Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

A Financial Market Model for the Netherlands

A Methodological Refinement

Sander Muns

DP 09/2015-028

A financial market model for the Netherlands: A methodological refinement*

Sander Muns†

Friday 4th September, 2015

1 Introduction

The Committee Parameters (Langejan et al. (2014)) advises to use the KNW-model (after Koijen et al. (2010)) to generate a representative scenario set for feasibility studies of pension funds. The scenario set enables a stochastic analysis of such feasibility studies. The underlying KNW-model is based on an affine factor model for the term structure. Stock returns, bond returns, interest rates, and inflation depend on observed factors and two latent factors. As such, the model contains relations between key financial risk factors of pension funds. CPB's task is to estimate the model on Dutch data and, if appropriate, to calibrate some parameters in order to fit the recommendations of the Committee Parameters. Draper (2014) describes the current methods for this estimation and calibration.

The calibration aims to adjust the Ultimate Forward Rate (UFR) and certain long-term expectations and covariances of the variables in the model. However, this calibration process introduces some arbitrariness. More specifically, the resulting parameter set may deviate substantially from the maximum likelihood set, even when taking the restrictions of the calibration into account. Instead of calibrating the model, we show how to impose restrictions in a continuous-time affine term structure model. In this way, the parameters correspond to the optimum of a constrained maximum likelihood estimation. The results suggest that the method in Draper (2014) provides suboptimal parameter estimates.

The main result of this paper is the derivation of closed-form expressions for the long-term (unconditional) expectations, covariances, and the term structure. The expressions are required for the constrained likelihood optimization, and replace simulations for a long-run analysis of parameter sets. Our results apply to a wide range of continuous-time affine term strucure models with the Markov property, including the models in Dai and Singleton (2002) and Koijen et al. (2010).

*I am grateful to Nick Draper and Bas Werker for discussing several issues, and Albert van der Horst and Marcel Lever for helpful comments. I would also like to thank DNB for sharing their parameter set.

†email:s.muns@cpb.nl

The model is outlined in Section 2. Section 3 provides expressions for the mean and covariance of possibly transformed variables in a VAR(1)-model. Section 4 presents closed-form expressions for some characteristics of the term structure in terms of the parameters. The estimation results are in Section 5. We draw conclusions in Section 6.

2 The model[1]

Affine term structure models are very common in the literature (see e.g., Dai and Singleton (2002), Koijen et al. (2010), and Gürkaynak and Wright (2012)). We outline a generalized version of the model in Koijen et al. (2010)[2], though our results apply to the VAR(1)-representation of any continuous-time affine term structure model.

As key determinants of pension risk, we consider inflation $\mathrm{d}\Pi/\Pi$, the stock return $\mathrm{d}S/S$, the bond portfolio return $\mathrm{d}P^B(\tau)/P^B(\tau)$ with a constant maturity τ, and the term structure $y(\tau)$. The assumed dynamics are[3]

Unobserved states	$\mathrm{d}X_t = -KX_t\mathrm{d}t + \mathrm{d}\tilde{Z}_t$	(1)
Instant. expected inflation	$\pi_t = \delta_{0\pi} + \delta'_{1\pi}X_t$	(2)
Price index process	$\frac{\mathrm{d}\Pi_t}{\Pi_t} = \pi_t\mathrm{d}t + \sigma'_\Pi\mathrm{d}Z_t$	(3)
Instant. nom. interest rate	$R_t(0) = \delta_{0R} + \delta'_{1R}X_t$	(4)
Stock return process	$\frac{\mathrm{d}S_t}{S_t} = (R_t(0) + \eta_S)\,\mathrm{d}t + \sigma'_S\mathrm{d}Z_t$	(5)
Bond return process	$\frac{\mathrm{d}P_t^B(\tau)}{P_t^B(\tau)} = \left(R_t(0) + B(\tau)'\tilde{\Lambda}_t\right)\mathrm{d}t + B(\tau)'\mathrm{d}\tilde{Z}_t$	(6)
Prices of risk	$\Lambda_t = \Lambda_0 + \Lambda_1 X_t$	(7)
Stochastic discount factor	$\frac{\mathrm{d}\phi_t}{\phi_t} = -R_t(0)\mathrm{d}t - \Lambda'_t\mathrm{d}Z_t$	(8)

where $\mathrm{d}Z \sim N(0_{(k+2)\times 1}, I_{(k+2)\times(k+2)})$, and

$$K \in \mathbb{R}^{k\times k} \qquad \sigma_\Pi, \sigma_S, \Lambda_0 \in \mathbb{R}^{k+2} \qquad \mathrm{d}\tilde{Z}_t = \begin{bmatrix} I_{k\times k} & 0_{k\times 2}\end{bmatrix}\mathrm{d}Z_t$$
$$\delta_{1\pi}, \delta_{1R} \in \mathbb{R}^k \qquad \Lambda_1 \in \mathbb{R}^{(k+2)\times k} \qquad \tilde{\Lambda}_t = \begin{bmatrix} I_{k\times k} & 0_{k\times 2}\end{bmatrix}\Lambda_t$$

To ensure identification of X and Z, Koijen et al. (2010) impose

(*i*) $\sigma_{\Pi(k+2)} = 0$: This excludes rotations (e.g., a switch) of the last two components of Z.

[1] Koijen et al. (2010) and Draper (2014) contain further details and references on the derivations in this section.

[2] This model is based on assumptions and derivations in Duffie and Kan (1996) and Duffee (2002).

[3] The assumptions are identical to Koijen et al. (2010), except that (*i*) here, the number of state variables is k instead of two, and (*ii*) we add the implied dynamics of a bond portfolio with constant duration as introduced in Draper (2014).

(ii) K is a lower triangular matrix: This excludes rotations (e.g., a switch) of the components of X, and thus of the first k components of Z.

More implicit identification restrictions are the absence of a drift term in (1) which excludes a translation of X, and the unit standard deviations of $\mathrm{d}Z$ which excludes a scaling of X. To exclude that $-X$ gives the same fit as X, an additional identification restriction should be imposed on the sign of X or on the sign of certain parameters. Of course, none of the identification restrictions changes any of the dynamics of the non-state variables, or any of the expressions we derive.

Using stochastic calculus, the aggregate model is a multivariate Ornstein-Uhlenbeck process:

$$\mathrm{d}\Upsilon_t = (\Theta_0 + \Theta_1 \Upsilon_t)\mathrm{d}t + \sigma_\Upsilon \mathrm{d}Z_t \tag{9}$$

where

$$\Upsilon_t = \begin{bmatrix} X_t \\ \log \Pi_t \\ \log S_t \\ \log P^B(\tau) \end{bmatrix} \qquad \sigma_\Upsilon = \begin{bmatrix} I_{k\times k} & 0_{k\times 3} \\ \multicolumn{2}{c}{\sigma_\Pi'} \\ \multicolumn{2}{c}{\sigma_S'} \\ B(\tau)' & 0_{1\times 3} \end{bmatrix}$$

$$\Theta_0 = \begin{bmatrix} 0_{k\times 1} \\ \delta_{0\pi} - \frac{1}{2}\sigma_\Pi'\sigma_\Pi \\ \delta_{0R} + \eta_S - \frac{1}{2}\sigma_S'\sigma_S \\ \delta_{0R} + B(\tau)'\tilde{\Lambda}_0 - \frac{1}{2}B(\tau)'B(\tau) \end{bmatrix} \qquad \Theta_1 = \begin{bmatrix} -K & 0_{k\times 3} \\ \delta_{1\pi}' & 0_{1\times 3} \\ \delta_{1R}' & 0_{1\times 3} \\ \delta_{1R}' + B(\tau)'\tilde{\Lambda}_1 & 0_{1\times 3} \end{bmatrix}$$

$$\tilde{\Lambda}_0 = \begin{bmatrix} I_{k\times k} & 0_{k\times 2} \end{bmatrix} \Lambda_0 \qquad \tilde{\Lambda}_1 = \begin{bmatrix} I_{k\times k} & 0_{k\times 2} \end{bmatrix} \Lambda_1$$

After discretization with step length h, a VAR(1)-model emerges from (9):

$$Y_t = \gamma + \Gamma Y_{t-h} + \varepsilon_t \qquad \varepsilon_t \sim N(0, V) \tag{10}$$

where $Y_t = \begin{bmatrix} X_t & \Delta \log \Pi_t & \Delta \log S_t & \Delta \log P^B(\tau) \end{bmatrix}'$ is a random vector, and[4]

$$\gamma \;\; = UFU^{-1}\Theta_0 \qquad \Gamma = U \exp(Dh) U^{-1} \qquad V = UWU' \tag{11}$$

with F a diagonal matrix, $\Theta_1 = UDU^{-1}$, and

$$F_{ii} = h\alpha(D_{ii}h)$$

$$\alpha(x) = \begin{cases} (\exp(x) - 1)/x & x \neq 0 \\ 1 & x = 0 \end{cases}$$

$$W_{ij} = \left[U^{-1}\sigma_Y \left(U^{-1}\sigma_Y\right)'\right]_{ij} h\alpha\left(\left[D_{ii} + D_{jj}\right] h\right)$$

[4] As in Dai and Singleton (2002), Sangvinatsos and Wachter (2005), Koijen et al. (2010), the expressions in (11) are for diagonalizable Θ_1. Though this class is dense in the class of complex matrices, Appendix A derives the expressions for the general case. This extension seems to be unknown in the literature.

