A Principal-Component-Based Affine Term Structure Model

R. Rebonato, I Saroka, V Putyatin
PIMCO, Oxford University

June 16, 2014

Abstract

We present an essentially affine model with pricipal components as state variables. We show that, once no-arbitrage is imposed, this choice of state variables imposes some unexpected constraints on the reversion-speed matrix, whose N^2 elements can be uniquely specified by its N eigenvalues. The requirement that some of its elements should be negative gives rise to a potentially complex dynamics, whose implications we discuss at length. We show how the free parameters of the model can be determined by combining cross-sectional information on bond prices with time-series information about excess returns and by enforcing a 'smoothness' requirement. The calibration in the $\mathbb{P}$ and $\mathbb{Q}$ measures does not require heavy numerical search, and can be carried out almost fully with elementary matrix operations. Once calibrated, the model recovers exactly the (discrete) yield cuirve shape, the yield covariance matrix, its eigenvalues and eigenvectors. The ability to recover yield volatilities well makes it useful for the estimation of convexity and term premia. The model also recovers well quantities to which it has not been calibrated, and offers an estimation of the term premia for yields of different maturities which we discuss in the last section.

1 Introduction and Motivation

The theory of affine and essentially affine models is well established. See, eg, Bolder (2001) for a survey that covers both theory and implementation issues, or Piazzesi (2010) for an up-to-date and comprehensive review.

Recently an interesting variation on this well-rehearsed theme has been introduced by Christensen, Diebold, Rudebush (2011), and developed in Diebold and Rudebush (2013) who show how to turn the 'static' (curve-fitting) Nelson and Siegel (1987) model into a dynamic affine model.[1] To the extent that the coefficients of the Nelson-Siegel model generate a close match to the observed term structure – and it is well known that they do –, the dynamic Nelson-Siegel

[1] See also Ungari and Turc (2012) for a closely related treatment.

formulation automatically ensures an easy calibration to the market bond prices. This is in itself a desirable result. There is, however, a more important positive feature to the approach: Diebold and Rudebush (2013) in fact show that, perhaps surprisingly, after a clever transformation the factors of the associated affine model lend themselves to an appealing interpretation as principal components. If one can identify the factors as principal components (or their proxies) one can draw on a wealth of econometric[2] and macrofinancial[3] studies to constrain their behaviour, and guide the parameter estimation ('calibration') process.

The appeal of this approach naturally raises the question if it is possible to assign a Gaussian affine behaviour *exactly* to the principal components, rather than to some proxies, and, at the same time, comply with the conditions of no-arbitrage.

The idea of harnessing together two of the most-commonly-used workhorses of term structure modelling – principal component analysis and affine (mean-reverting) modelling – is natural enough, and indeed has been exploited, more or less directly, in some recent approaches. (See, eg, Joslin, Ahn Le and Singleton (2013), Joslin, Singleton and Zhu (2011), Joslin, Priebsh and Singleton (2104) and references therein). Our work positions itself in this line of research. More precisely, for our purposes a useful starting point is the work by Dai and Singleton (2001), who show that, if N factors, $\overrightarrow{x}_t$,[4] follow a diffusive process of the form

$$d\overrightarrow{x}_t = \overrightarrow{a}\left(\overrightarrow{x}_t\right)dt + \underline{b}\left(\overrightarrow{x}_t\right)d\overrightarrow{z}_t \tag{1}$$

with

$$\begin{aligned} a\left(\overrightarrow{x}_t\right) &= \overrightarrow{a}_0 + \overrightarrow{a}_1\overrightarrow{x}_t, \qquad a_0 \in R^N, a_1 \in R^{N\times N} \\ \underline{b}\left(\overrightarrow{x}_t\right)\underline{b}\left(\overrightarrow{x}_t\right)^T &= \underline{b}_0 + b_1\overrightarrow{x}_t, \qquad b_0 \in R^{N\times N}, b_1 \in R^{N\times N\times N} \end{aligned} \tag{2}$$

and if the short rate, r_t, can be written as a linear combination of these N factors plus a constant,

$$r_t = c_0 + \overrightarrow{c}_1\overrightarrow{x}_t \tag{3}$$

then bond prices, P_t^T, can be written as exponentially affine functions of the factors,

$$P_t^T = e^{A_t^T + \left(\overrightarrow{B}_t^T\right)^T\overrightarrow{x}_t} \tag{4}$$

Following the notation in Dai and Singleton (2001), we focus in what follows on models for which $b_1 = 0$, in which case the factors follow an N-dimensional Ornstein-Uhlenbeck process.

Apart from the short-rate requirement that $r_t = c_0 + \overrightarrow{c}_1\overrightarrow{x}_t$ the factors can, up to this point, be totally general. However, given the exponentially affine nature of the bond pricing function, it is always the case that

$$\overrightarrow{y}_t^T = \overrightarrow{u}_t + \underline{U}_t\overrightarrow{x}_t \tag{5}$$

[2] For an early study relating to Treasuries, see Litterman and Scheinkman (1991).

[3] See, eg, Duffee (2002), Fama and Bliss (1987), Fama and French (1983, 1989), and the references in Joslin, Priebsch and Singleton (2014).

[4] See Section 00 for a description of our notation.

How the affine link between the factors and the yields is established provides a useful classification perspective for exponentially affine models. More precisely, in some approaches the factors are latent variables and the 'loadings' ($\overrightarrow{u}_t$ and $\underline{U}_t$) are not specified a priori, (see, eg, D'Amico et al (2004)), but are derived from the no-arbitrage conditions and the calibration (eg, via Kalman filtering) of the model. In other approaches, the modeller assigns *a priori* the link between the factors and the yields. For instance, Duffie and Kan (1996) simply identify the factors with the yields themselves ($\overrightarrow{u}_t = 0$ and $\underline{U}_t = 0$). More interestingly, macrofinancial models link the observable yields (or linear functions thereof) to macroeconomic observables via some structural models. We define *pre-specified models*[5] all models where the loadings $\overrightarrow{u}_t$ and $\underline{U}_t$ are assigned *a priori*.

The advantages of working with non-latent factors have been widely discussed[6]. However, once absence of arbitrage is imposed, an exogenous, *a priori* specification of the loadings places severe restrictions on the admissible coefficients of the process for the factors. Fortunately, Saroka (2014) presents a general expressions for the admissible parameters of the N-dimensional O-U process for the factors of pre-specified models, ie, when the loadings $\overrightarrow{u}_t$ and $\underline{U}_t$ are assigned *a priori*. In this paper we make use of these results for the special case when the factors are chosen to be principal components.

In so doing we discover some interesting results: indeed, we show in Section 3 that it is possible to specify *an infinity* of term structure models such that:

- the driving factors are principal components;
- they follow a men-reverting (generalized Ornstein Uhlenbeck) dynamics;
- an arbitrary exogenous covariance matrix among N yields can always be exactly recovered (and hence so are all the observed eigenvalues and eigenvectors);
- an arbitrary exogenous yield curve (also defined by N yields) is exactly recovered;
- no-arbitrage is satisfied.

