

The Risk Microstructure of Corporate Bonds: A Bayesian Analysis

Manfred Frühwirth *

Paul Schneider[†]

Leopold Sögner [‡]

Abstract

This article presents joint econometric analysis of interest rate risk, issuer-specific risk (especially credit risk) and bond-specific risk (especially liquidity risk) in a Lando (1998) type model within the Duffie/Singleton framework. Our model includes correlation between interest rate risk and issuer-specific risk. By means of data augmentation we develop a framework that allows for exact Bayesian analysis to estimate the model parameters and to separate the different components of risk. In particular we do not require an arbitrary benchmark bond that is free of any bond-specific risk. The estimation procedure is applied to coupon bond data from the German market.

Keywords: Credit risk, Liquidity risk, Duffie/Singleton framework, Markov Chain Monte Carlo estimation, Density approximation.

JEL: C52, G12, B13, E43

1 Introduction

Credit risk literature and industry measure the difference between risky bonds and risk-free bonds in the form of risk spreads. These spreads, in general, include both credit risk and liquidity risk and any market microstructure problems (see e.g. Campbell et al. (1996) or Gallmeyer et al. (2004)). Given empirical bond data, the standard procedure in literature (see e.g. Duffie et al. (2003) or Pan and Singleton (2005))

*Manfred Frühwirth, Weatherhead Center for International Affairs, Harvard University, 1737 Cambridge Street, Cambridge, MA 02138, USA, mfruehwirth@wcfa.harvard.edu

[†]Paul Schneider, Department of Corporate Finance, Vienna University of Economics and Business Administration, Nordbergstrasse 15, 1090 Wien, Austria, paul.schneider@wu-wien.ac.at

[‡]Leopold Sögner, Department of Management Science, Vienna University of Technology, Theresianumgasse 27, 1040 Wien, Austria, soegner@imw.tuwien.ac.at

assumes for each issuer the observation of exactly one bond without any liquidity risk, which obviously is a strong assumption. This article presents a refined estimation technique that allows identification and estimation of issuer-specific components and bond-specific components becomes without this strong assumption.

The main objective of this article is to model interest rate risk, issuer-specific and bond-specific risk and to develop an econometric methodology to separate and analyze these three types of risk. Our investigation is based on the Lando (1998), Duffie and Singleton (1999) and Feldhütter and Lando (2005) credit risk framework. For the risk-free term structure we employ a partially decoupled system of affine diffusion processes as formalized by Duffie and Kan (1996). We model the risk-free term structure by a representative of the $\mathbb{A}_1(3)$ class which was introduced by Dai and Singleton (2000) and, under a different probability measure by Collin-Dufresne and Goldstein (2002). In addition, we use for each issuer one latent issuer-specific factor (e.g. including credit and issuer-specific liquidity risk) and for each bond a latent bond-specific factor. One can imagine this bond-specific factor to represent bond-specific liquidity risk and other sources arising from market imperfections (see e.g. Campbell et al. (1996) or Gallmeyer et al. (2004)). Motivated by empirical evidence on the correlation of interest rate risk and credit risk (see e.g. Longstaff and Schwartz, 1995; Wei and Guo, 1997; Duffee, 1999; Frühwirth and Sögner, 2006), our model is able to account for correlation between the risk-free term structure and issuer-specific risk. After setting up the model we use coupon bond prices from the German market to estimate the model parameters. To find out more about the nature of the latent processes representing issuer-specific and bond-specific risk, the estimates are regressed against some possible determinants (including a liquidity proxy taken from literature).

Let us briefly review existing literature on the modeling (in a Duffie/Singleton type model) and estimation of different spread components. Especially, three papers have to be highlighted in this field: Duffie, Pederson and Singleton (see Duffie et al., 2003, henceforth DPS) estimate and separate credit risk and liquidity risk using Russian government bonds. They use simulated maximum likelihood to estimate the parameters. By contrast, Feldhütter and Lando (2005) focus on the swap and the corporate bond market. The authors adapt the Lando (1998) framework to decompose corporate bond yields into different

components (two factors for the risk free short rate, two factors for the credit spread, one factor treasury premium distinguishing the treasury rate from the riskless rate). For estimation they use the extended Kalman filter. Also ? used Kalman filter techniques to for parameter estimation. When working with yields, applying the Kalman filter is straightforward. However, when working with coupon bond data, the observation equation is highly non-linear in the system equations such that neither filtering is straightforward nor the distributional assumptions of Kalman filter correspond to the distributional assumptions of the term structure model. Although, filtering can be adapted to enable parameter estimation, parameter estimation of term structure models by means of maximum likelihood remains a nasty problem, especially, when no closed form solutions for the zero coupon prices are available. If not, like in the model structure used in this article, the maximization of the likelihood requires to solve a system of differential equations numerically (see Duffie and Kan (1996)) in each maximization iteration. In experiments with simulated data, Frühwirth et al. (2005) find out that maximum likelihood is unstable and MCMC clearly improves the quality of the estimation. Last but not least ? provide a lot of interesting results which are also tested and extended in this article. Driessen inferred a negative relationship between credit spreads and the risk free term structure, a liquidity component with a downward sloping term structure of the liquidity component and a significant impact of the bond age on the credit spread.

Our paper extends literature in the following ways:

In contrast to Feldhütter and Lando (2005) or ? we use a Markov Chain Monte Carlo estimation methodology instead of maximum likelihood estimation based on a Kalman filter. This enables us to use transition density approximations for the observed data instead of the normal assumption inherent in the Kalman filter technique. As these approximations in contrast to a normal density assumption are based on the characteristics of the underlying stochastic processes, these approximations are closer to the true density than the approximation by a normal density.

Let us come to the difference to DPS: The drawback of the methodology applied in DPS is that for each issuer they need to exogenously specify a benchmark bond (reference bond) that is absolutely and at each point in time free of any bond-specific risk (liquidity risk) in order to obtain a one-to-one relation between bonds available for estimation and driving factors. This assumption requires the non-existence

of idiosyncratic risk (which is in contrast to e.g. Pan and Singleton (2005), p. 32). Especially, it must be possible for each issuer analyzed to identify this benchmark bond even before starting the estimation procedure. Thus, altogether, such an assumption in general can be considered as too restrictive.

Therefore, in contrast to DPS, we take recourse to data augmentation releasing us from the assumption of a benchmark bond. The DPS framework can be seen as a special case of our methodology. With our methodology it is possible to evaluate for each issuer if the DPS assumption is appropriate and if this is the case our methodology can show which bond is the appropriate benchmark bond. Another difference between our paper and DPS is the estimation methodology: Whereas DPS use simulated maximum likelihood, we apply Markov Chain Monte Carlo (MCMC) estimation.

This paper is organized as follows: Section 2 describes the model. Section 3 outlines the estimation methodology. Section 4 describes the data used in the empirical part. Section 5 presents the estimation results received from the German bond data. Section 6 concludes.

2 The Model

We work in a frictionless and arbitrage-free market setting in continuous time t . A probability space with a filtration (F_t) with the usual properties (CADLAC) is underlying the model. The empirical probability measure and an equivalent martingale measure (risk-neutral measure) will be abbreviated by \mathcal{P} and \mathcal{Q} , respectively.

In the following, we define as “risk class” a homogenous set of bonds with identical issuer-specific and bond-specific risk. We consider one issuer with $j = 1, \dots, J$ coupon bonds on the market. We symbolize by $U_j(t)$ the set of coupon dates for bond j occurring between t and maturity (including maturity). Traded are risk-free zero-coupon bonds for all maturities and risky zero-coupon bonds for all maturities and all risk classes, both with a face value of 1. All zero-coupon bond prices satisfy the no-arbitrage condition. The time- t price of a zero-coupon bond with maturity T reflecting the risk of coupon bond j is abbreviated by $v_j(t, T)$.

The time t price of the risky coupon bond j , $p_j(t)$, is a linear combination of its remaining cash flows

$C_{pj}(u)$ and the risky zero-coupon bond prices $v_j(t, u)$:

$$p_j(t) = \sum_{u \in U_j(t)} v_j(t, u) C_{pj}(u). \quad (1)$$

The risk-free term structure, the issuer-specific risk and the bond-specific risk are modeled by the following latent stochastic processes $X(t)$ and discount rates $R(t)$:

Assumption 1. The risk-free area is modeled by a three factor model. Let us define the respective latent vector process $(X_{rf}(t))$ by $X_{rf}(t) = (X_1(t), X_2(t), X_3(t))^\top$. Based on recent literature (see e.g. Tang and Xia (2005)), we model $(X_{rf}(t))$ as a member of the $A_1(3)$ family introduced by Dai and Singleton (2000). This results in 14 identifiable parameters for the risk-free term structure (see Collin-Dufresne et al., 2004). This model structure turned out to cope with the volatility structure

of term structure models. From $(X_{rf}(t))$ we obtain the risk-free discount rate $R_{rf}(t) = \delta_{0,rf} + \delta_{x,rf}^\top X_{rf}(t)$, with $\delta_{x,rf} = (\delta_1, \delta_2, \delta_3)^\top$.

Assumption 2. For the issuer-specific risk we use one latent process denoted as $(X_4(t))$, independent of the risk-free area. We model $X_4(t)$ by means of a square root process since $X_4(t)$ is assumed to drive especially credit risk (besides some other issuer-specific sources of risk) and consequently should have a positive domain. From the latent processes driving the risk-free area and from $(X_4(t))$ we receive the discount rate for a fictitious bond with zero bond-specific risk: $R_I(t) = \delta_{0,I} + \delta_{x,I}^\top X_I(t)$, where $\delta_{x,I} = (\delta_1, \delta_2, \delta_3, \delta_4)^\top$ and $X_I(t) = (X_1(t), \dots, X_4(t))^\top$.

Assumption 3. To model the bond-specific risk we use one latent Ornstein-Uhlenbeck process for each bond. This process is symbolized by $(X_{5,j}(t))$, where j stands for the number of the corresponding bond ($j = 1, \dots, J$). The processes for different bonds j are assumed to be identical *in distribution* but not path-wise. X_4 and $X_{5,j}$ are jointly independent, where each $X_{5,j}$ follows a Gauss process. The bond-specific discount rates are $R_j(t) = \delta_{0,j} + \delta_{x,j}^\top X_j(t)$, where $\delta_{x,j} = (\delta_1, \dots, \delta_4, \delta_{5,j})^\top$ and $X_j(t) = (X_1(t), \dots, X_4(t), X_{5,j}(t))^\top$.

