

Can Interest Rate Dynamics Save Structural Models ?¹

Steven Simon

January 2005

¹I would like to acknowledge the Ente Luigi Einaudi for generously funding this research during my stay as a research fellow in Rome during the academic year 2003-2004.

Abstract

The results for structural models for corporate bonds in the recent empirical literature are at best mixed; in general, such models generate credit spreads considerably below empirically observed levels. However, several authors have found that, for realistic parameter values, such models do seem able to generate levels for default probabilities in line with empirically observed default frequencies. This seems to suggest that the link between the physical and risk-neutral dynamics in the existing structural models might be too tight, such that a model can be calibrated to either default frequencies or credit spreads, but not to both.

In this paper the structural model of Collin-Dufresne and Goldstein (2001) is extended to incorporate a two-factor essentially affine term structure model. The results, generated along the lines of Huang and Huang (2003), strongly suggest that allowing for essentially affine dynamics significantly improves the ability of a structural model to generate credit spreads of realistically high levels.

1 Introduction

In the recent empirical literature direct tests of structural models have lead to mixed results. Some authors have suggested that this is due to the fact that corporate yield spreads are not pure default spreads. Clearly, credit spreads do compensate holders of corporate bonds for their exposure to default risk, but also other factors, such as liquidity or asymmetrical taxation of Treasury and corporate bonds, contribute towards credit spreads. Therefore, any empirical test of structural models creates a joint hypothesis problem. If a structural model fails to explain observed credit spreads, this could be due to the fact that the model under consideration is wrong as well as the fact that observed credit spreads are not pure default spreads.

Leland (2002) shows that for realistic parameter values the Merton model and the Black and Cox model are both able to generate default frequencies in line with empirically observed levels, despite their failure to replicate 'realistic' levels for the credit spreads. More recently, Huang and Huang (2003) test a wide selection of structural models on their ability to replicate credit spreads when they have been calibrated to reproduce observed levels of default frequencies. They make two interesting observations. First, they find that under this condition the different structural models generate very similar credit spreads. Secondly, the level of the credit spreads generated by structural models only accounts for a fairly small fraction of the observed credit spreads, e.g. around 20% in the case of A-grade bonds. That is, when structural models are made to reproduce observed levels of default frequencies, they consistently generate 'too low' credit spreads. Huang and Huang (2003) point out that this seems to suggest that observed credit spreads are not pure default spreads, and that therefore structural models will fail to generate observed levels of credit spreads while at same time being able to replicate observed default frequencies.

However, in a recent paper Longstaff (2004) shows that exposure to default risk accounts for a fairly high percentage of the credit spread across rating classes. Using a reduced-form model with a liquidity factor he finds that the default component in credit spreads is significantly higher than is usually obtained by means of structural models, e.g. the results of Huang and Huang (2003). Therefore, dismissing the 'low' credit spreads generated by structural models on the basis of the existence of liquidity and tax effects might be a bit too easy.

The fact that structural models, also the more basic ones, are able to replicate default frequencies but not credit spreads, is similar to the observation made by Duffee (2002) that affine term structure models are not able to capture and certain features of bond yields (driven by the risk-neutral dynamics of term structure model) and certain features of the dynamics of the spot rate (driven by the physical dynamics). This is due to the fact that in an affine model there is a strong link between the physical and risk-neutral dynamics, and as such an affine model can not be calibrated accurately to both of them simultaneously. He suggests to allow for a more general specification of the price-of-risk vector in affine term structure model in order to loosen the tight link between the physical and the risk-neutral dynamics, thereby creating the class of essentially affine term structure models.

As in a structural model the default frequencies and credit spreads are respectively determined by the physical dynamics and the risk-neutral dynamics of the asset-value process we suggest a similar approach in addressing what has been labeled the 'credit-spread puzzle': the inability of structural models to generate realistic levels for credit spreads. Following Duffee (2002) we allow for an essentially affine specification of the price-of-risk vector for the term-structure model and for the asset-value process. Numerical results indicate that such an 'essentially affine' structural model is able to generate realistic levels of credit spreads.

The rest of the paper is organized as follows. Section 2 gives a short introduction to structural models. In Section 3 we present the model of Collin-Dufresne and Goldstein (2001) (CDF) and we discuss the importance of the term structure dynamics in a structural model. This takes us to Section 4, where we give a short discussion of essentially affine term structure models and we derive a specific two-factor model. We extend the original CDF model by incorporating an the essentially affine two-factor model derived in Section 4. Three different versions of the term structure model are calibrated in Section 6, and in Section 7 the structural model is tested for each of the three choices for the interest rate dynamics. Finally, Section 8 contain the conclusions as well as some suggestions for future research.

2 Structural Models

The literature on theoretical models for corporate bond prices, or equivalently for credit spreads, is divided into two rather distinct approaches: so-called structural models on the one hand and reduced form models on the other. With the latter approach, the default event

is modelled as a separate stochastic process, without any relation to the firm-value process. Therefore, a defaultable bond is not a contingent claim, since the default event is driven by an exogenous process, see Jarrow, Lando and Turnbull (1997) and Duffie and Singleton (1999). In contrast, in structural models default is defined in terms of the firm-value process and hence corporate bonds can be priced using contingent-claims analysis.

The literature on structural models for corporate bonds emerged almost together with the option-pricing literature itself, going back to the seminal "Merton model" of Merton (1974). In this model, a firm defaults on its debt at maturity if the firm-value is less than the principal due. Geske (1977) extends the model by allowing for the possibility of early default, triggered by coupons due, which turns corporate bonds into Bermudan/compound style derivatives. In another early contribution to the literature, Black and Cox (1976) assume that default is triggered if the value of the asset of the firm hits some lower boundary level. The motivation for a Merton-style model is that, as a result of the limited liability of equityholders, a firm will default on its debt when the value of equity becomes equal to zero. In contrast, Black and Cox (1976) assume that default is triggered by a breach of the bond covenants, rather than by the limited liability of equityholders, an assumption supported by the observation that quite often the equity value of a firm has not dropped to zero at the time of default. The last approach is usually considered to have two advantages over the Merton-style models. First, with the models of Merton (1974) and Geske (1977) the default-mechanism is tied to the coupon rate and the face value of the debt. Therefore, such models assume that the total debt of a firm exist of a single debt issue. A second disadvantage of this type of models is that Bermudan/American-style derivatives are notoriously hard to value, reducing hopes of obtaining closed form expressions for bond prices and making their actual use, e.g. calibration, more cumbersome.

Because of these advantages of barrier-style models most of the literature on structural models falls within this group. More recent contributions to this part of the literature have either extended the original Black and Cox model to a stochastic interest rate framework, such as Longstaff and Schwartz (1995), or have loosened the restriction of a constant default boundary. Observe that for any exogenous default boundary to have a meaningful interpretation it would have to be increasing in the level of outstanding debt. Therefore, the assumption of a constant default boundary implies a constant debt level, as a result of which expected leverage ratios would decrease over time. However, in practice leverage ratios show no such

trend but tend to be stable. Several papers have loosened the assumption of a constant default boundary, in turn allowing for the default boundary to be stochastic, see for example Nielsen, Saá-Requejo and Santa-Clara (1993) and Briys and de Varenne (1997). In such models the default boundary is assumed to be driven by the spot rate. However, as this implies a stationary debt level, it does not lead to a stationary leverage ratio. Recently Collin-Dufresne and Goldstein (2001) and Taurén (1999) have taken a different approach. They assume that default is driven by the leverage process rather than by the firm-asset value process. By further assuming that the amount of outstanding debt is driven by the firm-value, the leverage ratio follows a stable process. In the Section 3 we discuss the Collin-Dufresne and Goldstein (2001) model in more detail.

3 The Model of Collin-Dufresne and Goldstein (2001)

In this section, we present the structural model of model of Collin-Dufresne and Goldstein (2001) (CDG), including a discussion of some empirical results obtained by Huang and Huang (2004).

3.1 Stationary Leverage Ratio

In the CDG model it is assumed that the firm-value V_t follows a geometric Brownian motion and that the dynamics of the spot rate r_t are those of Vasicek (1977). Let us define the log-firm value $y_t \equiv \log(V_t)$. The risk-neutral dynamics of these factors are given by:

$$dy_t = \left(r_t - \delta - \frac{\sigma^2}{2} \right) dt + \sigma_1 dW_1(t) \quad (1)$$

and:

$$dr_t = k_r^Q (\bar{r}^Q - r_t) dt + \sigma_2 dW_2(t) \quad (2)$$

with $dW_1(t)dW_2(t) = \rho dt$.

In the CDG model the default boundary is given by the nominal debt level. Where the dynamics of the log-nominal debt k_t are assumed to be given by:

$$dk_t = \kappa [y_t - \nu - \phi(r_t - \bar{r}) - k_t] dt, \quad (3)$$

From the above equation we see that k_t is mean-reverting, where the mean-reversion level or target level is given by:

$$y_t - \nu - \phi(r_t - \bar{r}), \quad (4)$$

where it is assumed that $\phi \geq 0$. This last assumption makes the target level of the amount of outstanding debt k_t decreasing in the spot rate, in line with Malitz (1994) who finds that debt issuances decreased dramatically during the high interest rate period in the early 1980's.

As in most barrier-style structural models, default is triggered by the event that the log-firm value hits a stochastic lower boundary, of which in this case the dynamics are given by equation 3. However, restating the model in terms of the log-leverage ratio rather than log-firm value, the default boundary becomes constant, more precisely it is equal to zero. The log-leverage l_t process is given by:

$$l_t = k_t - y_t. \quad (5)$$

Applying Ito's lemma we find that the risk-neutral dynamics of l_t are given by:

$$dl_t = \lambda \left(\bar{l}^Q(r_t) - l_t \right) dt - \sigma_1 dW_1(t), \quad (6)$$

where $\bar{l}^Q(r_t)$ has been defined as:

$$\bar{l}^Q(r) \equiv \frac{\delta + \sigma^2/2}{\lambda} - \nu + \phi\theta - r \left(\frac{1}{\lambda} + \phi \right).$$

From the above equation one sees that the log-leverage ratio is mean-reverting, and hence stationary, with mean-reversion level $\bar{l}^Q(r_t)$.

The assumption that default is triggered when the log-firm value process y_t reaches the lower boundary determined by the amount of outstanding debt k_t , is clearly equivalent to default occurring when the log-leverage ratio l_t reaches the upper boundary level zero. Therefore, one way to think of the CDG model is to see it as an extension of or variation on the original Longstaff and Schwartz (1995) model.

3.2 Pricing Corporate Bonds

Default occurs at the first point in time at which $l(t)$ reaches zero. Therefore, the probability under the T -forward measure that the firm will default prior to time T is given by:

$$\mathbf{Q}^T(r_0, l_0, T) \equiv E_0^T [\mathbf{1}_{\{\tau \leq T\}}],$$

with τ the hitting time for l_t for the boundary value zero. When default occurs, bondholders receive a claim to the fraction $(1 - \omega)$ of the principal at the time of maturity T . Therefore, a defaultable discount bond with maturity date T can be seen as a security that has a pay-off $H(T)$ at time T given by:

$$\begin{aligned} H(T) &= \mathbf{1}_{\{\tau > T\}} + (1 - \omega)\mathbf{1}_{\{\tau \leq T\}} \\ &= 1 - \omega\mathbf{1}_{\{\tau \leq T\}}. \end{aligned}$$

As a result, the price of a corporate zero-coupon bond with maturity date T is equal to:

$$\begin{aligned} P(T, r_t, y_t) &= D(T, r_t) E_0^T [1 - \omega\mathbf{1}_{\{\tau \leq T\}}] \\ &= D(T, r_t) [1 - \omega\mathbf{Q}^T(r_0, l_0, T)]. \end{aligned} \quad (7)$$

Where, as earlier, the expectation is taken under the T -forward measure.

We see that in order to have an expression for the price of a defaultable bond, one essentially needs an expression for the default probability $\mathbf{Q}^T(r_0, l_0, T)$. Collin-Dufresne and Goldstein (2001) obtain the following result:

Proposition 1 *Discretize time into n_T equal intervals of length Δt and define $t_i = i\Delta t$. Discretize the r -space by dividing the interval between some chosen minimum \underline{r} and maximum \bar{r} into n_r equal intervals of length Δr and define $r_l = l\Delta r$. The price of a risky discount bond is given by equation 7, where:*

$$\mathbf{Q}^T(r_0, l_0, T) = \sum_{k=1}^{n_T} \sum_{j=1}^{n_r} q^T(r_j, t_k),$$

with $\forall j \in (1, 2, \dots, n_r)$:

$$q^T(r_j, t_1) = \Delta f \Delta r \Psi^T(r_j, t_1)$$

and $\forall k \in (2, \dots, n_T), \forall j \in (1, 2, \dots, n_r)$:

$$q^T(r_j, t_k) = \Delta r \left[\Psi^T(r_j, t_k) - \sum_{u=1}^{k-1} \sum_{w=1}^{n_r} q^T(r_w, t_u) \Phi^T(r_j, t_k | r_w, t_u) \right].$$

Where the functions $\Psi^T(r_t, t)$ and $\Phi^T(r_t, t | r_s, t_s)$ are given by:

$$\Psi^T(r_t, t) = \pi^T(r_t, t | r_0, 0) N \left(\frac{\mu^T(l_t, t | r_t, l_0, 0)}{\Sigma^T(l_t, t | r_t, l_0, 0)} \right),$$

and:

$$\begin{aligned} & \Phi^T(r_t, t | r_s, s) \\ = & \pi^T(r_t, t | r_s, s) N \left(\frac{\mu^T(l_t, t | r_t, l_s = 0, r_s, s)}{\Sigma^T(l_t, t | r_t, l_s = 0, r_s, s)} \right) \quad \forall t > s. \end{aligned}$$

Here $\pi^T(r_t, t | r_s, s)$ is the transition density for the interest rate, and $\mu^T(l_t, t | r_t, l_s, s)$ and $\Sigma^T(l_t, t | r_t, l_s, s)$ are the expected value and variance at time t_s of l_t conditional on r_s, l_s and r_t .

Having reached this point, we move on to a discussion of some empirical results for structural models in general and the CDG model in particular.

3.3 The Role of the Interest Rate Dynamics

As already mentioned in the introduction, structural models seem to generate too low values for credit spreads. A number of more recent extensions of the original models of Merton (1974) and Black and Cox (1976) are able to generate higher levels of credit spreads, among others the model of Anderson and Sundaresan (1996) incorporating strategic default by equityholders, the model of Leland and Toft (1996) allowing for an endogenous default boundary and the CDG model discussed earlier on. However, Huang and Huang (2003) calibrate a wide range of structural models to the physical default frequencies of corporate bonds of various rating categories and find that they generate similarly low values for credit spreads. This robust result seems to suggest that default risk might only account for a small fraction of the credit spread. But in a recent paper Longstaff et al. (2004), using a reduced-form model, show that default risk can account for the majority of the corporate bond spread. Hence, the result of Huang and Huang (2003) rather seems to underline the failure, so far, of structural models to generate adequate levels for credit spreads. The following analysis might shed some light on why this could be the case.