Accomplishing this, however, imposes some important constraints on the mean-reverting dynamics, the reason for which is rather subtle. An intuitive explanation of what these constraints entail goes along the following lines.

1.1 Parameter Constraints for PCA Pre-Specified Models

First of all, it is well known that, given an N-dimensional O-U process, diagonalizing via an orthogonal transformation *either* the diffusion *or* the 'reversion-

[5] Saroka (2014) calls them *observable affine-factor models.*

[6] As discused in Diebold and Rudebush (2013) amd Kim (2007), latent factors are difficult to interpret economically, make an assessment of the plausibility of their equation of motion arduous, fail to impose stringent constraints on the admissible values of the parameters, and tend to be produce models which are far from parsimonious.

speed' matrix is always possible (and, indeed, straightforward).[7] The associated 'rotation of axes' has no economic significance, and all the 'invariants' – bond prices, short rate, etc –[8] are recovered. However, we show that an affine model with diagonal diffusion *and* $\mathbb{Q}$-measure reversion-speed matrices is not compatible with absence of arbitrage.

This is somewhat surprising, and economically significant: if we want the factors to be principal components, the diffusion matrix must be diagonal. Because of the result in Saroka (2014), we show that the reversion-speed matrix cannot be diagonal as well, *and that some of its elements must be negative and of the same order of magnitude as the positive ones.*

This matters: if the reversion-speed matrix is forced to contain non-diagonal negative elements, the outcome is a rather 'complex' $\mathbb{Q}$-measure *deterministic* dynamics (even for asymptotically stable systems): each principal component not only reverts to its own fixed reversion levels, but is also attracted to, and repulsed by, the other dynamically moving principal components. Therefore the principal components state variables (and hence yields, to which they are linked via an affine transformation) are forced to follow a complex *deterministic* evolution, whereby a reversion level is not approached with a monotonic first derivative (a 'decaying-exponential' approach), but with unavoidable over- and/or undershoots during which the first derivative changes sign. This evolution may well be asymptotically stable, but can easily produce complex *predictable* oscillations of the expectations of yields many years into the future.

We find that the 'impossibility results' and the constraints they impose raise some interesting questions about what an affine description of the yield curve in terms of principal components entails – for instance, the interplay between the persistence of yields, the risk premia, and the 'complexity of the yield curve, or the ability to detect unit-root behaviour for rates and principal components with reasonable-size samples. We touch on these aspects in the concluding section of this paper.

2 Our Strategy to Link the $\mathbb{P}$ and $\mathbb{Q}$ Measures

Most work in affine term-structure modelling straddles the physical and risk-neutral measures. In one common approach (for a recent and popular example, see, eg, D'Amico et al (2004)), one starts from the estimation in the real-world ($\mathbb{P}$) measure of the statistical properties of some features of the state variables (say, the reversion level)[9]. In parallel, cross-sectional information about prices allows the determination in the risk-neutral ($\mathbb{Q}$) measure of the same 'risk-adjusted' statistical quantities. The market price of risk is then usually obtained as the 'difference' (change of measure) between the two set of quantities. Which

[7] tCherito, Filipovic and Kimmel (2010) shopw that, under loose conditions, it is possible to diagonalize the dioffusion matrix using a regular, but not necessarily orthogonal, transformation – see their Theorem 2.1 and Corollary 2.2.

[8] as defined in Cherito, Filipovic and Kimmel (2010)

[9] In this apprach, the $\mathbb{P}$-measure statistical estimation is sometime supplemented by survey data.

quantities require risk adjustment depends on the posited dependence (if any) of the market price of risk on the state variables.

In a complementary approach (see, eg Cochrane and Piazzesi (2008)) one links instead the two measures by looking at the excess returns produced by systematically investing in long-dated (n-period) bonds and by financing at the 1-period rate.

In our work we follow a variant of this approach. More precisely,

1. we start by determining the measure-invariant model parameters (the coefficients of the diagonal diffusion matrix) using real-world volatility data;

2. keeping these data fixed, we determine the measure-dependent reversion-speed matrix in the $\mathbb{Q}$ measure by *cross-sectional* fitting to the whole covariance matrix;

3. wit this information, we carry out a cross-sectional fit to the yield curve to determine the reversion-level vector in the $\mathbb{Q}$ measure;

4. we then carry out an empirical study of excess returns, and we establish (by multi-variate regression) a link between these excess returns and our state variables;

5. as a next step, we determine (see Section 00) the shape of the dependence of the reversion-level vector and the reversion-speed matrix (in the $\mathbb{P}$ measure) on the market prices of risk associated with our model and our chosen state variables – in order to accommodate the empirical findings in Duffee (2002) and the results of our own studies, at this point we allow for the market price of risk to depend in an affine manner on the state variables, (ie, we require our model to be essentially affine);

6. finally, we specialize the results in point 5. above so as to reflect the particular dependence determined in our empirical estimation of excess returns.

We stress that the last step is quite general, and does not rely on the specific empirical findings of our statistical estimation. For instance, a Cochrane-and-Piazzesi-like return-predicting factor (see Cochrane and Piazzesi (2005, 20008))[10] or a slope factor (as in Duffee, 2002) can be readily accommodated by our methodology.

So, for the avoidance of doubt: we start from the $\mathbb{Q}$ measure and we determine by cross-sectional fit to bond prices the $\mathbb{Q}$-measure model parameters; we carry out a statistical estimate of excess returns; with this information we distil the $\mathbb{P}$-measure model parameters.

[10] In order to accommodate exactly the Cochrane-Piazzesi 'tent' factor, five principal components would have to be used. Conceptually, our approach extends without difficulty to as many factors as desired. The uniqueness of parameters in the calibration phase may disappear if too many factors are used.

3 The Set-Up

3.1 Notation

In the following, we indicate by $\overrightarrow{x}$ a (column) vector in $\mathbb{R}^N$, and by $\overrightarrow{x}^T$ its transpose (a row vector). We do *not* employ the superscript-subscript convention for covariant and contravariant vectors.

A matrix in $\mathbb{R}^{N \times N}$ is denoted by $\underline{M}$. Its transpose and inverse are denoted by $\underline{M}^T$ and $\underline{M}^{-1}$, respectively. The symbol $[\underline{M}]_{ij}$ signifies the $[j,i]$th element of matrix $\underline{M}$.

The time-t price of a discount bond of maturity T is denoted by P_t^T, and its yield by y_t^T. The time-t value of the short rate is denoted by r_t.

We describe the time-t discrete yield curve by an $[N \times 1]$ vector of yields, $\overrightarrow{y_t}$, of elements $y_t^{T_i}$, $i = 1, 2..N$. The elements of the vector $\overrightarrow{y_t}$ are ordered with increasing maturity ($T_j > T_k$ if $j > k$). The first element of $\overrightarrow{y_t}$ is r_t: $y_t^{T_1} = r_t$.