The random vector Y is stationary if X is stationary, or equivalently, K is positive definite in (1).

The flexible step length h is useful if estimation and simulation have a different frequency. Without loss of generality, we restrict the analysis to the case $h = 1$.

Using (8) and a no arbitrage argument, the term structure of nominal continuously compounded (i.e., nominal log) yields y_t is affine in the k state variables:

$$y_t(\tau) = -\frac{1}{\tau}\left(A(\tau) + B(\tau)' X_t\right) + \xi_{\tau,t} \tag{12}$$

with τ the term to maturity, and $\xi_{\tau,t} \sim N(0, \sigma_\tau)$ an independent measurement error. The maximum likelihood estimate corresponds to the maximal loglikelihood sum of the disturbances ε_t in (10) and the measurement errors $\xi_{\tau,t}$ in (12).

The functions $A : \tau \to \mathbb{R}$, and $B : \tau \to \mathbb{R}^k$ satisfy

$$A(\tau) = \int_0^\tau \dot{A}(s)\mathrm{d}s \qquad A(0) = 0 \tag{13}$$

$$B(\tau) = M^{-1}\left[\exp\left(-M\tau\right) - I_{k\times k}\right]\delta_{1R} \qquad B(0) = 0_{k\times 1} \tag{14}$$

$$\dot{A}(\tau) = -\delta_{0R} - \tilde{\Lambda}_0' B(\tau) + \frac{1}{2}B(\tau)'B(\tau) \qquad \dot{A}(0) = -\delta_{0R} \tag{15}$$

$$\dot{B}(\tau) = -\delta_{1R} - MB(\tau) \qquad \dot{B}(0) = -\delta_{1R} \tag{16}$$

where $M = \left(K + \tilde{\Lambda}_1\right)'$ and $\dot{x}$ denotes the derivative of the function x. Define

$$\mathbf{b}_0 := \lim_{\tau\to\infty} B(\tau) = -M^{-1}\delta_{1R}.$$

We assume a positive definite M in (16) which ensures a finite $\mathbf{b}_0$.

The Ultimate Forward Rate ($\mathrm{UFR}_{\log}$) is the cross-sectional $\lim_{\tau\to\infty}\mathbb{E}[y_t(\tau)]$. Since X is stationary and $\lim_{\tau\to\infty} B(\tau)/\tau = 0$, we find for each t

$$\mathrm{UFR}_{\log} = \lim_{\tau\to\infty} -\frac{A(\tau)}{\tau} = \lim_{\tau\to\infty} -\dot{A}(\tau) = \delta_{0R} + \left(\tilde{\Lambda}_0 - \frac{1}{2}\mathbf{b}_0\right)'\mathbf{b}_0. \tag{17}$$

The UFR as an annually compounded rate of return is given by

$$\mathrm{UFR} = \exp\left(\mathrm{UFR}_{\log}\right) - 1. \tag{18}$$

Equation (12) implies that only measurement error $\xi_{\tau,t}$ can explain time variation in the observed UFR:

$$\lim_{\tau\to\infty} y_t(\tau) = \mathrm{UFR}_{\log} + \lim_{\tau\to\infty}\xi_{\tau,t}$$

Since $\mathbb{E}[X_t] = \mathbb{E}[\xi_{\tau,t}] = 0$ if $t \to \infty$, the unconditional expected term structure is determined by[5]

$$R(\tau) := \lim_{t\to\infty} \mathbb{E}[y_t(\tau)] = -\frac{A(\tau)}{\tau}. \tag{19}$$

The dynamics in the price of risk in (7) are completely determined by the k prices in $\tilde{\Lambda}$. To see this, (i) inflation, (ii) the stock return, and (iii) the bond return depend on the state variables in X as well as variable-specific shocks. None of these three variables affects the dynamics of priced risk:

(i) The price of unexpected inflation risk has no risk premium in (3) because the price of unexpected inflation risk cannot be identified on the basis of data on the nominal side of the economy alone. We follow Koijen et al. (2010) by assuming that the price of unexpected inflation risk equals zero.

(ii) The equity premium in (5) is assumed to be fixed at η_S, and thus time independent.

(iii) The risk premium of the bond price in (6) is derived from the affine term structure $y_t(\tau)$ in (12).

Hence, the dynamics in the price of risk are implied by the dynamics of the k linearly independent risks in $\tilde{\Lambda}$. The bottom two risks in Λ that are absent in $\tilde{\Lambda}$ follow from the restriction on the equity premium ($\sigma_S'\Lambda_0 = \eta_S$ and $\sigma_S'\Lambda_1 = 0_{1\times k}$), and the absence of priced unexpected inflation risk ($\Lambda_{0(k+1)} = \Lambda_{1(k+1,1)} = \ldots = \Lambda_{1(k+1,k)} = 0$). Therefore, the parameters that determine the price of risk are in $\tilde{\Lambda}_0$ and $\tilde{\Lambda}_1$.

It is straightforward to augment Y in (10) with additional bond portfolios having different maturities (augment Θ_0, Θ_1, and σ_Υ in (9)), or with yields from (12) (augment γ, Γ, and V directly in (10)). Adding bond portfolios or yields to Y does not add any new parameter to the VAR(1)-model. In any case, the model we estimate contains the same number of parameters.

In the next section, we derive the mean and covariance of Y in (10). Closed-form expressions for some characteristics of $R(\tau)$ in (19) are in Section 4.

3 Mean and covariance[6]

We derive expressions for the mean vector and covariance matrix of the variables Y in the VAR(1)-model in (10). Such expressions are needed for the implementation of certain constraints in an

[5] To keep our derivations manageable, we consider the term structure of *continuously compounded* yields. Otherwise, we need to consider yields following a lognormal distribution with long-run expectation

$$\lim_{t\to\infty} \mathbb{E}[\exp(y_t(\tau))] = \exp\left(-\tau^{-1}A(\tau) + \tau^{-2}B(\tau)'\mathrm{Var}(X)\,B(\tau)\right).$$

The additional term from the variance would complicate the analysis unnecessarily.

[6] The main derivations in this section are based on Cochrane (2005).

optimization procedure. In addition, they are helpful for informative purposes on the long-run behaviour of the model. The expressions are unconditional on realizations of Y which means that the expressions are long-run expectations. The random vector Y may contain any stationary series: inflation, GDP growth, returns (possibly continuously compounded), a latent stationary factor, etc. We derive arithmetic means and variances of the annually compounded returns $\tilde{Y} = \exp(Y) - 1$. We end this section with an expression for the geometric mean of Y.

Without loss of generality, we consider a step length of one period, $h = 1$. Induction on (10) gives

$$\begin{aligned} Y_t &= \gamma + \Gamma\left(\gamma + \Gamma Y_{t-2} + \varepsilon_{t-1}\right) + \varepsilon_t \\ &= \Gamma^{t-1} Y_0 + \sum_{s=0}^{t-1} \Gamma^s \left(\gamma + \varepsilon_{t-s}\right) \end{aligned}$$

It follows from the stationarity of Y that the eigenvalues of Γ are within the unit circle of the complex plane. Accordingly, $\Gamma^{t-1} Y_0 \to 0$ as $t \to \infty$, and for some μ and Σ,

$$Y_t \to Y^* \sim N(\mu, \Sigma) \tag{20}$$

$$Y^* \sim \gamma + \Gamma Y^* + \varepsilon \tag{21}$$

Using (21):

$$\mu = (I - \Gamma)^{-1} \gamma \tag{22}$$

$$\Sigma = \Gamma \Sigma \Gamma' + V \tag{23}$$

$$\operatorname{vec}(\Sigma) = (\Gamma \otimes \Gamma) \operatorname{vec}(\Sigma) + \operatorname{vec}(V)$$

This leads to an explicit expression for the elements in Σ:[7]

$$\operatorname{vec}(\Sigma) = (I - \Gamma \otimes \Gamma)^{-1} \operatorname{vec}(V) \tag{24}$$

Next, we consider the annually compounded rates of return $\tilde{Y}_t = \exp(Y_t) - 1$ where Y is a vector of annually compounded returns. Let D_Σ represent the diagonal matrix with the diagonal elements of Σ. Using well-known properties of the lognormal distribution and (20), we find as $t \to \infty$,

$$\mathbb{E}\left[\tilde{Y}_t\right] \to \exp\left(\mu + \frac{1}{2} D_\Sigma\right) - 1 \tag{25}$$

$$\operatorname{Var}\left[\tilde{Y}_t\right] \to \exp\left(D_\Sigma - 1\right) \exp\left(2\mu + D_\Sigma\right) \tag{26}$$

Equations (25)-(26) refer to the limiting distribution of continuously compounded returns over a *single* period t. Expressions for the cumulative mean arithmetic return $\frac{1}{t} \sum_{s=1}^{t} \tilde{Y}_s$ easily follow from the central limit theorem.

[7] Equation (24) differs from the incorrect equation (54) in Draper (2014), which is based on (A4) on p.253 in Banerjee et al. (1993). This equation is inappropriate here since it contains cross-covariances. In fact, our (23) is similar to (A3) in Banerjee et al. (1993).