In most structural models the physical dynamics of the firm-value process are given by an equation of the following form, possibly with an additional jump term:

$$dV_t = (\mu - \delta) dt + \sigma_1 dW_1(t), \quad (8)$$

As discussed earlier on, the credit spreads implied by a structural model are determined by the risk-neutral dynamics of $V(t)$, which are given by:

$$dV_t = (r_t - \delta) dt + \sigma_1 dW_1(t) \quad (9)$$

As the drift term in the above equation is driven by the spot rate, the term structure dynamics affect the risk-neutral dynamics of the firm-value process. If we assume that the dynamics of the spot rate are those of Vasicek (1977), then the physical dynamics of $r(t)$ are given by:

$$dr_t = k_r(\bar{r} - r_t)dt + \sigma_2 dW_2(t), \quad (10)$$

and for the risk-neutral dynamics one obtains:

$$dr_t = k_r(\bar{r}^Q - r_t)dt + \sigma_2 dW_2(t), \quad (11)$$

with:

$$\bar{r}^Q = \bar{r} + \sigma_2^2 \lambda,$$

with λ the price-of-risk for the spot rate. If we further assume that default occurs when the firm-value reaches a lower boundary level K , then the above set-up is that of Longstaff and Schwartz (1995).

Let us take the approach of Huang and Huang (2003) for calibrating the physical dynamics of V_t . For a corporate bond of a given maturity and rating category, the pay-out rate δ is set to some fixed value, and the value for the volatility σ_2 is chosen such that the default probability implied by the physical dynamics of $V(t)$ is equal to the observed historical default frequency for the appropriate rating class. Note that the physical dynamics of the spot rate play no role in this part of the calibration.

In order to obtain the implied credit spread for the given bond, we need the risk-neutral dynamics for the spot rate. Assuming that equation 10 has been calibrated to some observed proxy for the spot rate, we still need to obtain an estimate for the price-of-risk λ , in order to obtain an estimate for the mean-reversion level under the risk-neutral measure.

Observe that the price of a zero-coupon bond with maturity T is given by:

$$P(0, T) = E_0^Q \left[\exp \left(- \int_0^T r(u) du \right) \right] < E_0^P \left[\exp \left(- \int_0^T r(u) du \right) \right].$$

Where Q and P indicate that the expectations are taken under the risk-neutral and physical measure respectively, and the strict inequality is the result of the fact that interest rate risk is priced.

The only way this inequality can be obtained in the Vasicek model is by assuming a strictly positive value for λ . Which effectively increases the (conditional) mean of $\int_0^T r(u) du$ relative to that under the physical measure. However, as $r(t)$ drives the drift of $V(t)$ under the risk-neutral measure, this approach has an unwanted side-effect in the context of a structural model for corporate bonds. The higher the mean-reversion level under the risk-neutral measure, the lower the risk-neutral default probability will be, and hence the smaller the implied credit spread. Put differently, in the Vasicek model there is a very strong link between the physical and risk-neutral dynamics of the spot rate which carries over to the risk-neutral dynamics of the firm-value process $V(t)$. Therefore, a way to obtain the above inequality other than increasing the mean-reversion level would be more than welcome.

In the next section we introduce a class of term-structure models for which this link is less tight, and which allows for more flexibility between the physical and risk-neutral dynamics of the spot rate.

4 Essentially Affine Term Structure Models

In this section we first give a short introduction to the class of essentially affine term structure models of Duffee (2002) before moving on to the derivation a specific two-factor model.

4.1 General Discussion

In its most general form an affine term structure model is determined by n Brownian motions, $\widetilde{W}_t = (\widetilde{W}_{t,1}, \dots, \widetilde{W}_{t,n})$ and n state variables: $X_t = (X_{t,1}, \dots, X_{t,n})$. The spot rate is an affine function of the n state variables:

$$r_t = b_0 + bX_t,$$

with b_0 a scalar and b an n -vector. The dynamics of the state variables under the equivalent martingale measure are:

$$dX_t = [a^Q - B^Q X_t] dt + \Sigma S_t d\widetilde{W}_t,$$

with a^Q an n -vector and B^Q and Σ two $n \times n$ matrices. The matrix S_t is a diagonal matrix, with elements:

$$S_{t,ii} = \sqrt{\alpha_i + \beta'_i X_t},$$

where β_i is an n -vector and α_i is a scalar.

The description of an affine model is completed by the specification of the price of risk vector Λ_t . Given Λ_t the dynamics of X_t under the physical measure are given by:

$$dX_t = [a^Q - B^Q X_t] dt + \Sigma S_t \Lambda_t dt + \Sigma S_t dW_t,$$

In an affine model the market price of risk is of the form:

$$\Lambda_t = S_t \lambda. \tag{12}$$

This specification of Λ_t guarantees affine dynamics for X_t under both the physical and risk-neutral measure and it implies that also $\Lambda'_t \Lambda_t$, the instantaneous variance of the log state price deflator, is affine in X_t .

Duffee (2002) observes that the above specification of the price of risk vector creates a strong link between the risk-neutral and the physical dynamics. In order to increase the ability of affine models to fit certain features of Treasury yields, Duffee (2002) proposes a more general specification of the price of risk vector. Let us first define the matrix S_t^- as:

$$S_{t,ii}^- = \begin{cases} (\alpha_i + \beta'_i X_t)^{-1/2}, & \text{if } \inf(\alpha_i + \beta'_i X_t) > 0 \\ 0, & \text{otherwise} \end{cases}.$$

For an essentially affine model the price of risk vector is given by:

$$\Lambda_t = S_t \lambda_1 + S_t^- \lambda_2 X_t. \tag{13}$$

With such a specification for the price of risk vector, the tight link between the physical and the risk-neutral dynamics has been broken. This gives a term structure model more flexibility in capturing features of both the physical and the risk-neutral dynamics of the spot rate.

4.2 A Two-factor model with a Stochastic Risk-premium

Here we present a Gaussian version of a model presented by Duffee (2002). There are two state variables: the spot rate r_t and a second factor f_t of which the dynamics under the physical measure are given by:

$$\begin{aligned} df_t &= k_f(\bar{f} - f_t)dt + \sigma_f dW_{t,1} \\ dr_t &= k_r(\bar{r} - r_t)dt + \sigma_1 dW_{t,1} + \sigma_2 dW_{t,2} \end{aligned} \quad (14)$$

with W_1 and W_2 independent. From the above equation one sees that the factor f_t has no impact on the dynamics of the spot rate under the physical measure. If one would restrict the price of risk vector to the completely affine specification given by equation 12, then the factor f_t would not have any impact on the dynamics of r_t under the risk-neutral measure either. That is, in this case f_t would be irrelevant and the term structure model is the one of Vasicek (1977).

However, with an essentially affine specification of the price of risk vector as given by equation 13, the factor f_t can affect bond prices without affecting the dynamics of r_t under the physical measure. The specification for the price of risk vector is:

$$\Lambda_t = \begin{pmatrix} \sigma_f & 0 \\ \sigma_1 & \sigma_2 \end{pmatrix} \left[\begin{pmatrix} \lambda_1^{(1)} \\ \lambda_2^{(1)} \end{pmatrix} + \begin{pmatrix} \lambda_{11}^{(2)} & 0 \\ \lambda_{21}^{(2)} & \lambda_{22}^{(2)} \end{pmatrix} \begin{pmatrix} f_t \\ r_t \end{pmatrix} \right].$$

With this specification the price of risk vector is stochastic as it is a function of the spot rate itself and the factor f_t . The dynamics under the risk-neutral measure are given by:

$$\begin{aligned} df_t &= (\alpha_f - \beta_f f_t) dt + \sigma_f d\widetilde{W}_{t,1} \\ dr_t &= (\alpha_r - \beta_r f_t - \gamma_r r_t) dt + \sigma_1 d\widetilde{W}_{t,1} + \sigma_2 d\widetilde{W}_{t,2} \end{aligned} \quad (15)$$

with:

$$\begin{aligned} \alpha_f &= k_f \bar{f} - \sigma_f \lambda_1^1 \\ \beta_f &= k_f + \sigma_f \lambda_{11}^2 \\ \alpha_r &= k_r \bar{r} - \sigma_1 \lambda_1^1 - \sigma_2 \lambda_2^1 \\ \beta_r &= \sigma_1 \lambda_{11}^2 + \sigma_2 \lambda_{21}^2 \\ \gamma_r &= k_r + \sigma_2 \lambda_{22}^2. \end{aligned}$$

The value $D(T, r, f)$ of a zero-coupon bond with time to maturity T is given by:

$$D(T, r, f) = \exp(A(T) - rB(T) - fC(T)). \quad (16)$$

Formulas for the functions $A(T)$, $B(T)$ and $C(T)$ are given in Appendix A. Note that we indicate prices of Treasury bonds with a capital letter D . Throughout the paper, the capital letter P is used for corporate bonds.

Having derived a specific essentially affine two-factor term structure model, we now turn to incorporating it in the CDG model.

5 Extension of the Structural Model of CDG (2001)

In this section we derive an 'essentially affine' version of the structural model of Collin-Dufresne and Goldstein (2002). In our version the asset-risk premium is a function of the two factors that drive the stochastic price-of-risk vector in the term structure model of Section 3.2. Our approach is similar to that of Huang and Huang (2003). However, whereas in Huang and Huang (2003) the asset risk-premium is driven by a separate factor, in our model it is a function of the factors that drive the price-of-risk vector in the term structure model.

It is important to realize that the default-mechanism of the original CDG model is unchanged. Default is still triggered by the log-leverage ratio reaching the upper boundary zero, and the equations 3 to 5 still describe the default mechanism.

5.1 The Risk-neutral and Physical Dynamics

Here we assume that the dynamics of spot rate are those of the two-factor model described in Section 4.2. This introduces an extra factor f_t into the framework. Under the risk-neutral measure the firm-asset-value process has the following dynamics :

$$dV_t = (r_t - \delta)V_t dt + V_t \sum_{i=1}^3 \eta_i dW_{t,i}, \quad (17)$$

where δ is the pay-out ratio, r_t is the spot rate, η_1 , η_2 and η_3 are the three (constant) diffusion coefficients and with W_3 a third Brownian motion independent of the two Brownian motions which drive the term structure dynamics. The risk-neutral dynamics of the log-leverage ratio y_t are given by:

$$dy_t = (r_t - \delta - \sum_{i=1}^3 \eta_i^2 / 2) dt + \sum_{i=1}^3 \eta_i dW_{t,i}, \quad (18)$$

The physical dynamics are:

$$dy_t = (\pi_t + r_t - \delta - \sum_{i=1}^3 \eta_i^2/2)dt + \sum_{i=1}^3 \eta_i dW_{t,i}, \quad (19)$$

where the stochastic asset risk premium is given by:

$$\pi_t = \hat{\alpha}_y + \beta_y f_t + \gamma_y r_t. \quad (20)$$

That is, we also allow for the possibility that the risk-premium for the asset value process is a linear function of f_t and r_t , as the risk-premium for the spot rate process itself.

The above specification leads to the following physical dynamics for y_t :

$$dy_t = (\alpha_y + \beta_y f_t + (1 + \gamma_y)r_t) dt + \sum_{i=1}^3 \eta_i dW_{t,i},$$

with:

$$\alpha_y = \hat{\alpha}_y - \delta - \sum_{i=1}^3 \eta_i^2/2.$$

We now turn to deriving the physical and risk-neutral dynamics of l_t . Combining equation 5 and 18 we obtain that the risk-neutral dynamics of l_t are given by:

$$dl_t = \kappa [\bar{l}(r_t) - l_t] dt - \sum_{i=1}^3 \eta_i dW_{t,i}, \quad (21)$$

with the reversion level $\bar{l}(r_t)$ under the risk-neutral measure given by:

$$\bar{l}(r_t) = \phi \bar{r} + \frac{2\delta + \sum_{i=1}^3 \eta_i^2}{2\kappa} - \nu - \left(\phi + \frac{1}{\kappa} \right) r_t.$$

Substituting the risk-premium for y_t from equation 20 we can rewrite equation 21 as:

$$dl_t = [\alpha_l - \gamma_l r_t - \kappa l_t] dt - \sum_{i=1}^3 \eta_i dW_{t,i}, \quad (22)$$

with:

$$\begin{aligned} \alpha_l &= \delta + \frac{\sum_{i=1}^3 \eta_i^2}{2} - \kappa \nu + \kappa \phi \bar{r}, \\ \gamma_l &= \kappa \phi + 1. \end{aligned}$$

Note that equation 22 implies that l_t is mean-reverting with the reversion level $\bar{l}(r_t)$ under the risk-neutral measure given by:

$$\bar{l}(r_t) = \phi\bar{r} + \frac{2\delta + \sum_{i=1}^3 \eta_i^2}{2\kappa} - \nu - \left(\phi + \frac{1}{\kappa}\right) r_t.$$

Combining all of the above, we see that under the risk-neutral measure the dynamics of r_t, f_t, y_t and l_t are given by:

$$d \begin{pmatrix} f_t \\ r_t \\ y_t \\ l_t \end{pmatrix} = \left[a^Q + A^Q \begin{pmatrix} f_t \\ r_t \\ y_t \\ l_t \end{pmatrix} \right] dt + \Sigma d\widetilde{W}_t \quad (23)$$

with:

$$a^Q = \begin{pmatrix} \alpha_f \\ \alpha_r \\ \alpha_y \\ \alpha_l \end{pmatrix}$$

$$A^Q = \begin{bmatrix} -\beta_f & 0 & 0 & 0 \\ -\beta_r & -\gamma_r & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -\gamma_l & 0 & -\kappa \end{bmatrix}$$

Substituting the physical dynamics of the process y_t from equation 19, we obtain that the physical dynamics of l_t are given by:

$$dl_t = [\alpha_l - \widehat{\alpha}_y - \beta_y f_t - (\gamma_y + \gamma_l)r_t - \kappa l_t] dt - \sum_{i=1}^3 \eta_i dW_{t,i}, \quad (24)$$

Therefore, the dynamics of r_t, f_t, y_t and l_t under the physical measure are given by:

$$d \begin{pmatrix} f_t \\ r_t \\ y_t \\ l_t \end{pmatrix} = \left[a + A \begin{pmatrix} f_t \\ r_t \\ y_t \\ l_t \end{pmatrix} \right] dt + dW_t \quad (25)$$

with:

$$a = \begin{pmatrix} k_f \bar{f} \\ k_r \bar{r} \\ \alpha_y \\ \alpha_l - \widehat{\alpha}_y \end{pmatrix},$$

and:

$$A = \begin{bmatrix} -k_f & 0 & 0 & 0 \\ 0 & -k_r & 0 & 0 \\ \beta_y & (1 + \gamma_y) & 0 & 0 \\ -\beta_y & -(\gamma_y + \gamma_l) & 0 & -\kappa \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} \sigma_f & 0 & 0 \\ \sigma_1 & \sigma_2 & 0 \\ \eta_1 & \eta_2 & \eta_3 \\ -\eta_1 & -\eta_2 & -\eta_3 \end{bmatrix}$$

To obtain prices for corporate bonds we need the dynamics of f_t, r_t and l_t under the T -forward measure, we derive these dynamics in the next section.