Finally, we denote by $\overrightarrow{e}_1$ the column vector $[1, 0, 0, ...0]^T$, and by $\underline{I}$ the identity matrix. In particular,

$$r_t = \overrightarrow{e}_1^T \overrightarrow{y_t} \tag{6}$$

3.2 The Geometry of the Problem

Consider the following dynamics for the component yields of the $N \times 1$ vector $\overrightarrow{y}$:

$$\overrightarrow{dy} = [...]dt + \underline{\sigma}\overrightarrow{dw}^{\mathbb{Q},\mathbb{P}} \tag{7}$$

with

$$\underline{\sigma} = diag\left[\sigma_1, \sigma_2, ..., \sigma_n\right] \tag{8}$$

$$E\left[\overrightarrow{dw}\overrightarrow{dw}^T\right] = \underline{\rho}dt \tag{9}$$

and the drift term reflecting the no-arbitrage conditions when $\overrightarrow{dw}^{\mathbb{Q},\mathbb{P}} = \overrightarrow{dw}^{\mathbb{Q}}$, and the real-world deterministic dynamics when $\overrightarrow{dw}^{\mathbb{Q},\mathbb{P}} = \overrightarrow{dw}^{\mathbb{P}}$.

Thee covariance matrix among the yields is given by

$$E\left[\overrightarrow{dy}\overrightarrow{dy}^T\right] = \underline{\sigma}\underline{\rho}\underline{\sigma}^T = \underline{\Sigma}_{mkt}dt \tag{10}$$

This quantity is an exogenous market observable, which we assume to be known and constant. This is one of the key quantities that we would like our model to reproduce, linked as it is to the convexity contribution to the shape of the yield curve, and to the apportionment of the risk premia among different yields.

The real symmetric matrix $\underline{\Sigma}_{mkt}$ can always be diagonalized to give

$$\underline{\Sigma}_{mkt} = \underline{\Omega\ \Lambda\ \Omega}^T \tag{11}$$

with

$$\underline{\Lambda} = diag\left[\lambda_1, \lambda_2, ..., \lambda_N\right] \tag{12}$$

and $\underline{\Omega}$ is an orthogonal matrix:

$$\underline{\Omega}\underline{\Omega}^T = \underline{I} \tag{13}$$

To the extent that the exogenous matrix $\underline{\Sigma}_{mkt}$ is positive definite, all the eigenvalues λ_i are positive.

Given this diagonalization, we can define the principal components, $\overrightarrow{x}$, by

$$\overrightarrow{y}_t = \overrightarrow{\widetilde{y}} + \underline{\Omega}\overrightarrow{x}_t \tag{14}$$

where $\overrightarrow{\widetilde{y}}$ is a *constant* vector. Because of (**??**) and (6) we have

$$r_t = \overrightarrow{e}_1^T \overrightarrow{y_t} = \overrightarrow{e}_1^T \overrightarrow{\widetilde{y}} + \overrightarrow{e}_1^T \underline{\Omega}\overrightarrow{x} \tag{15}$$

To make the notation more compact we set

$$\omega_0 \equiv \overrightarrow{e}_1^T \overrightarrow{\widetilde{y}} \tag{16}$$

$$\overrightarrow{\omega}_1^T \equiv \overrightarrow{e}_1^T \underline{\Omega} \tag{17}$$

and therefore

$$r_t = \omega_0 + \overrightarrow{\omega}_1^T \overrightarrow{x} \tag{18}$$

3.3 The Dynamics of the Problem

We impose that the principal components, $\overrightarrow{x}_t$, should follow an affine diffusion of the form:

$$d\overrightarrow{x}_t = \underline{\mathcal{K}}\left(\overrightarrow{\theta} - \overrightarrow{x}_t\right)dt + \underline{S}\overrightarrow{dz} \tag{19}$$

and

$$E\left[\overrightarrow{dz}\overrightarrow{dz}^T\right] = \underline{I}dt \tag{20}$$

We refer to $\underline{\mathcal{K}}$ as the reversion-speed matrix, to $\underline{S}$ as the diffusion matrix, and to $\overrightarrow{\theta}$ as the reversion-level vector. For reasons that will become apparent in the following, we require the matrix $\underline{\mathcal{K}}$ to be invertible and full rank.[11] Since we want to interpret the factors, $\overrightarrow{x}_t$, as principal components, we require the matrix $\underline{S}$ to be diagonal:

$$\underline{S} = \text{diag}\left[s_1, s_2, ...s_N\right] \tag{21}$$

and we impose

$$s_i = \sqrt{\lambda_i} \tag{22}$$

For the reasons discussed in the introductory section, we would also like the reversion-speed matrix, $\underline{\mathcal{K}}$, to be diagonal, but, at this stage, we do not know whether this is possible (once a diagonal form is imposed for the diffusion matrix) – indeed, we shall see that it is not.

[11] Saroka (2014) shows how the full-rank requirement can be relaxed.

Absence of arbitrage then imposes that

$$P_t^T = E^{\mathbb{Q}}\left[e^{-\int_t^T r_s ds}\right] \tag{23}$$

and therefore, because of (**??**), we have

$$P_t^T = E^{\mathbb{Q}}\left[e^{-\int_t^T \left(\overrightarrow{\omega}_0 + \overrightarrow{\omega}_1^T \overrightarrow{x}_s\right) ds}\right] \tag{24}$$

3.4 Solution

It is well known[12] (see, eg, Dai and Singleton (2000)) that the solution to Equation (24) is given by

$$P_t^T = \exp^{A_t^T + \left(\overrightarrow{B}_t^T\right)^T \overrightarrow{x}_t} \tag{25}$$

with the vector $\overrightarrow{B}_t^T$ and the scalar A_t^T satisfying the ordinary differential equations (with $\tau = T - t$)

$$\frac{dA_\tau}{d\tau} = -\rho_0 + \left(\overrightarrow{B}_\tau\right)^T \underline{\mathcal{K}}\,\overrightarrow{\theta} + \frac{1}{2}\left(\overrightarrow{B}_\tau\right)^T \underline{SS}^T \overrightarrow{B}_\tau \tag{26}$$

$$\frac{d\overrightarrow{B}_\tau}{d\tau} = -\overrightarrow{\rho}_1^T - \underline{\mathcal{K}}\overrightarrow{B}_\tau \tag{27}$$

with boundary conditions

$$\overrightarrow{B}\,(\tau = 0) = 0, \quad A\,(\tau = 0) = 0 \tag{28}$$

The solution for $\overrightarrow{B}_\tau$ is given by

$$\overrightarrow{B}_\tau = -\int_0^\tau e^{-\underline{\mathcal{K}}\tau}\underline{\Omega}\,\overrightarrow{e}_1 d\tau \tag{29}$$

Not every square matrix can be diagonalized. In what follows, we consider the case where the matrix $\underline{\mathcal{K}}$ has distinct and real eigenvalues. When both these conditions are satisfied, the matrix $\underline{\mathcal{K}}$ can always be diagonalized, and the diagonalizing matrix is real. We refer to the reader to Saroka (2014) for a more general treatment. We find little difference between the the solutions we obtain assuming diagonalization and the more general treatment.