For notational completeness we introduce $X(t) = (X_1(t), \dots, X_3(t), X_4(t), X_{5,1}(t), \dots, X_{5,J}(t))^\top$, and $\delta_x = (\delta_1, \dots, \delta_3, \delta_4, \delta_{5,1}, \dots, \delta_{5,J})^\top$. From Assumptions 1, 2 and 3 the vector process $X(t)$ is of dimension

$M = 4 + J$, furthermore X is an affine stochastic process under \mathcal{P} and can be represented in the form

$$dX(t) = \beta(\alpha - X(t))dt + \Sigma\sqrt{S(t)}dW(t) \quad (2)$$

such that the Duffie and Kan (1996) methodology for affine term structure models can be applied. In (2) β and Σ are $(M \times M)$ matrices with positive elements, α is $M \times 1$ and $S(t)$ is a diagonal matrix with components

$$S_{ii}(t) = a_i + b_i^\top X(t) . \quad (3)$$

From the above assumptions $a_1 = 0, a_4 = 0, a_5, \dots, a_{4+J} = 1, b_{1(1)}, b_{4(4)} = 1$, and $b_{2(1)}, b_{3(1)}$ free. All other $b_i = 0$ and Σ diagonal. Note that $a_1 = 0, b_{1(1)} = 1$, and $b_{2(1)}, b_{3(1)}$ arbitrary, which follows from Assumption 1. The first three components of $X(t)$, i.e. $X_{rf}(t)$, are described by an $\mathbb{A}_1(3)$ model. The assumptions $a_1 = 0, a_4 = 0$ and $b_{1(1)}, b_{4(4)} = 1$ account for the fact that the first component of the risk free term structure and the issuer specific component are square root. The remaining restrictions arise from Assumption 3 where the bond specific components are independent Ornstein-Uhlenbeck processes.

Regarding parameter estimation, if the components a high dimensional diffusion system were independent, this would not cause big problems. However when considering term structure models, the results of e.g. Longstaff and Schwartz (1995); Wei and Guo (1997); Duffee (1999); Fröhwrth and Sögner (2006) provide strong arguments for factors, particularly for correlation between credit risk (which especially enters into our issuer-specific risk) and interest rate risk. Now, if some factors, i.e. components of $(X(t))$, are correlated with the risk-free term structure, the estimation problem becomes non-trivial. A natural and direct approach is to include these interdependencies by parameterizing the stochastic differential equations of the latent process $X(t)$. Unfortunately, due to current computing power, the calculation of closed-form likelihood expansions from Aït-Sahalia (2002) becomes a serious computational obstacle for order two expansions of higher dimensional interdependent diffusions. Thus, a direct application of a high dimensional Duffie and Singleton (1999) framework with correlated factors cannot be estimated.

Therefore, as an alternative we have to reduce the complexity of this problem. We do this by a separate treatment of the risk-free term structure by means of the $\mathbb{A}_1(3)$ model (see Assumption 1) and an application of the Lando (1998) generalized Markovian framework, also applied in Feldhütter and Lando (2005), for the issuer-specific risk. This framework enables us to construct risky zero coupon prices and a risky short rate correlated with the risk free short rate, while keeping the Assumptions 1 to 3 unchanged. Moreover, a separate estimation of the risk-free term structure parameters is possible.

Using this framework we are able to construct an affine term structure model where the issuer-specific risk is allowed to depend on the risk-free term structure *and* nevertheless a separate estimation of the risk-free term structure parameters and issuer-specific and bond-specific components is feasible.¹

We consider the most simple setting within the Lando (1998) framework with $K = 2$ states, where only a no-default and an absorbing default state are considered. The corresponding generator matrix is given by:

$$\Lambda = \begin{pmatrix} -\lambda_1 & \lambda_{12} \\ 0 & 0 \end{pmatrix} \quad (4)$$

From Lando (1998), if $(X_j(t))$ is a stochastic process satisfying the usual conditions, the risky zero-coupon bond prices for bond j in rating class i are given by:

$$v_j(t, T) = \sum_{l=1}^{K-1} \beta_{il} \mathbb{E}_{\mathcal{Q}} \left[\exp \left(\int_t^T (\mu(X_j(s)) - \nu(X_{rf}(s))) ds \right) | \mathcal{F}_t \right], \quad (5)$$

where $\mathbb{E}_{\mathcal{Q}}[.] | \mathcal{F}_t$ is the conditional expectation under the equivalent martingale measure \mathcal{Q} and $\mu(X_j(t)) -$

¹An alternative to the Lando (1998) framework would have been the Dai and Singleton (2000) framework. However, using the Dai and Singleton (2000) framework results in the following econometric problem: With interdependent factors the parameters of the issuer-specific and bond-specific components affect the risk-free zero-coupon bond prices, or more precisely the parameters of issuer-specific and/or bond-specific components enter into the ordinary differential equations of Duffie and Kan (1996), when solving for the zero-coupon bond prices. Thus, when calculating the risky zero-coupon bond prices $v_j(t, T) = \mathbb{E}_{\mathcal{Q}} \left[\exp \left(- \int_t^T R_j(s) ds \right) | \mathcal{F}_t \right]$, the risk-free zero-coupon bond prices would change. Thus, a separate treatment of the risk-free area and the risky area requires restrictions on the parameters that are very severe from an economic point of view. By contrast, the Lando (1998) approach provides a methodology where a separate treatment of the risk-free term structure is still possible, and only a small number of (at least one) extra parameter is necessary. For details see later in this section.

$\nu(X_{rf}(t))$ describes the risky stochastic discount rate. $\nu(X_{rf}(t))$ is a function describing the risk-free discount rate, i.e. by our above assumptions $\nu(X_{rf}) = \delta_{0,rf} + \delta_{x,rf}^\top X_{rf}(t) = R_{rf}(t)$. The risky discount rate can be derived from $R_j(t) = \beta_{11}[\mu(X_j(t)) - \nu(X_{rf}(t))]$.

Since K has been set to 2, the sum in (5) consists of only one element and the zero-coupon bond price and the discount rate are given by $\beta_{11}\mathbb{E}_{\mathcal{Q}}\left[\exp(\int_t^T(\mu(X_j(s)) - \nu(X_{rf}(s)))ds)\right]$ and $\beta_{11}[\mu(X_j(s)) - \nu(X_{rf}(t))]$, respectively. The coefficient β_{11} could be derived from $\beta_{11} = -b_{11}b_{12}^{-1}$, where b_{il} are the corresponding elements of the matrix of eigenvectors of Λ . However, since the zero-coupon bond price and the discount rate are proportional to $\mathbb{E}_{\mathcal{Q}}\left[\exp(\int_t^T(\mu(X_j(s)) - \nu(X_{rf}(s)))ds)\right]$ and $\mu(X_j(s)) - \nu(X_{rf}(t))$, β_{11} can be set to 1 without changing the structure of the model while keeping the model identified.²

Since, by our assumptions, the risk-free and the risky discount rate are affine (which is also assumed in the applied part of Lando (1998)), the stochastic discount rate $\mu(X_j(t)) - \nu(X_{rf}(t))$ has to be affine, too. In such a setting the correlation of risk-free term structure and the risky discount rates can be modeled as follows.³

$$R_j(t) = \delta_{0,j} + (1 - c) \sum_{l=1}^3 \delta_l X_l(t) + \delta_4 X_4(t) + \delta_{5,j} X_{5,j}(t) \quad (6)$$

where the parameter c controls the correlation of the risk-free and the risky discount rates and the risky zero-coupon bond prices are derived from

$$v_j(t, T) = \mathbb{E}_{\mathcal{Q}} \left[\exp(-\int_t^T R_j(X_j(s))ds) | \mathcal{F}_t \right]. \quad (7)$$

Thus, we can first estimate the risk-free area with the corresponding 14 parameters including $\delta_{x,rf}$ and the estimates of $(X_1(t), X_2(t), X_3(t))$, then with fixed estimates of $\delta_{x,rf} = (\delta_1, \delta_2, \delta_3)$ and estimates of

²Since β_{11} is derived from the eigenvalues of a generator matrix, β_{11} has to be in the interval $[0, 1]$. E.g. Feldhütter and Lando (2005) take the coefficients β_{il} from Moody's transition matrix for credit risk classes. However for a model with $K = 2$ we do not lose any information and flexibility of the model by setting $\beta_{11} = 1$. An estimation from the data would result in identification problems.

³Generalizations which are possible, e.g. of the form $R_j(t) = \delta_{0,j} + (1 - c_1)X_1(t) + (1 - c_2)X_2(t) + \dots$

$(X_1(t), X_2(t), X_3(t))$ ⁴, we estimate the remaining parameters – including c – of the risky term structure. In the ongoing analysis we set $\delta_{0,rf}, \delta_{0,j} = 0$.

In order to obtain an intuitive interpretation of the market prices of risk we employ a completely affine market price of risk specification $\lambda = \left(\lambda_1, \lambda_2, \lambda_3, \lambda_4 \right)^\top$. By this, $X_I(t)$ is an affine stochastic process also under \mathcal{Q} , such that the fundamental term structure PDE can be reduced to a system of ODE's as formalized by Duffie and Kan (1996). Due to the affine dynamics of $X(t)$, the price of a (fictitious) zero-coupon bond subject to the risk of bond j , $v_j(t, T)$, is exponentially affine in $X_j(t)$, resulting in zero-coupon bond prices

$$\begin{aligned} v_j(t, T) &= \mathbb{E}_{\mathcal{Q}} \left[\exp \left(- \int_t^T R_j(s) ds \right) \right] = \mathbb{E}_{\mathcal{Q}} \left[\exp \left(- \int_t^T \delta_{x,j}^\top X_j(s) ds \right) \right] \\ &= \exp \left(A_j(T-t) - B_j(T-t)^\top X_j(t) \right), \end{aligned} \quad (8)$$

where β_{11} has already been set to 1, $\delta_0=0$, $\delta_{0,j}=0$ and $j = 1, \dots, J$. In the following, the set of parameters defined in this section, including the parameters of the stochastic processes, δ_x and the market prices of risk λ , are abbreviated by ψ .