The geometric mean return Y_t^G is a cumulative return that depends as follows on the mean realized continuously compounded return $\bar{Y}_t := \frac{1}{t}\sum_{s=1}^{t} Y_s$:

$$\begin{aligned}
1 + Y_t^G &= \left(\prod_{s=1}^{t}\left(1+\tilde{Y}_s\right)\right)^{1/t} \\
&= \left(\prod_{s=1}^{t}\exp\left(Y_s\right)\right)^{1/t} \\
&= \exp\left(\frac{1}{t}\sum_{s=1}^{t} Y_s\right) \\
&= \exp\left(\bar{Y}_t\right)
\end{aligned}$$

Thus, the gross return $1 + Y_t^G$ follows a lognormal distribution $LN\left(\mu, D_\Sigma/t\right)$ with

$$\begin{aligned}
\mathbb{E}\left[1 + Y_t^G\right] &= \exp\left(\mu + \frac{D_\Sigma}{2t}\right) \\
\mathrm{Var}\left[1 + Y_t^G\right] &= \left(\exp\left(\frac{D_\Sigma}{t}\right) - 1\right)\exp\left(2\mu + \frac{D_\Sigma}{t}\right)
\end{aligned}$$

As such, Y_t^G converges to a degenerate distribution Y^G with full probability mass at

$$\mathbb{E}\left[Y^G\right] = \exp\left(\mu\right) - 1. \tag{27}$$

When optimizing parameter sets in the stationary model (10), the expressions in (22), (24)–(27) enable us to restrict the unconditional, i.e., long-run, means and covariances of the variables.

4 Restrictions on the term structure

In Section 3, the time series model in (10) led to relatively straightforward expressions for the mean and covariance of the variables in Y. In this section, we derive more complicated, though closed-form, expressions for the cross-sectional expressions of some characteristics of the unconditional, i.e. long-run, term structure, $R(\tau)$ in (19).

Referring to the functions $A(\tau)$ and $B(\tau)$ in (13) and (14), Duffee (2002, p.408) states: 'The functions $A(\tau)$ and $B(\tau)$ can be calculated numerically by solving a series of ordinary differential equations (ODEs).' Compared to numerical evaluations, the closed-form expressions we derive below are superior in terms of speed and accuracy. In a similar setting, Dai and Singleton (2002) also derive closed-form expressions for $A(\tau)$ and $B(\tau)$. Nonetheless, their result contains a matrix exponential and the trace of a complicated matrix product. As a consequence, the link to the underlying parameters is less obvious in their expressions.

Let us start with a straightforward restriction. Since the price of unexpected inflation risk equals zero (see p.5), the long-run instantaneous real interest rate is the difference between the long-run nominal interest rate δ_{0R} in (4) and the long-run instantaneous inflation $\delta_{0\pi}$ in (2):

$$\delta_{0r} = \delta_{0R} - \delta_{0\pi}. \tag{28}$$

A nonnegative long-run instantaneous real interest rate is thus equivalent to $\delta_{0R} \geq \delta_{0\pi}$.

It is straightforward from (13)–(16) that

$$\ddot{A}(\tau) = \left(B(\tau) - \tilde{\Lambda}_0\right)' \dot{B}(\tau) \qquad \ddot{A}(0) = \tilde{\Lambda}_0' \delta_{1R} \tag{29}$$

$$\ddot{B}(\tau) = -M\dot{B}(\tau) \qquad \ddot{B}(0) = M\delta_{1R} \tag{30}$$

$$\dddot{A}(\tau) = \left(B(\tau) - \tilde{\Lambda}_0\right)' \ddot{B}(\tau) + \dot{B}(\tau)'\dot{B}(\tau) \qquad \dddot{A}(0) = \left(-M'\tilde{\Lambda}_0 + \delta_{1R}\right)' \delta_{1R} \tag{31}$$

We must have $\lim_{\tau\to\infty} |B(\tau)| < \infty$ to ensure $\lim_{\tau\to\infty} \mathbb{P}(|y_t(\tau)| < \infty) = 1$. This corresponds to positive eigenvalues of M. Equivalently, the leading principal minors are positive

$$|M_i| > 0 \qquad i = 1, \ldots, k \tag{32}$$

where M_i is the upper-left $i \times i$ sub-matrix.

Expressions for the term structure $R(\tau)$ at $\tau = 0$ follow most easily from a Taylor series expansion of the integrand $\dot{A}(s)$ in (13) around $\tau = 0$:

$$\begin{aligned}
R(\tau) &= -\frac{1}{\tau} A(\tau) \\
&= -\frac{1}{\tau} \int_0^\tau \dot{A}(s)\mathrm{d}s \\
&= -\frac{1}{\tau} \int_0^\tau \left(\dot{A}(0) + \ddot{A}(0)s + \frac{1}{2}\dddot{A}(0)s^2 + O(s^3)\right) \mathrm{d}s \\
&= \frac{1}{\tau} \int_0^\tau \left(\delta_{0R} - \tilde{\Lambda}_0'\delta_{1R}s - \frac{1}{2}\left(-M'\tilde{\Lambda}_0 + \delta_{1R}\right)' \delta_{1R}s^2 + O(s^3)\right) \mathrm{d}s \\
&= \delta_{0R} - \frac{1}{2}\tilde{\Lambda}_0'\delta_{1R}\tau + \frac{1}{6}\left(M'\tilde{\Lambda}_0 - \delta_{1R}\right)' \delta_{1R}\tau^2 + O(\tau^3)
\end{aligned}$$

Therefore,

$$R(0) = \delta_{0R} \tag{33}$$

$$\dot{R}(0) = -\frac{1}{2}\tilde{\Lambda}_0'\delta_{1R} \tag{34}$$

$$\ddot{R}(0) = \frac{1}{3}\left(M'\tilde{\Lambda}_0 - \delta_{1R}\right)' \delta_{1R} \tag{35}$$

Next, we find closed-form expressions for $A(\tau)$, $\dot{A}(\tau)$, and $B(\tau)$ and $R(\tau)$ with $\tau > 0$. In Appendix B, we show that we can discard M with complex eigenvalues because the corresponding term structure has components with an oscillating pattern. In addition, we assume that M is diagonalizable since (i) this simplifies derivations significantly, (ii) we did not encounter non-diagonalizable (defective) M in the likelihood optimization when optimizing over K and $\tilde{\Lambda}_1$, and (iii) the class of defective matrices has measure zero in the class of matrices. Nonetheless, the Jordan matrix decomposition can extend the results to defective matrices.

Let D_λ denote a diagonal matrix with diagonal elements of the vector λ. For a diagonalizable matrix M with eigenvalue vector λ and eigenvalue decomposition $M = VD_\lambda V^{-1}$:

$$
\begin{aligned}
M^{-1}\exp(-M\tau)\,\delta_{1R} &= M^{-1}VD_{\exp(-\lambda\tau)}V^{-1}\delta_{1R} \\
&= VD_\lambda^{-1}D_{\exp(-\lambda\tau)}V^{-1}\delta_{1R} \\
&= \sum_{i=1}^{k}\mathbf{b}_i e^{-\lambda_i\tau}
\end{aligned}
$$

where

$$
V = \begin{bmatrix}\mathbf{v}_1 & \dots & \mathbf{v}_k\end{bmatrix} \qquad V^{-1} = \begin{bmatrix}\mathbf{v}_1^{-1} \\ \vdots \\ \mathbf{v}_k^{-1}\end{bmatrix} \qquad \mathbf{b}_i = \frac{1}{\lambda_i}\left(\mathbf{v}_i\mathbf{v}_i^{-1}\right)\delta_{1R} \qquad i = 1,\dots,k \tag{36}
$$

This gives for (14) the analytical expression

$$
B(\tau) = \mathbf{b}_0 + \sum_{i=1}^{k}\mathbf{b}_i e^{-\lambda_i\tau} \tag{37}
$$

where $\mathbf{b}_0 = -M^{-1}\delta_{1R}$. By substituting (37) in (15), we may find

$$
\begin{aligned}
\dot{A}(\tau) &= \left(\frac{1}{2}B(\tau) - \tilde{\Lambda}_0\right)' B(\tau) - \delta_{0R} \\
&= \left(\frac{1}{2}\left[\mathbf{b}_0 + \sum_{i=1}^{k}\mathbf{b}_i e^{-\lambda_i\tau}\right] - \tilde{\Lambda}_0\right)'\left(\mathbf{b}_0 + \sum_{i=1}^{k}\mathbf{b}_i e^{-\lambda_i\tau}\right) - \delta_{0R} \\
&= \left(\frac{1}{2}\mathbf{b}_0 - \tilde{\Lambda}_0\right)'\mathbf{b}_0 - \delta_{0R} + \left(\mathbf{b}_0 - \tilde{\Lambda}_0\right)'\sum_{i=1}^{k}\mathbf{b}_i e^{-\lambda_i\tau} + \frac{1}{2}\sum_{i=1}^{k}e^{-\lambda_i\tau}\mathbf{b}_i'\sum_{j=1}^{k}\mathbf{b}_j e^{-\lambda_j\tau} \\
&= a_0^{(1)} + \sum_{i=1}^{k}\left\{a_i^{(1)}e^{-\lambda_i\tau} + \sum_{j=1}^{k}a_{ij}^{(1)}e^{-(\lambda_i+\lambda_j)\tau}\right\}
\end{aligned} \tag{38}
$$

where for $i, j \in \{1,\dots,k\}$

$$
a_0^{(1)} = \left(\frac{1}{2}\mathbf{b}_0 - \tilde{\Lambda}_0\right)'\mathbf{b}_0 - \delta_{0R} \qquad a_i^{(1)} = \left(\mathbf{b}_0 - \tilde{\Lambda}_0\right)'\mathbf{b}_i \qquad a_{ij}^{(1)} = \frac{1}{2}\mathbf{b}_i'\mathbf{b}_j
$$