5.2 The dynamics of f_t, r_t and l_t under the T -forward measure

From the risk-neutral dynamics given by equation 23 and equation 16 one obtains that the risk-neutral dynamics of the price $D(t, T, f, r)$ of a discount bond with maturity date T are given by:

$$\frac{dD(t, T, f_t, r_t)}{D(t, T, f_t, r_t)} = r(t)dt - [\sigma_1 B(T-t) + \sigma_f C(T-t)] d\widetilde{W}_1 - \sigma_2 B(T-t) d\widetilde{W}_2.$$

If we fix a date T^* , we find that under the risk-neutral measure (induced by the spot market account as numéraire):

$$\begin{aligned} & d \left(\frac{D(t, T, f_t, r_t)}{D(t, T^*, f_t, r_t)} \right) \\ = & \frac{D(t, T, f_t, r_t)}{D(t, T^*, f_t, r_t)} \left[m(t, T, T^*) dt + \sigma_2 (B(T^* - t) - B(T - t)) d\widetilde{W}_2 \right. \\ & \left. + [\sigma_1 (B(T^* - t) - B(T - t)) + \sigma_f (C(T^* - t) - C(T - t))] d\widetilde{W}_1 \right], \end{aligned}$$

with:

$$\begin{aligned} m_D(t, T, T^*) &= (\sigma_1 B(T^* - t) + \sigma_f C(T^* - t))^2 + \sigma_2^2 B^2(T^* - t) - \sigma_2^2 B(T^* - t) B(T - t) \\ &\quad - (\sigma_1 B(T^* - t) + \sigma_f C(T^* - t)) (\sigma_1 B(T - t) + \sigma_f C(T - t)). \end{aligned}$$

For the firm-asset value one obtains:

$$\begin{aligned}
& d\left(\frac{V(t)}{D(t, T^*, f_t, r_t)}\right) \\
= & \frac{V(t)}{D(t, T^*, f_t, r_t)} \left[m_V(t, T^*) dt + (\eta_1 + \sigma_1 B(T^* - t) + \sigma_f C(T^* - t)) d\widetilde{W}_1 \right. \\
& \left. + (\eta_2 + \sigma_2 B(T^* - t)) d\widetilde{W}_2 + \eta_3 d\widetilde{W}_3 \right],
\end{aligned}$$

with:

$$\begin{aligned}
m_V(t, T^*) = & -\delta + \eta_1 (\sigma_1 B(T^* - t) + \sigma_f C(T^* - t)) + \eta_2 \sigma_2 B(T^* - t) \\
& + (\sigma_1 B(T^* - t) + \sigma_f C(T^* - t))^2 + \sigma_2^2 B^2(T^* - t).
\end{aligned}$$

We now determine the Radon-Nykodim derivative $d\mathbb{Q}^T/d\mathbb{Q}$ of the T -forward measure with respect to the risk-neutral measure induced by the spot market account as numéraire. We know that $d\mathbb{Q}^T/d\mathbb{Q}$ has the following form:

$$\frac{d\mathbb{Q}^T}{d\mathbb{Q}}(t) = \int_0^t \phi(u) du - \frac{1}{2} \int_0^t |\phi(u)|^2 du,$$

with $\phi(t) = (\phi_1(t), \phi_2(t), \phi_3(t))$. Observe that under the T^* -forward measure the process $D(t, T, f_t, r_t)/D(t, T^*, f_t, r_t)$ has a zero drift and the drift term of $V/D(t, T^*, f_t, r_t)$ is equal to $-\delta dt$. Using Girsanov's Theorem, this leads to two conditions on $\phi(t)$. The first condition is:

$$\begin{aligned}
\forall T \leq T^* \quad m_D(t, T, T^*) + \phi_1(t) [\sigma_1 (B(T^* - t) - B(T - t)) + \sigma_f (C(T^* - t) - C(T - t))] \\
+ \phi_2(t) \sigma_2 (B(T^* - t) - B(T - t)) \equiv 0.
\end{aligned}$$

From this condition we obtain:

$$\phi_1(t) = -(\sigma_1 B(T^* - t) + \sigma_f C(T^* - t)), \tag{26}$$

$$\phi_2(t) = \sigma_2 B(T^* - t). \tag{27}$$

The second condition is given by:

$$m_V(t) + \phi_1(t) (\eta_1 + \sigma_1 B(T^* - t) + \sigma_f C(T^* - t)) + \phi_2(t) (\eta_2 + \sigma_2 B(T^* - t)) + \phi_3(t) \eta_3 \equiv 0.$$

This leads to:

$$\phi_3(t) \equiv 0, \quad (28)$$

which might have been expected since the change of numéraire was from one type of term structure instrument to another. Combining equations 26 to 28 with definition 5 for l_t and Girsanov's Theorem leads to the following dynamics for f_t, r_t and l_t under the T^* -forward measure:

$$d \begin{pmatrix} f_t \\ r_t \\ l_t \end{pmatrix} = \left[a^T(t) + A^T \begin{pmatrix} f_t \\ r_t \\ l_t \end{pmatrix} \right] dt + \Sigma^T d\widehat{W}_t \quad (29)$$

where $(\widehat{W}_1, \widehat{W}_2, \widehat{W}_3)$ is a 3-dimensional standard Brownian motion and with:

$$a_f^T(t) = \alpha_f - \sigma_f \sigma_1 B(T^* - t) - \sigma_f^2 C(T^* - t) \quad (30a)$$

$$a_r^T(t) = \alpha_r - (\sigma_1^2 + \sigma_2^2) B(T^* - t) - \sigma_1 \sigma_f C(T^* - t) \quad (30b)$$

$$a_l^T(t) = \alpha_l + (\eta_1 \sigma_1 + \eta_2 \sigma_2) B(T^* - t) + \eta_1 \sigma_f C(T^* - t). \quad (30c)$$

The feedback matrix A^T is given by:

$$A^T = \begin{bmatrix} -\beta_f & 0 & 0 \\ -\beta_r & -\gamma_r & 0 \\ 0 & -\gamma_l & -\kappa \end{bmatrix} \quad (31)$$

and the matrix Σ^T is:

$$\Sigma^T = \begin{bmatrix} \sigma_f & 0 & 0 \\ \sigma_1 & \sigma_2 & 0 \\ -\eta_1 & -\eta_2 & -\eta_3 \end{bmatrix}. \quad (32)$$

Notice that the T -forward dynamics of l_t do not directly depend on the factor f_t . That is, the lower left element of the feedback matrix A^T is equal to zero. Having obtained the T -forward dynamics of f_t, r_t and l_t , we are ready to price corporate bonds.

5.3 Pricing Corporate Bonds

As in Section 3.2 the price of a corporate bond is determined by the default probability under the T -forward measure. A result similar to that of Proposition 1 is given below.

Proposition 2 *Discretize time into n_T equal intervals of length Δt and define $t_i = i\Delta t$. Discretize the f -space by dividing the interval between some chosen minimum \underline{f} and maximum \bar{f} into n_f equal intervals of length Δf and define $f_k = k\Delta f$. Similarly, discretize the r -space by dividing the interval between some chosen minimum \underline{r} and maximum \bar{r} into n_r equal intervals of length Δr and define $r_l = l\Delta r$.*

The default probability $\mathbf{Q}^{\mathbb{M}}(r_0, f_0, l_0, T)$ under a measure \mathbb{M} is given by:

$$\mathbf{Q}^{\mathbb{M}}(r_0, f_0, l_0, T) = \sum_{k=1}^{n_T} \sum_{i=1}^{n_f} \sum_{j=1}^{n_r} q^{\mathbb{M}}(f_i, r_j, t_k),$$

with $\forall i \in (1, 2, \dots, n_f), \forall j \in (1, 2, \dots, n_r)$:

$$q^{\mathbb{M}}(f_i, r_j, t_1) = \Delta f \Delta r \Psi^{\mathbb{M}}(f_i, r_j, t_1)$$

and $\forall k \in (2, \dots, n_T), \forall i \in (1, 2, \dots, n_f), \forall j \in (1, 2, \dots, n_r)$:

$$q^{\mathbb{M}}(f_i, r_j, t_k) = \Delta f \Delta r \left[\Psi^{\mathbb{M}}(f_i, r_j, t_k) - \sum_{u=1}^{k-1} \sum_{v=1}^{n_f} \sum_{w=1}^{n_r} q^{\mathbb{M}}(f_v, r_w, t_u) \Phi^{\mathbb{M}}(f_i, r_j, t_k | f_v, r_w, t_u) \right].$$

Where the functions $\Psi^{\mathbb{M}}(f_t, r_t, t)$ and $\Phi^{\mathbb{M}}(f_t, r_t, t | f_s, r_s, t_s)$ are given by:

$$\Psi^{\mathbb{M}}(f_t, r_t, t) = \pi^{\mathbb{M}}(f_t, r_t, t | f_0, r_0, 0) N \left(\frac{\mu^{\mathbb{M}}(l_t, t | f_t, r_t, l_0, f_0, 0)}{\Sigma^{\mathbb{M}}(l_t, t | f_t, r_t, l_0, f_0, 0)} \right),$$

and:

$$\begin{aligned} & \Phi^{\mathbb{M}}(f_t, r_t, t | f_s, r_s, s) \\ = & \pi^{\mathbb{M}}(f_t, r_t, t | f_s, r_s, s) N \left(\frac{\mu^{\mathbb{M}}(l_t, t | f_t, r_t, l_s = 0, f_s, r_s, s)}{\Sigma^{\mathbb{M}}(l_t, t | f_t, r_t, l_s = 0, f_s, r_s, s)} \right) \quad \forall t > s. \end{aligned}$$

Proof :

A proof of this proposition and expressions for $\pi^{\mathbb{M}}, \mu^{\mathbb{M}}$ and $\Sigma^{\mathbb{M}}$ are given in Appendix B. ■

If \mathbb{M} is equal to the physical measure, then $\mathbf{Q}^{\mathbb{P}}(r_0, f_0, l_0, T)$ is the physical default probability or default frequency. If, alternatively, \mathbb{M} is equal to the T -forward measure, then $\mathbf{Q}^T(r_0, f_0, l_0, T)$ determines the yield spread on a bond, as can be seen as follows:

$$-\frac{\ln(P(T, r_t, f_t, y_t))}{T} = -\frac{\ln(D(T, r_t, f_t))}{T} - \frac{\ln[1 - \omega \mathbf{Q}^T(r_0, f_0, l_0, T)]}{T}.$$

We see that the yield spread $sp(T, r_t, f_t, y_t)$ on a defaultable discount bond with maturity T over the yield on a default-free bond is given by:

$$sp(T, r_t, f_t, y_t) = -\frac{\ln [1 - \omega \mathbf{Q}^T(r_0, f_0, l_0, T)]}{T}.$$

Having derived these last expressions, we are almost ready to test the model. In the next section we calibrate two versions of the term structure model derived in Section 4, before we move on to testing the 'affine structural model' in Section 7.

6 Two Versions of the Term-Structure Model

Here we calibrate two term structure models which we will use in the next section to test the structural model. For the first term structure model we assume that $\lambda_{21}^2 = 0$. Under this assumption the two-factor model collapses into a one-factor one, with the price-of-risk now being driven by the spot rate itself only. For the second model we impose the restriction that $\lambda_{22}^2 = 0$, as a result of which the price-of-risk vector is only driven by the factor f_t .

6.1 The One-factor Version

6.1.1 The Model

As mentioned above, for the one-factor version of the model we assume that the price-of-risk for the spot rate r_t is a function of r_t only. However, because the factor f_t plays no role in the physical dynamics of r_t it drops out of the model altogether, and we are left with a one-factor Vasicek-style dynamics. The difference with the original Vasicek model is that the price-of-risk for the spot rate is still a function of r_t , where in the original model it is a constant.

The physical dynamics of r_t are given by:

$$dr_t = k_r^P (\bar{r}^P - r_t) dt + \sigma_2^P dW_{t,2}. \quad (33)$$

Where the absence of a second Brownian motion is the only difference with equation 14. Notice that in order to keep the notation similar to that used in Section 4, the sole Brownian motion has index 2 and not 1. The same applies to some of the parameters. The superscript 'p' indicates that these are the physical dynamics. The risk-neutral dynamics of r_t , given by equation 15, collapse into:

$$dr_t = k_r^Q (\bar{r}^Q - r_t) dt + \sigma_2^P dW_{t,2}, \quad (34)$$

with:

$$k_r^Q = k_r^P + \sigma_2^P \lambda_{22}^2 \quad (35)$$

and:

$$\bar{r}^Q = \frac{k_r^P \bar{r}^P - \sigma_2^P \lambda_2^1}{k_r^P + \sigma_2 \lambda_{22}^2}. \quad (36)$$

The price $D(t, T)$ at time t of a zero-coupon bond with maturity date T is given by the well known formula:

$$D(t, T) = e^{-A(T-t) - B(T-t)r_t},$$

with the functions $B(\tau)$ and $A(\tau)$ given by:

$$B(\tau) = \frac{1 - e^{-k_r^Q \tau}}{k_r^Q},$$

$$A(\tau) = \frac{(B(\tau) - 1) \left((k_r^Q)^2 \bar{r}^Q - (\sigma_r^P)^2 / 2 \right)}{(k_r^Q)^2} - \frac{(\sigma_r^P)^2 \tau B(\tau)^2}{4 (k_r^Q)}.$$

Notice that under the restriction $\lambda_{22}^2 = 0$ which is the case for the original Vasicek model, one has that $k_r^Q = k_r^P$.

6.1.2 Calibration of the Physical Dynamics of r_t

We calibrate the term structure model to yield data from the period September 1976 until December 1997, in total 256 months. As a proxy for the spot rate process we use the 6-month interest rate. Calibration of the spot rate process is done by means of the Method of Moments. For which we use the two following expressions for the conditional expected value and conditional variance of r_t :

$$E_t^{\mathbb{P}} [r_{t+1}] = \bar{r}^P + e^{-k_r^P \Delta t} (r_t - \bar{r}^P) \quad (37)$$

$$var_t^{\mathbb{P}} [r_{t+1}] = \frac{(\sigma_2^P)^2 \left(1 - e^{-2k_r^P \Delta t} \right)}{2k_r^P}. \quad (38)$$

Using the exact expressions for $E_t^{\mathbb{P}} [r_{t+1}]$ and $var_t^{\mathbb{P}} [r_{t+1}]$ instead of discretizing equation 33 has the advantage that no approximation or discretization error is introduced into the estimation. Next, define:

$$\epsilon_{1,t} = r_{t+1} - E_t^{\mathbb{P}}[r_{t+1}] \quad (39)$$

$$\epsilon_{2,t} = (r_{t+1})^2 - \text{var}_t^{\mathbb{P}}[r_{t+1}] + \left(E_t^{\mathbb{P}}[r_{t+1}]\right)^2. \quad (40)$$

Similar to Chan et. al (1992) the moment equations are given by:

$$E[\epsilon_{1,t}Z_{t-1}] = 0$$

$$E[\epsilon_{2,t}Z_{t-1}] = 0.$$

As instrumental variables we use a constant, (the proxy for) the spot rate and the yield on a 7-year discount bond. This results in six moment conditions in three parameters, k_r , \bar{r} and σ_r . The spectral density matrix S was estimated using the estimator of Newey-West (1987) with 12 lags. The results of the estimation are given in Table 1.

