Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Kalman Filter Estimation of the KNWModel

Technical note

Antoon Pelsser

DP 05/2019-047

Technical note:
Kalman Filter Estimation of the KNW Model

Antoon Pelsser*

Maastricht University & Netspar
Email: a.pelsser@maastrichtuniversity.nl

This version: May 28, 2019

Abstract

This technical note gives implementation notes for estimating the Koijen-Nijman-Werker model from historical data based on a Kalman filter. We provide an independent derivation of the KNW model. We propose a different implementation of the state-space formulation of the KNW model and we test the impact of two different specifications for the initialisation of the Kalman filter maximum-likelihood estimation. By doing so, we provide an independent verification of the parameter estimations provided by DNB for the Committee Parameters. We find that the parameter estimates reported by DNB and our own parameter estimates are very similar.

*The author would like to thank Bart Diris, Sander Muns and Bas Werker for comments and useful suggestions.

1 Introduction

This paper presents implementation notes for estimating the parameters in the Koijen et al. (2010) model (KNW) from historical data based on a Kalman filter. The implementation of the KNW model based on Draper (2014) and Muns (2015). For a comprehensive introduction to Kalman filter estimation of time series model with Kalman filters, we refer to Harvey (1991) and the papers by Babbs and Nowman (1999) and De Jong (2000).

2 KNW Model

Gaussian affine models for the economy are very common in the literature (see e.g., Dai and Singleton (2000), Koijen et al. (2010)). We present results on the estimation of a VAR(1)-representation of a continuous-time model. We apply these results to the estimation of the parameters in the KNW model.

As key determinants of pension risk, we consider the consumer price index Π_t, the stock-market index S_t, and (nominal, continuously compounded) interest rates $y_t(\tau)$ with different maturities τ.

The core of the KNW model consists of the specification of the instantaneous nominal interest rate r_t and the instantaneous expected inflation π_t. Both these processes are assumed to be driven by an (unobserved) vector-valued factor-process X_t, which contains k factors:

$$dX_t = -KX_t\,dt + d\tilde{W}_t^{\mathbb{P}} \tag{2.1}$$

$$\pi_t = \delta_{0\pi} + \delta_{1\pi}' X_t \tag{2.2}$$

$$r_t = \delta_{0r} + \delta_{1r}' X_t \tag{2.3}$$

where $\tilde{W}_t^{\mathbb{P}}$ denotes a k-dimensional Brownian Motion with respect to the “real world” probability measure $\mathbb{P}$. K is a $k \times k$ matrix of auto-regressive coefficients that controls the dynamics of X_t. Both the price-inflation π_t and the instantaneous interest rate r_t are assumed to be an affine transformation of X_t. Please note that π_t and r_t cannot be observed directly in the economy. We also note that the factor-process X_t has a multivariate Gaussian distribution at each point in time, is Markovian and has Gaussian transition densities. Hence, the processes π_t and r_t are also Gaussian. This implies that π_t and r_t can have negative values with a positive probability.

The observable processes in the economy are the consumer price index Π_t, the stock-market index S_t and the prices of (nominal) zero-coupon rates $y_t(\tau)$ for various maturities τ.[1] The observed processes Π_t and S_t follow the dynamics

$$d\Pi_t = \pi_t \Pi_t\,dt + \Pi_t \sigma_\Pi' dW_t^{\mathbb{P}} \tag{2.4}$$

$$dS_t = (r_t + \eta_S) S_t\,dt + S_t \sigma_S' dW_t^{\mathbb{P}} \tag{2.5}$$

[1] The Dutch government does not issue index-linked government bonds. However, other countries in the euro-area (in particular France) do issue government bonds linked to the euro-wide inflation index HICP.

where $W_t^{\mathbb{P}}$ denotes a $(k+2)$-dimensional $\mathbb{P}$-Brownian Motion, that extends $\tilde{W}_t$ by two extra Brownian components: $W_t' = \left(\tilde{W}_t', W_{t,(k+1)}, W_{t,(k+2)}\right)$.

2.1 Nominal Interest Rates

The nominal zero-coupon rates $y_t(\tau)$ are determined from the market prices of discount bonds. To determine the prices of discount bonds of different maturities τ we use a model that is free of arbitrage opportunities. We can mathematically prove that a model is arbitrage-free (and complete) by proving that there exists a (unique) probability measure $\mathbb{Q}$ for which the price processes of all traded assets divided by a given numéraire are martingales.[2]

A convenient choice for the numéraire is to choose the (nominal) money-market account M_t, which is the value of €1 invested in a (risk-free) money-market account earning the risk-free interest rate r_t at each moment in time:

$$M_0 = 1, \quad dM_t = r_t M_t \, dt \qquad \Longleftrightarrow \qquad M_t = e^{\int_0^t r_s \, ds}. \tag{2.6}$$

The money-market account is a traded asset in the economy with a strictly positive price, and is therefore a valid choice of numéraire.

We can specify the change of probability measure from $\mathbb{P}$ to $\mathbb{Q}$ via a strictly positive (one-dimensional) $\mathbb{P}$-martingale R_t, which is called the Radon-Nikodym derivative:

$$dR_t = -R_t \lambda_t' \, dW_t^{\mathbb{P}} \tag{2.7}$$

where λ_t is a $(k+2)$-dimensional (stochastic) process.[3] When we define the measure $\mathbb{Q}$ via $d\mathbb{Q}_t = R_t \, d\mathbb{P}_t$, then Girsanov's Theorem states that the process $dW_t^{\mathbb{P}} + \lambda_t \, dt$ becomes a $\mathbb{Q}$-Brownian motion $dW_t^{\mathbb{Q}}$ under the measure $\mathbb{Q}$. In other words, the change of measure from $\mathbb{P}$ to $\mathbb{Q}$ "removes" a drift-term $\lambda_t \, dt$ from the Brownian motion W_t.

We can remain within the affine class of models under $\mathbb{Q}$ by specifying the process λ_t as

$$\lambda_t = \lambda_0 + \Lambda_1 X_t \tag{2.8}$$

where λ_0 is a constant vector of length $(k+2)$ and Λ_1 is a constant matrix of size $(k+2) \times k$. These define λ_t as an affine transformation of the factor-process X_t.[4] For future reference, we split the matrix Λ_1 in two parts: a $k \times k$ matrix $\tilde{\Lambda}_1$ and a $2 \times k$ block of zeros.

Under the measure $\mathbb{Q}$, the prices of all assets divided by M_t become martingales. The relative price of a discount bond with maturity τ is therefore a martingale:

$$\frac{e^{-\tau y_t(\tau)}}{M_t} = \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{M_{t+\tau}} \,\middle|\, \mathscr{F}_t \right] \quad \Longrightarrow \quad e^{-\tau y_t(\tau)} = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{t+\tau} r_s \, ds} \,\middle|\, \mathscr{F}_t \right] \tag{2.9}$$

[2]This is the Fundamental Theorem of Asset Pricing, see e.g. Delbaen and Schachermayer (1994).