3 Estimation Methodology

For parameter estimation of continuous-time multi-factor term structure models in general, financial econometrics has recently created different approaches to make parameter estimation feasible or to improve the quality of estimators (e.g. Brandt and Santa-Clara (2002), Elerian et al. (2001), DPS). Onestrand of recent literature has developed tools based on discretization schemes, for both frequentist and Bayesian methodologies. Roughly speaking, the idea of these schemes is to augment the observations by latent terms. With this augmented set of

observations, either maximum likelihood based methods or Bayesian methods can be applied coming to an improved quality of the estimates. In a Bayesian setting, Elerian et al. (2001) approximate non-

⁴We use the means from the risk-free posterior.

linear stochastic differential equations by the Euler scheme. In Bayesian terminology – they augment the parameter space by latent spot rates between the actually observed data to improve the quality of the estimation methodology. The authors find that inserting approximately ten data points between two actually observed data points is sufficient. A similar approach with similar results is provided in Eraker (2001).

Alternatives such as simulated maximum likelihood (see e.g. Brandt and Santa-Clara, 2002) or some Bayesian analogs developed in Elerian et al. (2001) and Eraker (2001), as well as the methods of Singleton (2001), or Bates (2005), all suffer from a different curse of dimensionality. Consequently, these approaches turn out to be very time consuming, resulting in an insufficient number of MCMC steps within reasonable time.

We consider a problem with discretely sampled data, generated by diffusion processes. We have time gaps $\Delta_n = t_n - t_{n-1}$ and $n = 1, \dots, N$. In absence of holidays and weekends the step-width would be always $1/365$. Due to holidays and weekends the step-width is correspondingly higher between some successive observations. The measurements of the continuous-time stochastic process $(X(t))$ and the corresponding transformations $X_{rf}(t)$, $R(t)$, and $p_j(t)$ at t_n are X_n , $X_{rf,n}$, $R_{j,n}$, and $p_{j,n}$.

A well-known difficulty with this kind of estimation problem is that an estimation of the stochastic differential equations based on an Euler approximation results in poor performance for different estimation methodologies when Δ_n is too large.

From our assumptions in Section 2 the transition density $\pi(X_n | X_{n-1})$ corresponds to

$$\begin{aligned} \pi(X_n | X_{n-1}) &= \pi(X_{1,n}, X_{2,n}, X_{3,n} | X_{1,n-1}, X_{2,n-1}, X_{3,n-1}) \\ &\quad \cdot \pi(X_{4,n} | X_{4,n-1}) \prod_{j=1}^J \pi(X_{5,j,n} | X_{5,j,n-1}), \end{aligned} \quad (9)$$

where the closed-form approximations from Aït-Sahalia (2001), Aït-Sahalia (2002), and Aït-Sahalia and Kimmel (2002) are applied to obtain the first density on the right hand side.

For any one-to-one transformation $P_n = F(X_n)$, the transition density of P_n is derived by

$$\pi(P_n|P_{n-1}) = \pi(X_n|X_{n-1}) \frac{1}{\det|JF(.)|} \quad (10)$$

where $\det|JF(.)|$ is the determinant of the Jacobian of the function $F(.)$. This requires that the Jacobian of the transformation $F(.)$ has full rank.

3.1 Augmentation of the Parameter Space

Our goal is the estimation of the parameters without having to exogenously select a benchmark bond free of any bond-specific risk. The main problem that occurs without defining a benchmark bond is that the number of coupon bonds is too small compared to the number of latent stochastic processes.

Given that we have three factors for the risk-free area and one latent process reflecting issuer-specific risk, modeling and estimating the processes for J bonds of the same issuer requires a joint density of dimension $4 + J$. The processes for the risk-free area can be estimated from time series of risk-free interest rates for three different maturities. In addition, we have J time series from the J coupon bond prices, giving altogether $3 + J$ time series. Thus, there is a lack of one time series. E.g. if we consider two bonds of the same issuer ($J = 2$) we have one latent process driving the issuer-specific risk and two processes driving bond-specific risk. We therefore have three latent processes but only two bonds.

In formal terms, the main difficulty arises from the fact that the dimension of the observations $\bar{P}_n \in \mathbb{R}^L$ is smaller than the dimension of the vector of latent variables $X_n \in \mathbb{R}^M$. While due to the number of risk factors $M = 4 + J$, the observations \bar{P}_n include three risk-free interest rates and J coupon bond prices, such that the dimension is $L = 3 + J$. From our model assumption the components of $\bar{P} \in \mathbb{R}^L$ are linear combinations of zero-coupon bond prices. This fact has already been described by equations (1) and (8), where the risky zero-coupon bond prices $\in \mathbb{R}^M$ depend on the risk-free components $X_{rf,n}$. Next we abbreviate the map from $X_n \in \mathbb{R}^M$ to $\bar{P} \in \mathbb{R}^L$ by $\bar{F}(X_n)$. Since the dimension of the domain of $\bar{F}(X_n)$ is of higher dimension than its range, the mapping $\bar{P}_n = \bar{F}(X_n)$ cannot be one-to-one.

Due to this lack of a one-to-one relation between the latent factors and the data, the transition densities of the bond prices cannot be calculated by using the change of variables formula in (10) using e.g. the

closed-form likelihood expansions as developed in Aït-Sahalia (2001), Aït-Sahalia (2002) and Aït-Sahalia and Kimmel (2002).

One way to solve this technical problem is to assume that there is a benchmark bond without any bond-specific risk, as done by DPS. By this, one can reduce the dimension of X_n from $M = L + 1$ to $M = L$ which enables a one-to-one mapping.

In this paper we present an alternative resolution of this dilemma, namely *data augmentation*. To be more precise, we add the entire time series of an *artificial bond* to the parameter space. This artificial bond is free of any bond-specific risk.⁵ Thus, we augment the parameter vector by $\tilde{X}_{4,n}$, $n = 1, \dots, N$, which is a one-to-one transformation (parameterization) of $X_{4,n}$; this transformation will be abbreviated by $g(X_{4,n})$.

$$\begin{pmatrix} \bar{P}_n \\ \tilde{X}_{4,n} \end{pmatrix} = \begin{pmatrix} \bar{F}(X_n) \\ g(X_{4,n}) \end{pmatrix} := F(X_n) = P_n \quad (11)$$

This makes the vector of augmented bond prices $P_n \in \mathbb{R}^M$, such that the transformation $P_n = F(X_n)$ is one-to-one. P_n includes three risk-free yields (or bond prices), the artificial bond $\tilde{X}_{4,n}$ and the risky coupon bond prices p_j , $j = 1, \dots, J$.

A well-known fact with MCMC methods is that the parameterization of latent variables has an important impact on the convergence properties of the sampler (see e.g. ?, ?). For this reason we performed simulation experiments. In these experiments we found out that working directly with $X_4(t)$, assuming $\tilde{X}_4(t) = X_4(t)$, results in good sampling properties. Therefore, in the following we set $\tilde{X}_4(t)$ equal to $X_4(t)$.

⁵A similar methodology is also applied in Bayesian estimation of stochastic volatility models, where latent volatility paths are included into to the set of model parameters. Since the transition densities of the returns are usually known, adding latent volatility terms makes parameter estimation by means of MCMC feasible (see e.g. Jones (2003) or ?). Alternatively, in terms of filtering literature, we consider a dynamic model of dimension $4+J$, where the observation equation defines a vector of dimension $3+J$. The system equations are of dimension 1. By augmenting the parameter space, we add the system equations to the observation equation and derive a system enabling a one-to-one mapping between the bond prices and the stochastic processes on the one hand and the latent backward driving stochastic process ($X(t)$) on the other hand. Then the conditional densities of the bonds can be computed by applying the change of variables formula. As our approach is in principle an application of the well-developed missing-value or filtering literature, it is worth noting that our methodology offers an opportunity to investigate also other models where extra securities are necessary to complete the model and to perform parameter estimation.

An important aspect of our approach is that the Markov property is maintained⁶ such that the calculation of joint conditional densities remains straightforward. The key ingredient to our methodology is that given a specification (including its parametrization) for the dynamic evolution of the processes, we only look out for values of $X_4(t)$ and consequently $(X_{5,j}(t))$, that *most likely* have happened. Instead of, before estimation, defining one actually existing and observed time series as the time series that reflects a bond without bond-specific risk we add to the existing time series a time series of an artificial bond without bond-specific risk. This is done simultaneously to the estimation of all parameters. We point out that our estimation framework allows both the estimation of the sample path of *this* artificial bond and the estimation of the corresponding model parameters.

We point out that the methodology developed above can be easily extended to models with an arbitrary number n_f of factors for the risk-free area and $n_d > 1$ issuer-specific risk factors. In this case time series of risk-free interest rates for n_f maturities are necessary and n_d time series are missing which requires augmentation of the parameter space by n_d latent stochastic processes (i.e. n_d artificial bonds free of any bond-specific risk).

Our approach in one aspect resembles the DPS approach, as both approaches involve a bond without bond-specific risk. Therefore, we want to point out the major differences: First, with our approach no *observed* time series is defined as “without bond-specific risk”. Thus, (in our framework) the DPS technique involves a reduction of the number of latent processes (by one process) whereas our technique involves an increase in the size of the parameter space (by one complete time series of an artificial bond). Second, we have a simultaneous simulation of the posterior distribution of the artificial bond and estimation of the model parameters (including those generating the simulated time series). Third, we admit that our methodology produces a “softer” statement than DPS: We can only identify the *most likely* latent processes for bond-specific and issuer-specific risk instead of presenting unique latent processes (i.e. with probability 1).

To further elaborate the relation between the two approaches: The DPS approach can be seen as a

⁶See Appendix A for a more detailed coverage. Appendix A shows that our augmentation results in a well-defined model and that $\{P(t), \sigma(P(s), s < t)\}$ is Markov as well (since $L < M$). It is worth noting that $\tilde{P}(t)$ need not necessarily be Markov (see e.g. the comment in Ait-Sahalia and Kimmel (2002).)

special case of our methodology. If the benchmark bond assumption of the DPS approach does not hold, then our approach produces a "softer" statement than DPS, which however is more correct than a hard statement relying on a bond (without bond-specific risk) that does not exist. If the DPS benchmark bond assumption holds, our methodology should estimate for one of the bonds a bond-specific component insignificantly different from zero over the whole time span considered. This bond so turned out to be the appropriate benchmark bond. Thus, with our methodology it is possible to evaluate for each issuer if the DPS assumption is appropriate and if yes which bond can be used as a benchmark bond.