If one diagonalizes the reversion-speed matrix, $\underline{\mathcal{K}}$, one obtains:

$$\underline{\mathcal{K}} = \underline{a}\Lambda_{\mathcal{K}}\underline{a}^{-1} \tag{30}$$

with

$$\underline{\Lambda}_{\mathcal{K}} = \mathrm{diag}\,[l_j]\,, \quad j = 1, 2, .., N \tag{31}$$

One can then easily derive

$$\overrightarrow{B}_\tau = -a\mathrm{diag}\left[\frac{1}{l_j}\left(1 - e^{-l_j\tau}\right)\right] a^{-1}\underline{\Omega}^T\overrightarrow{e}_1 \tag{32}$$

[12] All the proofs are presented in the Appendices I to V.

Once the vector $\overrightarrow{B}_\tau$ has been obtained, the scalar A_τ can be obtained as

$$A_\tau = \int_0^\tau \left[-\rho_0 + \left(\overrightarrow{B}_\tau\right)^T \underline{\mathcal{K}} \overrightarrow{\theta} + \frac{1}{2} \left(\overrightarrow{B}_\tau\right)^T \underline{SS}^T \overrightarrow{B}_\tau \right] d\tau \tag{33}$$

In the case we consider, the integral can be carried out analytically, and the resulting expression is given in Appendix A.

3.5 Conditions for Identifiability

As mentioned above, we want to explore under which conditions it is possible to assign the mean-reverting dynamics (19) for the factors, and to identify them as principal components. In particular, we would like to see whether the choice of principal components as state variables admits a diagonal reversion-speed matrix, $\underline{\mathcal{K}}$. We call this the 'identifiability problem'.

From the no-arbitrage dynamics (19), and the solution (112), the yields vector, $\overrightarrow{y}_t$, has the expression

$$\overrightarrow{y}_t = -\underline{\beta} \overrightarrow{x}_t - \overrightarrow{\alpha} \tag{34}$$

with

$$\overrightarrow{\alpha} = \begin{bmatrix} \frac{A_{\tau_1}}{\tau_1} \\ \frac{A_{\tau_2}}{\tau_2} \\ \cdots \\ \frac{A_{\tau_N}}{\tau_N} \end{bmatrix} \tag{35}$$

and

$$\underline{\beta} = \begin{bmatrix} \frac{\left(\overrightarrow{B}_{\tau_1}\right)^T}{\tau_1} \\ \frac{\left(\overrightarrow{B}_{\tau_2}\right)^T}{\tau_2} \\ \cdots \\ \frac{\left(\overrightarrow{B}_{\tau_N}\right)^T}{\tau_N} \end{bmatrix} \tag{36}$$

At the same time, for identifiability of the factors with principal components, Equation (14) must also hold:

$$\overrightarrow{y}_t = \overrightarrow{\widetilde{y}} + \underline{\Omega} \overrightarrow{x}_t \tag{37}$$

For Equations (14) and (34) to be compatible for an arbitrary vector $\overrightarrow{x}_t$, one must therefore have

$$\underline{\Omega} = -\underline{\beta} \tag{38}$$

and

$$\overrightarrow{\widetilde{y}} = -\overrightarrow{\alpha} \tag{39}$$

From Equations (14) and (34) it also follows that

$$\widetilde{y}_1 = -\frac{A(+0)}{+0} \tag{40}$$

$$\widetilde{y}_k = -\frac{A(\tau_k)}{\tau_k}, \quad k = 2, 3, ..., N \tag{41}$$

In sum: if the vector $\overrightarrow{\widetilde{y}}$ is chosen as per Equations (40) to (41), the time-0 discrete yield curve is automatically and exactly recovered for any reversion-level vector, $\overrightarrow{\theta}$. (We discuss in the calibration section how a 'good' choice the vector $\overrightarrow{\theta}$ can be made.) If we can then find a reversion speed matrix, $\mathcal{K}$, such that Equation (38) is also satisfied, we can rest assured that the chosen yields will have a (discrete) model covariance matrix automatically identical to the exogenously assigned matrix, $\underline{\Sigma}_{mkt}$.

Assuming that a solution satisfying (38) and (39) can indeed be found, the extreme ease with which this usually nettlesome joint calibration problem can be tackled clearly shows at least one advantage from identifying the vector $\overrightarrow{x}_t$ with the principal components. We therefore turn in the next section to the showing that the identification is indeed possible.

Before that, we note in passing that the first element of the vector $\overrightarrow{\widetilde{y}}$ is at this point indeterminate, ie, any value can be chosen for it, and all N yields can be recovered exactly.[13] This can be seen as follows. Recall that the bond price is given by

$$P_t^T = \exp^{A_t^T + \left(\overrightarrow{B}_t^T\right)^T \overrightarrow{x}_t} \tag{42}$$

But we have from above that

$$\alpha_i = \frac{A_i}{\tau_i} \tag{43}$$

and $\overrightarrow{\widetilde{y}} = -\overrightarrow{\alpha}$ (with $A_i \equiv A_t^{T_i}$). Therefore $A_i = -\tau_i \widetilde{y}_i$, and $A_1 = -\tau_1 \widetilde{y}_1$ in particular. We know, however, that

$$\lim_{\tau \to 0} \frac{A(\tau)}{\tau} = 0 \tag{44}$$

$$\lim_{\tau \to 0} \frac{B(\tau)}{\tau} = 1 \tag{45}$$

and therefore we see from Equation (40) that any value can be assigned to $\widetilde{y}_1$, while retaining the property that the infinitesimally short yield be given by

$$\lim_{(T-t) \to 0} y_t^T = r_t \tag{46}$$

4 Results

4.1 Impossibility of Identification When $\mathcal{K}$ Is Diagonal

From Equation (38), and recalling that $\underline{\Omega}$ is an orthogonal matrix, it is clear that one must have

$$\underline{\beta\beta}^T = \underline{I} \tag{47}$$

[13] However, all the yield-recovering solutions associated with different values of $\overrightarrow{\widetilde{y}}_1$ will imply different model parameters: we will show in the following how this indeterminacy can be resolved.

To prove that the reversion-speed matrix $\mathcal{K}$ cannot be diagonal when the state variables are principal components, we proceed by reduction ad absurdum, ie, we show that, given Equation (47), an impossibility arises.

So, let's assume that $\mathcal{K}$ is diagonal and the state variables independent. In this setting it is straightforward to show that the matrix $\underline{\beta}$ is given by[14]

$$\underline{\beta} = \begin{bmatrix} [\omega_{11}, \omega_{12}, ..., \omega_{1N}] \\ \frac{1}{\tau_2}\left[\frac{1-e^{-\kappa_{11}\tau_2}}{\kappa_{11}}\omega_{11}, \frac{1-e^{-\kappa_{22}\tau_2}}{\kappa_{22}}\omega_{12}, ..., \frac{1-e^{-\kappa_{NN}\tau_2}}{\kappa_{NN}}\omega_{1N}\right] \\ ... \\ \frac{1}{\tau_N}\left[\frac{1-e^{-\kappa_{11}\tau_N}}{\kappa_{11}}\omega_{11}, \frac{1-e^{-\kappa_{22}\tau_N}}{\kappa_{22}}\omega_{12}, ..., \frac{1-e^{-\kappa_{NN}\tau_N}}{\kappa_{NN}}\omega_{1N}\right] \end{bmatrix} \tag{48}$$

with

$$[\omega_{11}, \omega_{12}, ..., \omega_{1N}] = \overrightarrow{e}_1^T \underline{\Omega} \tag{49}$$

ie, the row vector whose elements are the first row of the eigenvector matrix, $\underline{\Omega}$. (The rows 2 to N are straightforward. The first row is obtained by recalling from Equation (45) that the limit of $B(\tau)/\tau$ as τ goes to zero.)