[3]The process λ_t has to satisfy a condition to ensure that the process R_t is a true martingale for $0 \le t \le T$. A well-known sufficient condition is Novikov's condition: $\mathbb{E}^{\mathbb{P}}\left[e^{\frac{1}{2}\int_0^T |\lambda_t|^2 \, dt} \right] < \infty$.

[4]To ensure that Novikov's condition is satisfied, the elements in the matrix Λ_1 cannot be too large relative to the covariance matrix Σ_{X_t} of the process X_t for all t. Inspecting the quadratic forms in X_t in the expectation $\mathbb{E}[e^{\frac{1}{2}|\lambda_t|^2}] \propto \int e^{\frac{1}{2}(\lambda_0 + \Lambda_1 X_t)'(\lambda_0 + \Lambda_1 X_t)} e^{-\frac{1}{2}(X_t - \mu_{X_t})\Sigma_{X_t}^{-1}(X_t - \mu_{X_t})} \, dX_t$ shows that this expectation is finite only when $\Sigma_{X_t}^{-1} - \Lambda_1' \Lambda_1$ is a positive definite matrix.

where we use the fact that at the maturity date $t+\tau$ the price of the discount bond is always equal to 1. To evaluate the expectation, we consider the $(k+1)$-dimensional process (X_t, i_t) with $i_t := \int_0^t r_s\,ds$. The process i_t can be represented in differential form as $di_t = r_t\,dt = (\delta_{0r} + \delta_{1r}' X_t)\,dt$. Hence, the dynamics of the process (X_t, i_t) under the measure $\mathbb{Q}$ are given by

$$d\begin{pmatrix} X_t \\ i_t \end{pmatrix} = \left[\begin{pmatrix} -\lambda_0 \\ \delta_{0r} \end{pmatrix} + \begin{pmatrix} -(K+\tilde{\Lambda}_1) & 0 \\ \delta_{1r}' & 0 \end{pmatrix}\begin{pmatrix} X_t \\ i_t \end{pmatrix}\right] dt + \begin{pmatrix} I_k \\ 0_{1\times k} \end{pmatrix} d\tilde{W}_t^{\mathbb{Q}} \tag{2.10}$$

which is a vector-OU process of the form $dY = (a + AY)\,dt + C\,dW$. The multivariate Gaussian transition density of a vector-OU process is derived in Appendix A. The price of a discount bond $D_t(\tau)$ with maturity τ can be expressed as[5]

$$D_t(\tau) = \mathbb{E}_t^{\mathbb{Q}}\left[e^{-(i_{t+\tau}-i_t)}\right] = e^{-A(\tau)-B(\tau)'X_t} \tag{2.11}$$

$$B(\tau) = (K+\tilde{\Lambda}_1)'^{-1}\left(I_k - e^{-(K+\tilde{\Lambda}_1)'\tau}\right)\delta_{1r} \tag{2.12}$$

$$A(\tau) = \int_0^\tau \delta_{0r} - \lambda_0' B(s) - \tfrac{1}{2}B(s)'B(s)\,ds \tag{2.13}$$

Due to the time-homogeneity of the KNW model, the functions $A(\tau)$ and $B(\tau)$ are deterministic functions of the maturity τ only. The time-dependence on t only enters through the factors X_t.

The (continuously compounded) zero-coupon rate $y_t(\tau)$ for maturity τ is defined as

$$y_t(\tau) := \frac{\ln D_t(\tau)}{-\tau} = \frac{A(\tau)}{\tau} + \frac{B(\tau)'}{\tau}X_t. \tag{2.14}$$

The Ultimate Forward Rate (UFR) is defined as the limit for $\tau \to \infty$ of the zero-coupon rates. The (continuously compounded) UFR in the KNW model is given by

$$\text{UFR} = \lim_{\tau\to\infty}\frac{A(\tau)}{\tau} = \delta_{0r} - \lambda_0' B_\infty - \tfrac{1}{2}B_\infty' B_\infty \quad \text{with} \quad B_\infty = (K+\tilde{\Lambda}_1)'^{-1}\delta_{1r}. \tag{2.15}$$

The expression for the UFR can be decomposed into three terms: the $\mathbb{P}$-expectation of r_t given by δ_{0r}; a market-price of (interest) risk correction term $-\lambda_0' B_\infty$ and a convexity correction term $-\frac{1}{2}B_\infty' B_\infty$. The Dutch central bank (DNB) quotes the UFR in annual compounded terms:

$$\text{UFR}_{\text{ann}} = e^{\text{UFR}} - 1. \tag{2.16}$$

Hence, an annually compounded UFR_{ann} of 2.1% corresponds to a continuously compounded UFR of $\ln(1.021) \approx 0.020783$.

[5] The formulas presented here are consistent with the formulas given in Muns (2015). We also note that the two matrices $(K+\tilde{\Lambda}_1)'$ and $I_k - e^{-(K+\tilde{\Lambda}_1)'\tau}$ commute, which leads to the alternative expression $B(\tau) = \left(I_k - e^{-(K+\tilde{\Lambda}_1)'\tau}\right)(K+\tilde{\Lambda}_1)'^{-1}\delta_{1r}$.

2.1.1 Explicit Formulas for k=2

In this subsection we derive the explicit formulas for the case $k = 2$, which is the standard assumption in Koijen et al. (2010). In this case, the matrix K is a lower-triangular 2×2 matrix and $\tilde{\Lambda}_1$ is a (general) 2×2 matrix. To lighten the notation, we introduce the matrix $M := K + \tilde{\Lambda}_1$.

We want to impose a constraint on the dynamics of X_t such that the process converges to a well-defined stationary distribution under measure $\mathbb{Q}$. Furthermore, to avoid oscillations in the term-structure of interest rates, we must avoid cyclical behaviour in X_t under measure $\mathbb{Q}$. Both these conditions are satisfied whenever the matrix M has two positive real eigenvalues. For a general 2×2 matrix $M = \left(\begin{smallmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{smallmatrix}\right)$ we can explicitly compute the eigenvalues as

$$\lambda_{1,2} = \frac{(m_{11} + m_{22}) \pm s_M}{2} \quad \text{with} \quad s_M := \sqrt{(m_{11} - m_{22})^2 + 4 m_{12} m_{21}}. \tag{2.17}$$

We have two real eigenvalues if s_M is strictly positive. In that case, the square root is a strictly positive real number. Hence, both eigenvalues can only be positive if $m_{11} + m_{22} > 0$ and if $m_{11} + m_{22}$ is larger than s_M. We obtain the following three conditions on the matrix M:[6]

$$(m_{11} - m_{22})^2 + 4 m_{12} m_{21} > 0 \tag{2.18}$$

$$m_{11} + m_{22} > 0 \tag{2.19}$$

$$m_{11} m_{22} - m_{12} m_{21} > 0 \tag{2.20}$$

The last equation follows from the fact that s_M can also be represented in the form $s_M^2 = (m_{11} + m_{22})^2 - 4(m_{11} m_{22} - m_{12} m_{21}) = \mathrm{tr}(M)^2 - 4|M|$.