By integrating (38) and using $A(0) = 0$ (see (13)),

$$
\begin{aligned}
A(\tau) &= a_0^{(1)}\tau + \sum_{i=1}^{k}\left\{\frac{a_i^{(1)}}{\lambda_i}\left(1-e^{-\lambda_i\tau}\right) + \sum_{j=1}^{k}\frac{a_{ij}^{(1)}}{\lambda_i+\lambda_j}\left(1-e^{-(\lambda_i+\lambda_j)\tau}\right)\right\} \\
&= a_0 + a_0^{(1)}\tau + \sum_{i=1}^{k}\left\{a_i e^{-\lambda_i\tau} + \sum_{j=1}^{k} a_{ij}e^{-(\lambda_i+\lambda_j)\tau}\right\} \qquad (39)
\end{aligned}
$$

where for $i, j \in \{1, \ldots, k\}$

$$
a_0 = \sum_{i=1}^{k}\left\{\frac{a_i^{(1)}}{\lambda_i} + \sum_{j=1}^{k}\frac{a_{ij}^{(1)}}{\lambda_i+\lambda_j}\right\} \qquad a_i = -\frac{a_i^{(1)}}{\lambda_i} \qquad a_{ij} = -\frac{a_{ij}^{(1)}}{\lambda_i+\lambda_j}
$$

Substitution of (39) in (19) gives for the unconditional term structure:[8]

$$
R(\tau) = -a_0^{(1)} - \frac{1}{\tau}\left(a_0 + \sum_{i=1}^{k}\left\{a_i e^{-\lambda_i\tau} + \sum_{j=1}^{k} a_{ij}e^{-(\lambda_i+\lambda_j)\tau}\right\}\right) \qquad (40)
$$

The slope of the unconditional term structure is:

$$
\dot{R}(\tau) = -\frac{1}{\tau}\left(R(\tau) + a_0^{(1)} + \sum_{i=1}^{k}\left\{a_i^{(1)} e^{-\lambda_i\tau} + \sum_{j=1}^{k} a_{ij}^{(1)}e^{-(\lambda_i+\lambda_j)\tau}\right\}\right) \qquad (41)
$$

The expressions in (40) and (41) enable us to restrict the level and the slope of the term structure at certain τ.

To facilitate an easier optimization of the likelihood function, we optimize over $M = \left(K + \tilde{\Lambda}_1\right)'$ instead of $\tilde{\Lambda}_1$. Then, we can impose that M has k distinct positive eigenvalues and each eigenvector $\mathbf{v}_i$ of M has a positive first entry. The optimal (continuous) likelihood function is unaffected when matrices with identical real eigenvalues are excluded because the class of matrices with k distinct real eigenvalues λ_i and $\mathbf{v}_{i(1)} \neq 0$ is dense in the class of matrices with real eigenvalues. Similarly, the sign restriction does not restrict the optimal likelihood since $\mathbf{v}_i$ and $-\mathbf{v}_i$ are both eigenvectors of M.

Therefore, an appropriate k-dimensional matrix M has a unique representation $M = VD_\lambda V^{-1}$ with $(i = 1, \ldots, k)$

$$
0 < \lambda_1 < \ldots < \lambda_k \quad V = \begin{bmatrix}\mathbf{v}_1 & \ldots & \mathbf{v}_k\end{bmatrix} \quad \mathbf{v}_{i(1)} > 0 \quad \|\mathbf{v}_i\| = 1
$$

[8]Indeed, $\mathrm{UFR}_{\log} := \lim_{\tau\to\infty} R(\tau) = -a_0^{(1)} = \delta_{0R} + \left(\tilde{\Lambda}_0 - \frac{1}{2}\mathbf{b}_0\right)' \mathbf{b}_0$ which corresponds to (17).

To deal with the inequality constraints on λ_i, the auxiliary parameters $\tilde{\lambda}_i = \log(\lambda_i - \lambda_{i-1})$ with $\lambda_0 = 0$ are useful in the optimization since each $\tilde{\lambda}_i$ is unrestricted and $\lambda_j = \sum_{i=1}^{j} \exp\left(\tilde{\lambda}_i\right)$. The vectors $\mathbf{v}_i$ represent directions, which are efficiently captured by a polar coordinate system.

For the two-dimensional case ($k = 2$), one can verify that distinct real eigenvalues are equivalent to the condition

$$\operatorname{tr}(M)^2 > 4 \det(M) \tag{42}$$

5 Estimation results

We estimate the model in Section 2 with $k = 2$ and certain restrictions on the variables. The optimal parameter set is compared with other parameter sets: the estimated set in Draper (2014), and two calibrations of this set. We use the same time series in Van den Goorbergh et al. (2011) to facilitate a fair comparison with Draper (2014). This dataset contains quarterly time series for inflation, stock returns, and swap yields with maturities of 3 months, 1, 2, 3, 5, and 10 years. We have updated this dataset with 2014 data.[9]

We follow Draper (2014) by deriving the latent state variables in X_t from the two-year and five-year yields. More specifically, we assume that both yields $y_t(2)$ and $y_t(5)$ are measured without error ($\xi_{2,t} = \xi_{5,t} \equiv 0$) such that the state vector $X_t \in \mathbb{R}^2$ at each time t follows uniquely from (12):[10]

$$\begin{bmatrix} -2y_t(2) \\ -5y_t(5) \end{bmatrix} = \begin{bmatrix} A(2) \\ A(5) \end{bmatrix} + \begin{bmatrix} B(2)' \\ B(5)' \end{bmatrix} X_t$$

where $A(\tau) \in \mathbb{R}$ is the intercept, and $B(\tau) \in \mathbb{R}^2$ captures time variation in the risk premiums.

The functions A and B depend on the model parameters. The optimal set of these parameters maximizes the sum of the loglikelihoods of the observed series in (10) and (12).[11] That is, the likelihood of the time series of inflation, stock returns, and the term structure. The likelihood of the term structure is based on the measurement errors $\hat{\xi}_{\tau,t}$ (residuals) of the three-month, one-year, three-year, and ten-year yields. Since the yields with 2 and 5 years to maturity are measured without error, they completely determine the two-dimensional state X_t at each time t. The state in turn determines in (10) the errors at the other maturities.

Rather than optimizing the parameter set without any constraint and calibrating the model afterwards, we add the constraints in Table 1 to the optimization procedure. The constraints force us to reconsider the optimization procedure of Goffe et al. (1994) as employed in Draper (2014). This procedure iterates over potentially optimal sets of parameters, and should reduce the likelihood

[9] I thank Peter Vlaar for providing the dataset. A detailed data description is in Appendix C.

[10] This differs from the Kalman procedure described in Appendix C of Koijen et al. (2010). However, Table 5 in their paper shows a zero measurement error for the one-year and five-year interest rates. As such, Koijen et al. (2010) employ a similar procedure as in Draper (2014).

[11] When feasible, Duffee and Stanton (2012) advocate a maximum likelihood approach in favor of the efficient method of moments due to the finite-sample properties of the estimators.

restriction	restriction on variables	restriction on parameters	equation
long-run inflation 2%	$\lim_{t\to\infty} \mathbb{E}[\pi_t] = \log(1.02)$	$\exp(\mu_\pi) - 1 = 0.02$	(27)
UFR 4.2%	$R(\tau) = \log(1.042)$	$\delta_{0R} + \left(\tilde{\Lambda}_0 - \frac{1}{2}\mathbf{b}_0\right)' \mathbf{b}_0 = \log(1.042)$	(17)
converging ts	$\lim_{\tau\to\infty} \lvert R(\tau)\rvert < \infty$	$\lvert M_i\rvert > 0$	(32)
		$k = 2$: $M_{11} > 0$, $M_{11}M_{22} > M_{12}M_{21}$	(32)
non-oscillating ts		$\mathrm{Im}(\lambda_i(M)) = 0$	(51)
		$k = 2$: $\mathrm{tr}(M)^2 > 4\det(M)$	(42)
nonnegative real int. rate at $\tau = 0$	$r(0) \geq 0$	$\delta_{0R} \geq \delta_{0\pi}$	(28)
increasing ts at $\tau = 0$	$\dot{R}(0) \geq 0$	$-\frac{1}{2}\tilde{\Lambda}_0'\delta_{1R} \geq 0$	(34)
concave ts at $\tau = 0$	$\ddot{R}(0) \leq 0$	$\frac{1}{3}\left(M'\tilde{\Lambda}_0 - \delta_{1R}\right)' \delta_{1R} \leq 0$	(35)
increasing ts at $\tau = 120$	$\dot{R}(120) \geq 0$	$0 \leq -\frac{1}{120}(R(120) + a_0^{(1)} + \ldots$ $+ \sum_{i=1}^{k} e^{-120\lambda_i} \left\{a_i^{(1)} + \sum_{j=1}^{k} a_{ij}^{(1)} e^{-120\lambda_j}\right\})$	(41)

Table 1: Restrictions, ts = term structure

of finding a local, non-global, optimum. The latter is highly relevant for our large scale optimization with 23 parameters.[12]

As the method of Goffe et al. (1994) cannot deal with restrictions, we rewrote both variable restrictions of the equality form $a'x = b$ (long-run inflation at 2%, and UFR at 4.2%) in such a way that one parameter is uniquely determined by the other parameters. The substitutions ensure that one free parameter is dropped for each restriction. The remaining parameter set is without any restriction of the equality form $a'x = b$. A candidate parameter set x which violates any of the restrictions in the inequality form $a'x \leq b$ or $a'x \geq b$ is simply discarded.