Consider now $\underline{\beta\beta}^T$. The element $\left[\underline{\beta\beta}^T\right]_{11}$ is indeed equal to 1 (as it should if Equation (47) is to be satisfied). Consider, however, a generic element $[r, s]$ with $r \neq 1, s = 1$. For identifiability, we we should have

$$\left[\overrightarrow{\beta}\,\overrightarrow{\beta}^T\right]_{r1} = \delta_{r1} \tag{50}$$

In reality we have:

$$\left[\overrightarrow{\beta}\,\overrightarrow{\beta}^T\right]_{r1} = \frac{1}{\tau_r}\sum_j \frac{1-e^{-\kappa_{jj}\tau_s}}{\kappa_{jj}}\omega_{1j}^2 \tag{51}$$

But this term cannot possibly be zero for $r \neq s$, because we know that $\sum_j \omega_{1j}^2 = 1$, *and all the terms* $\frac{1-e^{-\kappa_{jj}\tau_s}}{\kappa_{jj}}$ *are strictly positive.* Therefore the matrix $\mathcal{K}$ cannot be diagonal.

We can summarize the first result as follows.

Conclusion 1 *If the factors $\overrightarrow{x}_t$ are principal components, and hence the diffusion matrix $\underline{S}$ is diagonal, absence of arbitrage is not compatible with a diagonal reversion-speed matrix, $\mathcal{K}$.*

4.2 Constraints on $\mathcal{K}$ for Identifiability

We have ascertained that, if the factors $\overrightarrow{x}_t$ are principal components (and we want to preclude the possibility of arbitrage), the matrix $\mathcal{K}$ cannot be diagonal. This raises the question of whether absence of arbitrage is compatible with *some* reversion-speed matrix, $\underline{\mathcal{K}}$, for factors $\overrightarrow{x}_t$ that behave like principal components.

[14] Given the decoupling of the variables in this setting, each term $\frac{1-e^{-\kappa_{ii}\tau_2}}{\kappa_{ii}}$ is simply a 'Vasicek'-like term.

The answer is affirmative. More precisely, in Appendix VI we prove the following.

Proposition 2 *Given N yields as above, let the reversion speed matrix, $\mathcal{K}$, be diagonalizable as in*

$$\underline{\mathcal{K}} = \underline{a}\Lambda_{\mathcal{K}}\underline{a}^{-1} \tag{52}$$

$$\underline{\Lambda}_{\mathcal{K}} = diag\left[l_j\right] \tag{53}$$

with the eigenvalues $\{l_j\}$ distinct and real. Let $\underline{F}$ be the $[N \times N]$ matrix of elements $[F]_{ij}$ given by

$$F_{ij} \equiv \frac{1}{\tau_j}\frac{1-e^{-l_i\tau_j}}{l_i}. \tag{54}$$

Then for any τ_2, τ_3, .., τ_N and for any l_1, l_2, ..., l_N such that all the distinct and real eigenvalues are also positive (so as to ensure stability of the dynamical system)[15]*, there always exists a non-diagonal matrix, $\underline{\mathcal{K}} = \underline{\mathcal{K}}\left(\overrightarrow{l}\right)$, given by*

$$\underline{\mathcal{K}} = \underline{\Omega}^T\underline{F}^{-1}\underline{\Lambda_{\mathcal{K}}F\Omega} \tag{55}$$

such that

$$\underline{\beta\beta}^T = \underline{I} \tag{56}$$

We have therefore concluded that, for any reversion speed vector , $\overrightarrow{\theta}$, it is possible to find an N-tuple infinity of possible solutions (each indexed by the distinct eigenvalues, $\{l_j\}$, $j = 1, 2, ...N$) such that i) any exogenous discrete yield curve is perfectly recovered (condition (39)), and ii) any discrete exogenous covariance matrix is exactly recovered (condition (55)).

As we shall see, each choice for the eigenvalues $\{l_j\}$ gives rise to very different model behaviour. It also gives rise to different yield curves and covariance matrices for yields other than the N reference yields. We discuss in Section 00 some criteria to strongly bound the acceptable values for the eigenvalues $\{l_j\}$. These criteria will also give a precise meaning to the idea of 'behaviour complexity' which has so far been repeatedly, but hand-wavingly, mentioned.

4.3 Consequences of the $\mathbb{Q}$-measure Reversion-Speed Matrix

In the approach we present in this paper we require the observed yields to be rotated in such a way as to obtain orthogonal principal components. Once this 'rotation of axes' has been made, it permeates every aspect of the resulting mean-reverting dynamics. This is obvious enough, if one looks at Equations

[15] Saroka (2014) shows that the result can be generalized to the case when the eigenvalues are real, positive but not distinct. We do no pursue this angle here because, apart from numerical issues (arising from matrix inverrsion), the case can be approximated arbitrarily closely by having two or more eigenvalues becoming closer and closer. Saroka (2014) also deals with the case where the eigenvalues are imaginary, but with positive real part.

(32) and (36), that play a central role in determining the prices and yields of bonds, and in ensuring the orthogonality. But the rotation of axes imposed by the principal-component interpretation of the factors also affects, in a less obvious way, the admissible reversion-speed matrices, which, we recall, are given by

$$\underline{\mathcal{K}} = \underline{\Omega}^T \underline{F}^{-1} \underline{\Lambda}_{\mathcal{K}} \underline{F} \underline{\Omega} \tag{57}$$

This link between the reversion speed matrix and the particular rotation singled out by $\underline{\Omega}$ entails a rather complex mean-reverting deterministic dynamics. This can be see as follows.

Once the state variables have been chosen to be principal components, and they have been assigned a mean-reverting behaviour as in Equation (19), no-arbitrage only leaves as 'degrees of freedom' the N eigenvalues, $\left\{\overrightarrow{l}\right\}$, of the reversion speed-matrix, and, as we have seen, requires the $\underline{\mathcal{K}}$ matrix to be non-diagonal. Furthermore, the negative entries must be large enough to cancel completely the contributions coming from the positive-sign reversion speeds (see Equation (56)). So, the negative entries of the reversion-speed matrix (which can give rise to a locally mean-fleeing behaviour), are not a 'small correction', but must be of the same order of magnitude as the positive matrix elements.

It is this feature (and the locally-mean-fleeing behaviour between some state variables it implies) that causes the behaviour complexity:[16] these negative reversion speeds simultaneously generate attraction to and *significant* repulsion from, the various state variables and their fixed reversion levels. The overall system *is* , of course, asymptotically stable (because we have required all the eigenvalues l_k to be real and positive[17]), but, as shown in detail below, given a set of initial conditions, $\overrightarrow{x}_0$, for the state variables, the *deterministic* path to their equilibrium is forced to display an oscillatory behaviour, with over- and undershoots of their reversion levels. A similar oscillatory behaviour is inherited by the yields, which are linear combination of the factors.