For the numerical calculations, we do not impose any constraints on the matrix M (nor the matrix K). We check that the eigenvalues of M are both real and positive for the estimated model parameters. Furthermore, we will not use the expressions (2.13) and (2.12). We will instead extract $A(\tau)$ and $B(\tau)$ from the mean and variance of the process (2.10) as $A(\tau) = m_0(\tau)_{[3]} - \frac{1}{2} V(\tau)_{[3,3]}$ and $B(\tau)' = M_1(\tau)_{[3,1:2]}$ where the mean $m_0(\tau), M_1(\tau)$ and variance $V(\tau)$ are defined in equations (A.10, A.11, A.12). This approach has the advantage that we avoid numerical instabilities in (2.12) and (2.13) when the matrix $M = K + \tilde{\Lambda}_1$ is close to singular.

2.2 Stock Prices

The (total return) stock-market index S_t is also a traded asset in the KNW economy. Therefore, the relative price process S_t / M_t must be a martingale under measure $\mathbb{Q}$. Application of Itô's Lemma on S_t / M_t and the change of measure to $\mathbb{Q}$ leads to the dynamics

$$d\frac{S_t}{M_t} = \left(\eta_S - \sigma_S'(\lambda_0 + \Lambda_1 X_t)\right) \frac{S_t}{M_t}\, dt + \frac{S_t}{M_t} \sigma_S'\, dW_t^{\mathbb{Q}}. \tag{2.21}$$

[6] The first and third conditions are the same conditions as in Muns (2015). However, Muns imposes $m_{11} > 0$ as the second condition, which is not sufficient to guarantee positive eigenvalues. A counter-example is given by the matrix $\left(\begin{smallmatrix} 1.0 & 2.1 \\ -1.0 & -2.0 \end{smallmatrix}\right)$ which has two negative eigenvalues -0.11 and -0.89 but satisfies Muns's constraints.

This process can only be a martingale if the drift-term is identically equal to zero. Hence, we must impose the conditions

$$\sigma_S'\lambda_0 = \eta_S \tag{2.22}$$

$$\sigma_S'\Lambda_1 = 0 \tag{2.23}$$

on λ_0 and Λ_1. This can be used to identify the "remaining" parameters in the extended market-price of risk vectors λ_0 and Λ_1. For the estimation of the model parameters in the KNW model, we only use the $\mathbb{P}$-dynamics of S_t and Π_t and we do not use the equations (2.22) and (2.23) during the estimation calculations.

2.3 Inflation-Linked Bonds

Even though the Dutch government does not issue any inflation-linked bonds, we offer the following observation on the pricing of inflation-linked bonds in the KNW model. The payoff of an inflation-linked discount bond at maturity is equal to €1 times the accumulated price-inflation over the life of the bond. We can express the price $P_t(\tau)$ of an inflation-linked bond under measure $\mathbb{Q}$ as

$$\frac{P_t(\tau)}{M_t} = \mathbb{E}^{\mathbb{Q}}\left[\frac{\Pi_{t+\tau}/\Pi_t}{M_{t+\tau}}\,\middle|\,\mathscr{F}_t\right] \implies P_t(\tau) = \mathbb{E}^{\mathbb{Q}}\left[e^{(\ln\Pi_{t+\tau}-i_{t+\tau})-(\ln\Pi_t-i_t)}\,\middle|\,\mathscr{F}_t\right]. \tag{2.24}$$

When we define the auxiliary process $p_t := \ln\Pi_t - i_t$, then the dynamics of p_t can be expressed (under $\mathbb{P}$) as

$$\begin{aligned} dp_t &= \left(\pi_t - r_t - \tfrac{1}{2}\sigma_\Pi'\sigma_\Pi\right)dt + \sigma_\pi' dW_t^{\mathbb{P}} \\ &= \left((\delta_{0\pi} - \delta_{0r} - \tfrac{1}{2}\sigma_\Pi'\sigma_\Pi) + (\delta_{1\pi} - \delta_{1r})'X_t\right)dt + \sigma_\Pi' dW_t^{\mathbb{P}}. \end{aligned} \tag{2.25}$$

Hence, the dynamics of the process (X_t, p_t) under the measure $\mathbb{Q}$ are given by

$$d\begin{pmatrix} X_t \\ p_t \end{pmatrix} = \left[\begin{pmatrix} -\lambda_0 \\ \delta_{0\pi} - \delta_{0r} - \frac{1}{2}\sigma_\Pi'\sigma_\Pi - \lambda_0'\sigma_\Pi \end{pmatrix} + \begin{pmatrix} -(K+\tilde{\Lambda}_1) & 0 \\ (\delta_{1\pi} - \delta_{1r} - \Lambda_1'\sigma_\Pi)' & 0 \end{pmatrix}\begin{pmatrix} X_t \\ p_t \end{pmatrix}\right] dt + \begin{pmatrix} [I_k \; 0_{k\times 2}] \\ \sigma_\Pi' \end{pmatrix} dW_t^{\mathbb{Q}}. \tag{2.26}$$

These dynamics imply that we can also obtain closed-form expressions for the prices of index-linked bonds of the form $P_t(\tau) = \exp\{-A_P(\tau) - B_P(\tau)'X_t\}$ in the KNW model.

3 State-Space Formulation and the Kalman Filter

In this section we represent the KNW model in state-space form. This allows us to impute the unobservable factor process X_t using a Kalman filter. The idea behind the Kalman filter is that we have observed data and we have unobserved state variables. Using the Kalman filter we can "filter" the value for X_t from the observed data. For a more elaborate derivation of the Kalman filter and its applications in time series estimation we refer to Harvey (1991).

To formulate the KNW model in state-space form, we will first augment[7] the state-vector $\tilde{X}_t = (X_t, \ln \Pi_t, \ln S_t)$. The dynamics of the process $\tilde{X}_t$ under measure $\mathbb{P}$ are given by

$$d\tilde{X}_t = \left[\begin{pmatrix} 0_{1\times k} \\ \delta_{0\pi} - \frac{1}{2}\sigma_\Pi'\sigma_\Pi \\ \delta_{0r} + \eta_S - \frac{1}{2}\sigma_S'\sigma_S \end{pmatrix} + \begin{pmatrix} -K & 0_{k\times 2} \\ \delta_{1\pi}' & 0_{1\times 2} \\ \delta_{1r}' & 0_{1\times 2} \end{pmatrix} \begin{pmatrix} X_t \\ \ln \Pi_t \\ \ln S_t \end{pmatrix} \right] dt + \begin{pmatrix} [I_k \; 0_{k\times 2}] \\ \sigma_\Pi' \\ \sigma_S' \end{pmatrix} dW_t^{\mathbb{P}}. \tag{3.1}$$

The dynamics of $\tilde{X}_t$ are of the form $d\tilde{X}_t = (a + A\tilde{X}_t)\,dt + C\,dW_t$. For a time-step Δt, we derive in Appendix A the expression for the (multivariate Gaussian) transition density

$$f\left(\tilde{X}_t \,|\, \tilde{X}_{t-\Delta t}\right) \sim N\left(e^{A\Delta t}\tilde{X}_{t-\Delta t} + \int_0^{\Delta t} e^{Au} a\,du\,;\, \int_0^{\Delta t} e^{Au}CC'e^{A'u}\,du\right). \tag{3.2}$$