In addition to the substitutions introduced above, we supplement the optimization algorithm of Goffe et al. (1994) with a constrained interior point optimization routine.

Table 2 presents the estimates and the loglikelihood of four different parameter sets:

(i) the maximum likelihood estimate in Draper (2014), which is based on the sample 1973-2013.

(ii) the calibrated parameter set in Draper (2014).

(iii) the calibrated parameter set of DNB based on Draper (2014), and employed for the 2015Q2 feasibility study for Dutch pension funds.

(iv) our maximum likelihood estimate using the full sample 1973-2014 subject to the restrictions in Table 1.

Taking account of the wide standard errors of the estimates (i) and (iv), the different parameter estimates are close to each other.[13] Nonetheless, the parameter sets differ in a number of ways. The constraint of a nonnegative real interest rate at $\tau = 0$ is binding for estimate (iv) since $\delta_{0r} = \delta_{0R} - \pi_0 = 0$ in (28). The other parameter estimates correspond to a strictly positive instantaneous real interest rate. Notably, the likelihood LL_{2013} of our optimal estimate (iv) exceeds the loglikelihood of the estimate (i) from Draper (2014). This is surprising as our sample includes 2014 while LL_{2013} is based on the sample 1973-2013. Apparently, the optimization algorithm in Draper (2014) found a local optimum.

Table 3 reports the UFR and the long-run statistics of each parameter set. It turns out that the relatively small difference in parameters have a large impact on (a) the equity risk premium, as measured by $R^S - R(0)$, and (b) the term structure, as measured by $\text{UFR}_{\log}$, $R(\tau)$, and $R^B(\tau)$ with $\tau > 0$.

First, the equity premium is mainly explained by differences in $\eta_S - \delta_{0R}$. By (4) and (5), a high η_S results in a high $\mathbb{E}\big[R^S\big]$ in estimate (ii). In estimate (iv), a low η_S and a low δ_{0R} imply a low $\mathbb{E}\big[R^S\big]$ and a low $\mathbb{E}[R(0)]$, respectively. Second, following (33) and (34), the relatively large differences in the estimates of δ_{0R} and Λ_0 explain differences in levels, slopes and curvatures of the term structure. This translates into large differences of the UFR, $R(\tau)$, and $B(\tau)$, particularly

[12] Koijen et al. (2010) do not refer to a specific optimization procedure, though the same large number of parameters is estimated.

[13] The standard errors of (i) slightly differ from Draper (2014) because that paper approximates the derivatives numerically with a relative step length of 0.01. Our smaller relative step length of 10^{-6} gives more accurate estimates.

Par	(*i*) Estimate	(*i*) Std. error	(*ii*) Estimate	(*iii*) Estimate	(*iv*) Estimate	(*iv*) Std. error
Instantaneous expected inflation $\pi_t = \delta_{0\pi} + \delta'_{1\pi} X_t$						
$\delta_{0\pi}$	1.81%	(3.19%)	1.98%	2.00%	1.98%	(4.05%)
$\delta_{1\pi(1)}$	-0.63%	(0.12%)	-0.63%	-0.63%	-0.60%	(0.20%)
$\delta_{1\pi(2)}$	0.14%	(0.41%)	0.14%	0.14%	0.27%	(0.41%)
Instantaneous nominal interest rate $R_t(0) = \delta_{0R} + \delta'_{1R} X_t + \eta_t(0)$						
δ_{0R}	2.40%	(6.94%)	2.40%	2.40%	1.98%	(10.40%)
$\delta_{1R(1)}$	-1.48%	(0.35%)	-1.48%	-1.48%	-1.44%	(0.38%)
$\delta_{1R(2)}$	0.53%	(0.96%)	0.53%	0.53%	0.56%	(0.97%)
State variables nominal term structure $\mathrm{d}X_t = -KX_t\mathrm{d}t + \Sigma'_X \mathrm{d}Z_t$						
$K_{(1,1)}$	7.63%	(14.63%)	7.63%	7.63%	6.15%	(17.06%)
$K_{(2,1)}$	-19.00%	(20.71%)	-19.00%	-19.00%	-22.23%	(20.13%)
$K_{(2,2)}$	35.25%	(19.54%)	35.25%	35.25%	31.90%	(21.85%)
Realized inflation process $\frac{\mathrm{d}\Pi_t}{\Pi_t} = \pi_t \mathrm{d}t + \sigma'_\Pi \mathrm{d}Z_t$						
$\sigma_{\Pi(1)}$	0.02%	(0.07%)	0.02%	0.02%	0.02%	(0.08%)
$\sigma_{\Pi(2)}(10^{-4})$	-0.568	(6.30)	-0.568	-0.568	-1.93	(6.47)
$\sigma_{\Pi(3)}$	0.61%	(0.04%)	0.61%	0.61%	0.61%	(0.04%)
Stock return process $\frac{\mathrm{d}S_t}{S_t} = (R_t(0) + \eta_S)\,\mathrm{d}t + \sigma'_S \mathrm{d}Z_t$						
η_S	4.52%	(3.68%)	6.57%	4.52%	4.20%	(3.77%)
$\sigma_{S(1)}$	-0.53%	(1.44%)	-0.53%	-0.53%	-0.54%	(1.44%)
$\sigma_{S(2)}$	-0.76%	(1.53%)	-0.76%	-0.76%	-0.78%	(1.54%)
$\sigma_{S(3)}$	-2.11%	(1.51%)	-2.11%	-2.11%	-2.23%	(1.46%)
$\sigma_{S(4)}$	16.59%	(0.96%)	17.69%	16.59%	16.39%	(0.93%)
Prices of risk $\Lambda_t = \Lambda_0 + \Lambda_1 X_t$						
$\Lambda_{0(1)}$	0.403	(0.337)	0.242	0.280	0.187	(0.513)
$\Lambda_{0(2)}$	0.039	(0.294)	0.039	0.027	0.137	(0.624)
$\Lambda_{1(1,1)}$	0.149	(0.231)	0.149	0.149	0.142	(0.184)
$\Lambda_{1(1,2)}$	-0.381	(0.052)	-0.381	-0.381	-0.355	(0.037)
$\Lambda_{1(2,1)}$	0.089	(0.178)	0.089	0.089	0.144	(0.192)
$\Lambda_{1(2,2)}$	-0.083	(0.233)	-0.083	-0.083	-0.100	(0.211)
LL_{2013}	6525.6		6450.7	6471.0	6549.2	
LL_{2014}	6696.7		6619.6	6640.5	6720.4	

Table 2: Parameters and standard errors of (*i*) max. likelihood estimate 2013 in Draper (2014), (*ii*) calibrated estimate in Draper (2014), (*iii*) calibrated DNB parameter set, and (*iv*) max. likelihood estimate 2014 with the restrictions in Table 1. Standard errors are determined using the outer product gradient estimator of the likelihood function, which is only feasible at a max. likelihood estimate. The parameter symbols are identical to Koijen et al. (2010). LL_y refers to the loglikelihood with the sample ending at year y.

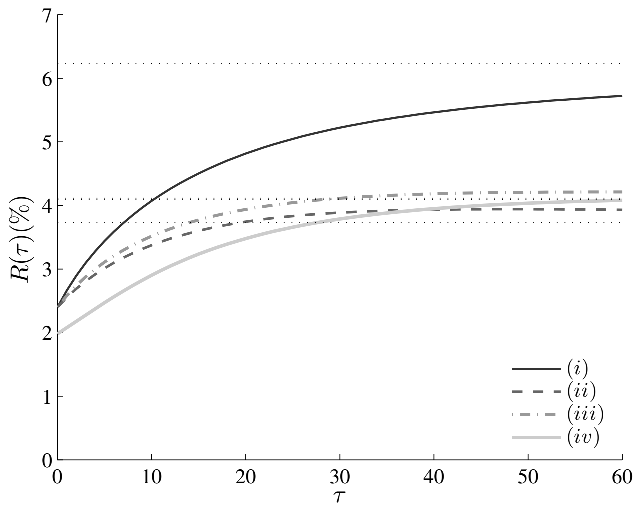

Figure 1: Nominal term structure of the parameter sets with the corresponding levels of $\mathrm{UFR}_{\log}$. (i) max. likelihood estimate 2013 in Draper (2014), (ii) calibrated estimate in Draper (2014), (iii) calibrated DNB parameter set, and (iv) max. likelihood estimate 2014.

for large τ. The long-run standard deviations (volatilities) of the different parameter estimates are remarkably similar.

