

Most of this error - the total square error over the 88 bond prices is 0.01744 - is in the fitting of the B and C rated bonds. There were not many of these in the original sample and their zero-coupon prices are the most suspect since they are not underpinned by prices of lower rated bonds. The parameters can be chosen so that the total square error over the 66 BB and higher rated bond prices is 0.002739 which corresponds to a mean square error of 0.000034.

TABLE 7: Risk premia when optimizing over both risk and interest rate parameters

π^G	96	97	98	99	0	1	2	3	4	5
AAA	0	0	0	0.0111	0.0309	0.0265	0.0109	0.0348	0	0

AA	0.4139	0	0	0	0	0	0.5266	0	0
A	0.2771	0.2906	0.0135	0	0	0	0.2518	0.0129	0
BBB	0	0	0	0	0	0	0	0.0156	0
BB	0.2600	0.2614	0.2607	0.2640	0.2701	0.2885	0.0067	0.1982	0.1513
B	0	0	0	0	0	0	0.5230	1.0000	1.0000
C	0.2876	0.5665	0.9773	1.0000	1.0000	0.0000	1.0000	0.2201	1.0000
π^B									
AAA	0.0371	0.0604	0.2568	0	0.1554	0.2114	0.0299	0.2438	0.0186
AA	0.0333	0.0610	0.2761	0	0.1098	0.2451	0	0	0.7921
A	0.0510	0.0444	0.2742	0.0295	0.1862	0.1455	0.5496	0	0
BBB	0	0.3284	0	0.3350	0	0	0	0.4111	0
BB	0	0.4091	0	0.4530	0	0	1.0000	1.0000	1.0000
B	0	0.2138	0	0.3097	0	0	1.0000	1.0000	1.0000
C	0.6006	0	0	0	0	1.0000	1.0000	0.3708	0.6164
c(t)	1.0000	0.9810	0.9704	0.9517	0.9485	0.9177	0.9819	0.9659	0.9477
$\cdot p_g(t)$	0.9500	0.0500	0.0500	0.0500	0.0500	0.2495	0.9161	0.4390	0.0500
$\cdot p_b(t)$	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.2620
r(0)%	2.9548	3.8136	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
g=	0.5000	b=	0.5556	f=	0.5072	MSE= 0.000198			

Comparing the risk premia in Tables 5,6 and 7 the only noticeable feature in table 5 , where all but one of the interest rate parameters were fixed is that the premia suggest the market is overly optimistic about the survival of C-rated bonds. This over-optimism of the market is much more marked in Table 6 where the interest rate parameters are those that best fitted the bond prices. 38 of the 70 possible $\pi^G(j,t)$ are non-zero including the C-rated ones for all times t while only 8 of the 70 $\pi^B(j,t)$ are non-zero. So on balance the prices reflect a market that is much more likely to accept there will be no change in bond ratings than one that is worrying that they will default. The interest rate parameters show the market expects a large rise in interest rate in the 1999-2001 period. When as in Table 7, one allows both the interest rates and the risk premium to be moving at the same time to find a best fit to the prices, one gets a fit where the error decreases by 60%. However what happens is that the risk premium seek to describe not just the spread between the differently rated bonds but also the term structure of all bonds. Thus the interest rate parameters in this case suggest an interest rate structure that is essentially deterministic and flat. All the uncertainty in it has been translated into a much more complex risk premium structure.

One can recover the model with only one underlying economic state (which is akin to the Jarrow model (Jarrow et al 1997)) by setting g=1 and starting in state $E_0 = G$. The results of doing this and finding the best fit to the bond prices over f, c(t) and $p_g(t)$, $p_b(t)$ are given in

Table 8. Comparing the results with the comparable 2-state model in Table 4 one finds the MSE is now 0.001231 whereas the 2-state model has a MSE of 0.000856 which is 30% lower.

TABLE 8: Parameter values for 1-state model with optimal f , c and p

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
$c(t)$	1	1	1	1	1	0.97	0.99	0.98	0.97	0.99	0.99
$p_g(t)$	0.95	0.95	0.95	0.57	0.13	0.05	0.05	0.28	0.56	0.59	0.6
$r(0)\%$	2.95	5.73	4.07	6.36	6.96	1	1	1	1	1	1
$g=$	1										
	$f=$ 0.40669					MSE= 0.001231					

The improvement the 2-state model makes over the 1-state model is even more marked in the results with risk premia, where we follow the Jarrow model and only allow the more optimistic extreme view in the good state. π^G is the probability in our model that the market is taking the no change view in the good state. Table 9 gives the risk premia values and all the other parameters which best fit the bond prices in the 1-state model obtained by taking $g=1$. In this case the mean square error (MSE) of the 1-state model is 0.001003 while that for the 2-state model is 0.000198 – a cut in the error of 80%.