From the transition density, we obtain the following vector-AR(1) specification for the state-vector:

$$\tilde{X}_t = \phi + \Phi\tilde{X}_{t-\Delta t} + \varepsilon_t \qquad \mathbb{V}\text{ar}[\varepsilon_t] = Q. \tag{3.3}$$

This *transition equation* describes how the state-vector $\tilde{X}_t$ evolves in time. The vector ϕ and the matrices Φ and Q in the transition equation are given by

$$\phi := \int_0^{\Delta t} e^{Au} a\,du, \qquad \Phi := e^{A\Delta t}, \qquad Q := \int_0^{\Delta t} e^{Au}CC'e^{A'u}\,du. \tag{3.4}$$

We can obtain information about the state-vector $\tilde{X}_t$ by observing zero-rates $y_t(\tau_i)$ for different maturities τ_i and the price index Π_t and the stock-index S_t. In the KNW model we have that the zero-rates y_t, $\ln \Pi_t$ and $\ln S_t$ are an affine function of $\tilde{X}_t$. We can represent this relation in the so-called *measurement equation*

$$\tilde{y}_t = \begin{pmatrix} y_t \\ \ln \Pi_t \\ \ln S_t \end{pmatrix} = a + B\tilde{X}_t + \eta_t \qquad \mathbb{V}\text{ar}[\eta_t] = H. \tag{3.5}$$

Here the vector $\tilde{y}_t$ has length $m+2$, consisting of m zero-rates with maturities $\tau_1, \ldots, \tau_m$ augmented with $\ln \Pi_t$ and $\ln S_t$. The $(m+2)$ vector a and the $(m+2)\times(k+2)$ matrix B are given by

$$a := \begin{pmatrix} A(\tau_1)/\tau_1 \\ \vdots \\ A(\tau_m)/\tau_m \\ 0 \\ 0 \end{pmatrix} \qquad B := \begin{pmatrix} B(\tau_1)'/\tau_1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ B(\tau_m)'/\tau_m & 0 & 0 \\ 0_{1\times k} & 1 & 0 \\ 0_{1\times k} & 0 & 1 \end{pmatrix}. \tag{3.6}$$

[7] Augmenting the state-vector in this way, allows us to estimate the KNW model with a standard Kalman filter, unlike Draper (2014) or Muns (2015) who use a non-standard version of the Kalman filter.

The vector η_t represents the measurement errors. The measurement errors η_t are assumed to follow an i.i.d. multivariate Gaussian distribution $N(0; H)$, where the $(m+2)\times(m+2)$ matrix H is partitioned in an $m \times m$ diagonal matrix and a 2×2 block as

$$H := \begin{pmatrix} \operatorname{diag}(h_m^2) & 0_{m\times 2} \\ 0_{2\times m} & 0_{2\times 2} \end{pmatrix}. \tag{3.7}$$

This structure for H implies that we observe $\ln \Pi_t$ and $\ln S_t$ without measurement error. Also, we assume that the measurement errors η_t are independent from the random vectors $\tilde{X}_t$ and ε_t.

If we combine equations (3.5) and (3.3) together with the independence assumption of the errors, we obtain that the joint distribution of $\tilde{y}_t$ and $\tilde{X}_t$ is a multivariate Gaussian distribution

$$f\left(\begin{pmatrix} \tilde{X}_t \\ \tilde{y}_t \end{pmatrix} \middle| \tilde{X}_{t-\Delta t}\right) \sim N\left(\begin{pmatrix} \phi + \Phi \tilde{X}_{t-\Delta t} \\ a + B\big(\phi + \Phi \tilde{X}_{t-\Delta t}\big) \end{pmatrix}; \begin{pmatrix} Q & QB' \\ BQ & BQB' + H \end{pmatrix}\right). \tag{3.8}$$

There is however a problem at this point: we do not (fully) know $\tilde{X}_{t-\Delta t}$. The best possible information we have about $\tilde{X}_{t-\Delta t}$ is the estimated state $\hat{X}_{t-\Delta t}$. Let $P_{t-\Delta t}$ denote the covariance matrix of the estimation error $(\hat{X}_{t-\Delta t} - \tilde{X}_{t-\Delta t})$. The conditional distribution $f\left(\tilde{X}_t \mid \hat{X}_{t-\Delta t}\right)$ is given by

$$f\left(\tilde{X}_t \mid \hat{X}_{t-\Delta t}\right) \sim N\left(\phi + \Phi \hat{X}_{t-\Delta t}\,; P_{t|t-\Delta t}\right) \qquad \text{with} \qquad P_{t|t-\Delta t} := \Phi P_{t-\Delta t} \Phi' + Q. \tag{3.9}$$

Hence, our knowledge at time t is summarised by the joint distribution

$$f\left(\begin{pmatrix} \tilde{X}_t \\ \tilde{y}_t \end{pmatrix} \middle| \hat{X}_{t-\Delta t}\right) \sim N\left(\begin{pmatrix} \phi + \Phi \hat{X}_{t-\Delta t} \\ a + B\big(\phi + \Phi \hat{X}_{t-\Delta t}\big) \end{pmatrix}; \begin{pmatrix} P_{t|t-\Delta t} & P_{t|t-\Delta t}B' \\ BP_{t|t-\Delta t} & V_t \end{pmatrix}\right)$$
$$\text{with } V_t := BP_{t|t-\Delta t}B' + H. \tag{3.10}$$

Given that we observe $\tilde{y}_t$ at time t, we can compute the conditional distribution of $\tilde{X}_t$ given $\tilde{y}_t$ and $\hat{X}_{t-\Delta t}$:

$$f\left(\tilde{X}_t \mid \tilde{y}_t, \hat{X}_{t-\Delta t}\right) \sim N\left(\phi + \Phi \hat{X}_{t-\Delta t} + K_t u_t\,; P_t\right) \tag{3.11}$$

where

$$u_t := \tilde{y}_t - \Big(a + B\big(\phi + \Phi \hat{X}_{t-\Delta t}\big)\Big), \tag{3.12}$$

$$K_t := P_{t|t-\Delta t} B' V_t^{-1}, \tag{3.13}$$

$$P_t := P_{t|t-\Delta t} - P_{t|t-\Delta t} B' V_t^{-1} B P_{t|t-\Delta t} = (I - K_t B) P_{t|t-\Delta t}. \tag{3.14}$$

The best estimate, in the least-squares sense, for $\tilde{X}_t$ is the conditional expectation

$$\hat{X}_t := \mathbb{E}\left[\tilde{X}_t \mid \tilde{y}_t, \hat{X}_{t-\Delta t}\right] = \phi + \Phi \hat{X}_{t-\Delta t} + K_t u_t. \tag{3.15}$$

With this last step we have completed the specification of the Kalman filter and we can proceed to the next time-step $t + \Delta t$.