In sum: we can assign and recover exactly an exogenous (discrete) covariance matrix and we can assign and recover exactly an exogenous (discrete) time-0 yield curve. However, once no-arbitrage is enforced, we can only imperfectly specify how the short rate will evolves from 'here to there'. For large eigenvalues of the reversion-speed matrix (which, we recall, can be arbitrarily assigned under the only constraint that they should be positive and distinct) there can be significant overshoots *and* undershoots, *even if all the N reference 'market' yields are exactly recovered* – the larger the trace of the $\underline{\mathcal{K}}$ matrix, the 'wilder' the over- and undershootings in between the 'nodes' of the market yields.

This can be seen more precisely as follows. Consider the first yield, which is just the short rate. The time-0 $\mathbb{Q}$-measure expectation over the paths of the

[16] For the moment we call a determinsitic behaviour 'complex' if, in the absence of stochastic shocks, the expectation of the state variables approaches the relative reversion levels with a non-monotonic behaviour. The larger the ampliutde of these oscillations is, and the more numerous the oscillations are, the more complex the resulting behaviour. See the results in Section 00.

[17] As far as stability is concerned, we could allow imaginary eigenvalues with a positive real part. We do not explore this route, for which we see little *a priori* justification.

short rate out to a given maturity, T, is straightforwardly related to the time-0 yield of a discount bond of that maturity:

$$y_0^T = -\frac{1}{T}\log E^{\mathbb{Q}}\left[e^{-\int_0^T r_s ds}\right] \tag{58}$$

Neglecting for the moment convexity effects (which anyhow plays a very small role for maturities out to 5 years), one can approximately write

$$y_0^T \simeq E^{\mathbb{Q}}\left[\frac{1}{T}\int_0^T r_s ds\right] \tag{59}$$

Therefore, simply by averaging out to a horizon T the values of the short rate along a deterministic path, one can immediately relate the path of the short rate to the current model yield curve. At the reference points, by construction, one will observe a match between the market and the model yields (again, within the limits of the approximation above); in between the reference points, however, every choice of eigenvalues will determine, by affecting the path of the state variables to their reversion levels, the values of the intermediate-maturity yields. The observed market yields therefore behave as fixed 'knots' through which the yield curve has to move: how smoothly it goes from point to point will depend on the eigenvalues of the reversion speed matrix. This is clearly shown in Fig (1):

The three lines in Fig (1) show the averages of the short rate out to five years for the eigenvalues of the reversion-speed matrix used in the case study (curve labelled "Base"); for eigenvalues twice as large (curve labelled "Base*2"); and for eigenvalues four times as large (curve labelled "Base*4"). In all cases, to within the accuracy of the approximation, the average of the short rate out to five years is indeed 2.00% (the exogenously assigned 'market' value of y_0^2); however, intermediate yields can assume values which strongly depend on the eigenvalues of the reversion-speed matrix. The larger these eigenvalues, the more 'complex' the behaviour in between the 'knots'.

This interpretation also makes more precise the concept of 'complexity' for the yield behaviour', to which we have frequently alluded to above: for instance, as in the case of splines, the integral of the non-convexity-induced curvature of the model yield curve between reference points can be taken as a measure of complexity.

Similar considerations apply to the interpolated covariance matrix. Also in this case, the choice of the eigenvalues $\left\{\overrightarrow{l}\right\}$ strongly influences the values of the 'interpolated' covariance elements. Indeed, some choices for the vector $\overrightarrow{l}$ can even produce negative correlations for yields in between the exactly recovered covariances.

Figure (2) shows the square root of the entries of the model (top panel) and market (bottom panel) covariance matrix for yields from 1 to 30 years obtained with the optimal choice of the eigenvalues $\left\{\overrightarrow{l}\right\}$.[18] The overall quality of the fit

[18] Unless otherwise stated, in all our calibration studies we used $N = 3$, ie, 3 yields, and 3

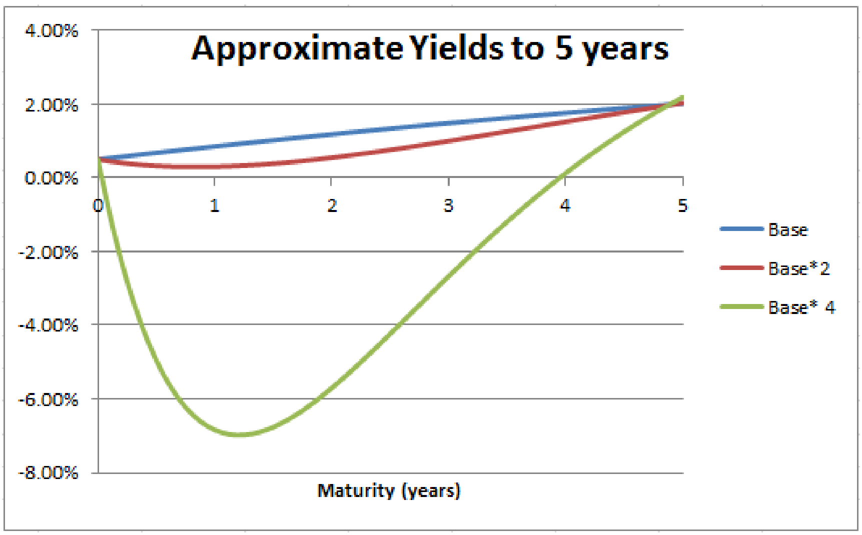

Figure 1: The averages of the short rate out to five years for the three cases discussed in the text, namely, for the eigenvalues of the reversion-speed matrix used in the case study (curve labelled "Base"); for eigenvalues twice as large (curve labelled "Base*2"); and for eigenvalues four times as large (curve labelled "Base*4"). In all cases, to within the accuracy of the approximation, the average of the short rate out to five years is indeed 2.00% (the exogenously assigned 'market' today value of the five-year yield).

for the inter- and extrapolated covariance matrix is excellent, with a maximum error of 6 basis points (in units equivalent to absolute volatility) and an average absolute error of 1.5 basis points (in the same units).