Figure 1 shows the four term structures together with the corresponding asymptotic level, $\mathrm{UFR}_{\log}$. The latter is appropriate here since $R(\tau)$ is also a continuously compounded yield (see (12) and (19)). The term structure of the parameter sets (ii) and (iii) starts to exceed the corresponding UFR around $\tau = 20$ and $\tau = 30$, respectively. This indicates a decreasing term structure at some (large) maturity τ, which is at odds with the empirical evidence of a higher risk premium at longer horizons.[14] Nonetheless, the overshooting remains small in both cases.

6 Conclusions

This paper provides closed-form expressions for certain characteristics of the VAR(1)-representation of a continuous-time affine term structure model such as the models in Dai and Singleton (2002) and Koijen et al. (2010). We provide analytical expressions for the long-run, i.e., unconditional, expectations and variances of inflation, stock returns, bond portfolio returns, and nominal yields. In addition, we derived closed-form expressions for some other characteristics of the term structure.

The analytical expressions are useful in the following ways. First, the expressions are essential when maximizing a likelihood function with constraints on long-run expectations or covariances. The

[14] This is not to say that in practice, the forward rate may decrease for large maturities.

					(i)	(ii)	(iii)	(iv)
annually compounded returns			(18)	UFR	6.43%	3.80%	4.18%	4.20%
	arithmetic	mean	(25)	π	1.84%	2.01%	2.03%	2.01%
				R^S	7.22%	9.44%	7.22%	6.44%
				$R(0)$	2.48%	2.48%	2.48%	2.05%
				$R^B(5)$	4.48%	3.70%	3.86%	3.17%
	arithmetic	st.dev	(26)	π	1.59%	1.59%	1.59%	1.45%
				R^S	18.43%	20.01%	18.43%	18.10%
				$R(0)$	3.29%	3.29%	3.29%	3.29%
				$R^B(5)$	5.96%	5.91%	5.92%	6.16%
	geometric	mean	(27)	π	1.83%	2.00%	2.02%	2.00%
				R^S	5.67%	7.65%	5.67%	4.93%
				$R(0)$	2.43%	2.43%	2.43%	2.00%
				$R(5)$	3.50%	3.06%	3.16%	2.50%
				$R(30)$	5.36%	3.96%	4.20%	3.86%
				$R^B(5)$	4.31%	3.53%	3.69%	2.99%
				$R^B(30)$	6.33%	4.22%	4.54%	4.55%
continuously compounded returns			(17)	$\mathrm{UFR}_{\log}$	6.23%	3.73%	4.09%	4.11%
	mean		(22)	π	1.81%	1.98%	2.00%	1.98%
				R^S	5.51%	7.37%	5.51%	4.81%
				$R(0)$	2.40%	2.40%	2.40%	1.98%
				$R^B(5)$	4.22%	3.47%	3.63%	2.94%
	st.dev	(volatility)	(24)	π	1.56%	1.56%	1.56%	1.42%
				R^S	17.06%	18.14%	17.06%	16.89%
				$R(0)$	3.21%	3.21%	3.21%	3.22%
				$R^B(5)$	5.70%	5.70%	5.70%	5.97%

Table 3: Long-run statistics of inflation π, stock return R^S, nominal interest rate $R(\tau)$, and bond portfolio return $R^B(\tau)$. The bond portfolios have a constant maturity. The parameter sets are as follows: (i) max. likelihood estimate 2013 in Draper (2014), (ii) calibrated estimate in Draper (2014), (iii) calibrated DNB parameter set, and (iv) max. likelihood estimate 2014 with the restrictions in Table 1.

analytical expressions obviates calibrations which necessarily involve some arbitrariness. Second, the analytical expressions enable an ex-post analysis of different parameter sets in terms of long-run values for expectations, covariances, and the term structure, without any need for simulation.

To illustrate our results, we estimated a parameter set with an additional year of data, and improved the optimization method. While our optimal parameter set is very similar to previous sets in terms of long-run standard deviations, it differs substantially in the long-run expectation and the term structure. This indicates that a thorough long-run analysis is crucial when evaluating a parameter set for a long-run scenario analysis.

Appendix

A Discretization

We discretize the multivariate Ornstein-Uhlenbeck process in (9). The expressions in (11) refer to the case with diagonalizable Θ_1. Here, we find expressions that hold for any Θ_1, including non-diagonalizable (defective) Θ_1. This may happen when $K_{(1,1)} = K_{(2,2)}$ in K. To the best of our knowledge, this extension is unknown in the literature.

By applying Ito's lemma to the process $f(t, \Upsilon_t) = e^{-\Theta_1 t}\Upsilon_t$ in (9), one may find

$$\mathrm{d}\left(\exp\left(-\Theta_1 t\right)\Upsilon_t\right) = \exp\left(-\Theta_1 t\right)\left(\Theta_0 \mathrm{d}t + \sigma_\Upsilon \mathrm{d}Z_t\right).$$

Therefore,

$$\begin{aligned}
&\exp\left(-\Theta_1 t\right)\Upsilon_t \\
&= \exp\left(-\Theta_1\left(t-h\right)\right)\Upsilon_{t-h} + \int_{u=t-h}^{t} \mathrm{d}\left(\exp\left(-\Theta_1 u\right)\Upsilon_u\right) \\
&= \int_{v=-t}^{h-t} \exp\left(\Theta_1 v\right)\Theta_0 \mathrm{d}v + \exp\left(\Theta_1\left(h-t\right)\right)\Upsilon_{t-h} + \int_{v=-t}^{h-t} \exp\left(\Theta_1 v\right)\sigma_\Upsilon \mathrm{d}Z_u
\end{aligned}$$

Some rewriting gives

$$\Upsilon_t = \left[\exp\left(\Theta_1 h\right) - I\right]\Theta_1^{-1}\Theta_0 + \exp\left(\Theta_1 h\right)\Upsilon_{t-h} + \int_{v=0}^{h} \exp\left(\Theta_1 v\right)\sigma_\Upsilon \mathrm{d}Z_v \tag{43}$$

It follows from the Jordan matrix decomposition $\Theta_1 = UJU^{-1}$ that for (10):

$$\gamma = U\left[\exp\left(Jh\right) - I\right]J^{-1}U^{-1}\Theta_0 \tag{44}$$

$$\Gamma = U\exp\left(Jh\right)U^{-1} \tag{45}$$

Next, we derive the disturbance covariance V from the integral in (43). We have for an $n_b \times n_b$ Jordan block J_b of an $n \times n$ Jordan matrix J

$$\exp\left(J_b\right) = e^{\lambda_b} \sum_{d=0}^{n_b-1} \tilde{M}^{(d)} \qquad J_b = \begin{bmatrix} \lambda_b & 1 & & \\ & \lambda_b & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_b \end{bmatrix}$$

where each $\tilde{M}^{(d)}$ is an $n_b \times n_b$ matrix with nonzero entries on the dth superdiagonal:

$$\tilde{M}_{ij}^{(d)} = \begin{cases} \frac{1}{(j-i)!} & \text{if } j = i + d \\ 0 & \text{else} \end{cases}$$

This extends to the Jordan matrix J as

$$\exp\left(J\right) = \exp\left(D\right) \sum_{d=0}^{n_B} M^{(d)} \tag{46}$$

where D is a diagonal matrix with the same diagonal as J, $n_B = \max_b\left(n_b\right) - 1$, and each $M^{(d)}$ is an $n \times n$ matrix with nonzero entries on the dth superdiagonal:

$$M_{ij}^{(d)} = \begin{cases} \frac{1}{k!} & \text{if } j = i + d \, \text{and} \, (i, j) \, \text{is in a Jordan block of} \, J \\ 0 & \text{else} \end{cases}$$

Applying (46) to the Jordan matrix decomposition $\Theta_1 = UJU^{-1}$ gives for the disturbance covariance in (43):

$$\begin{aligned} V &= \int_{v=0}^{h} \exp\left(\Theta_1 v\right) \sigma_\Upsilon \sigma_\Upsilon' \exp\left(\Theta_1' v\right) \mathrm{d}v \\ &= \int_{v=0}^{h} U \exp\left(Jv\right) U^{-1} \sigma_\Upsilon \left(U^{-1}\sigma_\Upsilon\right)' \exp\left(J'v\right) \mathrm{d}v U' \\ &= U \int_{v=0}^{h} \exp\left(Dv\right) \left(\sum_{a=0}^{n_B} M^{(a)}\right) U^{-1}\sigma_\Upsilon \left(U^{-1}\sigma_\Upsilon\right)' \left(\sum_{b=0}^{n_B} M^{(b)}\right)' \exp\left(Dv\right) \mathrm{d}v U' \\ &= UWU' \end{aligned} \tag{47}$$

where

$$\begin{aligned} W &= \sum_{a=0}^{n_B} \sum_{b=0}^{n_B} W^{(a,b)} \\ W^{(a,b)} &= \int_{v=0}^{h} \exp\left(Dv\right) M^{(a,b)} \exp\left(Dv\right) \mathrm{d}v \\ M^{(a,b)} &= M^{(a)} U^{-1} \sigma_\Upsilon \left(M^{(b)} U^{-1} \sigma_\Upsilon\right)' \end{aligned}$$

The entries of $W^{(a,b)}$ are

$$W_{ij}^{(a,b)} = \begin{cases} M_{ij}^{(a,b)} h & \text{if } D_{ii} + D_{jj} = 0 \\ M_{ij}^{(a,b)} \left(D_{ii} + D_{jj}\right)^{-1} \left[\exp\left(\left[D_{ii} + D_{jj}\right] h\right) - I_{n\times n}\right] & \text{else} \end{cases}$$

A diagonalizable matrix corresponds to $J = D$ and $n_B = 0$. Then, the expressions for γ, Γ and V in (11) coincide with (44), (45) and (47), respectively.