3.2 MCMC estimation

As already noted, the complex model structure makes a direct application of maximum likelihood infeasible. Therefore we apply Bayesian simulation methods to estimate the posterior distribution of the model parameters. The simulation methods concerned with this simulation task are Markov Chain Monte Carlo methods (see Robert and Casella (1999)). Using D for the data observed, by means of the Bayes theorem the posterior distribution of a parameter θ , $\pi(\theta|D)$ is proportional to the likelihood $f(D|\theta)$ times the prior $\pi(\theta)$, i.e.

$$\pi(\theta|D) \propto f(D|\theta)\pi(\theta) ,$$

In our model D corresponds to the bond prices observed $\bar{P} = (\bar{P}_1, \dots, \bar{P}_n, \dots, \bar{P}_N)$, $n = 1, \dots, N$, where N is the number of periods considered. As outlined in Section 3.1 we add the artificial bond process or issuer-specific spread process $X_{4,n} = \tilde{X}_{4,n}$, $n = 1, \dots, N$, to the set of parameters. A full Bayesian analysis also requires to calculate the density $\pi(X_1|X_0)$, where $X_0 = (X_{1,0}, X_{2,0}, X_{3,0}, X_{4,0}, \dots, X_{4+J,0})'$ has to be included to the set of unknown parameters.⁷ Thus with the augmented set of parameters $\theta = (X_4, X_0, \psi)$ and with our model structure, the a-posteriori distribution fulfills

⁷By contrast, Aït-Sahalia and Kimmel (2002) work with $N - 1$ conditional densities and marginal density $\pi(X_1)$, where the impact of this last density is asymptotically negligible; X_0 and X_1 the first and the second element of (X_n) . For an exact Bayesian analysis we can either work with the marginal density - which has to be known explicitly - or with data augmentation. Since this marginal is not available, we proceed with the latter approach.

$$\begin{aligned}
\pi(\theta|D) &\propto f(D|\theta)\pi(\theta) \\
&\propto \pi(\bar{P}|X_4, X_0, \psi)\pi(X_4|X_0, \psi)\pi(X_0|\psi)\pi(\psi) \\
&\propto f(P|X_0, \psi)\pi(X_0|\psi)\pi(\psi) .
\end{aligned} \tag{12}$$

where $P = (P_1, \dots, P_N)$. The definition $P_n = (\bar{P}_n, \tilde{X}_{4,n})^\top$, $\tilde{X}_{4,n} = X_{4,n}$ and by the Bayes theorem $\pi(\bar{P}|X_4, X_0, \psi)\pi(X_4|X_0, \psi) = f(P|X_0, \psi)$ we derive the likelihood $f(P|X_0, \psi)$, which is the product of N densities derived in equation (9).

Equipped with these conditional densities an exact Bayesian analysis can be performed to estimate all model parameters θ . $\pi(X_0|\psi)$ and $\pi(\psi)$ are the priors of the initial value of X and the unknown parameters of the stochastic processes ψ . These priors are chosen by the econometrician, while by the fact that $\pi(\bar{P}|X_4, X_0, \psi)\pi(X_4|X_0, \psi) = f(P|X_0, \psi)$, the "prior" $\pi(X_4|X_0, \psi)$ is completely determined by our model assumptions.

Priors: We use flat priors for all the remaining parameters X_0 and ψ . The prior for X_0 is multivariate normal with true unconditional expectation and five times the covariance, where each element of the variance-covariance matrix is multiplied by this constant factor. The expectation and the covariance can be computed as limits of the exact conditional expectation and covariance, that are readily available for affine diffusions. Since this limits only converge for stationary diffusions a consequence of this choice of prior is stationarity throughout the estimation. I.e. measure zero is put on parameter constellations where the Feller condition does not hold for any of the square root processes.

Moreover, from an economic point of view it is plausible that the spread induced by bond and issuer specific risk is non-negative, i.e. $R_{j,n} \geq R_{rf,n}$. From equation (6) we therefore derive $-c \sum_{l=1}^3 \delta_l X_{l,n} + X_{4,n} + \delta_{5,j} X_{5,j,n}(t) \geq 0 \ \forall n = 1, \dots, N$.⁸ Measure zero is put to all parameters where this restriction does not hold. Despite the fact that this restriction is based on economic considerations, the sampling properties improve and .. hier kommt noch Text dazu

MCMC: Since all conditional distributions are well-defined, Markov Chain Monte Carlo methods can

⁸Here we already use $\delta_{0,j} = 0$ and $\delta_4 = 1$.

be applied. By decomposing an updating sweep m into updating steps (which usually result from the structure of the joint density of the model), we construct an ergodic Markov chain $(\theta^{[m]})$. This chain converges to its invariant distribution which is the desired posterior distribution of the model parameters. In an application of MCMC, the updating procedure is repeated until the Markov chain has reached its invariant distribution. In the current application the autocorrelation of the sample paths are high. To cope with this, we generated runs with 1,000,000 simulation sweeps. For more detailed information on Markov chain Monte Carlo methods the reader is referred to Robert and Casella (1999) and Albert and Chib (2003)).

For the underlying model the updating sweep m , from $\theta^{[m-1]}$ to $\theta^{[m]}$, is split up into three steps:

Step 1: $X_4^{[m]}$ from $\pi(X_4|\bar{P}, X_0^{[m-1]}, \psi^{[m-1]})$

Step 2: $X_0^{[m]}$ from $\pi(X_0|\bar{P}, X_4^{[m]}, \psi^{[m-1]})$

Step 3: $\psi^{[m]}$ from $\pi(\psi|\bar{P}, X_4^{[m]}, X_0^{[m]})$

Samples of the conditionals $\pi(X_4|\bar{P}, X_0^{[m-1]}, \psi^{[m-1]})$, $\pi(X_0|\bar{P}, X_4^{[m]}, \psi^{[m-1]})$ and $\pi(\psi|\bar{P}, X_4^{[m]}, X_0^{[m]})$ are derived by means of the Metropolis/Hastings algorithm. Let ζ be a subset of θ , i.e. elements of X_4 , X_0 or ψ , then a proposal ζ^{new} from the proposal density $q(\zeta^{new}|\zeta)$ is accepted with probability $\min(1, r_p)$, where

$$r_p = \frac{\pi(P^{new}|\psi^{new}, X_0^{new})\pi(X_0^{new}|\psi^{new})\pi(\psi^{new})}{\pi(P|\psi, X_0)\pi(X_0|\psi)\pi(\psi)} \frac{q(\zeta^{new}|\zeta)}{q(\zeta|\zeta^{new})}. \quad (13)$$

ζ^{new} replaces the corresponding elements of θ in the case of acceptance, in the case of rejection corresponding parameters remain equal to ζ . The following comments show the use of the Metropolis/Hastings algorithm in each of the three steps of the respective updating sweeps in the MCMC estimation, described above.

ad Step 1: When X_4 is updated we split the path $(X_{4,n})$, $n = 0, 1, \dots, N$ into random blocks, with expected block size 3. This number is small compared to the proposals suggested in Elerian et al. (2001), but accounts for the fact that we are confronted with daily data which results in very informative densities.

Given the block indices n_0 and n_1 , $\zeta = (X_{4,n0}, \dots, X_{4,n1})$ is going to be updated. X_4 is a component of P in equation (13), such that the application of the Metropolis/Hastings algorithm is straightforward. We adopt the proposal strategy from Eraker (2001). As in step 1 only elements of X_4 are updated, the densities $\pi(X_0|\psi)\pi(\psi)$ cancel out in equation (13).

ad Step 2: The update of X_0 works equivalently to the update of X_4 . In updating X_0 we update the elements in one block jointly with X_4 . Note that when X_0 is updated the conditional densities referring to X_1 still remain in equation (13).

ad Step 3: Updating of ψ is once again a straightforward application of the Metropolis/Hastings algorithm. When comparing the results with different expected block sizes, it turned out that blocks with mean expected block size 3 resulted in good properties.

Proposal densities: For all the parameters we use a normal proposal e.g. $\log X_{4,n}$: $X_{4,n}^{new} = X_{4,n} + c_x \varepsilon$ and $\varepsilon \sim iid \mathcal{N}(0, 1)$. The term $\log q(\zeta^{new}|\zeta) - \log q(\zeta|\zeta^{new}) = 0$ in (13). Admissibility of the parameters is ensured through the prior of X_0 which implies stationarity.

To attain a sampler capable of “switching” between different regions of the parameter space, we vary c . In 10% of the proposals for the parameters ψ , c , is large, while for the remaining parameters we use a small variance. Particularly with the log-normal proposals we switch between $c = 0.1$. and $c = 0.5$.

4 Data

The data used in our study are daily observations from January 1st, 2004 to August 31st, 2005. Excluding holidays and weekends the observation period includes 428 days with data.

For the default-free area we exclusively used data provided by the Deutsche Bundesbank which can be downloaded from <http://www.bundesbank.de>. For the maturities 1 month, 3 months and 6 months we took recourse to the respective EURIBOR data. For the maturities 1, 2, ..., 10 years we used estimates of the parameters of the Svensson (1994) model. The Svensson parameters have been estimated by the Deutsche Bundesbank from German government bonds (“Federal bonds”) and government notes (“Federal notes”) with residual maturities of at least three months, using a non-linear parametric approach.⁹

⁹The interest rates described are interest rates with annual compounding (for the Svensson interest rates) or compounding

The default-risky bond data set comprises 7 German Mark (DEM) or Euro (EUR) denominated fixed-rate senior unsecured bank¹⁰ or corporate bonds¹¹ without sinking fund provisions or embedded options. Bonds issued by a financing subsidiary and guaranteed by the mother were considered as issued by the guaranteeing mother. From the Bloomberg database we extracted for each issuer the rating history and for each bond the bond features.

As regards rating, we used the long-term domestic issuer rating from S&P. All issuers selected had a stable rating (both the coarse rating and the fine rating reflected by - or +) throughout the observation period. 5 bonds are issued by Bayerische Hypo- und Vereinsbank with an A- rating, 2 bonds are issued by METRO with a BBB rating. Issuer, maturity, coupon rate and instrument code (ISIN) of all bonds are listed in Appendix B.

For each bond and each trading day, we obtained the gross price (dirty price) from the Datastream database with the prices of the HVB bonds derived from Munich stock exchange and those of the METRO bonds from Frankfurt stock exchange.