4 Maximum Likelihood Estimation

In this section we describe how to estimate the parameters in the KNW model using a maximum likelihood technique. The derivation of the Kalman filter in the previous section makes the implicit assumption that the system matrices ϕ, Φ, Q, a, B, H are known. In reality, we want to estimate the parameters of the KNW model from the observed data by maximising the likelihood of the observed data. The likelihood can be formulated as follows. From (3.10) we can deduce immediately that the distribution for $\left(\tilde{y}_t \mid \hat{X}_{t-\Delta t}\right)$ is given by

$$f\left(\tilde{y}_t \mid \hat{X}_{t-\Delta t}\right) \sim N\left(a + B\left(\phi + \Phi \hat{X}_{t-\Delta t}\right);\, V_t\right). \tag{4.1}$$

Therefore, the likelihood L_t of $\left(\tilde{y}_t \mid \hat{X}_{t-\Delta t}\right)$ is a multivariate Gaussian density function and the log-likelihood can be expressed as

$$\ln L_t = C - \tfrac{1}{2} \ln |V_t| - \tfrac{1}{2} u_t' V_t^{-1} u_t, \tag{4.2}$$

where C is a constant independent of the model parameters and u_t is given in (3.12). This representation of the log-likelihood is known as the *prediction error decomposition*. In the tables below, we will report log-likelihoods where we omit the constant C from the likelihood calculation.

One final point is the initialisation of the Kalman filter at time 0. The standard approach is to use the unconditional mean and covariance matrix of the $\tilde{X}$-process for $\hat{X}_0$ and P_0. However, our extended state-vector $\tilde{X}$ has two non-stationary components: $\ln \Pi_t$ and $\ln S_t$. Therefore, we consider two different initialisations:

Diffuse Prior: Initialise the Kalman filter with a "diffuse" prior $P_0 = I_{4\times 4}$ and $\hat{X}_0 = 0$ and delete the first two values from the log-likelihood to account for the two non-stationary state variables.[8] We have $N = 241$ monthly observations for the time period December 1998 to December 2018. The log-likelihood $\ln L$ for the whole sample is therefore computed as $\ln L = \sum_{n=3}^{N} \ln L_{n\Delta t}$, where the first term is the likelihood of the observed data of February 1999.

Stationary: Given that we can observe the non-stationary components $\ln \Pi_t$ and $\ln S_t$ without measurement errors, we can alternatively initialise the Kalman filter as follows. We set $\hat{X}_0 = \left(\mathbb{E}[X_\infty], \ln \Pi_0, \ln S_0\right)$ where $t = 0$ represents the observations for December 1998. For the initial estimation error we take

$$P_0 := \begin{pmatrix} \mathbb{V}\mathrm{ar}[X_\infty] & 0_{2\times 2} \\ 0_{2\times 2} & 0_{2\times 2} \end{pmatrix}, \tag{4.3}$$

where the 0-matrices signal that we know $(\ln \Pi_0, \ln S_0)$ exactly. We then run the Kalman filter on the remaining 240 observations starting at $t = 1$, i.e. January 1999. This initialisation is equivalent to replacing the components $\ln \Pi_t$ and $\ln S_t$ in the state-vector with their first

[8] We follow here the recommendation in Harvey (1991, Chapter 3).

differences and using the unconditional mean and variance of the adjusted (differenced) state-vector.

The likelihood L is a function of the model parameters of the KNW model. Therefore, we can estimate the model parameters of the KNW model by (numerically) maximising the likelihood L for the observed data-set of N observations.

When we have converged to an optimum, we compute the asymptotic standard errors as follows. At the optimal parameter values, we compute (numerically) the matrix $\mathscr{I}_L$ of second-order partial derivatives of $\ln L$ w.r.t. each of the model-parameters. This matrix is known as the *information matrix*. The inverse matrix $-\mathscr{I}_L^{-1}$ is the (asymptotic) covariance matrix for the maximum-likelihood parameter estimates. The (asymptotic) standard errors are computed as the square-roots of the diagonal elements $\text{diag}\left(-\mathscr{I}_L^{-1}\right)$.

4.1 Estimation Results: Unconstrained Model

In this subsection we present our estimation results and compare the outcomes to the results reported by DNB. Because we use numerical optimisation routines that are not exact, we expect to find small differences in the numerical values of the estimated parameters. We want to use a simple rule-of-thumb to assess whether the numerical differences in the estimated parameter values are large or small. The rule we use here is to compare the numerical differences to the standard error of the estimated parameter values in the form of a "z-score", which is the difference in estimated parameter value divided by the standard error. An absolute value of the z-score larger than 2 indicates that the differences in estimated parameter values are "large" and could be statistically significant.

Our estimation results are reported in Table 1 below, where the second column in the table reports the estimation results by DNB. The remainder of the table is split into two parts, which represent the "diffuse prior" initialisation and the "stationary" initialisation of the Kalman filter.

Diffuse Prior

We first discuss the "diffuse prior" results. The third column of the table reports our estimation results, the fourth column shows the (asymptotic) standard error for each parameter estimate. The fifth column reports the "z-score" to asses the difference from the values reported by DNB.

We highlight several features from the estimation results. First, we note that the measurement error standard deviations are very close to zero for y_5, y_{10}, y_{15}. This means that the Kalman filter mostly relies on these rates to estimate the state-vector X_t. It also means that the KNW model is capable of fitting the observed 5, 10 and 15 year rates very well. The 20 year rate has a measurement error st.dev. of 8bp, which indicates a slightly poorer fit. For the 1 year and 30 year rates, the model needs a measurement error st.dev. of 38bp and 22bp, which means that the KNW model cannot fit the observed 1 and 30 year rates very well using a 2-factor state-process.