B Complex eigenvalues

Things are slightly different if $M \in \mathbb{R}^{k\times k}$ has complex eigenvalues. We consider $k = 2$ and M with eigenvalues $\lambda_1 = a + bi$ and $\lambda_2 = a - bi$ ($b \neq 0$ and $a, b \in \mathbb{R}$). We show that although the complex terms cancel out, the term structure $R(\tau)$ contains an oscillating component which is undesirable from an empirical perspective. It is easy to extend the analysis to $k > 2$.

To rule out an oscillating term structure, we restrict the estimation of M to real eigenvalues. This restriction is important as for instance in the two-dimensional case the complex eigenvalue set $\{(a + bi, a - bi) : a, b \in \mathbb{R}\}$ has the same measure as the real eigenvalue set $\{(a, b) : a, b \in \mathbb{R}\}$.

Letting $M\mathbf{v}_j = \lambda_j \mathbf{v}_j$ ($j = 1, 2$) gives $\mathrm{Re}(M\mathbf{v}_1) = \mathrm{Re}(M\mathbf{v}_2)$, $\mathrm{Im}(M\mathbf{v}_1) = -\mathrm{Im}(M\mathbf{v}_2)$ and $V^{-1}V = I_{2\times 2}$. Hence, we may write for $\mathbf{p}, \mathbf{q}, \mathbf{u}, \mathbf{v} \in \mathbb{R}^2$

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{u} + \mathbf{v}i & \qquad \mathbf{v}_1^{-1} &= \mathbf{p}' + \mathbf{q}'i \\ \mathbf{v}_2 &= \mathbf{u} - \mathbf{v}i & \qquad \mathbf{v}_2^{-1} &= \mathbf{p}' - \mathbf{q}'i. \end{aligned}$$

This implies for the vector $\mathbf{b}_1$ in (36)

$$\begin{aligned} \mathbf{b}_1 &= \frac{1}{\lambda_1}\left(\mathbf{v}_1\mathbf{v}_1^{-1}\right)\delta_{1R} \\ &= \frac{1}{a + bi}\left(\mathbf{u} + \mathbf{v}i\right)\left(\mathbf{p}' + \mathbf{q}'i\right)\delta_{1R} \\ &= \frac{a - bi}{a^2 + b^2}\left(\mathbf{u}\mathbf{p}' - \mathbf{v}\mathbf{q}' + \left(\mathbf{u}\mathbf{q}' + \mathbf{v}\mathbf{p}'\right)i\right)\delta_{1R} \\ &= \left[\left(\mathbf{u}\mathbf{p}' - \mathbf{v}\mathbf{q}'\right)a + \left(\mathbf{u}\mathbf{q}' + \mathbf{v}\mathbf{p}'\right)b + \left\{-\left(\mathbf{u}\mathbf{p}' - \mathbf{v}\mathbf{q}'\right)b + \left(\mathbf{u}\mathbf{q}' + \mathbf{v}\mathbf{p}'\right)a\right\}i\right]\frac{\delta_{1R}}{a^2 + b^2} \\ &= \mathbf{f} + \mathbf{g}i \end{aligned}$$

where

$$\begin{aligned} \mathbf{f} &= \left[\left(\mathbf{u}\mathbf{p}' - \mathbf{v}\mathbf{q}'\right)a + \left(\mathbf{u}\mathbf{q}' + \mathbf{v}\mathbf{p}'\right)b\right]\frac{\delta_{1R}}{a^2 + b^2} \\ \mathbf{g} &= \left[-\left(\mathbf{u}\mathbf{p}' - \mathbf{v}\mathbf{q}'\right)b + \left(\mathbf{u}\mathbf{q}' + \mathbf{v}\mathbf{p}'\right)a\right]\frac{\delta_{1R}}{a^2 + b^2} \end{aligned}$$

Considering $\mathbf{v}_j = \mathbf{u} \pm \mathbf{v}i$, $\mathbf{v}_j^{-1} = \mathbf{p}' \pm \mathbf{q}'i$, and $\lambda_j = a \pm bi$ for $j = 1, 2$, only b, $\mathbf{q}$, and $\mathbf{v}$ have different signs for $\mathbf{b}_1$ and $\mathbf{b}_2$. Therefore,

$$\begin{aligned}\mathbf{b}_2 &= \left[\left(\mathbf{u}\mathbf{p}' - \mathbf{v}\mathbf{q}'\right) a + \left(\mathbf{u}\mathbf{q}' + \mathbf{v}\mathbf{p}'\right) b + \left(\left(\mathbf{u}\mathbf{p}' - \mathbf{v}\mathbf{q}'\right) b - \left(\mathbf{u}\mathbf{q}' + \mathbf{v}\mathbf{p}'\right) a\right) i\right] \frac{\delta_{1R}}{a^2 + b^2} \\ &= \mathbf{f} - \mathbf{g}i\end{aligned}$$

This gives

$$\begin{aligned}\mathbf{b}_1 e^{-\lambda_1 \tau} + \mathbf{b}_2 e^{-\lambda_2 \tau} &= e^{-a\tau}\left[(\mathbf{f} + \mathbf{g}i)\, e^{-b\tau i} + (\mathbf{f} - \mathbf{g}i)\, e^{b\tau i}\right] \\ &= e^{-a\tau}\left[(\mathbf{f} + \mathbf{g}i)\left(\cos(b\tau) - i\sin(b\tau)\right) + (\mathbf{f} - \mathbf{g}i)\left(\cos(b\tau) + i\sin(b\tau)\right)\right] \\ &= 2e^{-a\tau}\left[\mathbf{f}\cos(b\tau) + \mathbf{g}\sin(b\tau)\right]\end{aligned}$$

Equation (37) becomes

$$B(\tau) = \mathbf{b}_0 + 2e^{-a\tau}\left[\mathbf{f}\cos(b\tau) + \mathbf{g}\sin(b\tau)\right] \tag{48}$$

where $\mathbf{b}_0 = -M^{-1}\delta_{1R}$. That is, $B(\tau)$ is a real-valued oscillating function around $\exp(-a\tau)$.

Substituting (48) into (15) gives after some algebra,

$$\begin{aligned}\dot{A}(\tau) &= -\left(\mathbf{b}_0 + 2e^{-a\tau}\left(\mathbf{f}\cos(b\tau) + \mathbf{g}\sin(b\tau)\right)\right)' \Lambda_0 \\ &\quad + \frac{1}{2}\left(\mathbf{b}_0 + 2e^{-a\tau}\left(\mathbf{f}\cos(b\tau) + \mathbf{g}\sin(b\tau)\right)\right)'\left(\mathbf{b}_0 + 2e^{-a\tau}\left(\mathbf{f}\cos(b\tau) + \mathbf{g}\sin(b\tau)\right)\right) - \delta_{0R} \\ &= \left(\frac{1}{2}\mathbf{b}_0 - \tilde{\Lambda}_0\right)' \mathbf{b}_0 - \delta_{0R} + 2\left(\mathbf{b}_0 - \tilde{\Lambda}_0\right)' e^{-a\tau}\left(\mathbf{f}\cos(b\tau) + \mathbf{g}\sin(b\tau)\right) \\ &\quad + 2e^{-2a\tau}\left(\mathbf{f}'\mathbf{f}\cos^2(b\tau) + 2\mathbf{f}'\mathbf{g}\cos(b\tau)\sin(b\tau) + \mathbf{g}'\mathbf{g}\sin^2(b\tau)\right) \\ &= \left(\frac{1}{2}\mathbf{b}_0 - \tilde{\Lambda}_0\right)' b_0 - \delta_{0R} + 2\left(\mathbf{b}_0 - \tilde{\Lambda}_0\right)' e^{-a\tau}\left(\mathbf{f}\cos(b\tau) + \mathbf{g}\sin(b\tau)\right) \\ &\quad + 2e^{-2a\tau}\left(\frac{\mathbf{f}'\mathbf{f}}{2}\left[1 + \cos(2b\tau)\right] + \mathbf{f}'\mathbf{g}\sin(2b\tau) + \frac{\mathbf{g}'\mathbf{g}}{2}\left[1 - \cos(2b\tau)\right]\right) \\ &= \left(\frac{1}{2}\mathbf{b}_0 - \tilde{\Lambda}_0\right)' \mathbf{b}_0 - \delta_{0R} + 2\left(\mathbf{b}_0 - \tilde{\Lambda}_0\right)' e^{-a\tau}\left[\mathbf{f}\cos(b\tau) + \mathbf{g}\sin(b\tau)\right] \\ &\quad + e^{-2a\tau}\left[\mathbf{f}'\mathbf{f} + \mathbf{g}'\mathbf{g} + \left(\mathbf{f}'\mathbf{f} - \mathbf{g}'\mathbf{g}\right)\cos(2b\tau) + 2\mathbf{f}'\mathbf{g}\sin(2b\tau)\right]\end{aligned}$$