5 Estimation Results

5.1 Estimated Spread Processes

From our model assumptions the total risk spread of a bond consists of both issuer-specific and bond-specific risk. From equation (6) (where $\delta_{0,j} = 0$) we can derive the instantaneous issuer specific spread

at the end of the maturity (for EURIBOR). As in our model continuously compounded interest rates are required, the (sub-)annually compounded rates were converted into continuously compounded rates before entering into our model.

¹⁰One argument sometimes raised against the analysis of credit risk of banks is that banks usually are bailed out and therefore cannot go bankrupt. However, bailing out is no default in the sense of credit risk models, as no claims are reduced. One popular example of a bank default, corresponding to the usual definition, is the Barings case in 1995. In addition, analysts forecast that intensifying competitive pressures in the banking sector will provoke an increase in the number of bank defaults. Furthermore, if banks could not default it would have to be argued why banking laws contain insolvency provisions for banks, why secured senior bonds or subordinated bonds are issued by banks and why there are price differences between government bonds and identical bank bonds and within identical bank bonds of different seniority. It is hard to believe that the total of these price differences is attributable to liquidity differences. Also, Kiesel et al. (2003) cannot support the argument, that bank bonds are less risky than non-bank bonds with the same rating.

¹¹As the liquidity of bank and corporate bonds usually is smaller than that of government bonds, sometimes researchers use credit derivatives instead of or in addition to default-risky bonds to estimate or evaluate credit risk models (see Cossin and Hricko (2001) or Houweling and Vorst (2003)). However, as the credit derivatives market in Germany is only in its infancy and therefore liquidity of credit derivatives is not satisfactory, we decided to use bonds. The same is done by e.g. Düllmann et al. (2000) and Houweling P and Kleiberger (2001) who use German bond data of different rating classes for credit risk analysis.

at time-step n , $ISPR_n = R_I(t_n) - R_{rf}(t_n)$ and the instantaneous total spread for bond j at time-step n , $TSPR_{n,j} = R_j(t_n) - R_{rf}(t_n)$:

$$ISPR_n = -c \sum_{l=1}^3 \delta_l X_l(t) + X_4(t) , \quad (14)$$

$$TSPR_{n,j} = -c \sum_{l=1}^3 \delta_l X_l(t) + X_4(t) + \delta_{5,j} X_{5,j,n} . \quad (15)$$

From the MCMC output we can estimate these spreads. \widehat{ISPR}_n is an estimate of $ISPR_n$ taken from the posterior distribution. \widehat{ISPR}_1 is derived by taking the sample mean of the MCMC samples of the issuer-specific process (1,000,000 MCMC steps with 500,000 burn in), $n = 0, 1, \dots, N$; $\widehat{TSPR}_{j,n}$, $j = 1, \dots, J$ is derived in the same way as the estimates of the model parameters. Although the Feller condition was used to check for stationary model parameters for square root processes, the standard Dickey-Fuller test rejects the zero hypothesis of a unit root at a 5% significance level, while an augmented Dickey-Fuller test (with a constant) does not reject the zero hypothesis of a unit root when using the estimates of the issuer-specific risk process. Table 1 presents some descriptive statistics, with a multiplication of these statistics by 10,000 being required to receive the corresponding statistics for spreads in terms of basis points.

Figure 1 shows samples of the posterior distribution of the total spread for the two METRO bonds. We observe that the estimated paths for both bonds have a downward trend and that the spread for the second METRO bond is higher. We observe the same behavior with the HVB bonds.

Table 2 presents estimates of the model parameters.

The following subsection will provide a more in depth analysis on the properties of these estimates.

5.2 Determinants of the Spreads

The goal of this subsection is to find out the drivers of corporate bond spreads. To this end we first present plausible candidates and afterwards show the estimation results.

	MEAN	MEDIAN	MAX	MIN	STD
HVB					
\widehat{ISPR}	0.01989	0.01989	0.002077	0.001859	3.89E-05
\widehat{TSPR}_1	0.006677	0.006583	0.011362	0.001710	0.002377
\widehat{TSPR}_2	0.003856	0.003961	0.007969	5.40E-05	0.001653
\widehat{TSPR}_3	0.007254	0.007088	0.013124	0.000872	0.002890
\widehat{TSPR}_4	0.007550	0.007669	0.013056	0.001255	0.003090
\widehat{TSPR}_5	0.013421	0.014303	0.022633	0.003011	0.005580
METRO					
\widehat{ISPR}	0.002036	0.002046	0.002139	0.001852	5.34E-05
\widehat{TSPR}_1	0.002235	0.002066	0.005022	0.000243	0.001267
\widehat{TSPR}_2	0.001938	0.001577	0.005694	0.000140	0.001311

Table 1: Descriptive statistics of issuer specific spreads and total spreads.

5.2.1 Candidates for Determinants of the Spreads

The candidates used in our analysis as explanatory variables for the issuer-specific spread and the bond-specific spread are the DAX index, the market-value debt ratio, the distance to default, the default-free term structure level, the term spread of the default-free term structure, and the age of the respective bond.

We include the DAX 30 Xetra Performance Index, extracted from Datastream, to measure economic activity. An economic upturn reflected by rising stock prices (DAX) should reduce the credit risk perceived by the market participants, by this reducing especially the (credit risk related) issuer-specific spread. Including a stock market index (instead of the less frequently observed macroeconomic indicators GNP or GDP) has been suggested by e.g. Jarrow and Turnbull (2000). Additionally, we include the share price, STP_n , for which similar arguments are supposed to hold.

A further candidate for the credit spread is the market value debt ratio ("debt to value ratio")

$$DVR_n = \frac{D_n}{S_n + D_n} , \quad (16)$$

where S_n is the daily market capitalization at stock exchange and D_n is the market value of a firm's debt. Since the difference between book and the market values with debt is smaller than with equity,

	HVB		METRO	
α_4	0.00205523	(4.92E-01)	0.00201856	(5.85E-05)
β_{44}	4.79083	(0.65898255)	5.40477	(0.32670972)
λ_{1d}	0.189665	(0.03556379)	0.0699962	(0.0223327)
Σ_{44}	0.0100979	(3.59E-05)	0.0100061	(3.78E-05)
λ_{2d}	0.00597845	(0.0025578)	-0.00031647	(0.00180545)
$\lambda_{5,1}$	0.415789	(0.12063633)	-0.114041	(0.051889)
$\lambda_{5,2}$	-0.415164	(0.11363376)	-0.455991	(0.03734826)
$\lambda_{5,3}$	-0.395363	(0.05975682)		
$\lambda_{5,4}$	-0.242079	(0.07161819)		
$\lambda_{5,5}$	-0.377617	(0.05298538)		
β_{55}	1.20426	(0.16118358)	1.63966	(0.18226555)
β_{66}	1.17691	(0.09890209)	1.45526	(0.17783472)
β_{77}	1.20585	(0.1094075)		
β_{88}	1.22573	(0.09176127)		
β_{99}	1.1274	(0.18987861)		
$\delta_{5,1}$	0.00890408	(0.00029579)	0.00314502	(0.00011315)
$\delta_{5,2}$	0.00593328	(0.00038141)	0.00241247	(0.00012731)
$\delta_{5,3}$	0.00778648	(0.00070022)		
$\delta_{5,4}$	0.00631431	(0.00103032)		
$\delta_{5,5}$	0.00923439	(0.00343892)		
c	0.573638	(0.01918452)	0.118908	(0.01479916)
Drift under \mathcal{Q}				
β_{55}	0.000643764	(0.01334669)	-1.63733	(0.02269867)
β_{66}	0.0293178	(0.04795671)	-1.45183	(0.02318898)
β_{77}	0.140446	(0.05384456)		
β_{88}	0.00792883	(0.04865862)		
β_{99}	0.143731	(0.11206857)		

Table 2: Parameter estimates (and standard deviations in parantheses). β_{ii} are elements of the $(4 + J) \times (4 + J)$ matrix β (see equation (2)).

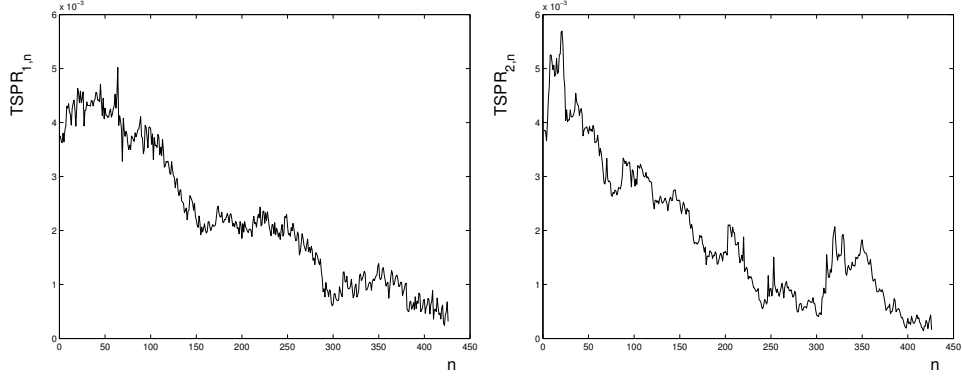


Figure 1: Estimates of the total spreads for the METRO Bonds (left subfigure: Bond METRO 1, right subfigure: Bond METRO 2) estimated from 1,000,000 MCMC steps (500,000 burn in steps).

we use book values. For the dates at the end of the respective quarters we take the quarterly debt data from the quarterly balance sheets. For all other dates, we derive D_n by linear interpolation. We expect, that the higher the debt to value ratio, the higher the probability of default and therefore, the higher the corresponding spreads.

In Merton type models, the distance to default is a major ratio to describe the conditional probability of default. In industry practice the KMV distance to default (see e.g. ?) is often used:

$$DD_n^j = \frac{S_n + D_n - DP}{(S_n + D_n)\sigma_V^j} \quad (17)$$

$S_n + D_n$ is the value of the firm. σ_V is the standard deviation of the firm value. In this article we calculate three different distances to default, depending on the methodology σ_V is calculated. First, based on the model assumptions σ_V^j is a constant parameter. Therefore, we estimate this parameter by means of unlevering the equity volatility, derived from the market capitalization; this results in DD_n^I . I.e. we calculated $\sigma_V^2 = N_S \mathbb{V}(STP_n)365$, where STP_n is the share price and N_S is the number of shares outstanding. As often done in applied literature, the standard deviation of the firm value is derived by an implicit estimation from the Black-Scholes formula. KMV used a constant interest rate of five percent in this calibration scheme. This results in DD_n^{II} . Third, we used the one year risk-free spot rate instead

of a constant interest rate, resulting in DD_n^{III} . Although, there are even problems with the distance to default to predict default probabilities (see ?), the KMV distance to default is one commonly accepted indicator of default risk. Thus, we expect that the higher the distance to default, the smaller the actual (issuer-specific) spread.