Table 1: The Physical dynamics of r_t

	k_r	\bar{r}	σ_r		TJ
value	0.2769	0.0681	0.0206	value	1.11
s.e.	0.173	0.0313	$6.53E - 03$	p -value	0.76

First, we observe that the diffusion coefficient σ_2^P and the reversion-level \bar{r}^P are both significant at the 2.5% level, but that the mean-reversion speed k_r^P is only significant at the 5% level. Secondly, the p -value for the χ^2 -test on overidentifying restrictions is equal to 0.78, so the model seems able to match the imposed moment conditions.

6.1.3 Calibration of the Price-of-Risk Vector

The coefficients of the price-of-risk are estimated by calibrating the risk-neutral dynamics to the 5-year yield (for the same 257 months as above), given the (proxy for the) spot rate. For the essentially affine version of the Vasicek model one obtains the following estimates:

$$\lambda_2^1 = -0.3337$$

$$\lambda_{22}^2 = -4.4321$$

Which in turn leads to:

$$\begin{aligned}k_r^Q &= 0.1856 \\ \bar{r}^Q &= 0.1051.\end{aligned}$$

For the original Vasicek model (with $\lambda_{22}^2 = 0$) one obtains:

$$\lambda_2^1 = -1.0402,$$

which yields:

$$\bar{r}^Q = 0.1455.$$

6.2 The Two-factor Version

Here we calibrate a two-factor version of the term structure model derived in Section 4, with the restriction that $\lambda_{22}^2 = 0$. In this case, the price-of-risk vector in the two-factor model is no longer a function of r_t but only of f_t . Under this restriction the two-factor model is essentially the original Vasicek model to which an exogenous price-of-risk has been added.

6.2.1 The Econometric Model

Before moving on to the actual estimation, we first make the following observation. If one assumes that $\lambda_{22}^2 = 0$, one can back out the unobserved factor f_t from observed bond prices or yields. Let us select two bonds with fixed maturities T_1 and T_2 , of which the prices are given by:

$$D_1(T, r, f) = \exp(A(T_1) - rB(T_1) - fC(T_1))$$

and:

$$D_2(T, r, f) = \exp(A(T_2) - rB(T_2) - fC(T_2)).$$

A bit of algebra shows that the following equality holds:

$$B(T_1) \log(P(T_2)) - B(T_2) \log(P(T_1)) \tag{41}$$

$$= [B(T_1)A(T_2) - B(T_2)A(T_1)] - f [B(T_2)C(T_1) - B(T_1)C(T_2)]. \tag{42}$$

As in Balduzzi et. al (2000), the above equality allows us to construct a proxy for f_t . Rewriting equation 41 one obtains the following expression for the factor f :

$$f = a_0 + a_1 [T_1 B(T_2) y(T_1) - T_2 B(T_1) y(T_2)], \quad (43)$$

with $y(T_1)$ and $y(T_2)$ the yields to maturity of the two bonds and a_0 and a_1 two unknown constants. Because the function $B(T)$ is completely determined by the mean-reversion speed k_r of the spot rate, obtaining a linear transformation of the process f_t , once the physical dynamics of r_t have been estimated, is a straightforward exercise. Of course, one still needs estimates for the constant a_0 and a_1 in order to obtain a proxy for the process f_t itself. However, as demonstrated in the next paragraph, we only need to determine the process f_t up to a linear transformation.

Remember that the physical dynamics of the 2-factor model are given by:

$$\begin{aligned} df_t &= k_f(\bar{f} - f_t)dt + \sigma_f dW_{t,1} \\ dr_t &= k_r(\bar{r} - r_t)dt + \sigma_1 dW_{t,1} + \sigma_2 dW_{t,2} \end{aligned} \quad (44)$$

with W_1 and W_2 independent Brownian motions, and that the specification for the price of risk vector is:

$$\begin{aligned} \Lambda_t &= \begin{pmatrix} \sigma_f & 0 \\ \sigma_1 & \sigma_2 \end{pmatrix} \left[\begin{pmatrix} \lambda_1^{(1)} \\ \lambda_2^{(1)} \end{pmatrix} + \begin{pmatrix} \lambda_{11}^{(2)} & 0 \\ \lambda_{21}^{(2)} & \lambda_{22}^{(2)} \end{pmatrix} \begin{pmatrix} f_t \\ r_t \end{pmatrix} \right] \\ &= \begin{bmatrix} \sigma_f (\lambda_1^1 + \lambda_{11}^2 f_t) \\ \sigma_1 (\lambda_1^1 + \lambda_{21}^2 f_t) + \sigma_2 (\lambda_2^1 + \lambda_{22}^2 r_t) \end{bmatrix}. \end{aligned} \quad (45)$$

Let us first turn to the physical dynamics. Note that if f_t follows a Ornstein-Uhlenbeck process then so will any non-trivial linear transformation of it, and vice versa. Therefore, if equation 44 is an adequate description of the dynamics of the process (f_t, r_t) , then it is also and adequate description of the dynamics of $(a_0 + a_1 f_t, r_t)$, albeit for different values of \bar{f} and σ_f .

Let us now assume that we are given a non-trivial linear transformation of f_t instead of the real process. Note that both components of the price of the risk vector are linear in f_t . A bit of algebra shows that the values of $\lambda_1^1, \lambda_2^1, \lambda_{11}^2$ and λ_{21}^2 can always be chosen such that the transformation of f_t is off-set as far as the dynamics of r_t are concerned. That is, the effect of the linear transformation of f_t on the risk-neutral dynamics of r_t is "undone".

The parameter λ_{22}^2 plays no role in all of this. Therefore, the above result also holds if one imposes the additional restriction that $\lambda_{22}^2 \equiv 0$.

6.2.2 The Data

Again, we calibrate the term structure model to yield data from the period: September 1976 until December 1997, in total 256 months. As earlier on, we use the 6-month interest rate as a proxy for the spot rate. The two maturities T_1 and T_2 used to obtain the non-physical factor f_t are one year and seven years. To estimate the four components of the price of risk vector we use yields on discount bonds with maturities of 1,2,5,7 and 10 years. To obtain those yields, we used the following procedure. For each month we started from the par-yields for the same five maturities, which are available from the website of the Federal Reserve. Using cubic-spline interpolation we obtained the par-yield curve at 6-month intervals. From this last curve we obtained the zero-coupon curve using a bootstrapping procedure.

6.2.3 The physical Dynamics of r_t and f_t

Because the factor f_t plays no role in the physical dynamics of the model, i.e. under the physical measure the model is a one-factor model, the results from the calibration of the physical dynamics of r_t in the previous section are also valid here, and we can immediately move on to the estimation of the f_t factor.

Using the approach described above, we can obtain an estimate (up to a linear transformation) for the second factor from the yields on any two government bonds. To check for robustness, two different specifications for f_t are calibrated, one estimate is based on the yields on the 1-year and 7-year discount bond, the other uses the 2-year and 10-year discount bond.

From equation 43 we get a series for f_t , which we use to estimate the remaining four parameters k_f, \bar{f}, σ_f and ρ . Again, we use the analytical expressions for the moments involved:

$$\begin{aligned} E_0^{\mathbb{P}} [f_t] &= \bar{f} + e^{-k_f t} (f(0) - \bar{f}) \\ var_0^{\mathbb{P}} [f_t] &= \frac{\sigma_f^2 (1 - e^{-2k_f t})}{2k_f} \\ covar_0^{\mathbb{P}} [f_t, r_t] &= \frac{\sigma_f \sigma_1 (1 - e^{-(k_f + k_r)t})}{k_f + k_r}. \end{aligned}$$

Where we need to keep in mind that for given values of σ_r, σ_f and ρ , the parameters σ_1 and σ_2 are respectively equal to $\sigma_r \rho$ and $\sigma_r \sqrt{1 - \rho^2}$. Using the same three instruments as in the previous section leads to nine moment conditions in four parameters, k_f, \bar{f}, σ_f and ρ . Again, the spectral density matrix S was estimated using the estimator of Newey-West (1987) with 12 lags. The results of the estimation are given in Table 2.

Table 2: The Physical dynamics of f_t

	k_f	\bar{f}	σ_f	ρ		TJ
	$T_1 = 1$ and $T_2 = 7$					
value	0.0803	0.0504	0.0179	-0.6291	value	6.47
s.e.	0.1035	0.0640	$2.33E - 03$	0.4421	p -value	0.26
	$T_1 = 2$ and $T_2 = 10$					
value	0.0755	0.0333	0.0166	-0.4808	value	5.87
s.e.	0.1001	0.0613	$2.26E - 03$	0.4688	p -value	0.32

We see that in contrast to the diffusion coefficient, the estimates for two parameters of the drift term and the correlation ρ are not statistically significant for both specifications of f_t . The χ^2 -test on over-identifying restrictions leads to a p -value of 0.26 and 0.32 respectively, so the model is able to match the moment conditions for both estimates for f_t .

6.2.4 The Price-of-risk Vector

Having estimated the parameters for the physical dynamics of the two factors of the term-structure model, we need to obtain estimates for the values of the four parameters of the price-of-risk vector Λ_t for both specifications of f_t . To estimate $\lambda_1^1, \lambda_2^1, \lambda_{11}^2$ and λ_{21}^2 we take the same approach as Balduzzi et al. (2000). We minimize the root mean squared price prediction error (RMSE) for discount bonds with maturities of 1, 2, 5, 7 and 10 years. For every month in our sample we compare the actual price of each of the five bonds with the price predicted by the model. We have five observations for each month for 256 months, which results in 1280 observations. Table 3 gives the estimated values for the four coefficients $\lambda_1^1, \lambda_2^1, \lambda_{11}^2$ and λ_{21}^2 that minimize the RMSE for either specification of the factor f_t . The values in Table 3 are comparable in magnitude to those found by Duffee (2002).

From Table 2 and Table 3 one sees that the calibration of the two-factor term structure model is relatively insensitive to the choice for T_1 and T_2 , the two maturities used to estimate

Table 3: The Elements of the Price-of-Risk Vector

λ_1^1	λ_2^1	λ_{11}^2	λ_{22}^2
$T_1 = 2$ and $T_2 = 10$			
0.9604	1.5615	2.7390	7.1661
$T_1 = 2$ and $T_2 = 10$			
0.9218	1.2056	3.4473	6.8694

the factor f_t . Both the parameter values for the physical dynamics of f_t , given by Table 2, as the estimated values for the price-of-risk vector, in Table 3, do not change drastically from one specification to the other.

7 Results for the Structural Model

In this section we test whether allowing for 'essentially affine' interest rate dynamics allows the structural model to do a better job in matching observed yield spreads than the original model of Collin-Dufresne and Goldstein (2002) with the Vasicek interest rate dynamics. More precisely, we test whether they generate 'higher' credit spreads for reasonable parameter values.

Here in the empirical part we assume that the risk-premium for the firm-value process is a constant, i.e. $\gamma_y = \beta_y = 0$. There are two reasons for which we impose this restriction. The first one is that we lack empirical data on the correlation between the asset-value process and the term structure of interest rates; and secondly, we have only estimated the factor f_t up to a linear transformation.

We follow a three-step procedure to test the different versions of the structural model for four classes of bonds. In a first step, we pick the coefficients for the dynamics of the process y_t such that the dynamics of y_t are comparable to those given by Huang and Huang (2003) for the original Collin-Dufresne and Goldstein (2002) model. In the second step we calibrate the dynamics of the debt process k_t such that both the observed historical default frequency and the mean-reversion level of the leverage ratio, as given by Huang and Huang (2003), are matched. Finally we compute the spreads implied by the three different term structure models: the original Vasicek model, the Vasicek model with an essentially affine specification of the price-of-risk and the two-factor model.

Because a Matlab routine based the results in Proposition 1 turned out to be fairly slow, the default probabilities and credit spreads in this section are generated using Monte Carlo simulation instead of the semi-analytic expression given by Proposition 1. Simulations use 120 time steps per year and are based on 10000 sample paths (5000 paths + 5000 antithetic paths). Because of the small time step the downward bias that can result from simulating barrier-style pay-offs should be fairly small. Moreover, the results in this section indicate that under-estimation of default probabilities/credit spreads due to Monte Carlo simulation is probably not a problem here.

For all tables we assumed that the value for the spot rate $r(t_0)$ is equal to 0.0516 and, where applicable, that the value for the second factor at $f(t_0)$ is 0.0492.

Table 4 gives the results from the calibration of the (log) firm-value process y_t for four classes of bonds: the maturity is equal to ten years or four years and the credit rating is Aa or Ba. For the risk-premium we have taken the same values as Huang and Huang (2003), the values for the two diffusion coefficients have been chosen so that the standard deviation of the monthly return and the correlation between the monthly return and the monthly change in the spot rate have the value given in Table 4.

Table 4: The Dynamics of the firm-value process y_t

Credit Rating	α_y	η_2	η_3	sdev(Δy_{t+1})	$\rho(\Delta y_{t+1}, \Delta r_{t+1})$
Maturity = years 10					
Aa	0.0495	-0.0686	0.26	0.268	-0.252
Ba	0.0547	-0.0686	0.31	0.318	-0.214
Maturity = years 4					
Aa	0.0491	-0.0686	0.293	0.300	-0.226
Ba	0.0526	-0.0686	0.285	0.294	-0.233

The results from the second step are given in Table 5. Again, for reasons of comparability, the parameter values for the dynamics of k_t have been chosen in line with Huang and Huang (2003).

The values for the initial leverage ratio are identical to those of Huang and Huang (2003), and the values for κ and ν are chosen to match the values of the mean-reversion level of the leverage ratio¹ and the historical default frequency. We have set the value of ϕ equal to 2 for

¹The leverage ratio at time t of a firm is given by $\exp(l_t)$. Therefore, the mean-reversion level of the

Table 5: The Dynamics of the Debt-level Process k_t

Credit Rating	l_0 (%)	κ	ϕ	ν	\bar{l}_∞	PD (%)
Maturity = 10 years						
Aa	21.2	0.4300	2.0	0.9177	0.38	0.98
Ba	53.5	0.3053	2.0	0.9270	0.38	20.03
Maturity = 4 years						
Aa	21.2	0.4000	2.0	0.9380	0.38	0.22
Ba	53.5	0.2670	2.0	0.9012	0.38	8.42

the four classes of bonds, a value close to that used by Collin-Dufresne and Goldstein (2002).

The credit spreads generated under the different term-structure dynamics are given by Table 6.