Therefore,

$$\dot{A}(\tau) = a_0^{(1)} + \sum_{j=1}^{5} a_j^{(1c)} h_j(\tau) \tag{49}$$

where

$$a_0^{(1)} = \left(\frac{1}{2}\mathbf{b}_0 - \tilde{\Lambda}_0\right)' \mathbf{b}_0 - \delta_{0R}$$

$$
\begin{aligned}
a_1^{(1c)} &= 2\left(\mathbf{b}_0 - \tilde{\Lambda}_0\right)' \mathbf{f} & h_1(\tau) &= \exp(-at)\cos(b\tau) \\
a_2^{(1c)} &= 2\left(\mathbf{b}_0 - \tilde{\Lambda}_0\right)' \mathbf{g} & h_2(\tau) &= \exp(-at)\sin(b\tau) \\
a_3^{(1c)} &= \mathbf{f}'\mathbf{f} + \mathbf{g}'\mathbf{g} & h_3(\tau) &= \exp(-2at) \\
a_4^{(1c)} &= \mathbf{f}'\mathbf{f} - \mathbf{g}'\mathbf{g} & h_4(\tau) &= \exp(-2at)\cos(2b\tau) \\
a_5^{(1c)} &= 2\mathbf{f}'\mathbf{g} & h_5(\tau) &= \exp(-2at)\sin(2b\tau)
\end{aligned}
$$

Because M is positive definite, we have without loss of generality $a, b > 0$. Using the identities[15]

$$
\begin{aligned}
\int_{t=0}^{\tau} \exp(-at)\cos(bt)\,\mathrm{d}t &= \frac{a}{a^2+b^2} + \frac{\exp(-a\tau)}{a^2+b^2}\left[b\sin(b\tau) - a\cos(b\tau)\right] \\
\int_{t=0}^{\tau} \exp(-at)\sin(bt)\,\mathrm{d}t &= \frac{b}{a^2+b^2} - \frac{\exp(-a\tau)}{a^2+b^2}\left[a\sin(b\tau) + b\cos(b\tau)\right]
\end{aligned}
$$

and integrating (49) leads to another real-valued oscillating function:

$$
A(\tau) = a_0^{(1)}\tau + \sum_{j=1}^{5} a_j^{(1c)} k_j(\tau) \tag{50}
$$

where

$$
\begin{aligned}
k_1(\tau) &= \frac{1}{a^2+b^2}\left[a + \exp(-a\tau)\left(b\sin(b\tau) - a\cos(b\tau)\right)\right] \\
k_2(\tau) &= \frac{1}{a^2+b^2}\left[b - \exp(-a\tau)\left(a\sin(b\tau) + b\cos(b\tau)\right)\right] \\
k_3(\tau) &= \frac{1}{2a}\left(1 - \exp(-2a\tau)\right) \\
k_4(\tau) &= \frac{1}{2\left(a^2+b^2\right)}\left[a + \exp(-2a\tau)\left(b\sin(2b\tau) - a\cos(2b\tau)\right)\right] \\
k_5(\tau) &= \frac{1}{2\left(a^2+b^2\right)}\left[b - \exp(-2a\tau)\left(a\sin(2b\tau) + b\cos(2b\tau)\right)\right]
\end{aligned}
$$

Again $A(0) = 0$ as $k_j(0) = 0$ for $j \in \{1, \ldots, 5\}$. We obtain from (50) an explicit expression for the unconditional term structure $R(\tau) = -A(\tau)/\tau$:

$$
R(\tau) = -a_0^{(1)} - \frac{1}{\tau}\sum_{j=1}^{5} a_j^{(1c)} k_j(\tau) \tag{51}
$$

[15] Both equations are easily verified by differentiating both hand sides and requiring that the expressions are zero at $\tau = 0$.

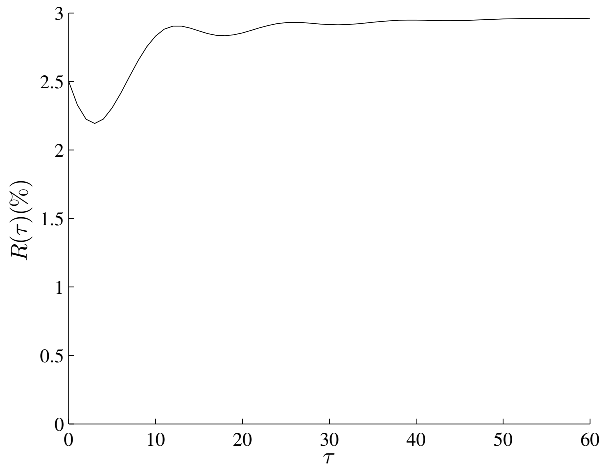

Figure 2: Oscillating term structure with parameters as in (53). The matrix M has complex eigenvalues.

with derivative

$$\dot{R}(\tau) = -\frac{1}{\tau}\left(R(\tau) + a_0^{(1)} + \sum_{j=1}^{5} a_j^{(1c)}\dot{k}_j(\tau)\right) \tag{52}$$

The expressions in (51) and (52) are more complicated than (40) and (41) which represent the case of two (possibly distinct) real eigenvalues. More important, the term structure $R(\tau)$ in (51) exhibits oscillations in τ if $b \neq 0$. This motivates us to restrict the estimation of M to matrices with real positive eigenvalues.

As as example, Figure 2 shows the implied term structure for

$$M = \begin{bmatrix} -\frac{1}{10} & 1 \\ -\frac{1}{4} & \frac{1}{5} \end{bmatrix} \quad \delta_{0R} = 0.025 \quad \delta_{1R} = \tfrac{1}{100}\begin{bmatrix}1\\1\end{bmatrix} \quad \tilde{\Lambda}_0 = \tfrac{1}{5}\begin{bmatrix}1\\1\end{bmatrix} \tag{53}$$

The eigenvalues of M are $0.05 \pm 0.477i$ and the term structure does indeed have an oscillating pattern.

C Data

Following Draper (2014) we use the following time series from Van den Goorbergh et al. (2011):

- Inflation: From 1999 on, the Harmonized Index of Consumer Prices for the euro area from the European Central Bank data website[16] is used. Before then, German (Western German

[16]http://sdw.ecb.europa.eu

until 1990) consumer price index figures published by the International Financial Statistics of the International Monetary Fund are included.

- Yields: Six yields are used in estimation: yields with a three-month, one-year, two-year, three-year, five-year, and ten-year maturity.
  - Short nominal yields: three-month money market rates are taken from the Bundesbank.[17] For the period 1973:I to 1990:II, end-of-quarter money market rates reported by Frankfurt banks are taken, whereas thereafter three-month Frankfurt Interbank Offered Rates are included.
  - Long nominal yields: From 1987:IV on, zero-coupon rates are constructed from swap rates published by De Nederlandsche Bank.[18] For the period 1973:I to 1987:III, zero coupon yields with maturities of one to 15 years (from the Bundesbank website) based on government bonds were used as well (15-year rates start in June 1986). No adjustments were made to correct for possible differences in the credit risk of swaps, on the one hand, and German bonds, on the other. The biggest difference in yield between the two term structures (for the two-year yield) in 1987:IV was only 12 basis points.
- Stock returns: MSCI index from FactSet. Returns are in euros (Deutschmark before 1999) and hedged for US dollar exposure.

References

Banerjee, A., Dolado, J. J., Galbraith, J. W., and Hendry, D. (1993). Co-integration, error correction, and the econometric analysis of non-stationary data. *OUP Catalogue*.

Cochrane, J. H. (2005). Time series for macroeconomics and finance.

Dai, Q. and Singleton, K. J. (2002). Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics*, 63(3):415–441.

Draper, N. (2014). A financial market model for the netherlands. *CPB Background Document*.

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. *The Journal of Finance*, 57(1):405–443.

Duffee, G. R. and Stanton, R. H. (2012). Estimation of dynamic term structure models. *The Quarterly Journal of Finance*, 2(02).

Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates. *Mathematical finance*, 6(4):379–406.

[17] www.bundesbank.de
[18] www.dnb.nl

Goffe, W. L., Ferrier, G. D., and Rogers, J. (1994). Global optimization of statistical functions with simulated annealing. *Journal of Econometrics*, 60(1-2):65–99.

Gürkaynak, R. S. and Wright, J. H. (2012). Macroeconomics and the term structure. *Journal of Economic Literature*, 50(2):331–367.

Koijen, R. S., Nijman, T. E., and Werker, B. J. (2010). When can life cycle investors benefit from time-varying bond risk premia? *Review of Financial Studies*, 23(2):741–780.

Langejan, T., Gelauff, G., Nijman, T., Sleijpen, O., and Steenbeek, O. (2014). Advies commissie parameters. Technical report, SZW.

Sangvinatsos, A. and Wachter, J. A. (2005). Does the failure of the expectations hypothesis matter for long-term investors? *The Journal of Finance*, 60(1):179–230.

Van den Goorbergh, R., Molenaar, R., Steenbeek, O. W., and Vlaar, P. (2011). Risk models with jumps and time-varying second moments.