The default-free term structure level and the term spread are selected to check for a dependence of the spread process on the default-free term structure. This is important because there is plenty of literature showing a relation between credit risk (which is part of our issuer-specific spread process) and the risk-free term structure. As indicated by several articles (see e.g. Litterman and Scheinkman (1991) or Duffee (1998)), most of the variation in the default-free term structure can be captured by its level and its slope (also referred to as "term spread").

Literature on the influence of the level of the default-free term structure is mixed. Longstaff and Schwartz (1995), Wei and Guo (1997), Duffee (1998), Alessandrini (1999), Düllmann et al. (2000) and Annaert et al. (2000) come to the result that the credit spread is inversely related to the default-free term structure level. According to Duffee (1999) the influence of the risk-free rate level is not significant. Jaffee (1975) and Fridson and Jonsson (1995) find that the influence of the default-free term structure on credit spreads is negligible. Bühler et al. (2001) find an insignificant dependence on the default-free interest rate level (for most of the maturities investigated). Morris et al. (1999) show a negative short-run relationship and a positive long-run relationship. Arak and Corcoran (1996) find that parameters and significance depend on credit quality.

There is plenty of literature relating to the influence of the term spread, too. E.g. Duffee (1998) shows at least in part a significant influence of the default-free term spread. In Alessandrini (1999) or Annaert et al. (2000) a significant dependence of credit spreads for alternative credit risk classes on the term spread can be found - at least for some maturities. Düllmann et al. (2000) find that a decrease in the slope of the default-free term structure leads to a decrease in short-term credit spreads and an increase in medium- and long-term credit spreads.

Our proxy for the default-free term structure level is the one-year spot rate (denoted as *RFLEVEL* in the regression models), calculated from the Svensson parameters. Basically, the choice of the maturity

is arbitrary and we could equally select a long-term spot rate as a level proxy. Our proxy for the term spread (symbolized by $RFSLOPE$ in the regression models) is the difference between the ten year spot rate and the one year spot rate, both computed from the Svensson parameters. We do not select a spot rate with a maturity of more than ten years as both number and liquidity of government bonds is very low in the ">10 years" segment. Moreover, the 10-year/1-year spread is also used as a proxy for the interest rate differential in the German bond market by the Deutsche Bundesbank.

Finally, we use age as a liquidity proxy in our analysis. Two arguments are raised for including liquidity proxies in our regression. First, it is plausible to assume that both the issuer-specific spread and (especially) the bond-specific spread reflect liquidity risk. Second, Ericsson and Renault (2001) find a positive correlation between liquidity risk and credit risk. These facts should be reflected in higher issuer-specific and bond-specific spreads for less liquid bonds. The use of age (AGE) as a liquidity proxy is consistent with literature (see e.g. Sarig and Warga (1989), Warga (1992), Ericsson and Renault (2001) and Schultz (2001)). Newly issued bonds in general are considered to be more liquid than older bonds. Thus, we would expect a positive coefficient for the age variable. Last but not least we also include the trading volume, VOL_n . Unfortunately, volume data were only available for the second Metro bond.

5.2.2 Results - Determinants of the Spreads

Given the candidates discussed in Subsection 5.2.1, we estimate the regression model $\widehat{ISPR}_n = \beta_0 + \beta_1 AGE_n + \beta_2 RFLEVEL_n + \beta_3 RFSLOPE_n + \beta_4 DAX_n + \beta_5 STP_n + \beta_6 I DD_n^I + \beta_6 II DD_n^{II} + \beta_6 III DD_n^{III} + \beta_7 DVR_n + \beta_9 \widehat{ISPR}_{n-1} + \beta_{10} \widehat{ISPR}_{n-2} + \varepsilon_n$, where n is the number of the observation, AGE_n in the context of the issuer-specific spread is the age of the *second* bond issued by this issuer (HVB 2 or METRO 2, respectively).

Performing standard model selection for the \widehat{ISPR}_n time series, results in the significant variables AGE , $RFLEVEL$, $RFSREAD$ and \widehat{ISPR}_{n-1} for the HVB when using a 5% significance level. For METRO the variables AGE , DAX and the first order lagged term remained significant. I.e. only the variable AGE has a significant influence. The sign is negative for both issuers. Thus, the older the bonds (to be more precise the second bond of an issuer), the lower the issuer-specific spread.

For HVB the signs for the variables $RFLEVEL$ and $RFSLOPE$ are significant and positive, i.e. the higher the level or the slope of the risk-free term structure, the higher the issuer-specific spread. The positive influence of the term structure slope on the spread is unexpected, as a bank borrows on a short-term basis and lends on a longer-term basis resulting in higher earnings if the term structure slope is higher. By contrast, for METRO $RFLEVEL$ and $RFSLOPE$ are highly insignificant. We point out that as we have already included the parameter c to cope with the interdependence between the risk-free and the risky term structure, the regression results especially for HVB indicate that the model is not flexible enough to describe the complete interdependence between the risky and the risk-free area observed with the data.

The DAX index is significant only for METRO. The different influence of the risk-free interest rates and the equity markets for the two issuers seems plausible: HVB is a financial institution where one would expect a stronger influence of the interest rate environment than for the (non-financial) METRO. Obviously, by contrast the bond prices of METRO are driven to a larger extent by the overall equity market than by the interest rate environment.

It is interesting to see that the both the distance to default and the debt ratio, that seem to be very obvious candidates, turn out to be insignificant for both issuers.

Table 4 presents the regression analysis for the total spreads, with the regression setting $\widehat{TSR}_{j,n} = \beta_0 + \beta_1 AGE_n + \beta_2 RFLEVEL_n + \beta_3 RFSLOPE_n + \beta_4 DAX_n + \beta_5 STR_n + \beta_6 I DD_n^I + \beta_6 II DD_n^{II} + \beta_6 III DD_n^{III} + \beta_7 DVR_n + \beta_8 VOL_n + \beta_9 \widehat{ISPR}_{n-1} + \beta_{10} \widehat{ISPR}_{n-2} + \varepsilon_{TSR,n}$. For all bonds considered, the variable AGE is significant with the sign of the corresponding regression parameter being negative. This negative impact of age on the spread has already been observed with the issuer-specific spread. We conclude from our regression analysis that the liquidity proxy AGE has a significant impact on both bond-specific and issuer-specific spreads. However, the negative signs of the estimates are in contrast to our expectations (see Subsection 5.2.1). As the influence of size is negative for all spreads and all bonds, the reason may be that a time trend that is automatically included in the variable AGE, as both age and time evolve deterministically, may outweigh the liquidity effect in the AGE variable. E.g. if there was a trend like a reduction of credit and/or liquidity risk perceived by the market, this could explain the

inverse relation between age and spread.

An interesting structure of significance of the parameters can be observed. While a lot of the regression parameters are insignificant for the METRO bonds and the HVB bonds with a relatively short time to maturity, the credit proxies become significant for those bonds with longer time to maturity. For the first METRO bond, which matures in 2006, only *RFLEVEL*, *DAX* and *DRG* are significant in addition to *AGE* and the first order autoregressive term, while for the second METRO bond (with maturity 2008) *DD* and *DRG* are also significant. For the first HVB bond it was *RFLEVEL*, *DAX* and *DD^j* and *DRG*, while it was *DAX* and *DRV* for the second one. For the HVB bonds with the longer time to maturity only *RFSLOPE* is insignificant, while the other parameters considered are significant. For the long term bonds all proxies have a significant impact.

What is interesting concerning the influence of the risk-free term structure is the difference in the influence of the risk-free interest rate level on the issuer-specific risk on the one hand and total risk on the other hand. For the issuer-specific risk, the significance level of parameters is relatively low, however, the higher the risk-free interest rate level, the larger the issuer-specific spread (which is in line with the results of Longstaff and Schwartz (1995), Wei and Guo (1997), Duffee (1998), Alessandrini (1999), Düllmann et al. (2000) and Annaert et al. (2000)). For the total spread the results are mixed. This may be a possible explanation of the ambiguous findings of Arak and Corcoran (1996) and Morris et al. (1999) on the influence of the risk-free term structure on the spread.

A further interesting finding is the fact that the remaining impact of *RFLEVEL* measured by $\hat{\beta}_2$ is significant, i.e. the model setup cannot completely cope with the interdependence of the risky and the risk-free term structure. For the second HVB bond and the METRO bonds the sign of $\hat{\beta}_2$ is negative, for the other bonds the sign is positive. For those bonds with a positive (negative) $\hat{\beta}_2$, the Lando (1998) model underestimates (overestimates) the impact of the risk-free term structure. The mixed results with respect to the slope of the risk-free term structure correspond to the ambiguous results one can find in literature.

As regards the DAX, we observe only positive regression parameter estimates, i.e. if β_4 is significant, a growth in the stock price index increases the spread. Thus, against the intuition an increase in the general

stock market increases the total spread. This counter-intuitive result leaves room for future research. Especially, it has to be checked whether a disaggregation using the stock price of the respective issuer instead of a broad market index (as in the Jarrow and Turnbull (2000) suggestion) leads to more intuitive results. For the variables DD and DRV , the signs of the parameter estimates, when significant, meet our expectations. A higher distance to default decreases the total spread, while a higher debt to value ratio increases the spread.

Altogether, when comparing the two issuers we can observe that for the BBB-rated METRO bonds a lot of variables are insignificant while for the A-rated HVB bonds many variables are significant. Especially, it is plausible that the influence of the risk-free term structure is more dominant/significant for the HVB bank bonds than for the non-bank bonds of METRO or alternatively, that this impacts are caused by the time to maturity of the bonds.