		Diffuse Prior			Stationary		
Param	**DNB**	**Pelsser**	**Std.Err**	**z-sc**	**Pelsser**	**Std.Err**	**z-sc**
$\delta_{0\pi}$	0.0158	0.0232	(0.0076)	1.0	0.0158	(0.0067)	0.0
$\delta_{1\pi,1}$	-0.0028	-0.0028	(0.0010)	0.0	-0.0028	(0.0011)	0.0
$\delta_{1\pi,2}$	-0.0014	-0.0014	(0.0021)	0.0	-0.0014	(0.0021)	0.0
δ_{0r}	0.0097	0.0375	(0.0270)	1.0	0.0099	(0.0248)	0.0
$\delta_{1r,1}$	-0.0094	-0.0092	(0.0006)	0.3	-0.0093	(0.0007)	0.1
$\delta_{1r,2}$	-0.0024	-0.0029	(0.0016)	-0.3	-0.0025	(0.0023)	0.0
K_{11}	0.0479	-0.0728	(0.0765)	-1.6	0.0395	(0.0475)	-0.2
K_{22}	1.2085	1.2177	(0.2005)	0.0	1.2090	(0.3785)	0.0
K_{21}	0.5440	0.4709	(0.2105)	-0.3	0.5355	(0.3214)	0.0
$\sigma_{\Pi,1}$	-0.0010	-0.0010	(0.0004)	0.1	-0.0009	(0.0005)	0.2
$\sigma_{\Pi,2}$	0.0013	0.0012	(0.0006)	-0.1	0.0013	(0.0006)	-0.1
$\sigma_{\Pi,3}$	0.0055	0.0055	(0.0003)	0.2	0.0055	(0.0003)	0.2
η_S	0.0451	0.0365	(0.0313)	-0.3	0.0473	(0.0306)	0.1
$\sigma_{S,1}$	-0.0483	-0.0504	(0.0096)	-0.2	-0.0475	(0.0095)	0.1
$\sigma_{S,2}$	0.0078	0.0046	(0.0150)	-0.2	0.0071	(0.0177)	0.0
$\sigma_{S,3}$	0.0010	0.0016	(0.0089)	0.1	0.0009	(0.0067)	0.0
$\sigma_{S,4}$	0.1335	0.1328	(0.0062)	-0.1	0.1339	(0.0062)	0.1
$\lambda_{0,1}$	0.6420	0.4889	(0.1363)	-1.1	0.6475	(0.1515)	0.0
$\lambda_{0,2}$	-0.0240	-0.0139	(0.0717)	0.1	-0.0217	(0.1457)	0.0
$\tilde{\Lambda}_{1,1}$	0.1710	0.2751	(0.1052)	1.0	0.1788	(0.1010)	0.1
$\tilde{\Lambda}_{1,2}$	0.3980	0.4103	(0.0371)	0.3	0.4002	(0.0509)	0.0
$\tilde{\Lambda}_{2,1}$	-0.5140	-0.4360	(0.2266)	0.3	-0.5058	(0.3508)	0.0
$\tilde{\Lambda}_{2,2}$	-1.1470	-1.1381	(0.1788)	0.1	-1.1466	(0.3544)	0.0
h_1	0.0038	0.0038	(0.0002)	0.0	0.0038	(0.0002)	0.1
h_5	0.0003	0.0003	(0.0001)	0.1	0.0003	(0.0001)	0.0
h_{10}	0.0003	0.0003	(0.0000)	-0.2	0.0003	(0.0000)	0.0
h_{15}	0.0000	0.0000	(0.0001)	0.0	0.0000	(0.0000)	0.1
h_{20}	0.0008	0.0008	(0.0000)	0.1	0.0008	(0.0000)	0.1
h_{30}	0.0021	0.0022	(0.0001)	0.0	0.0022	(0.0001)	0.1
$\ln L$ (Diff)	10910.4	10912.4					
$\ln L$ (Stat)	10950.6				10951.3		
min.ev(K)	0.0479	-0.0728			0.0395		
min.ev(M)	0.0055	0.0065			0.0064		
UFR	-86.64%	-75.37%			-76.96%		
Π-return	1.59%	***			1.59%		
S-return	4.57%	***			4.82%		

Table 1: Parameter Estimates for Unconstrained KNW Model

Considering the differences in estimated values, we find no large differences in the estimates of the parameters. When we evaluate our log-likelihood for the DNB-parameters, we obtain (within our model setup) a value $\ln L = 10\,910.4$, which is not much lower than the optimised value of 10 912.4.

At the bottom of the table we report some additional diagnostic values. We present the smallest eigenvalue of the matrix K and M. The smallest eigenvalue for the matrix M is real and positive. For the matrix K, we find a negative eigenvalue of -0.0728 for the matrix K (this corresponds to the parameter K_{11} = -0.0728). Please note that the standard error for this estimate is 0.0765, hence the difference between our parameter estimate and the DNB-value of 0.0479 is not "large" with a z-score = 1.6, which is less than 2.

Economically, a negative eigenvalue for the matrix K means that the X_t-process is non-stationary. This means that the mean and variance of the interest rates, inflation rate and stock prices diverge for long horizons. As a consequence, the unconditional expected return for Π and S do not exist.

Stationary

We now discuss the "stationary" results, as shown in the right panel of the table. Considering the differences in estimated values, we find again no large differences with the DNB estimates. When we evaluate our log-likelihood for the DNB-parameters, we obtain (within our model setup) a value $\ln L = 10\,950.6$, which is not much lower than the optimised value of 10 951.3.

The biggest difference is the value for the parameter $K_{11} = 0.0395$ which is now positive, which also implies that the X_t-process is now stationary. This also means that the unconditional (geometric) expected return for Π and S exist and are equal to 1.59% and 4.82%, respectively.

We can explain the difference in the estimated value for K_{11} by noting that when we start the Kalman filter with stationary values, we impose an implicit restriction on the K-matrix that it should have positive eigenvalues. When the smallest eigenvalue of K approaches zero, this implies that the (unconditional) mean and variance of the X-process become ever larger. This has an adverse effect on the likelihood of the first observation, as the Kalman filter has to make an ever larger adjustment to bring $\hat{X}_1$ back in line with the first observation $\tilde{y}_1$.

At the end of this section, we want to emphasise that both sets of estimates are *statistically indistinguishable*. However, economically it is more plausible to work with a model that is stationary for modelling economic variables over a long time-horizon. We will therefore only report the "stationary" estimation results in the following sections.

4.2 Estimation Results: Return Constraints

The Dutch Parameter Committee 2019 has decided it wants to impose restrictions on the parameters of the KNW model, such that the UFR and the unconditional expected returns for inflation and stocks are equal to specific values.

The UFR value within the KNW model is given in equation (2.15). For the geometric return of Π_t and S_t, we make the following observation: the unconditional geometric expected return of S_t over a period of 1 year is given by

$$\ln\left(1+r_S^g\right) = \lim_{t\to\infty} \mathbb{E}\left[\ln\left(\frac{S_{t+1}}{S_t}\right)\right] = \delta_{0r} + \eta_S - \tfrac{1}{2}\sigma_S'\sigma_S \tag{4.4}$$

and the unconditional geometric expected return of Π_t over a period of 1 year is given by

$$\ln\left(1+r_\Pi^g\right) = \lim_{t\to\infty} \mathbb{E}\left[\ln\left(\frac{\Pi_{t+1}}{\Pi_t}\right)\right] = \delta_{0\pi} - \tfrac{1}{2}\sigma_\Pi'\sigma_\Pi. \tag{4.5}$$

These unconditional expectations are only well-defined for a stationary model, i.e. for a K-matrix with strictly positive eigenvalues.

When we impose the values UFR = 2.1% and $r_S^g = 5.6\%$ and $r_\Pi^g = 1.9\%$, this is equivalent to imposing the following three linear equality constraints upon the model-parameters:

$$\delta_{0r} = \ln(1.021) + \lambda_0' B_\infty + \tfrac{1}{2} B_\infty' B_\infty \tag{4.6}$$

$$\eta_S = \ln(1.056) - \delta_{0r} + \tfrac{1}{2}\sigma_S'\sigma_S \tag{4.7}$$

$$\delta_{0\pi} = \ln(1.019) + \tfrac{1}{2}\sigma_\Pi'\sigma_\Pi. \tag{4.8}$$

This means that we can estimate the KNW model with 29-3 = 26 free parameters plus the three constraints to identify the remaining three parameters. The results of this constrained estimation are reported in Table 2.

The first conclusion is that there are no large differences between the DNB estimates and our estimates. We can also confirm that imposing the constraints leads to a large drop in log-likelihood value of 10 737.8 - 10 951.3 = -213.5. This drop is mainly caused by imposing that the UFR should be 2.1%. A potential explanation is that such a high UFR-value is not supported by the interest rate data. This explanation is corroborated by the increase of measurement error standard deviations for the 20 and 30 year rates. The measurement st.dev. for y_{30} increases from 22bp to 35bp due to the UFR constraint.