Table 6: Implied Credit Spreads

Credit Rating	Avg. Yield Spread (bps)	Org. Vas. Calc. Spread (bps)	Ess. Aff. Vas. Calc. Spread (bps)	2-fact. Model Calc. Spread (bps)
Maturity = 10 years				
Aa	91	5	14	40
Ba	320	135	169	250
Maturity = 4 years				
Aa	65	8	9.0	14
Ba	320	175	190	230

The second column gives the historical average yield spreads, columns four to six give the credit spreads implied by the three term structure models. The spreads calculated using the original Vasicek model are, as expected, very similar to those given obtained by Huang and Huang (2003) for the Collin-Dufresne and Goldstein (2002) model. The results in column four show that allowing for essentially affine interest rate dynamics leads to higher implied credit spreads for all four classes of bonds. With the biggest difference for the 10 year Aa bonds, where the credit spread is increased by a factor three. Finally, from the last column we see that with the two-factor term structure model implied credit spreads increase still

leverage ratio is not equal to that of the process l_t .

further. Here the choice for the two maturities T_1 and T_2 for the estimation of f_t are $T_1 = 1$ and $T_2 = 7$.

To check for the robustness of this last result, Table 7 compares the implied credit spreads for the two different proxies for f_t .

Table 7: Robustness w.r.t. the specification of f_t

Credit Rating	$T_1 = 1, T_2 = 7$ (bps)	$T_1 = 2, T_2 = 10$ (bps)
Maturity = 10 years		
Aa	40.89	45.80
Ba	245.79	257.89
Maturity = 4 years		
Aa	14.35	14.96
Ba	226.64	230.23

We observe that the credit spreads generated under the two specifications for f_t are very similar. Therefore, the result that the two-factor term structure model gives rise to significantly higher the credit spreads seems to be robust to the precise specification of the factor f_t .

8 A Second Look at the Factor f_t

In the previous sections we have seen that the factor f_t improves the ability of the term structure model to match observed Treasury prices and it enables the CDG model to generate higher levels for credit spreads. Moreover, both effects seem robust with respect to the specification used to estimate the factor f_t . Can we say anything more about this factor? In attempt get some feeling for what the factor f_t might be a proxy for we compare it with a number of factors from the asset pricing literature.

8.1 Asset Pricing factors

A first set of factors are the equity based factors from Fama-French (1993): the market factor, the HML and the SMB factor. To this we add three factors derived from bond returns.

The first bond factor: TERM is the difference between the par yield on the 10-year Treasury bond and the par-yield for a maturity of two years. Fama and French (1993) show

that the three equity-based factors together with the a term-factor are not able to explain the cross-section of corporate bond returns, and that a seperate default factor needs to be added.

In order to obtain a factor that captures default risk, we construct two portfolios of corporate bonds based on the Lehman Brothers Fixed Income Database. For the period April 1990 to May 1995 the two portfolios are constructed by ranking corporate bonds on rating. For a bond to be included, the issuer needs to belong to one of the following industry sectors: industrial, utilities or transportation. Every month, bonds are devided between A-grade and non-A-garde bonds. The factor DEF is the unweighted average return on the non-A-grade portfolio minus the unweighted average return on the A-grade portfolio. This factor should capture the 'default-premium' in corporate bond returns.

The third bond factor is created in the spirit of Longstaff (2004), who shows that Refcorp bonds have the same default risk as Treasury bonds, but are in general less liquid than Treasuries. To construct this factor for the period April 1990 to May 1995 we pair each RefCorp strip in the Lehman Brothers Fixed Income Database with a maturity date after April 30th 1995 with the two Treasury strips that have their maturity date closest to the RefCorp strip. Then, we construct a position in these two Treasury strips such that the duration of this position is equal to the duration of the RefCorp strip. The LIQ factor is the equally weighted return on the portfolio of RefCorp strips minus the equally weighted return on the portfolio of Treasury strip positions. This leads to a total of six factors, of which three are equity based and the three others are based on bond data.

8.2 The Factor f_t

Table 8 gives the correlations between the two estimates for the factor f_t and the six factors described in the previous section. The factor f_1 is the estimate of f_t with $T_1 = 1$ and $T_2 = 7$ and f_2 is the estimate based on $T_1 = 2$ and $T_2 = 10$. We see that the two estimates of f_t are strongly correlated and that they show very similar correlations with the six other factors, in line with the robustness results obtained earlier on.

We see from Table 8 that the two estimates for f_t are only very weakly correlated with both the default factor DEF and the HML and SMB factors. Notice that the fact that the factor f_t has been extracted from Treasury prices does not nessecarily imply a weak correlation between f_t and DEF. For instance Fama and French (1993) find that a default

Table 8: Descriptive statistics for f_t

	Mkt-Rf	SMB	HML	DEF	TERM	LIQ	f_1
SMB	0.19						
HML	-0.35	-0.14					
DEF	0.21	0.58	-0.01				
TERM	-0.01	0.20	0.36	0.14			
LIQ	0.18	-0.20	-0.12	0.02	-0.24		
f_1	0.11	0.05	-0.05	-0.09	-0.18	-0.25	
f_2	0.15	0.10	-0.02	-0.01	-0.15	-0.26	0.98

factor is more significant for explaining Treasury returns than either the HML or SMB factor.

The results that the TERM factor is more strongly correlated with the two estimates of f_t is to be expected, f_t being one of the two factors in a term structure model. Most significant however, is the observation that the strongest correlations between the two estimates for f_t and the six other factors are found for the liquidity factor LIQ. This seems to suggest that the factor f_t is at least partially related to liquidity. This result is in line with the results of Longstaff (2004) who finds that a factor derived from default-free bonds and with a strong link to liquidity is relevant for explaining both the cross-section of Treasury bond returns and the cross-section of corporate bond returns.

9 Conclusions

The results in Section 7 strongly suggest that one way to deal with the well known failure of structural models to generate 'realistically' high credit spreads could be to allow for more general term structure dynamics. In this paper we were able to generate levels of credit spreads considerably higher than usually found in the existing literature, assuming essentially affine dynamics for the term structure model. More precisely, we allowed the risk-neutral dynamics of the spot rate to be driven by a factor f_t which has no effect on its physical dynamics.

These results seem to suggest that the performance of a structural model is strongly related to the choice of term structure model. As discussed in Section 3.3, the link between the physical and the risk-neutral dynamics might be of particular importance. However, whether or not an essentially affine specification is needed to improve the performance of

structural models, or whether similar results can be obtained under some 'completely affine' term structure models is on open question.

Appendix A. The Dynamics of

Prices of zero-coupon bonds Here we derive equation 16 and expressions for the functions $A(T)$, $B(T)$ and $C(T)$ are obtained. Let us start from the risk-neutral dynamics given by equation 23. Then, the following substitution:

$$m_t = -\frac{\beta_r f_t}{\gamma_r}$$

leads to:

$$\begin{aligned} dm_t &= (a + bm_t) dt + \sigma_m d\widetilde{W}_{t,1} \\ dr_t &= q(m_t - r_t - \lambda_r \sigma_r) dt + \sigma_r d\widetilde{W}_{t,2} \end{aligned} \quad (46)$$

with: $d\widetilde{W}_1 d\widetilde{W}_2 = \rho dt$ and:

$$\begin{aligned} a &= -\frac{\beta_r \alpha_f}{\gamma_r} \\ b &= -\beta_f \\ \sigma_m &= -\frac{\beta_r \sigma_f}{\gamma_r} \\ q &= \gamma_r \\ \lambda_r &= -\frac{\alpha_r}{\sigma_r \gamma_r} \\ \sigma_r &= \sqrt{\sigma_1^2 + \sigma_2^2} \\ \rho &= \frac{\sigma_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}. \end{aligned}$$

Equation 46 has precisely the same structure as that for the risk-neutral dynamics of a stochastic-mean model. Using the results from Balduzi et al. (2000) one obtains:

$$D(\tau, r, f) = \exp(A(\tau) - rB(\tau) - fC(\tau))$$

where the functions $A(\tau)$, $B(\tau)$ and $C(\tau)$ are given by:

$$A(\tau) = -\frac{\gamma_r^3 \sigma_f^2}{4\beta_r^3 \beta_f^3 (\beta_r \beta_f + 1)^2} + e^{2\frac{\beta_r \beta_f}{\gamma_r} \tau} \frac{\sigma_f^2 \gamma_r^3}{4\beta_r^3 \beta_f^3 (\beta_r \beta_f + 1)^2}$$

$$\begin{aligned}
& \frac{\gamma_r \left(\beta_r^2 \alpha_f \beta_f - \sigma_f^2 \gamma_r^2 + \beta_r \beta_f \sigma_f \rho \sigma_r \gamma_r^2 \right)}{\beta_r^3 \beta_f^3 (\beta_r \beta_f + 1)} \\
& + e^{\frac{\beta_r \beta_f}{\gamma_r} \tau} \frac{\left(\beta_r^2 \alpha_f \beta_f - \sigma_f^2 \gamma_r^2 + \beta_r \beta_f \sigma_f \rho \sigma_r \gamma_r^2 \right) \gamma_r}{\beta_r^3 \beta_f^3 (\beta_r \beta_f + 1)} + \frac{\gamma_r^3 \sigma_f (-\sigma_f + \beta_r \beta_f \rho \sigma_r + \rho \sigma_r)}{\beta_r \beta_f (\beta_r \beta_f - 1) (\beta_r \beta_f + 1)^2} \\
& \frac{\gamma_r \left(-\beta_r^2 \alpha_f \beta_f + \sigma_f^2 \gamma_r^2 - 2\beta_r \beta_f \sigma_f \rho \sigma_r \gamma_r^2 - \sigma_f \rho \sigma_r \gamma_r^2 + \beta_r^2 \beta_f^2 \sigma_r^2 \gamma_r^2 + \beta_r \beta_f \sigma_r^2 \gamma_r^2 \right)}{\beta_r \beta_f (\beta_r \beta_f + 1)} \\
& + e^{-\frac{\tau}{\gamma_r}} \frac{\gamma_r^3 \left(\sigma_f^2 - 2\beta_r \beta_f \sigma_f \rho \sigma_r + 2\sigma_f \rho \sigma_r + \beta_r^2 \beta_f^2 \sigma_r^2 + 2\beta_r \beta_f \sigma_r^2 + \sigma_r^2 \right)}{4\sigma_f (\beta_r \beta_f + 1)^2} \\
& + \frac{\gamma_r^3 \left(\sigma_f^2 - 2\beta_r \beta_f \sigma_f \rho \sigma_r - 2\sigma_f \rho \sigma_r + \beta_r^2 \beta_f^2 \sigma_r^2 + 2\beta_r \beta_f \sigma_r^2 + \sigma_r^2 \right)}{4(\beta_r \beta_f + 1)^2} \\
& + e^{-\frac{\tau}{\gamma_r}} \frac{\gamma_r \left(-\beta_r^2 \alpha_f \beta_f + \sigma_f^2 \gamma_r^2 - 2\beta_r \beta_f \sigma_f \rho \sigma_r \gamma_r^2 - \sigma_f \rho \sigma_r \gamma_r^2 + \beta_r \beta_f \sigma_r^2 \gamma_r^2 \right)}{\beta_r \beta_f (\beta_r \beta_f + 1)} \\
& + \frac{-2\beta_r^2 \alpha_f \beta_f + \sigma_f^2 \gamma_r^2 - 2\beta_r \beta_f \sigma_f \rho \sigma_r \gamma_r^2 + \beta_r^2 \beta_f^2 \sigma_r^2 \gamma_r^2}{2\beta_r^2 \beta_f^2} \tau \\
& - e^{-\frac{\tau}{\gamma_r}} \frac{\gamma_r^3 \sigma_f (-\sigma_f + \beta_r \beta_f \rho \sigma_r + \rho \sigma_r)}{\beta_r \beta_f (\beta_r \beta_f - 1) (\beta_r \beta_f + 1)^2} \\
& \frac{\alpha_r \left(\beta_r \beta_f \left(e^{-\frac{1}{\gamma_r} \tau} - 1 \right) + e^{\beta_r \frac{\beta_f}{\gamma_r} \tau} - 1 \right)}{(\beta_r \beta_f + 1) \beta_r \beta_f}
\end{aligned}$$

$$B(\tau) = \frac{1 - e^{-\gamma_r \tau}}{\gamma_r},$$

$$C(\tau) = \frac{\gamma_r (e^{-\beta_f \tau} - 1) - \beta_f (e^{-\gamma_r \tau} - 1)}{\beta_f (\beta_f - \gamma_r)}.$$

First and second moments of f_t, r_t, y_t and l_t under the physical dynamics Here we derive the expectation function $m(t)$ and the variance function $V(t)$ for $(f_t, r_t, y_t, l_t)^T$ under the physical dynamics. The expectation function $m(t)$ is equal to the vector of expected values of the vector $(f_t, r_t, y_t, l_t)^T$ conditional on $(f_0, r_0, y_0, l_0)^T$. The variance function $V(t)$ is equal to the variance-covariance matrix of $(f_t, r_t, y_t, l_t)^T$ conditional on $(f_0, r_0, y_0, l_0)^T$.

The physical dynamics of $(f_t, r_t, y_t, l_t)^T$ are given by equation 25. Because this is a linear stochastic differential equation, $m(t)$ and $V(t)$ are the solutions of two matrix differential equations, see Karatzas and Shreve (1991, Ch. 5). More precisely, $m(t)$ and $V(t)$ are the solutions of:

$$\begin{aligned}\dot{m}(t) &= Am(t) + a \\ \dot{V}(t) &= AV + VA^T + \Sigma\Sigma^T,\end{aligned}$$

with the matrix A , the vector a and the matrix Σ as in equation 25. The initial conditions are:

$$\begin{aligned}m(0) &= (f_0, r_0, y_0, l_0)^T \\ V(0) &= 0^{4 \times 4}.\end{aligned}$$

Because the feedback matrix A is triangular, the solutions of these two matrix differential equations can be obtained by first solving for the moments of f_t and r_t and then substituting these expressions into the equations for the moments of y_t and l_t . Taking this approach one obtains after a some tedious but straightforward algebra:

$$\begin{aligned}E_0^{\mathbb{P}}[f_t] &= \bar{f} + e^{-k_f t}(f(0) - \bar{f}) \\ E_0^{\mathbb{P}}[r_t] &= \bar{r} + e^{-k_r t}(r(0) - \bar{r}) \\ E_0^{\mathbb{P}}[y_t] &= y(0) + (\alpha_y + \beta_y \bar{f} + (1 + \gamma)\bar{r})t \\ &\quad + \frac{\beta_y}{k_f} (1 - e^{-k_f t})(f(0) - \bar{f}) + \frac{1 + \gamma_y}{k_r} (1 - e^{-k_r t})(r(0) - \bar{r}) \\ E_0^{\mathbb{P}}[l_t] &= \frac{\alpha_l - \hat{\alpha}_y - \beta_y \bar{f} - (\gamma_y + \gamma_l)\bar{r}}{\kappa} (1 - e^{-\kappa t}) - \frac{\beta_y}{k_f} (e^{-k_f t} - e^{-\kappa t})(f(0) - \bar{f}) \\ &\quad - \frac{\gamma_l + \gamma_y}{k_r} (e^{-k_r t} - e^{-\kappa t})(r(0) - \bar{r}) + l(0)e^{-\kappa t}.\end{aligned}$$