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_{6I}$	$\hat{\beta}_{6II}$	$\hat{\beta}_{6III}$	$\hat{\beta}_7$	$\hat{\beta}_9$	$\hat{\beta}_{10}$
0.000679 (0.001069)	1.18E-05 (1.13E-05)	0.000103 (0.001138)	0.000907 (0.001250)	-5.70E-05 (9.02E-05)	5.10E-05 (4.82E-05)	-5.19E-05 (6.00E-05)	2.37E-06 (4.84E-06)	-9.27E-07 (4.96E-06)	-0.000440 (0.001084)	0.893553 (0.050253)	-0.045291 (0.050293)
-1.93E-05 (0.000163)	1.05E-05 (1.10E-05)	0.000739 (0.000858)	-0.000105 (0.001356)	-0.000118 (0.000103)	2.39E-07 (8.48E-05)	1.09E-05 (1.53E-05)	5.40E-06 (1.49E-05)	2.86E-08 (2.22E-06)	0.000157 (0.000234)	0.943038 (0.049079)	-0.044906 (0.049259)

Table 3: Regression estimates (and standard deviations in parenthesis) of the issuer-specific spreads. $\widehat{ISPR}_n = \beta_0 + \beta_1 AGE_n + \beta_2 RFLEVE_{L_n} + \beta_3 RFSLOPE_n + \beta_4 DAX_n + \beta_5 STP_n + \beta_{6I} DD_n^I + \beta_{6II} DD_n^{II} + \beta_{6III} DD_n^{III} + \beta_7 DV R_n + \beta_9 \widehat{ISPR}_{n-1} + \beta_{10} \widehat{ISPR}_{n-2} + \varepsilon_n$.

%

#	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_{6I}$	$\hat{\beta}_{6II}$	$\hat{\beta}_{6III}$	$\hat{\beta}_7$	$\hat{\beta}_8$	$\hat{\beta}_9$	$\hat{\beta}_{10}$
1	-0.063477 (0.029125)	-0.001832 (0.000342)	0.543546 (0.052401)	0.047711 (0.036404)	-0.000433 (0.002642)	-0.002214 (0.001393)	HVB 2.45E-05 (4.44E-05)	-0.000724 (0.000144)	0.000673 (0.000146)	0.061285 (0.029607)		0.366222 (0.048341)	0.124240 (0.045562)
2	-0.043548 (0.020568)	-4.87E-05 (0.000222)	0.392687 (0.038416)	0.076150 (0.025828)	0.001406 (0.001867)	-0.000699 (0.000980)	1.91E-05 (3.13E-05)	-0.000587 (0.000102)	0.000549 (0.000104)	0.036037 (0.020884)		0.691598 (0.048598)	0.031563 (0.044833)
3	-0.043489 (0.024019)	-0.000909 (0.000276)	0.582772 (0.044659)	0.162727 (0.031010)	0.000269 (0.002185)	-0.001070 (0.001148)	-8.44E-06 (3.68E-05)	-0.000635 (0.000117)	0.000591 (0.000119)	0.035130 (0.024427)		0.560237 (0.045433)	0.044562 (0.041985)
4	-0.044322 (0.020530)	-0.001368 (0.000246)	0.631789 (0.041233)	0.251876 (0.028449)	0.000833 (0.001860)	-0.001335 (0.000983)	-3.80E-05 (3.16E-05)	-0.000394 (9.94E-05)	0.000364 (0.000102)	0.036363 (0.020874)		0.427848 (0.044383)	0.016058 (0.040087)
5	-0.001147 (0.027347)	-0.003483 (0.000376)	0.827732 (0.050658)	0.462874 (0.040161)	0.001952 (0.002464)	-0.000842 (0.001300)	-0.000124 (4.30E-05)	-8.99E-06 (0.000135)	1.38E-05 (0.000137)	-0.004777 (0.027822)		0.406480 (0.042775)	-0.022135 (0.039003)
1	0.002007 (0.001637)	-0.000533 (0.000156)	-0.031553 (0.009548)	-0.012107 (0.013724)	-0.000215 (0.001001)	-0.000267 (0.000824)	9.22E-05 (0.000149)	-9.83E-05 (0.000146)	-2.82E-05 (2.24E-05)	0.000505 (0.002260)		0.621678 (0.047846)	0.223468 (0.047813)
2	-0.041409 (0.023014)	-0.000153 (0.000250)	0.378269 (0.042684)	0.060256 (0.029298)	0.001730 (0.002076)	-0.000854 (0.001100)	2.31E-05 (3.55E-05)	-0.000552 (0.000116)	0.000515 (0.000119)	0.034639 (0.023410)	2.15E-10 (3.45E-10)	0.682547 (0.053345)	0.046786 (0.049633)

Table 4: Regression estimates (and standard deviations in parenthesis) of the bond-specific spreads: $\widehat{TSPR}_{i,n} = \beta_0 + \beta_1 AGE_n + \beta_2 RFLEVE_{L_n} + \beta_3 RFSLOPE_n + \beta_4 DAX_n + \beta_5 STR_n + \beta_6 I DD_n^I + \beta_{6II} DD_n^{II} + \beta_{6III} DD_n^{III} + \beta_7 DV R_n + \beta_8 VOL_n + \beta_9 \widehat{TSPR}_{n-1} + \beta_{10} \widehat{TSPR}_{n-2} + \varepsilon_{TSPR,n}$.

#	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_{6I}$	$\hat{\beta}_{6II}$	$\hat{\beta}_{6III}$	$\hat{\beta}_7$	$\hat{\beta}_8$	$\hat{\beta}_9$	$\hat{\beta}_{10}$
HVB													
<i>ISPR</i>	-0.000488 (0.000272)	6.08E-06 (2.47E-06)									0.000758 (0.000301)	0.865962 (0.024848)	
<i>TSPR₁</i>	-0.026917 (0.011294)	-0.002413 (0.000211)	0.434400 (0.043154)						-8.55E-05 (2.01E-05)		0.029771 (0.011986)	0.507977 (0.039397)	
<i>TSPR₂</i>	-0.005613 (0.000648)		0.258464 (0.030425)	0.089848 (0.011687)					-3.72E-05 (9.87E-06)			0.764481 (0.025430)	
<i>TSPR₃</i>	-0.013027 (0.011178)	-0.000898 (0.000233)	0.460643 (0.039503)	0.171262 (0.028381)					-7.61E-05 (1.96E-05)			0.633703 (0.028386)	
<i>TSPR₄</i>	-0.035186 (0.009700)	-0.001147 (0.000202)	0.536661 (0.036875)	0.265154 (0.026559)					-7.28E-05 (1.65E-05)			0.028152 (0.010649)	
<i>TSPR₅</i>	-0.048447 (0.013315)	-0.002526 (0.000285)	0.743170 (0.046684)	0.470037 (0.038076)				-8.34E-05 (2.01E-05)			0.04123 (0.014576)	0.457548 (0.030516)	
METRO													
<i>ISPR</i>	-7.51E-05 (0.000111)	1.06E-05 (3.77E-06)					4.78E-05 (1.83E-05)				0.000235 (0.000129)	0.904365 (0.020190)	
<i>TSPR₁</i>	0.001882 (0.000308)	-0.000440 (6.71E-05)	-0.034451 (0.008471)									0.829379 (0.025490)	
<i>TSPR₂</i>	4.50E-06 (1.29E-05)											0.993225 (0.005517)	

Table 5: Regression estimates (and standard deviations in parenthesis) of the issuer and bond-specific spreads after stepwise deletion of insignificant variables: $\widehat{ISPR}_n = \beta_0 + \beta_1 AGE_n + \beta_2 RFLEVEL_n + \beta_3 RFSLOPE_n + \beta_4 DAX_n + \beta_5 STR_n + \beta_{6I} DD_n^I + \beta_{6II} DD_n^{II} + \beta_{6III} DD_n^{III} + \beta_7 DV R_n + \beta_8 VOL_n + \beta_9 ISPR_{n-1} + \beta_{10} ISPR_{n-2} + \varepsilon_n$ and $\widehat{TSPR}_{i,n} = \beta_0 + \beta_1 AGE_n + \beta_2 RFLEVEL_n + \beta_3 RFSLOPE_n + \beta_4 DAX_n + \beta_5 STR_n + \beta_{6I} DD_n^I + \beta_{6II} DD_n^{II} + \beta_{6III} DD_n^{III} + \beta_7 DV R_n + \beta_8 VOL_n + \beta_9 TSPR_{n-1} + \beta_{10} TSPR_{n-2} + \varepsilon_{TSPR,n}$.

6 Conclusions

The goal of this article is to separate and estimate the components of the spread between corporate bonds and the risk-free term structure. In our model the risk-free area is modeled by a representative of the $\mathbb{A}_1(3)$ family, the issuer-specific spread is driven by a square root process and the bond-specific spreads are modeled by Gauss processes. We apply and adapt the Lando (1998) framework. This setting enables a separate estimation of the risk-free term structure parameters, along with integrating a correlation between issuer-specific risk and interest rate risk.

For this model standard maximum likelihood estimation cannot be applied for lack of the likelihood. As regards the estimation technology, in contrast to existing literature (Duffie et al. (2003), Feldhütter and Lando (2005)) we use MCMC estimation based on density approximations from Aït-Sahalia (2002). We apply Bayesian simulation methods to estimate the posterior distribution of the model parameters. To enable an exact Bayesian analysis we augment the parameter space by a latent process, reflecting the issuer-specific spread (artificial bond without any bond-specific risk). This procedure makes the vector of "observed" variables the same dimension as the vector of latent stochastic factors. Together with standard assumptions on affine term structure models this results in a one-to-one correspondence between the observations and the latent stochastic factors such that the density transformation formula can be used to calculate the densities required for Bayesian estimation. Therefore, our approach sheds light on new applications that can be engineered with the well-developed tools of the missing-value literature.

Extensions of our methodology to models with rating-specific and/or industry-specific risk are straightforward, given sufficient computing power for higher dimensional density approximations as well as to perform a sufficient number of MCMC steps. Also, the methodology developed in this article allows to build cascades of factors, e.g. rating-specific factors on an upper (more aggregate) level, industry-specific factors on a lower (more disaggregated) level, etc. An analysis of this kind is hardly possible if one has to rely on benchmark bonds (with one risk factor equal to zero) or other approaches currently documented in literature.

One major advantage of the MCMC sampler is that also the posterior distribution of the latent issuer-specific and bond-specific processes is estimated. Therefore, we can use these estimates and check for

further determinants of the bond-specific and issuer-specific spread. We find that the age of the bond (inversely), in part the level of the risk-free term structure (inversely for the issuer-specific spread and positively correlated for the bond-specific spread) and in part the slope to the risk-free term structure (positively correlated) have a significant impact on the issuer-specific and bond-specific spreads. The influence of the risk-free term structure is more dominant/significant for the HVB bank bonds than for the non-bank bonds issued by METRO. The DAX index has no significant impact on the spreads. This provides an argument against Jarrow and Turnbull (2000) type models that model explicitly the dependence of the spread on a stock index.