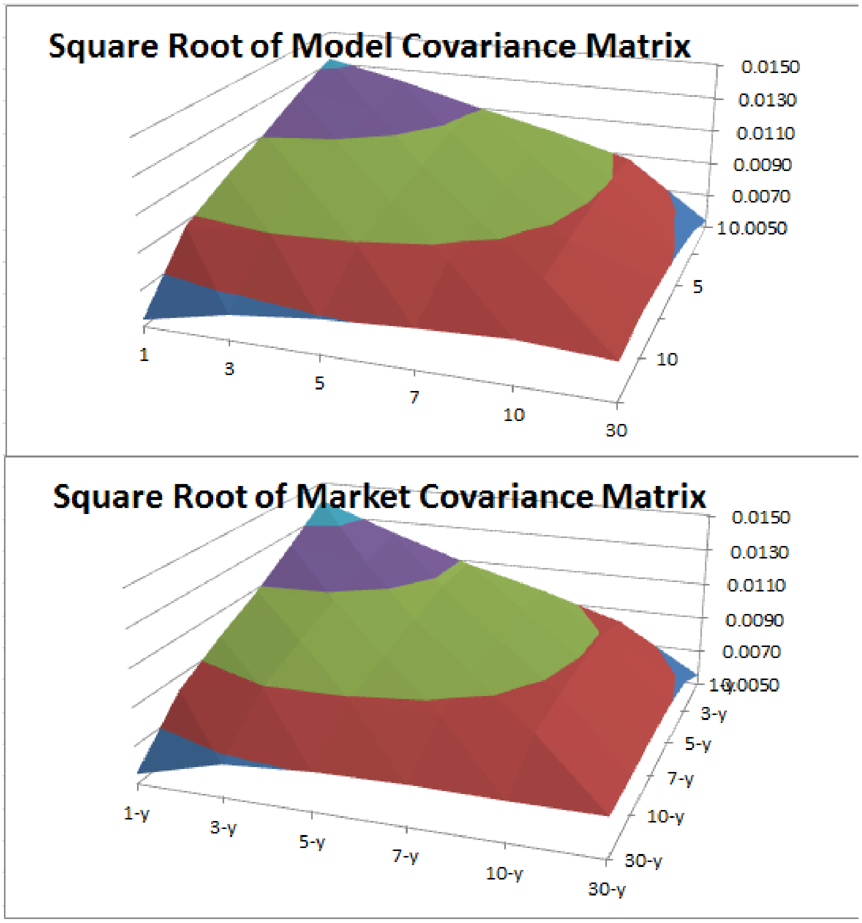

Figure 2: The square root of the entries of the model (top panel) and market (bottom panel) covariance matrix for yields from 1 to 30 years for the optimal choice of the eigenvalues $\left\{\overrightarrow{l}\right\}$. (Note that the intervals along the x and y axes are not equally spaced – units in years.)

One might think that, since the values of the covariance matrix in correspondence with the reference yields are exactly recovered by construction, the errors in interpolation (and possibly extrapolation) should be small. If this were true, little information about the eigenvalues $\left\{\overrightarrow{l}\right\}$ could be gleaned from the

principal components.

non-reference covariance elements. This is not the case, as displayed clearly in Fig (3), which shows that an injudicious (but, at first blush, reasonable) choice of eigenvalues $\left\{ \overrightarrow{l} \right\}$ can give extrapolated covariance elements wrong by a factor of 5 (even if the covariance elements among the reference yields are still exactly recovered)!

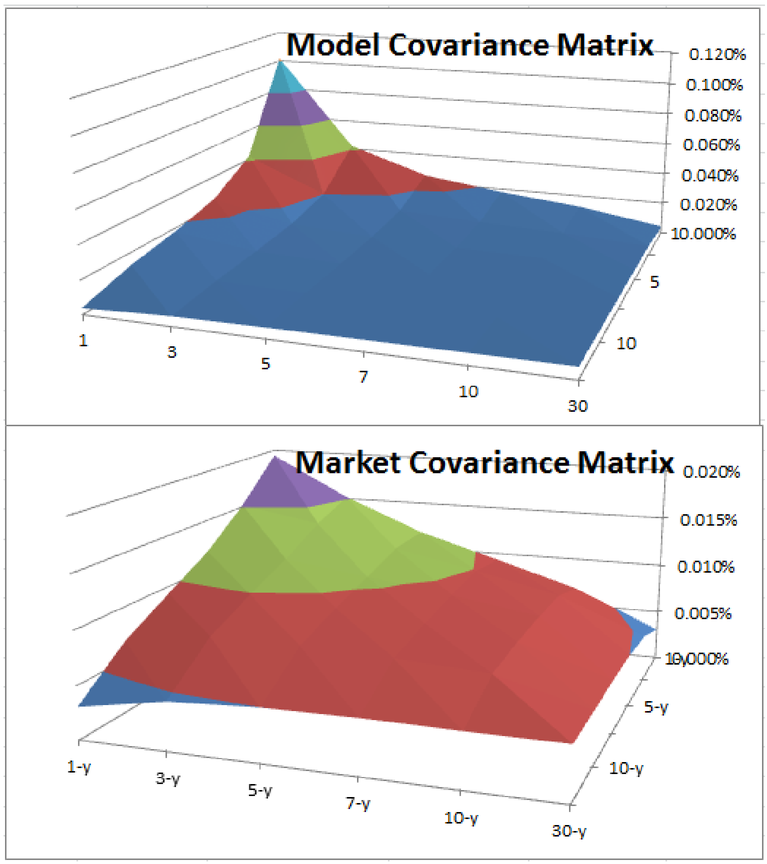

Figure 3: Same above for a poor choice of the eigenvalues $\left\{ \overrightarrow{l} \right\}$. Note that the covariance matrix elements between each reference yield are still exactly recovered.

This dependence of the 'intermediate' yields and of the 'intermediate' covariance matrix elements on the eigenvalues $\left\{ \overrightarrow{l} \right\}$ is not a drawback, but one of the most appealing features of the model. As we shall show in the calibration section this set of dependences will provide very useful guidance to determine the acceptable values of the eigenvalues $\left\{ \overrightarrow{l} \right\}$.

5 Moving from the $\mathbb{Q}$ to the $\mathbb{P}$ Measure

We want to show in this section how the model behaviour can be specified both in the $\mathbb{Q}$ (risk-neutral) and in the $\mathbb{P}$ (data-generating) measures.

Let's go back to the $\mathbb{Q}$-measure dynamics (19) – which we re-write for ease of reference:

$$d\overrightarrow{x}_t = \underline{\mathcal{K}}\left(\overrightarrow{\theta} - \overrightarrow{x}_t\right) dt + \underline{S}\overrightarrow{dz}^{\mathbb{Q}} \tag{60}$$

As discussed in Section 2, we are going to assign to the market price of risk, $\overrightarrow{T}_t$, one of the affine forms discussed in the literature, nested in the following general formulation:

$$\overrightarrow{T}_t = \overrightarrow{q}_0 + \underline{\mathcal{R}}\overrightarrow{x}_t \tag{61}$$

For instance, if we embraced the Duffee (2002) specification (according to which the magnitude of the market price of risk depends on the slope of the yield curve) we would have[19] for the matrix $\underline{\mathcal{R}}$

$$\underline{\mathcal{R}} = \begin{bmatrix} 0 & a & 0 \\ 0 & b & 0 \\ 0 & c & 0 \end{bmatrix} \tag{62}$$

In general, for *any* specification of the dependence of the market price of risk on the state variables, we have

$$d\overrightarrow{x}_t = \underline{\mathcal{K}}\left(\overrightarrow{\theta} - \overrightarrow{x}_t\right) dt + \underline{S}\overrightarrow{dz}^{\mathbb{P}} + \underline{S}\left(\overrightarrow{q}_0 + \underline{\mathcal{R}}_t\overrightarrow{x}_t\right) dt \tag{63}$$