The 10 elements determining $V(t)$ are given by:

$$\begin{aligned}var_0^{\mathbb{P}}[f_t] &= \frac{\sigma_f^2 (1 - e^{-2k_f t})}{2k_f} \\ var_0^{\mathbb{P}}[r_t] &= \frac{(\sigma_1^2 + \sigma_2^2) (1 - e^{-2k_r t})}{2k_r} \\ covar_0^{\mathbb{P}}[f_t, r_t] &= \frac{\sigma_f \sigma_1 (1 - e^{-(k_f + k_r)t})}{k_f + k_r} \\ covar_0^{\mathbb{P}}[f_t, y_t] &= C_1 - C_2 e^{-k_f t} + C_3 e^{-2k_f t} + C_4 e^{-(k_f + k_r)t}\end{aligned}$$

$$\begin{aligned}
\text{covar}_0^{\mathbb{P}} [r_t, y_t] &= D_1 - D_2 e^{-k_r t} + D_3 e^{-2k_r t} + D_4 e^{-(k_f+k_r)t} \\
\text{var}_0^{\mathbb{P}} [y_t] &= (2\beta_y C_1 + 2(1 + \gamma_y) D_1 + \eta_1^2 + \eta_2^2 + \eta_3^2) t + \frac{2\beta_y}{k_f} C_2 (e^{-k_f t} - 1) \\
&\quad + \frac{2(1 + \gamma_y)}{k_r} C_2 (e^{-k_r t} - 1) - \frac{\beta_y}{k_f} C_2 (e^{-2k_f t} - 1) \\
&\quad - \frac{(1 + \gamma_y)}{k_r} C_2 (e^{-2k_r t} - 1) - \frac{2\beta_y C_4 + 2(1 + \gamma_y) D_4}{k_f + k_r} (e^{-(k_f+k_r)t} - 1) \\
\text{covar}_0^{\mathbb{P}} [f_t, l_t] &= G_1 - G_2 e^{-(\kappa+k_f)t} + G_3 e^{-2k_f t} + G_4 e^{-(k_f+k_r)t} \\
\text{covar}_0^{\mathbb{P}} [r_t, l_t] &= H_1 - H_2 e^{-(\kappa+k_r)t} + H_3 e^{-2k_r t} + H_4 e^{-(k_f+k_r)t} \\
\text{var}_0^{\mathbb{P}} [l_t] &= \left(\frac{\eta_1^2 + \eta_2^2 + \eta_3^2}{2\kappa} - \frac{\beta_y}{\kappa} G_1 - \frac{\gamma_l + \gamma_y}{\kappa} H_1 \right) (1 - e^{-2\kappa t}) \\
&\quad + \frac{2\beta_y}{\kappa - k_f} G_2 (e^{-(\kappa+k_f)t} - e^{-2\kappa t}) - \frac{\beta_y}{\kappa - k_f} G_3 (e^{-2k_f t} - e^{-2\kappa t}) \\
&\quad + \frac{2(\gamma_l + \gamma_y)}{\kappa - k_r} H_2 (e^{-(\kappa+k_r)t} - e^{-2\kappa t}) - \frac{\gamma_l + \gamma_y}{\kappa - k_r} H_3 (e^{-2k_r t} - e^{-\kappa t}) \\
&\quad - \frac{2\beta_y G_4 + 2(\gamma_l + \gamma_y) H_4}{2\kappa - (k_f + k_r)} (e^{-(k_f+k_r)t} - e^{-2\kappa t}) \\
\text{covar}_0^{\mathbb{P}} [y_t, l_t] &= \frac{\eta_1^2 + \eta_2^2 + \eta_3^2}{\kappa} (1 - e^{-\kappa t}) + I_1(t) + I_2(t) + I_3(t) + I_4(t)
\end{aligned}$$

where the functions $I_1(t)$, $I_2(t)$, $I_3(t)$ and $I_4(t)$ are given by:

$$\begin{aligned}
I_1(t) &= -\frac{\beta_y}{\kappa} C_1 (1 - e^{-\kappa t}) + \frac{\beta_y}{\kappa - k_f} C_2 (e^{-k_f t} - e^{-\kappa t}) \\
&\quad - \frac{\beta_y}{\kappa - 2k_f} C_3 (e^{-2k_f t} - e^{-\kappa t}) - \frac{\beta_y}{\kappa - (k_f + k_r)} C_4 (e^{-(k_f+k_r)t} - e^{-\kappa t}) \\
I_2(t) &= -\frac{\gamma_y + \gamma_l}{\kappa} D_1 (1 - e^{-\kappa t}) + \frac{\gamma_y + \gamma_l}{\kappa - k_r} D_2 (e^{-k_r t} - e^{-\kappa t}) \\
&\quad - \frac{\gamma_y + \gamma_l}{\kappa - 2k_r} D_3 (e^{-2k_r t} - e^{-\kappa t}) - \frac{\gamma_y + \gamma_l}{\kappa - (k_f + k_r)} D_4 (e^{-(k_f+k_r)t} - e^{-\kappa t}) \\
I_3(t) &= \frac{\beta_y}{\kappa} G_1 (1 - e^{-\kappa t}) + \frac{\beta_y}{k_f} G_2 (e^{-(\kappa+k_f)t} - e^{-\kappa t}) + \frac{\beta_y}{\kappa - 2k_f} G_3 (e^{-2k_f t} - e^{-\kappa t}) \\
&\quad + \frac{\beta_y}{\kappa - (k_f + k_r)} G_4 (e^{-(k_f+k_r)t} - e^{-\kappa t}) \\
I_4(t) &= \frac{1 + \gamma_y}{\kappa} H_1 (1 - e^{-\kappa t}) + \frac{1 + \gamma_y}{k_r} H_2 (e^{-(\kappa+k_r)t} - e^{-\kappa t}) \\
&\quad + \frac{1 + \gamma_y}{\kappa - 2k_r} H_3 (e^{-2k_r t} - e^{-\kappa t}) + \frac{1 + \gamma_y}{\kappa - (k_f + k_r)} D_4 (e^{-(k_f+k_r)t} - e^{-\kappa t})
\end{aligned}$$

The constants C_1, C_2, C_3 and C_4 are given by:

$$\begin{aligned}
C_1 &= \frac{\beta_y \sigma_f^2}{2k_f^2} + \frac{(1 + \gamma_y)(\sigma_1^2 + \sigma_2^2)}{(k_f + k_r)k_f} + \frac{\sigma_f \eta_1}{k_f} \\
C_2 &= \frac{\beta_y \sigma_f^2}{k_f^2} + \frac{(1 + \gamma_y)(\sigma_1^2 + \sigma_2^2)}{(k_f + k_r)k_f} + \frac{(1 + \gamma_y)(\sigma_1^2 + \sigma_2^2)}{(k_f + k_r)k_r} + \frac{\sigma_f \eta_1}{k_f} \\
C_3 &= \frac{\beta_y \sigma_f^2}{2k_f^2} \\
C_4 &= \frac{(1 + \gamma_y)(\sigma_1^2 + \sigma_2^2)}{(k_f + k_r)k_r}.
\end{aligned}$$

The constants D_1, D_2, D_3 and D_4 are given by:

$$\begin{aligned}
D_1 &= \frac{\beta_y \sigma_f \sigma_1}{(k_f + k_r)k_f} + \frac{(1 + \gamma_y)(\sigma_1^2 + \sigma_2^2)}{2k_r^2} + \frac{\sigma_1 \eta_1 + \sigma_2 \eta_2}{k_r} \\
D_2 &= \frac{\beta_y \sigma_f \sigma_1}{(k_f + k_r)k_f} + \frac{\beta_y \sigma_f \sigma_1}{(k_f + k_r)k_r} + \frac{(1 + \gamma_y)(\sigma_1^2 + \sigma_2^2)}{2k_r^2} + \frac{\sigma_1 \eta_1 + \sigma_2 \eta_2}{k_r} \\
D_3 &= \frac{(1 + \gamma_y)(\sigma_1^2 + \sigma_2^2)}{2k_r^2} \\
D_4 &= \frac{\beta_y \sigma_f \sigma_1}{(k_f + k_r)k_f}
\end{aligned}$$

The constants G_1, G_2, G_3 and G_4 are given by:

$$\begin{aligned}
G_1 &= -\frac{\beta_y \sigma_f^2}{2k_f(k_f + \kappa)} - \frac{(\gamma_y + \gamma_l)\sigma_f \sigma_1}{(k_f + k_r)(k_f + \kappa)} - \frac{\sigma_f \eta_1}{k_f + \kappa} \\
G_2 &= \frac{\beta_y \sigma_f^2}{2k_f(k_f + \kappa)} + \frac{(\gamma_y + \gamma_l)\sigma_f \sigma_1}{(k_f + k_r)(k_f + \kappa)} + \frac{\sigma_f \eta_1}{k_f + \kappa} \\
&\quad - \frac{\beta_y \sigma_f^2}{2k_f(\kappa - k_f)} - \frac{(\gamma_y + \gamma_l)\sigma_f \sigma_1}{(k_f + k_r)(\kappa - k_r)} \\
G_3 &= \frac{\beta_y \sigma_f^2}{2k_f(\kappa - k_f)} \\
G_4 &= \frac{(\gamma_y + \gamma_l)\sigma_f \sigma_1}{(k_f + k_r)(\kappa - k_r)}
\end{aligned}$$

The constants H_1, H_2, H_3 and H_4 are given by:

$$H_1 = -\frac{\beta_y \sigma_f \sigma_1}{(k_f + k_r)(k_r + \kappa)} - \frac{(\gamma_y + \gamma_l)(\sigma_1^2 + \sigma_2^2)}{2k_r(k_r + \kappa)} - \frac{\sigma_1 \eta_1 + \sigma_2 \eta_2}{k_r + \kappa}$$

$$\begin{aligned}
H_2 &= -\frac{\beta_y \sigma_f \sigma_1}{(k_f + k_r)(k_r + \kappa)} - \frac{(\gamma_y + \gamma_l)(\sigma_1^2 + \sigma_2^2)}{2k_r(\kappa - k_r)} - \frac{(\gamma_y + \gamma_l)(\sigma_1^2 + \sigma_2^2)}{2k_r(\kappa - k_r)} - \frac{\sigma_1 \eta_1 + \sigma_2 \eta_2}{k_r + \kappa} \\
H_3 &= -\frac{(\gamma_y + \gamma_l)(\sigma_1^2 + \sigma_2^2)}{2k_r(\kappa - k_r)} \\
H_4 &= -\frac{\beta_y \sigma_f \sigma_1}{(k_f + k_r)(\kappa - k_f)}
\end{aligned}$$

First and second moments of f_t, r_t and l_t under the T -forward measure Here we derive the expectation function $m^T(t)$ and the variance function $V^T(t)$ of f_t, r_t and l_t under the T -forward measure. Because equation 29 is a linear stochastic differential equation, $m^T(t)$ and $V^T(t)$ are again the solutions of two matrix differential equations. More precisely, $m^T(t)$ and $V(t)$ are the solution of:

$$\dot{m}^T(t) = A^T m^T(t) + a^T(t) \quad (47)$$

$$\dot{V}^T(t) = A^T V^T + V^T (A^T)' + \Sigma^T (\Sigma^T)', \quad (48)$$

with the matrix A^T , the vector $a^T(t)$ and the matrix Σ as in equation 29. The initial conditions are again given by:

$$m^T(0) = (f_0, r_0, y_0)'$$

$$V^T(0) = 0^{3 \times 3}.$$

Note that a superscript T with a matrix does not indicate the transpose but that it refers to the T -forward measure. The transpose of a matrix M is indicated as M' .

The moments of f_t and r_t Observe that the matrix A^T is again triangular, and as such we can first solve for the moments of f_t and r_t , and afterwards derive the moments involving l_t . Let us denote the upper left 2×2 sub-matrix of A^T by \bar{A}^T , that is:

$$\bar{A}^T = \begin{bmatrix} -\beta_f & 0 \\ -\beta_r & -\gamma_r \end{bmatrix}.$$

The vector of means $m(t)$ and the covariance matrix $V(t)$ of f_t and r_t are the solutions of:

$$\dot{m}^T(t) = \bar{A}^T m^T(t) + \bar{a}^T(t) \quad (49)$$

$$\dot{V}^T(t) = \bar{A}^T V^T + V^T (\bar{A}^T)' + \bar{\Sigma}^T (\bar{\Sigma}^T)', \quad (50)$$

where the definitions of $\bar{a}^T(t)$ and $\bar{\Sigma}^T$ are analogue to that of \bar{A}^T . Since the feedback matrix \bar{A}^T is not diagonal, we use a linear transformation of the variables f_t and r_t such that the feedback matrix of the transformed process is diagonal. That is, we need two matrices E_1 and E_2 that meet the two following conditions:

$$E_1 A^T E_2 = \Lambda \quad (51)$$

and:

$$E_1 E_2 = \Omega. \quad (52)$$

With Λ and Ω both diagonal matrices. Note that the matrix $E_2 \Omega^{-1}$ is the inverse of E_1 . It is well known that the matrix E_1 is the matrix which has the column-eigenvectors of \bar{A}^T as its rows and the matrix E_2 is the matrix which has the row-eigenvectors of \bar{A}^T as its columns. The two eigenvalues of \bar{A}^T are:

$$\lambda_1 = -\beta_f$$

$$\lambda_2 = -\gamma_r.$$

The matrices E_1 and E_2 are given by:

$$E_1 = \begin{bmatrix} 1 & 0 \\ \beta_r & \gamma_r - \beta_f \end{bmatrix}, \quad (53)$$

$$E_2 = \begin{bmatrix} \beta_f - \gamma_r & 0 \\ \beta_r & 1 \end{bmatrix}. \quad (54)$$

Let us denote the diagonal matrix with diagonal elements λ_1 and λ_2 by Γ that is:

$$\Gamma = \begin{bmatrix} -\beta_f & 0 \\ 0 & -\gamma_r \end{bmatrix}.$$

Because we have not normalized the eigenvectors we obtain:

$$\Omega \equiv E_1 E_2 = \begin{bmatrix} \beta_f - \gamma_r & 0 \\ 0 & \gamma_r - \beta_f \end{bmatrix}. \quad (55)$$

Finally, the diagonal matrix Λ in equation 51 is given by:

$$\Lambda = \Omega \cdot \Gamma = \begin{bmatrix} \beta_f (\gamma_r - \beta_f) & 0 \\ 0 & \gamma_r (\beta_f - \gamma_r) \end{bmatrix}.$$

The fact that the matrix \bar{A}^T is diagonalizable, as shown by equation 51, allows us to decouple both of the two matrix differential equations 47 and 48. More precisely, let us define a new process (X_1, X_2) as:

$$\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = E_1 \begin{pmatrix} f_t \\ r_t \end{pmatrix}.$$