On the other hand, imposing a constraint on the poorly identified parameters $\delta_{0\pi}$ and η_S is helpful for the estimation of other model parameters. The standard errors of the blocks of parameters $K, \tilde{\Lambda}_1$ have dropped considerably.

The bottom part of the table confirms that the eigenvalues of K and M are strictly positive for the constrained model.

4.3 Estimation Results: Negative Interest Rates

The last line in Table 2 reports the 2.5% quantile (under measure $\mathbb{P}$) of the 10 year zero-coupon rate $y_{10}(60)$ at the horizon $T = 60$. When this quantile is negative, it means that the probability of a negative 10-year rate at $T = 60$ is larger than 2.5%.

In this subsection, we investigate the impact of imposing one more constraint upon the KNW-model: to restrict the probability $\mathbb{P}[y_{10}(60) < 0] \leq 0.0250$. This constraint is a bit more challenging than the previous constraints, as it represents a non-linear constraint on the parameters of the KNW-model. Hence, we now have to impose a non-linear constraint on the

Param	DNB	Pelsser	Std.Err	z-sc
$\delta_{0\pi}$	0.0188	0.0188		
$\delta_{1\pi,1}$	-0.0022	-0.0023	(0.0015)	0.0
$\delta_{1\pi,2}$	0.0002	0.0003	(0.0011)	0.1
δ_{0r}	0.0209	0.0211		
$\delta_{1r,1}$	-0.0077	-0.0077	(0.0005)	0.1
$\delta_{1r,2}$	0.0004	0.0010	(0.0023)	0.3
K_{11}	0.0327	0.0386	(0.0749)	0.1
K_{22}	0.2627	0.2629	(0.1783)	0.0
K_{21}	0.3180	0.3774	(0.2158)	0.3
$\sigma_{\Pi,1}$	-0.0090	-0.0007	(0.0005)	0.3
$\sigma_{\Pi,2}$	0.0007	0.0008	(0.0004)	0.3
$\sigma_{\Pi,3}$	0.0055	0.0055	(0.0003)	0.1
η_S	0.0436	0.0434		
$\sigma_{S,1}$	-0.0558	-0.0549	(0.0104)	0.1
$\sigma_{S,2}$	-0.0026	0.0044	(0.0183)	0.4
$\sigma_{S,3}$	0.0005	0.0002	(0.0032)	-0.1
$\sigma_{S,4}$	0.1300	0.1305	(0.0064)	0.1
$\lambda_{0,1}$	0.6491	0.6377	(0.1675)	-0.1
$\lambda_{0,2}$	0.0080	-0.0745	(0.2054)	-0.4
$\tilde{\Lambda}_{1,1}$	0.1563	0.1717	(0.0626)	0.2
$\tilde{\Lambda}_{1,2}$	0.1902	0.1875	(0.0602)	0.0
$\tilde{\Lambda}_{2,1}$	-0.3077	-0.3887	(0.2485)	-0.3
$\tilde{\Lambda}_{2,2}$	-0.2201	-0.2429	(0.1516)	-0.2
h_1	0.0033	0.0032	(0.0002)	-0.5
h_5	0.0007	0.0007	(0.0001)	0.0
h_{10}	0.0004	0.0004	(0.0000)	0.1
h_{15}	0.0000	0.0000	(0.0001)	0.0
h_{20}	0.0011	0.0011	(0.0001)	0.4
h_{30}	0.0034	0.0035	(0.0002)	0.3
$\ln L$ (Stat)	10736.6	10737.8		
min.ev(K)	0.0327	0.0386		
min.ev(M)	0.0303	0.0319		
$\mathbb{P}_{2.5\%}[y_{10}(60)]$	-2.67%	-2.88%		

Table 2: Parameter Estimates for Constrained KNW Model

parameters of the model during the numerical optimisation of the log-likelihood. The results of this optimisation are reported in Table 3.

The first conclusion is that there are no large differences between the DNB estimates and our estimates. We can also confirm that imposing the non-negative rates constraint leads to a very small drop in log-likelihood value of 10 737.3 - 10 737.8 = -0.5. The main effect of imposing a non-negativity constraint on $y_{10}(60)$ is the increase in the lowest eigenvalue of K

Param	DNB	Pelsser	Std.Err	z-sc
$\delta_{0\pi}$	0.0188	0.0188		
$\delta_{1\pi,1}$	-0.0021	-0.0022	(0.0007)	-0.1
$\delta_{1\pi,2}$	0.0000	0.0000	(0.0001)	0.0
δ_{0r}	0.0212	0.0215		
$\delta_{1r,1}$	-0.0077	-0.0077	(0.0004)	0.0
$\delta_{1r,2}$	-0.0008	-0.0003	(0.0017)	0.3
K_{11}	0.0656	0.0738	(0.0352)	0.2
K_{22}	0.3032	0.3293	(0.1029)	0.3
K_{21}	0.2366	0.2954	(0.2312)	0.3
$\sigma_{\Pi,1}$	-0.0010	-0.0009	(0.0004)	0.3
$\sigma_{\Pi,2}$	0.0006	0.0007	(0.0004)	0.2
$\sigma_{\Pi,3}$	0.0055	0.0055	(0.0003)	0.1
η_S	0.0433	0.0431		
$\sigma_{S,1}$	-0.0528	-0.0525	(0.0101)	0.0
$\sigma_{S,2}$	-0.0114	-0.0054	(0.0155)	0.4
$\sigma_{S,3}$	0.0005	0.0001	(0.0070)	-0.1
$\sigma_{S,4}$	0.1307	0.1314	(0.0064)	0.1
$\lambda_{0,1}$	0.6730	0.6774	(0.1408)	0.0
$\lambda_{0,2}$	0.1180	0.0475	(0.1592)	-0.4
$\tilde{\Lambda}_{1,1}$	0.0910	0.1010	(0.0393)	0.3
$\tilde{\Lambda}_{1,2}$	0.2080	0.2147	(0.0302)	0.2
$\tilde{\Lambda}_{2,1}$	-0.2090	-0.2797	(0.2549)	-0.3
$\tilde{\Lambda}_{2,2}$	-0.2280	-0.2738	(0.1437)	-0.3
h_1	0.0033	0.0032	(0.0002)	-0.5
h_5	0.0007	0.0006	(0.0001)	-0.1
h_{10}	0.0004	0.0004	(0.0000)	0.1
h_{15}	0.0000	0.0000	(0.0001)	0.0
h_{20}	0.0011	0.0011	(0.0001)	0.4
h_{30}	0.0034	0.0035	(0.0002)	0.3
$\ln L$ (Stat)	10735.8	10737.3		
min.ev(K)	0.0656	0.0738		
min.ev(M)	0.0299	0.0319		
$\mathbb{P}_{2.5\%}[y_{10}(60)]$	0.07%	0.03%		

Table 3: Parameter Estimates for Constrained KNW Model with bound on negative rates

from 0.0386 to 0.0738. Increasing the lowest eigenvalue of K reduces the variance of X_{60} and this reduces the variance of $y_{10}(60)$ to satisfy the probability constraint. The 2.5% quantile of $y_{10}(60)$ is 0.07% for the DNB parameter values and 0.03% for our parameter values.[9] Both

[9]We obtain slightly positive values for the quantile, because we implement the non-linear inequality constraint via a "barrier function" that imposes a negative penalty on the likelihood when the inequality constraint is violated. The numerical optimum is therefore located in the strict interior of the feasible parameter-set.

these values are slightly positive, indicating that $\mathbb{P}[y_{10}(60) < 0] < 0.0250$.