A Markov Property of the Augmented Observations

First we check whether the augmentation results in a well-defined model. In (11), the function \bar{F} is extended such that $P(t) = F(Y(t))$ and $F^{-1}(\cdot)$ exists (a.s.). An augmentation of this kind of course requires that the Jacobian of \bar{F} has full rank, i.e. $\text{rank}(J(\bar{F}(\cdot))) = L$ (a.s.), which is fulfilled by the model assumptions of Section 2. If $\text{rank}(J(\bar{F}(\cdot))) < L$ extra augmentations would be necessary. Now a function $F(\cdot)$ can be constructed such that the Jacobian of F has rank M (a.s.). E.g. if $\text{rank}(\bar{F}) = L = M - 1$, then an augmentation with the component $\tilde{X}_{4,n} = g(X_{4,n})$ provides us with the desired result. In simulation experiments it turned out that working with $\tilde{X}_{4,n} = X_{4,n}$ results in good sampling properties.

Second, we have to show if $\{X(t), \sigma(X(s), s < t)\}$ and $\bar{F}(\cdot)$ fulfills the properties described in Section 2 with the augmentation constructed in Section 3, then $\{P(t), \sigma(P(s), s < t)\}$ is Markov as well.

Proposition 1 (Joint Markovianity of P). *If $\{X(t), \sigma(X(s), s < t)\}$ is Markov, $P(t) = F(Y(t))$ and is continuous and one-to-one, then $\{P(t), \sigma(P(s), s < t)\}$ is a joint Markovian system.*

Proof. We need to show that

$$\mathbb{E}[P(t) \mid \sigma(P(s), s < t)] = \mathbb{E}[P(t) \mid P(s)] \ .$$

Since X is Markov, $\bar{P}(t)$ given the sigma field generated by $X(t)$ is Markov. F is one-to-one, i.e. $F^{-1}(\cdot)$ exists and continuous (and therefore measurable). Therefore every open set in $\sigma(P(s), s < t)$ has a open

corresponding set in $\sigma(X(s), s < t)$, i.e. both sigma fields are equivalent. □

B List of bonds used for estimation

HVB bonds (AA bank):

Bond#	Abbreviation	Issuer	Maturity	Coupon	ISIN
			(MM/DD/YYYY)	(p.a.)	
1	HVB 1	Bayerische Hypo- und Vereinsbank	02/13/2006	4.75%	DE0002515590
2	HVB 2	Bayerische Hypo- und Vereinsbank	01/08/2007	4.5%	DE0002516416
3	HVB 3	Bayerische Hypo- und Vereinsbank	08/11/2008	3.875%	DE0008087834
4	HVB 4	Bayerische Hypo- und Vereinsbank	11/26/2010	5.75%	DE0002515566
5	HVB 5	Bayerische Hypo- und Vereinsbank	03/27/2012	5.625%	DE0002516556

METRO bonds (BBB non-bank):

Bond#	Abbreviation	Issuer	Maturity	Coupon	ISIN
			(MM/DD/YYYY)	(p.a.)	
6	METRO 1	Metro Finance BV	03/09/2006	5.75%	DE0006111909
7	METRO 2	Metro AG	02/13/2008	5.125%	DE0002017217

References

- Aït-Sahalia, Y. (2001). Closed-form likelihood expansions for multivariate diffusions. Working paper, Princeton University and NBER.
- Aït-Sahalia, Y. (2002). Maximum-likelihood estimation of discretely-sampled diffusions: A closed-form approximation approach. *Econometrica*, 70:223–262.
- Aït-Sahalia, Y. and Kimmel, R. (2002). Estimating affine multifactor term structure models using closed-form likelihood expansions. Working paper, Princeton University and NBER.
- Albert, J. and Chib, S. (2003). Bayes inference via gibbs sampling of autoregressive time series subject to markov mean and variance shifts. *Journal of Business and Economics Statistics*, 11(1):1–15.
- Alessandrini, F. (1999). Credit risk, interest rate risk and the business cycle. *Journal of Fixed Income*, 10(2).
- Annaert, J., DeCeuster, M., and DeJonghe, F. (2000). Modeling european credit spreads. Technical report, Deloitte & Touche.
- Arak, M. and Corcoran, P. (1996). Yields on privately placed debt: Examining the behavior. *Journal of Portfolio Management*, 22(3):88–96.
- Bates, D. (2005). Maximum likelihood estimation of latent affine processes. *Review of Financial Studies*. forthcoming.
- Brandt, M. W. and Santa-Clara, P. (2002). Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets. *Journal of Financial Economics*, 63:161–210.
- Bühler, W., Düllmann, K., and Windfuhr, M. (2001). Affine models, credit spreads and the delivery option of a multi-issuer bond future. Working Paper, University of Mannheim.
- Campbell, J. W., Lo, A. W., and MacKinlay, C. (1996). *The Econometrics of Financial Markets*. Princeton University Press, Oxford, UK.

- Collin-Dufresne, P. and Goldstein, R. (2002). Do bonds span the fixed income markets? theory and evidence for unspanned stochastic volatility. *Journal of Finance*, 57(4):1685–1730.
- Collin-Dufresne, P., Goldstein, R. S., and Jones, C. S. (2004). Can interest rate volatility be extracted from the cross section of bond yields? an investigation of unspanned stochastic volatility. Working Paper, Haas School of Business.
- Cossin, D. and Hricko, T. (2001). Exploring for the determinants of credit risk in credit default swap transaction data. Working Paper, HEC and University of Lausanne.
- Dai, Q. and Singleton, K. J. (2000). Specification analysis of affine term structure models. *Journal of Finance*, 55(5):1943–1978.
- Düllmann, K., Uhrig-Homburg, K., and Windfuhr, M. (2000). Risk structure of interest rates: An empirical analysis for deutschemark-denominated bonds. *European Financial Management*, 5:367–388.
- Duffee, G. R. (1998). The relation between treasury yields and corporate bond yield spreads. *Journal of Finance*, 53:2225–2241.
- Duffee, G. R. (1999). Estimating the price of default risk. *Review of Financial Studies*, 12(1):197–226.
- Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance*, 6(4):379–406.
- Duffie, D., Pedersen, L. H., and Singleton, K. J. (2003). Modeling sovereign yield spreads: A case study of russian debt. *Journal of Finance*, 53(1):119–159.
- Duffie, D. and Singleton, K. (1999). Modelling term structures of defaultable bonds. *Review of Financial Studies*, 12(4):687–720.
- Elerian, O., Chibb, S., and Shephard, N. (2001). Likelihood inference for discretely observed nonlinear diffusions. *Econometrica*, 69(4):959–993.
- Eraker, B. (2001). Mcmc analysis of diffusion models with application to finance. *Journal of Business and Economic Statistics*, 19(2):177–191.

- Ericsson, J. and Renault, O. (2001). Liquidity and credit risk. Working Paper, LSE.
- Feldhütter, P. and Lando, D. (2005). Decomposing swap spreads. *mimeo, Copenhagen Business School*.
- Frühwirth, M., Schneider, P., and Sögner, L. (2005). Bayesian versus maximum likelihood estimation of term structure models driven by latent diffusions. Working Paper, Vienna University of Economics and BA.
- Frühwirth, M. and Sögner, L. (2006). The jarow/turnbull default risk model - evidence from the german market. *European Journal of Finance*. forthcoming.
- Fridson, M. and Jonsson, J. (1995). Spread vs. treasuries and the riskiness of high-yield bonds. *Journal of Fixed Income*, 5(3):79–88.
- Gallmeyer, M., Hollifield, B., and Seppi, D. (2004). Liquidity discovery and asset pricing. Working paper, Carnegie Mellon University.
- Houweling, P. and Vorst, T. (2003). Pricing default swaps: Empirical evidence. Working Paper, Erasmus University Rotterdam, forthcoming in *Journal of International Money and Finance*.
- Houweling P, J. H. and Kleibergen, F. (2001). The joint estimation of term structures and credit spreads. *Journal of Empirical Finance*, 8(3):297–323.
- Jaffee, D. (1975). Cyclical variations in the risk structure of interest rates. *Journal of Monetary Economics*, 3(1):309–325.
- Jarrow, R. and Turnbull, S. (2000). The intersection of market and credit risk. *Journal of Banking and Finance*, 24(1,2):271–299.
- Jones, C. (2003). The dynamics of stochastic volatility: evidence from underlying and options markets. *Journal of Econometrics*, 116:181–224.
- Kiesel, R., Perraudin, W., and Taylor, A. (2003). Pricing derivatives on financial securities subject to credit risk. *Journal of Risk*, 6(1).

- Lando, D. (1998). On cox processes and credit risky securities. *Review of Derivatives Research*, 2:99–120.
- Litterman, R. and Scheinkman, J. A. (1991). Common factors affecting bond returns. *Journal of Fixed Income*, 1:54–61.
- Longstaff, F. A. and Schwartz, E. (1995). A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance*, 50(3):789–819.
- Morris, C., Neal, R., and Rolph, D. (1999). Credit spreads and interest rates: A cointegration approach. Working Paper, Federal Reserve Bank of Kansas City.
- Pan, J. and Singleton, K. (2005). Default and recovery implicit in the term structure of sovereign cds spreads. working paper, MIT Sloan School of Management.
- Robert, C. and Casella, G. (1999). *Monte Carlo Statistical Methods*. Springer, New York.
- Sarig, O. and Warga, A. (1989). Some empirical estimates of the risk structure of interest rates. *Journal of Finance*, 44:1351–1360.
- Schultz, P. (2001). Corporate bond trading costs and practices: A peek behind the curtain. *Journal of Finance*, 56(2):677–698.
- Singleton, K. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics*, 102(6):111–141.
- Svensson, L. (1994). Estimating and interpreting forward interest rates : Sweden 1992-1994. IMF Working Papers 94/114, International Monetary Fund.
- Tang, H. and Xia, Y. (2005). An international examination of affine term structure models and the expectations hypothesis. Working Paper, Wharton School.
- Warga, A. (1992). Bond returns, liquidity, and missing data. *Journal of Financial and Quantitative Analysis*, 27:605–617.

Wei, D. and Guo, D. (1997). Pricing risky debt, an empirical comparison of the longstaff and schwartz and merton models. *Journal of Fixed Income*, 7(2):8–28.