TABLE 9: Risk premia for 1-state model optimizing over all parameters

Year	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
AAA	0	0	0	0	0	0	0	0	0	0
AA	0	0	0	0	0	0.0002	0	0	0	0
A	0	0	0	0	0	0	0.2576	0	0	0
BBB	0	0	0	0	0	0	0	0	0	0
BB	0	0	0	0	0	0	0	0	0	0
B	0.4040	0	0	0	0	0	0	0	0	0
C	0	0.0580	0.4239	0.1446	0.2304	0	0	0	0	0
$c(t)$	1	1	1	1	1	0.9541	0.9891	0.9728	0.9541	0.9880
$p_g(t)$	0.05	0.05	0.95	0.9497	0.05	0.05	0.05	0.05	0.05	0.05
$r(0)\%$	2.95	5.73	4.07	6.35	6.96	1.00	1.00	1.00	1.00	1.00
$g=$	1									
	$F=$ 0.407538					MSE= 0.001003				

5. Conclusions

The previous sections develop a hidden Markov chain model for the term structure of credit risk spreads further extending the ideas in Lando (19994), Jarrow and Turnbull (1995) and Jarrow, Lando and Turnbull (1997). This model allows dependency between the rating process and the interest rate process through their joint dependency on a state of the economy process. The paper also provides a reinterpretation of the idea of risk premia introduced therein as the chance the markets view of the rating changes is more extreme than has been the case in the past. The paper also uses linear programming to provide a way of stripping the coupons for bonds in such a way as to minimise the mean absolute errors and at the same time ensure there is no mis-pricing of the zero-coupon bond prices

References

- Alderson M.J., Zivney T.L., (1994), On computing bond returns –the evaluation of low-grade debt. *J. Financial Research*, 17, 403-415
- Benninga S., (1997), *Financial Modeling*, MIT Press, Cambridge, Mass.
- Black F., Cox J., (1976), Valuing corporate securities; some effects of bond indenture provisions, *J. Finance*, 31, 351-367.
- Duffee G.R., (1996), On measuring credit risks of derivative instruments, *J. Banking and Finance*, 20, 805-833
- Frydman H., Kallberg J.G., Kao D-L, (1985), Testing the adequacy of Markov chains and mover-stayer models as representations of credit behavior, *Operations Research* 33, 1203-1214.
- Hull J., White A., (1991), The Impact of default risk on option prices, Working paper, University of Toronto
- Jarrow R.A., Lando D, Turnbull S.M., (1997) A Markov model for the term structure of credit risk spreads, *Review of Financial Studies*, 10, 481-523.
- Jarrow R.A., Turnbull S., (1995), Pricing derivatives on financial securities subject to credit risk, *J. Finance*, 50, 53-86.
- Kijima M., Komoribayashi K., (1998), A Markov chain model for valuing credit risk derivatives, *Journal of Derivatives*, 5, 97-108.
- Kim I., Ramaswamy K., Sundaresan S., (1993), Does default risk in the coupons affect the valuation of corporate bonds? A contingent claims model, *Financial Management*, 22, 117-131
- Lando D., (1994), Three Essays on contingent claims pricing, Ph.D. thesis, Cornell University. Ithaca.
- Leland H.E., (1994), Corporate Debt Value, Bond Covenants, and Optimal Capital Structure, *Journal of Finance*, 49, 1213-1252.
- Leland H.E., Toft K.B., (1996), Optimal capital structure, endogenous bankruptcy and the term structure of credit spreads, *J. Finance*, 51, 987-1019
- Litterman R., Iben T., (1991), Corporate Bond Valuation and the term structure of credit spreads, *Financial Analysis Journal*, Spring 52-64
- Longstaff F.A., Schwartz, E.S., (1995), A simple approach to valuing risky fixed and floating rate debt, *J. Finance*, 50, 789-819.

Madan D.B., Unal H., (1994), Pricing the risks of default, Working Paper, Wharton Business School, Philadelphia.

Merton R.C., (1974), On the pricing of corporate debt: the risk structure of interest rates, J. Finance, 2, 449-470.

Pliska S.R., (1997), Introduction to Mathematical Finance, Discrete time models, Blackwell, Oxford.

Shimko D., Teejima N., van Deventer D., (1993), The pricing of risky debt when interest rates are stochastic, J. Fixed Income, September, 58-66.

Standard and Poor (1997a), Standard and Poor's Ratings Performance 1996 Stability and Transition, New York.

Standard and Poor (1997b), Standard and Poor's Rating Roundup; Corporate, Financial Institution and Sovereign Rating Changes for 1996, New York.