5 Conclusion

In this paper, we have provided an independent derivation of the KNW model. We propose a different implementation of the state-space formulation of the KNW model and we tested the impact of two different specifications for the initialisation of the Kalman filter maximum-likelihood estimation.

We therefore have provided an independent verification of the parameter estimations provided by DNB for the Committee Parameters. We find no large differences between the parameters estimates reported by DNB and our own parameter estimates.

A Properties of vector-OU processes

In this appendix we present several useful results concerning vector-OU processes. More elaborate derivations and proofs of the results can be found in Chapter 8 of Arnold (1974).

Let us consider a vector linear stochastic differential equation of the form

$$dY_t = \left(a + AY_t\right)dt + C\,dW_t \tag{A.1}$$

where Y_t is d-dimensional stochastic process, a is a constant vector of length d and A is a constant matrix of size $d \times d$. We also have a k-dimensional Brownian Motion process W_t and the constant matrix C has size $d \times k$.

We briefly digress on the solution of systems of linear ordinary differential equations, which can be denoted in matrix notation as

$$\frac{dy(t)}{dt} = Ay(t) \tag{A.2}$$

with initial condition $y(0) = y_0$ and where the vector $y(t)$ has length d. The solution to (A.2) can be expressed as

$$y(t) = e^{At}y_0, \tag{A.3}$$

where e^{At} denotes the matrix exponential of the matrix At. The matrix exponential is defined by the power series

$$e^{At} = \sum_{n=0}^{\infty} \frac{A^n t^n}{n!}. \tag{A.4}$$

The above series always converges, so the matrix exponential is well-defined. From this definition follows that $e^{At}e^{As} = e^{A(t+s)}$ for all t, s. For $s = -t$ we get $e^{At}e^{-At} = e^0 = I$. If the matrix A can be diagonalised, i.e. when it has the representation $A = VDV^{-1}$ with D diagonal, then we can represent it's matrix exponential as

$$e^{At} = Ve^{Dt}V^{-1} \tag{A.5}$$

where e^{Dt} is also a diagonal matrix with elements $e^{d_{ii}t}$. This result follows directly by substituting $A = VDV^{-1}$ into the definition (A.4).

To calculate the derivative with respect to t we can differentiate term-by-term to obtain

$$\frac{de^{At}}{dt} = \sum_{n=1}^{\infty} \frac{A^n n t^{n-1}}{n!} = Ae^{At} = e^{At}A. \tag{A.6}$$

This result shows that (A.3) is indeed a solution to (A.2).

To find an explicit solution for (A.1), let us consider the auxiliary process $e^{-At}Y_t$. If we apply Ito's Lemma we obtain

$$\begin{aligned} de^{-At}Y_t &= \left(-Ae^{-At}Y_t + e^{-At}(a + AY_t)\right)dt + e^{-At}C\,dW_t \\ &= e^{-At}a\,dt + e^{-At}C\,dW_t \end{aligned} \tag{A.7}$$

and we see that this equation only depends on deterministic functions. Hence, we can explicitly represent the solution to (A.7) for $T > t$ as

$$Y_T = e^{A(T-t)} Y_t + \int_t^T e^{A(T-u)} a\, du + \int_t^T e^{A(T-u)} C\, dW_u. \tag{A.8}$$

We can interpret equation (A.1) as the "auto-regressive" representation of the process Y_t. The solution (A.8) can be interpreted as the "moving average" representation of the process Y_t.

From this explicit solution we can deduce that the transition density $f(Y_T \mid Y_t)$ is the multivariate Gaussian distribution

$$f(Y_T \mid Y_t) \sim N\Big(e^{A(T-t)} Y_t + \int_t^T e^{A(T-u)} a\, du\,;\, \int_t^T e^{A(T-u)} CC' e^{A'(T-u)}\, du\Big). \tag{A.9}$$

Due to the time-homogeneity of the model, the mean and variance depend only on the time difference $\tau := T - t$. We will therefore denote the mean and variance of the transition density by

$$m_0(\tau) := \int_0^\tau e^{Au} a\, du \tag{A.10}$$

$$M_1(\tau) := e^{A\tau} \tag{A.11}$$

$$V(\tau) := \int_0^\tau e^{Au} CC' e^{A'u}\, du. \tag{A.12}$$

The integrals over the matrix-exponentials can be solved numerically by noting that $m_0(t)$ and $V(t)$ solve the following (systems of) ode's:

$$\frac{dm_0(t)}{dt} = a + A\, m_0(t) \qquad m_0(0) = 0_{d\times 1} \tag{A.13}$$

$$\frac{dV(t)}{dt} = A\, V(t) + V(t)\, A' + CC' \qquad V(0) = 0_{d\times d} \tag{A.14}$$

For a small time-step Δt we can use an Euler-discretisation to obtain the following approximation to (A.1):

$$Y_{t+\Delta t} - Y_t = \big(a + AY_t\big)\Delta t + C(W_{t+\Delta t} - W_t) + \mathcal{O}(\Delta t^{3/2}). \tag{A.15}$$

Hence, we can approximate the transition density for small Δt as

$$f(Y_{t+\Delta t} \mid Y_t) \approx N\big((I + A\Delta t) Y_t + a\Delta t\,;\, CC'\Delta t\big). \tag{A.16}$$

References

Arnold, L. (1974). *Stochastic Differential Equations*. John Wiley and Sons. 17

Babbs, S. and Nowman, B. K. (1999). Kalman filtering of generalized Vasicek term structure models. *Journal of Financial and Quantitative Analysis*, 34(1). 2

Dai, Q. and Singleton, K. J. (2000). Specification analysis of affine term structure models. *The Journal of Finance*, 55(5):1943–1978. 2

De Jong, F. (2000). Time series and cross-section information in affine term-structure models. *Journal of Business & Economic Statistics*, 18(3):300–314. 2

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(3):463–520. 3

Draper, N. (2014). A financial market model for The Netherlands. Technical report, Centraal Planbureau. 2, 7

Harvey, A. (1991). *Forecasting, structural time series models and the Kalman filter*. Cambridge university press. 2, 7, 9

Koijen, R. S. J., Nijman, T. E., and Werker, B. J. M. (2010). When can life cycle investors benefit from time-varying bond risk premia? *Review of Financial Studies*, 23(2):741–780. 2, 5

Muns, S. (2015). A financial market model for The Netherlands: A methodological refinement. Technical report, Centraal Planbureau. 2, 4, 5, 7