The dynamics of $X(t)$ under the T -forward measure are given by:

$$(a_X^T(t) + \Gamma X(t)) dt + \Sigma_X^T dW(t), \quad (56)$$

with:

$$a_X^T(t) = E_1 a^T(t) = \begin{bmatrix} p_0 + p_1 e^{-\beta_f(T-t)} + p_2 e^{-(T-t)\gamma_r} \\ q_0 + q_1 e^{-\beta_f(T-t)} + q_2 e^{-(T-t)\gamma_r} \end{bmatrix},$$

with:

$$\begin{aligned} p_0 &= \alpha_f - \frac{\sigma_f \sigma_1}{\gamma_r} - \frac{\sigma_f^2}{\beta_f} \\ p_1 &= \frac{-\sigma_f^2 \gamma_r}{\beta_f (\beta_f - \gamma_r)} \\ p_2 &= \frac{\sigma_f \sigma_1}{\gamma_r} + \frac{\sigma_f^2}{\beta_f - \gamma_r} \\ q_0 &= \beta_r \alpha_f + (\gamma_r - \beta_f) \alpha_r - (\beta_r \sigma_f \sigma_1 + (\gamma_r - \beta_f) (\sigma_f^2 + \sigma_1^2)) + \frac{\beta_r \sigma_f^2 + (\gamma_r - \beta_f) \sigma_1 \sigma_f}{\beta_f} \\ q_1 &= \frac{\gamma_r}{\beta_f - \gamma_r} \frac{\beta_r \sigma_f^2 + (\gamma_r - \beta_f) \sigma_1 \sigma_f}{\beta_f} \\ q_2 &= \frac{\beta_r \sigma_f \sigma_1 + (\gamma_r - \beta_f) (\sigma_f^2 + \sigma_1^2)}{\gamma_r} + \frac{\gamma_r}{\beta_f (\gamma_r - \beta_f)} \beta_r \sigma_f^2 + (\gamma_r - \beta_f) \sigma_1 \sigma_f \end{aligned}$$

The matrix Σ_X^T is given by:

$$\Sigma_X^T = \begin{bmatrix} \sigma_f & 0 \\ \beta_r \sigma_f + (\gamma_r - \beta_f) \sigma_1 & (\gamma_r - \beta_f) \sigma_2 \end{bmatrix}$$

The mean and variance functions $m_X^T(t)$ and $V_X^T(t)$ of the process $X(t)$ are the solutions of the two following matrix differential equations.

$$\dot{m}_X^T(t) = \Gamma m_X^T(t) + a_X^T(t) \quad (57)$$

$$\dot{V}_X^T(t) = \Gamma V_X^T + V_X^T \Gamma + \Sigma_X \Sigma_X', \quad (58)$$

with initial conditions:

$$m_X^T(0) = X(0),$$

$$V_X^T(0) \equiv 0.$$

Where $X(0)$ is given by:

$$X(0) = E_1 \begin{bmatrix} f(0) \\ r(0) \end{bmatrix} = \begin{bmatrix} f(0) \\ \beta_r f(0) + (\gamma_r - \beta_f) r(0) \end{bmatrix}.$$

Because the feedback matrix in equation 56 is diagonal, the two above matrix differential equations are now both decoupled. As such, each of the components of $m_X^T(t)$ and $V_X^T(t)$ can be obtained by solving an ordinary linear first order differential equation.

Some straightforward calculations lead to:

$$m_{X_1}^T(t) = p_0 \frac{1 - e^{-\beta_f t}}{\beta_f} + p_1 t + p_2 \frac{e^{-\beta_f t} - e^{-\gamma_r t}}{\gamma_r - \beta_f} + e^{-\beta_f t} X_1(0)$$

$$m_{X_2}^T(t) = q_0 \frac{1 - e^{-\gamma_r t}}{\gamma_r} + q_2 t + q_1 \frac{e^{-\gamma_r t} - e^{-\beta_f t}}{\beta_f - \gamma_r} + e^{-\gamma_r t} X_2(0)$$

For the matrix $\Sigma_X \Sigma_X'$ we obtain:

$$\Sigma_X \Sigma_X' = \begin{bmatrix} \sigma_f^2 & \beta_r \sigma_f^2 + (\gamma_r - \beta_f) \sigma_1 \sigma_f \\ \beta_r \sigma_f^2 + (\gamma_r - \beta_f) \sigma_1 \sigma_f & (\beta_r \sigma_f + (\gamma_r - \beta_f) \sigma_1)^2 + (\gamma_r - \beta_f)^2 \sigma_2^2 \end{bmatrix}$$

Substituting the above in equation 58 one finds that the three elements of the variance function $V_X^T(t)$ are given by:

$$var_0^T [X_1(t)] = \frac{\sigma_f^2 (1 - e^{-2\beta_f t})}{2\beta_f} \quad (59)$$

$$var_0^T [X_2(t)] = \left[(\beta_r \sigma_f + (\gamma_r - \beta_f) \sigma_1)^2 + (\gamma_r - \beta_f)^2 \sigma_2^2 \right] \frac{1 - e^{-2\gamma_r t}}{2\gamma_r} \quad (60)$$

$$covar_0^T [X_1(t), X_2(t)] = \left[\beta_r \sigma_f^2 + (\gamma_r - \beta_f) \sigma_1 \sigma_f \right] \frac{1 - e^{-(\beta_f + \gamma_r)t}}{(\beta_f + \gamma_r)} \quad (61)$$

Given $m_X^T(t)$ and $V_X^T(t)$ the mean function $m^T(t)$ and the variance function $V^T(t)$ of the original process (f_t, r_t) are given by:

$$m^T(t) = E_1^{-1} m_X^T(t) = E_2 \Omega^{-1} m_X^T(t).$$

From this one obtains:

$$\begin{aligned} E_0^T [f(t)] &= p_0 \frac{1 - e^{-\beta_f t}}{\beta_f} + p_1 t + p_2 \frac{e^{-\beta_f t} - e^{-\gamma_r t}}{\gamma_r - \beta_f} + e^{-\beta_f t} f(0) \\ E_0^T [r(t)] &= c_0 + c_1 t + c_2 e^{-\beta_f t} + c_3 e^{-\gamma_r t}, \end{aligned}$$

with:

$$\begin{aligned} c_0 &= \frac{\beta_r}{\beta_f - \gamma_r} \frac{p_0}{\beta_f} - \frac{1}{\beta_f - \gamma_r} \frac{q_0}{\gamma_r} \\ c_1 &= \frac{\beta_r}{\beta_f - \gamma_r} p_1 - \frac{q_1}{\beta_f - \gamma_r} \\ c_2 &= \frac{q_1}{(\beta_f - \gamma_r)^2} - \frac{\beta_r}{\beta_f - \gamma_r} \left(\frac{p_0}{\beta_f} - \frac{p_2}{\gamma_r - \beta_f} - f(0) \right) \\ c_3 &= \frac{p_2}{(\beta_f - \gamma_r)^2} - \frac{1}{\beta_f - \gamma_r} \left(\frac{q_1}{\beta_f - \gamma_r} - \frac{q_0}{\gamma_r} + \beta_r f(0) + (\gamma_r - \beta_f) r(0) \right). \end{aligned}$$

Similarly, the conditional variance-covariance matrix $V^T(t)$ of f_t and r_t is given by:

$$\begin{aligned} V^T(t) &= E_1^{-1} V_X^T(t) (E_1^{-1})' \\ &= E_2 \Omega^{-1} V_X^T(t) \Omega^{-1} E_2'. \end{aligned}$$

Substituting for E_2, Ω and $V_X^T(t)$, from equations 54, 55 and 59 to 61 yields:

$$\text{var}_0^T [f(t)] = \frac{\sigma_f^2 (1 - e^{-2\beta_f t})}{2\beta_f},$$

$$\begin{aligned} \text{var}_0^T [r(t)] &= \frac{1}{(\gamma_r - \beta_f)^2} \left(\frac{\beta_r}{2\beta_f} \sigma_f^2 (1 - e^{-2\beta_f t}) + [\beta_r \sigma_f^2 + (\gamma_r - \beta_f) \sigma_1 \sigma_f] \frac{1 - e^{-(\beta_f + \gamma_r)t}}{(\beta_f + \gamma_r)} \right. \\ &\quad \left. - \frac{\beta_r}{\gamma_r} [(\beta_r \sigma_f + (\gamma_r - \beta_f) \sigma_1)^2 + (\gamma_r - \beta_f)^2 \sigma_2^2] (1 - e^{-2\gamma_r t}) \right), \end{aligned}$$

$$\text{covar}_0^T [f(t), r(t)] = \frac{\beta_r \sigma_f^2 + (\gamma_r - \beta_f) \sigma_1 \sigma_f}{(\gamma_r - \beta_f)(\beta_f + \gamma_r)} (1 - e^{-(\beta_f + \gamma_r)t}) - \frac{\beta_r \sigma_f^2}{2\beta_f(\beta_f - \gamma_r)} (1 - e^{-2\beta_f t}).$$

The moments involving l_t To obtain the conditional expected value of l_t , we first rewrite $a_l^T(t)$, given by equation 30c, as:

$$a_l^T(t) = h_0 + h_1 e^{\gamma_r t} + h_2 e^{-\beta_f t},$$

with:

$$\begin{aligned} h_0 &= \alpha_l + \frac{\eta_1 \sigma_1 + \eta_2 \sigma_2}{\gamma_r} + \frac{\eta_1 \sigma_f}{\beta_f} \\ h_1 &= -\frac{\eta_1 \sigma_1 + \eta_2 \sigma_2}{\gamma_r} - \frac{\eta_1 \sigma_f}{\beta_f - \gamma_r} \\ h_2 &= -\frac{\gamma_r \eta_1 \sigma_f}{\beta_f (\beta_f - \gamma_r)}. \end{aligned}$$

From equation 47 we see that $E_0^T[l(t)]$ is the solution of:

$$y'(t) = -\gamma_l E_0^T[r(t)] - \kappa y(t) + a_l^T(t).$$

Therefore, $E_0^T[l(t)]$ is given by:

$$\begin{aligned} E_0^T[l(t)] &= \left(h_0 - \gamma_l - \frac{c_1}{\kappa}\right) \frac{1 - e^{-\kappa t}}{\kappa} + (h_1 - \gamma_l c_3) \frac{e^{-\gamma_r t} - e^{-\kappa t}}{\kappa - \gamma_r} + (h_2 - \gamma_l c_2) \frac{e^{-\beta_f t} - e^{-\kappa t}}{\kappa - \beta_f} \\ &\quad + \frac{c_1}{\kappa} t + e^{-\kappa t} l(0). \end{aligned}$$

We now turn to deriving the three (co-)variances involving l_t . From equation 50 we see that the covariance $covar_0^T[f_t, l_t]$ between f_t and l_t is the solution of the following differential equation:

$$y'(t) = -(\beta_f + \kappa)y(t) - \gamma_l covar_0^T[f_t, r_t] - \sigma_f \eta_1.$$

The solution of this equation is given by:

$$covar_0^T[f_t, l_t] = g_0 + g_1 e^{-(\beta_f + \gamma_r)t} + g_2 e^{-(\beta_f + \kappa)t} - g_3 e^{-2\beta_f t},$$

with:

$$g_0 = \frac{\gamma_l}{\beta_f + \kappa} \left[\frac{\beta_r \sigma_f^2 + (\gamma_r - \beta_f) \sigma_1 \sigma_f}{(\gamma_r - \beta_f)(\beta_f + \gamma_r)} - \frac{\beta_r \sigma_f^2}{2\beta_f(\beta_f - \gamma_r)} \right]$$

$$\begin{aligned}
g_1 &= \frac{\gamma_l}{\kappa - \gamma_r} \\
g_2 &= \left(\frac{\gamma_l}{\kappa - \beta_f} + \frac{\gamma_l}{\beta_f + \kappa} \right) \frac{\beta_r \sigma_f^2}{2\beta_f(\beta_f - \gamma_r)} - \left(\frac{\gamma_l}{\kappa - \gamma_r} + \frac{\gamma_l}{\beta_f + \kappa} \right) \frac{\beta_r \sigma_f^2 + (\gamma_r - \beta_f)\sigma_1\sigma_f}{(\gamma_r - \beta_f)(\beta_f + \gamma_r)} \\
g_3 &= \frac{\gamma_l}{\kappa - \beta_f}.
\end{aligned}$$

Going back to equation 50, we see that the covariance $covar_0^T [r_t, l_t]$ between r_t and l_t is the solution of the following differential equation:

$$y'(t) = -\beta_f covar_0^T [f_t, l_t] - \gamma_l var_0^T [r_t] - (\gamma_r + \kappa)y(t) - (\sigma_1\eta_1 + \sigma_2\eta_2).$$

The solution of this equation is given by:

$$covar_0^T [r_t, l_t] = j_0 + j_1 e^{-(\gamma_r + \kappa)t} + j_2 e^{-(\gamma_r + \beta_f)t} + j_3 e^{-(\kappa + \beta_f)t} + j_4 e^{-2\gamma_r t} + j_5 e^{-2\beta_f t},$$

with:

$$\begin{aligned}
j_0 &= -\frac{\gamma_l(c_1 + c_2 + c_3) - \beta_f g_0 - \sigma_1\eta_1 - \sigma_2\eta_2}{\gamma_r + \kappa} \\
j_1 &= \frac{\beta_f g_0 + \sigma_1\eta_1 + \sigma_2\eta_2 + \gamma_l(c_1 + c_2 + c_3)}{\gamma_r + \kappa} + \frac{\beta_f g_1}{\kappa - \beta_f} + \frac{\beta_f g_2}{\gamma_r - \beta_f} - \frac{\beta_f g_3 + \gamma_l c_1}{\gamma_r + \kappa - 2\beta_f} \\
&\quad - \frac{\gamma_l c_2}{\kappa - 2\beta_f} - \frac{\gamma_l c_3}{\kappa - \gamma_r} \\
j_2 &= \frac{\gamma_l c_2 - \beta_f g_1}{\kappa - \beta_f} \\
j_3 &= \frac{\beta_f g_2}{\gamma_r - \beta_f} \\
j_4 &= \frac{\gamma_l c_3}{\kappa - \gamma_r} \\
j_5 &= \frac{\beta_f g_3 + \gamma_l c_1}{\gamma_r + \kappa - 2\beta_f}.
\end{aligned}$$

Finally, we see that $var_0^T [l_t]$ the variance of l_t is the solution of:

$$y'(t) = -2\gamma_l covar_0^T [r_t, l_t] - 2\kappa y(t) + \sum_{i=1}^3 \eta_i^2.$$

Therefore, $var_0^T [l_t]$ is given by:

$$\begin{aligned}
var_0^T [l_t] &= 2\gamma_l \left[j_1 \frac{e^{-(\kappa + \gamma_r)t} - e^{-2\kappa t}}{\gamma_r - \kappa} + j_2 \frac{e^{-(\gamma_r + \beta_f)t} - e^{-2\kappa t}}{2\kappa - \gamma_r + \beta_f} + j_3 \frac{e^{-(\kappa + \beta_f)t} - e^{-2\kappa t}}{\beta_f - \kappa} \right. \\
&\quad \left. + j_4 \frac{e^{-2\gamma_r t} - e^{-2\kappa t}}{2(\gamma_r - \kappa)} + j_5 \frac{e^{-2\beta_f t} - e^{-2\kappa t}}{2(\beta_f - \kappa)} \right] + \left(\sum_{i=1}^3 \eta_i^2 - \gamma_l j_0 \right) \frac{1 - e^{-2\kappa t}}{\kappa}
\end{aligned}$$

Appendix B. The derivation of $Q^{\mathbb{M}}(r_0, f_0, l_0, T)$

Our approach to derive the first passage time density in our three-dimensional Markovian set-up is based on Collin-Dufresne and Goldstein (2001) who derive this density in a two-dimensional Markovian framework.

The Fortet equation for a three-dimensional Markov process Let us define $g^{\mathbb{M}}[l_s = \underline{l}, f_s, r_s, s | l_0, f_0, r_0, 0]$ as the probability density (under the measure \mathbb{M}) that the first passage time for a given constant boundary \underline{l} is at time s . If we would know this density, then the probability that the process l_t has reached the barrier level \underline{l} before (or at) time t is given by:

$$\int_0^t ds \int_{-\infty}^{\infty} dr_s \int_{-\infty}^{\infty} df_s g^{\mathbb{M}}[l_s = \underline{l}, f_s, r_s, s | l_0, f_0, r_0, 0]. \quad (62)$$

Observe that the following equality holds for all $l_t, l_0, \underline{l} : l_0 < \underline{l} < l_t$:

$$\begin{aligned} & \pi^{\mathbb{M}}(l_t, f_t, r_t, t | l_0, f_0, r_0, 0) \\ &= \int_0^t ds \int_{-\infty}^{\infty} dr_s \int_{-\infty}^{\infty} df_s g^{\mathbb{M}}[l_s = \underline{l}, f_s, r_s, s | l_0, f_0, r_0, 0] \pi^{\mathbb{M}}(l_t, f_t, r_t, t | l_s = \underline{l}, f_s, r_s, s). \end{aligned} \quad (63)$$

As Collin-Dufresne and Goldstein (2001) point out, the interpretation of this equality is straightforward. Let three levels $l_0 < \underline{l} < l_t$ be given. For the process l to reach the level l_t at time t from a lower level l_0 at time 0 there is a point in time s at which it reaches the intermediary value \underline{l} for the first time. Having this principle in mind, the above equation is a direct result of the Law of Total Probability.

Integrating both sides of the equation by:

$$\int_{-\infty}^{\infty} dr_t \int_{-\infty}^{\infty} df_t,$$

leads to the following equation:

$$\begin{aligned} & \pi^{\mathbb{M}}(l_t, t | l_0, f_0, r_0, 0) \\ &= \int_0^t ds \int_{-\infty}^{\infty} dr_s \int_{-\infty}^{\infty} df_s g^{\mathbb{M}}[l_s = \underline{l}, f_s, r_s, s | l_0, f_0, r_0, 0] \pi^{\mathbb{M}}(l_t, t | l_s = \underline{l}, f_s, r_s, s). \end{aligned} \quad (64)$$

Let us introduce the following functions:

$$\Psi^{\mathbb{M}}(f_t, r_t, t) = \int_{\underline{l}}^{\infty} dl_t \pi^{\mathbb{M}}(l_t, t | l_0, f_0, r_0, 0) \quad (65)$$

$$\Phi^{\mathbb{M}}(f_t, r_t, t | f_s, r_s, s) = \int_{\underline{l}}^{\infty} dl_t \pi^{\mathbb{M}}(l_t, t | l_s = l, f_s, r_s, s) \quad (66)$$

$$g^{\mathbb{M}}(f_s, r_s, s) = g^{\mathbb{M}}[l_s = \underline{l}, f_s, r_s, s | l_0, f_0, r_0, 0]. \quad (67)$$

Integrating both sides of equation 63 by $\int_0^{+\infty} dt_t$ we see it can be rewritten as:

$$\Psi^{\mathbb{M}}(f_t, r_t, t) = \int_0^t ds \int_{-\infty}^{\infty} df_s \int_{-\infty}^{\infty} dr_s g^{\mathbb{M}}(f_s, r_s, s) \Phi^{\mathbb{M}}(f_t, r_t, t | f_s, r_s, s). \quad (68)$$

The derivation of $\mathbf{Q}^{\mathbb{M}}(r_0, f_0, l_0, T)$ Following the approach of Collin-Dufresne and Goldstein (2001), we discretize equation 68 to obtain an algorithm for calculating the probability $\mathbf{Q}^{\mathbb{M}}(r_0, f_0, l_0, T)$. Discretize time into n_T equal intervals of length Δt and define $t_i = i\Delta t$. Discretize the f -space by dividing the interval between some chosen minimum \underline{f} and maximum \bar{f} into n_f equal intervals of length Δf and define $f_k = k\Delta f$. Similarly, discretize the r -space by dividing the interval between some chosen minimum \underline{r} and maximum \bar{r} into n_r equal intervals of length Δr and define $r_l = l\Delta r$.

Remember that $\mathbf{Q}^{\mathbb{M}}(r_0, f_0, l_0, T)$ is the probability, under the measure \mathbb{M} , that the process l_t has hit the lower boundary value zero before time T . That is, here we have $\underline{l} = 0$. The discrete version of equation 68 is given by:

$$\Psi^{\mathbb{M}}(f_i, r_j, t_k) = \sum_{w=1}^k \sum_{u=1}^{n_f} \sum_{v=1}^{n_r} q^{\mathbb{M}}(f_u, r_v, t_w) \Phi^{\mathbb{M}}(f_i, r_j, t_k | f_u, r_v, t_w), \quad (69)$$

where:

$$q^{\mathbb{M}}(f_u, r_v, t_w) = \Delta t \Delta f \Delta r g^{\mathbb{M}}(f_u, r_v, t_w).$$

Under the measure \mathbb{M} the probability that the barrier $\underline{l} = 0$ has been reached by time T is:

$$\mathbf{Q}^{\mathbb{M}}(r_0, f_0, l_0, T) = \sum_{k=1}^{n_T} \sum_{i=1}^{n_f} \sum_{j=1}^{n_r} q^{\mathbb{M}}(f_i, r_j, t_k). \quad (70)$$

The $q^{\mathbb{M}}(f_i, r_j, t_k)$'s are obtained recursively as follows:

$$\forall i \in (1, \dots, n_f), \forall j \in (1, \dots, n_r) :$$

$$q^{\mathbb{M}}(f_i, r_j, t_1) = \Delta f \Delta r \Psi^{\mathbb{M}}(f_i, r_j, t_1) \quad (71)$$

and:

$$\forall i \in (1, \dots, n_f), \forall j \in (1, \dots, n_r), \forall k \in (2, \dots, n_T) :$$

$$q^{\mathbb{M}}(f_i, r_j, t_k) = \Delta f \Delta r \left[\Psi^{\mathbb{M}}(f_i, r_j, t_k) - \sum_{w=1}^{k-1} \sum_{u=1}^{n_f} \sum_{v=1}^{n_r} q^{\mathbb{M}}(f_u, r_v, t_w) \Phi^{\mathbb{M}}(f_i, r_j, t_k | f_u, r_v, t_w) \right]. \quad (72)$$

Now, we still need to obtain expressions for the functions Ψ and Φ . Note that the following relation holds between the three densities $\pi^{\mathbb{M}}(l_t, f_t, r_t, t | l_s, f_s, r_s, s)$, $\pi^{\mathbb{M}}(f_t, r_t, t | l_s, f_s, r_s, s)$ and $\pi^{\mathbb{M}}(l_t, t | f_t, r_t, l_s, f_s, r_s, s)$:

$$\pi^{\mathbb{M}}(l_t, f_t, r_t, t | l_s, f_s, r_s, s) = \pi^{\mathbb{M}}(f_t, r_t, t | l_s, f_s, r_s, s) \pi^{\mathbb{M}}(l_t, t | f_t, r_t, l_s, f_s, r_s, s).$$

Using this in the definition of $\Psi^{\mathbb{M}}$ and $\Phi^{\mathbb{M}}$, equations 65 and 66 respectively, we obtain:

$$\Psi^{\mathbb{M}}(f_t, r_t, t) = \pi^{\mathbb{M}}(f_t, r_t, t | l_s, f_s, r_s, s) \int_0^\infty dl_t \pi^{\mathbb{M}}(l_t, t | f_t, r_t, l_0, f_0, r_0, 0) \quad (73)$$

$$\Phi^{\mathbb{M}}(f_t, r_t, t | f_s, r_s, s) = \pi^{\mathbb{M}}(f_t, r_t, t | l_s = 0, f_s, r_s, s) \int_0^\infty dl_t \pi^{\mathbb{M}}(l_t, t | f_t, r_t, l_s = 0, f_s, r_s, s) \quad (74)$$

As such, $\Psi^{\mathbb{M}}(f_t, r_t, t)$ and $\Phi^{\mathbb{M}}(f_t, r_t, t | f_s, r_s, s)$ can be obtained from the various transition densities related to the three processes l_t , f_t and r_t . Note that this approach to obtain the probability $\mathbf{Q}^{\mathbb{M}}(r_0, f_0, l_0, T)$ does not depend on the specific dynamics of the three processes l_t , f_t and r_t .

Implementation Here we obtain more explicit expressions for the functions $\Psi^{\mathbb{M}}(f_t, r_t, t)$ and $\Phi^{\mathbb{M}}(f_t, r_t, t | f_s, r_s, s)$ for the case that the dynamics of f_t and r_t are as specified in Section 4.2. More generally, the same approach works as long as f_t and r_t are jointly Gaussian. In our case we obtain:

$$\begin{aligned} \Psi^{\mathbb{M}}(f_t, r_t, t) &= \pi^{\mathbb{M}}(f_t, r_t, t | l_s, f_s, r_s, s) N \left(\frac{E_0^{\mathbb{M}}[l_t, t | f_t, r_t, l_0, f_0, r_0, 0]}{\text{var}_0^{\mathbb{M}}[l_t, t | f_t, r_t, l_0, f_0, r_0, 0]} \right) \\ \Phi^{\mathbb{M}}(f_t, r_t, t | f_s, r_s, s) &= \pi^{\mathbb{M}}(f_t, r_t, t | l_s = 0, f_s, r_s, s) N \left(\frac{E_0^{\mathbb{M}}[l_t, t | f_t, r_t, l_s = 0, f_s, r_s, s]}{\text{var}_0^{\mathbb{M}}[l_t, t | f_t, r_t, l_s = 0, f_s, r_s, s]} \right) \end{aligned}$$

Where $E^{\mathbb{M}} [l_t, t | f_t, r_t, l_s, f_s, r_s, s]$ and $var_s^{\mathbb{M}} [l_t, t | f_t, r_t, l_s, f_s, r_s, s]$ are given by:

$$E^{\mathbb{M}} [l_t, t | f_t, r_t, l_s, f_s, r_s, s] = E_s^{\mathbb{M}} [l_t] + \begin{bmatrix} covar_s^{\mathbb{M}} [l_t, f_t] \\ covar_s^{\mathbb{M}} [l_t, r_t] \end{bmatrix}' \Sigma_s^{\mathbb{M}}(f_t, r_t)^{-1} \begin{bmatrix} f_t - E_s^{\mathbb{M}} [f_t] \\ r_t - E_s^{\mathbb{M}} [r_t] \end{bmatrix} \quad (s < t),$$

and;

$$var_s^{\mathbb{M}} [l_t, t | f_t, r_t, l_s, f_s, r_s, s] = var_s^{\mathbb{M}} [l_t] - \begin{bmatrix} covar_s^{\mathbb{M}} [l_t, f_t] \\ covar_s^{\mathbb{M}} [l_t, r_t] \end{bmatrix}' \Sigma_s^{\mathbb{M}}(f_t, r_t)^{-1} \begin{bmatrix} covar_s^{\mathbb{M}} [l_t, f_t] \\ covar_s^{\mathbb{M}} [l_t, r_t] \end{bmatrix} \quad (s < t)$$

with:

$$\Sigma_s^{\mathbb{M}}(f_t, r_t) = \begin{bmatrix} var_s^{\mathbb{M}} [f_t] & covar_s^{\mathbb{M}} [f_t, r_t] \\ covar_s^{\mathbb{M}} [f_t, r_t] & var_s^{\mathbb{M}} [r_t] \end{bmatrix},$$

with $\pi^{\mathbb{M}}(f_t, r_t, t | l_s, f_s, r_s, s)$ the well-known transition density of a Gaussian process.

References

- P. BALDUZZI, S.R. DAS, S. FORESI and R.K. SUNDARAM (2000) Stochastic Mean Models of the Term-Structure of Interest Rates, Chapter 5 in *Advanced Tools for the Fixed-Income Professional*, B. Tuckman and N. Jegadeesh, Eds, John Wiley and Sons, New York, 128-161.
- F. BLACK and J. COX (1976) Valuing corporate securities: Some effects of bonds indentures provisions, *Journal of Finance* **31**, 351-367.
- E. BRIYS and F. DE VARENNE (1997) Valuing risky fixed rate debt: An extension, *Journal of Financial and Quantitative Analysis* **32**, 239-248.
- J.H. COCHRANE (2001) *Asset Pricing*, Princeton University Press
- P. COLLIN-DUFRESNE and B. GOLDSTEIN (2001) Do credit spreads reflect stationary leverage ratios?, *The Journal of Finance*, **56**, 2177-2208.
- G.R. DUFFEE (2002) Term premia and interest rate forecasts in affine models, *The Journal of Finance*, **58**, 405-443.
- D. DUFFIE and K. SINGLETON (1999) Modeling term structures of defaultable bonds, *The Review of Financial Studies* **12**, 687-720.
- E. FAMA and K. FRENCH (1993) Common risk factors in returns on bonds and stocks, *Journal of Financial Economics* **33**, 3-56.
- R. GESKE (1977), The Valuation of Corporate Liabilities as Compound Options, *Journal of Financial and Quantitative Analysis* **12**, 541-552.
- J. HUANG and M. HUANG (2003) How much of the corporate-Treasury yield spread is due to credit risk?, Working paper Stanford University.
- R. JARROW, D. LANDO and S. TURNBULL (1997) A Markow model for the term structure of credit spreads, *The Review of Financial Studies* **10**, 481-523.
- I. KARATZAS and S. SHREVE (1991) *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York.
- H. LELAND (2002) Predictions of expected default frequencies in structural models for debt, Working paper Haas Business School, University of California, Berkley.

- F. LONGSTAFF (2001) The flight to liquidity premium in US Treasury bond prices, Working paper, Anderson School of Business, UCLA.
- F. LONGSTAFF and E. SCHWARTZ (1995) Valuing Risky Debt: A New Approach, *Journal of Finance* **50** 789-820.
- I. MALITZ (1994) The modern role of bond covenants, The Research Foundation of the Institute of Chartered Financial Analysts.
- R.C. MERTON (1974), On the pricing of corporate debt: the risk structure of interest rates, *The Journal of Finance* **29**, 429-470.
- L. NIELSEN, J. SAA-REQUEJO and P. SANTA-CLARA (1993) Default risk and interest rate risk: The term structure of default spreads, Working paper, Insead.
- O. VASICEK (1977), An equilibrium characterization of the term structure, *Journal of Financial Economics* **5**, 177-188.