This can be rewritten as

$$d\overrightarrow{x}_t = \underline{\mathcal{K}}^{\mathbb{P}}\left(\overrightarrow{\theta}^{\mathbb{P}} - \overrightarrow{x}_t\right) dt + \underline{S}\overrightarrow{dz}^{\mathbb{P}} \tag{64}$$

with

$$\underline{\mathcal{K}}^{\mathbb{P}} = \underline{\mathcal{K}} - \underline{S}\underline{\mathcal{R}} \tag{65}$$

and

$$\overrightarrow{\theta}^{\mathbb{P}} = \left(\underline{\mathcal{K}} - \underline{S}\underline{\mathcal{R}}_t\right)^{-1}\left(\underline{\mathcal{K}}\overrightarrow{\theta} + \underline{S}\overrightarrow{q}_0\right) \tag{66}$$

Equations (65) and (66) define the reversion-speed matrix and the reversion-level vector, respectively, as a function of the corresponding $\mathbb{Q}$-measure quantities, and of the market-price-of risk vector, $\overrightarrow{q}_0$, and matrix, $\underline{\mathcal{R}}$, respectively. We therefore show in the next section how we propose to estimate these quantities.

[19] This is not strictly correct. We find that the single regressor that most effectively explains excess returns is the second principal component from the orthopgonalization of the covariance matrix of yields (not yield differences). The second factor of our model is the second principal component from the orthogonalization of the covariance matrix of yeild *differences*. After the required transofrmation is applied, the matrix $\underline{\mathcal{R}}$ is similar to, but no longer exactly equal to, the simple single-column matrix displayed in Equation (00).

6 Estimating the Parameters of $\overrightarrow{q}_0$ and $\underline{\mathcal{R}}$

Using fifty years of data from the data base provided by the Federal Reserve Board of Washington, DC, (Gurkaynak et al, 2006), we have statistically estimated the excess returns from holding bonds up to 10 years. We have regressed these excess returns against our state variables, ie, the principal components.[20] If we call $\overrightarrow{rx}$ the vector of excess returns, the OLS estimation gives

$$\overrightarrow{rx_t} = \overrightarrow{a} + \underline{b}\,\overrightarrow{x}_t \tag{67}$$

At time t, the excess return vector is given by

$$\mathbb{E}\left[\overrightarrow{rx_t}\right] dt = \mathbb{E}\left[\overrightarrow{\frac{dP_t^T}{P_t^T} - r_t dt}\right] = \underline{Dur}\ \underline{S}\ \left[\overrightarrow{q}_0 + \underline{\mathcal{R}}\,\overrightarrow{x}_t\right] dt \tag{68}$$

where

$$[Dur]_{ij} = \frac{1}{P_t^{T_i}} \frac{\partial P_t^{T_i}}{\partial x_j} \tag{69}$$

Equating the coefficients of Equations (67) and (68) one gets

$$\overrightarrow{a} = \underline{Dur}\ \underline{S}\ \overrightarrow{q}_0 \tag{70}$$

and

$$\underline{b} = \underline{Dur}\ \underline{S}\ \underline{\mathcal{R}} \tag{71}$$

and therefore

$$\overrightarrow{q}_0 = \left(\underline{Dur}\ \underline{S}\right)^{-1} \overrightarrow{a} \tag{72}$$

$$\underline{\mathcal{R}} = \left(\underline{Dur}\ \underline{S}\right)^{-1} \underline{b} \tag{73}$$

Next, we note that

$$\frac{1}{P_t^{T_i}} \frac{\partial P_t^{T_i}}{\partial x_j} \simeq \frac{\partial \log P_t^{T_i}}{\partial x_j} \tag{74}$$

Recalling that

$$P_t^T = \exp^{A_t^T + \left(\overrightarrow{B}_t^T\right)^T \overrightarrow{x}_t} \tag{75}$$

we have

$$\underline{Dur} = \left(\underline{B}_t\right)^T \tag{76}$$

and therefore

$$\overrightarrow{q}_0 = \left(\left(\underline{B}_t\right)^T\ \underline{S}\right)^{-1} \overrightarrow{a} \tag{77}$$

$$\underline{\mathcal{R}} = \left(\left(\underline{B}_t\right)^T\ \underline{S}\right)^{-1} \underline{b} \tag{78}$$

[20] The method used is presented in a separate paper. The results we use here are independent of the specific findings.

Note that the 'duration' matrix ($\underline{B}_t^T$) clearly depends on the maturity of the yield under consideration; so does the matrix of regression coefficient, $\underline{b}$, and the vector of 'intercepts', $\overrightarrow{a}$. However, the market price of risk must be independent of the maturity of the yields. Therefore the maturity dependence in $\underline{b}$ and $\overrightarrow{a}$, on the one hand, and on the 'duration' matrix ($\underline{B}_t^T$) on the other must neatly cancel out. This means that, within the precision of the statistical estimate of the regressors, the market price of risk vector and matrix, $\overrightarrow{q}_0$ and $\underline{\mathcal{R}}$, must be independent of the N yields used in the regression. *This condition imposes a powerful internal consistency check on the model and on the statistical estimate of the coefficients in the excess return regression.*

The results derived so far complete the formal specification of the model. For a given set of exogenous market yields and covariance matrix, we have a $(2N+1)$-ple infinity of solutions (each exactly recovering the reference exogenous yield and covariance elements), parametrized by the vector $\overrightarrow{l}$, (the eigenvalues of the reversion-speed matrix), the vector $\overrightarrow{\theta}$, and the first element of the vector $\overrightarrow{\widetilde{y}}$. Each of these solutions gives rise to economically different behaviour for important quantities such as the market price of risk. The next section shows the criteria on the basis of which the number of degrees of freedom, or their acceptability range, can be reduced virtually to zero. We call this part of the project the 'calibration of the model'.

7 Calibration of the Model

7.1 Estimating the Values of the Eigenvalues $\overrightarrow{l}$

As we have shown above, the model automatically recovers the N exogenous market yields, and the $N \times N$ covariance matrix between the same yields. This does not mean, however, that that the yields or the covariance elements 'in between' the reference maturities will be similar to the corresponding market quantities. We therefore choose the eigenvalues of the $\mathbb{Q}$-measure reversion speed matrix in such a way that the covariance matrix and the yield curve in between the reference yield should be closely recovered.

7.2 Estimating the Values of the Reversion Levels $\overrightarrow{\theta}^{\mathbb{Q}}$

Next, we want to determine from statistical information the possible value for the reversion-level vector, $\overrightarrow{\theta}^{\mathbb{Q}}$. Our strategy is to estimate time averages of yields or principal components using a very-long-term historical record; to equate these quantities to the reversion levels, $\overrightarrow{\theta}^{\mathbb{P}}$, in the $\mathbb{P}$ measure; to translate this vector to the $\mathbb{Q}$ measure using Equation (66) (which 'contains' the vector $\overrightarrow{\theta}^{\mathbb{Q}}$). More precisely we proceed as follows.

We assume that, from the match to the covariance matrix the reversion-speed matrix, $\mathcal{K}^{\mathbb{Q}}$, has already been determined as described in the previous section. We also choose a 'trial' value for the $\mathbb{Q}$-measure reversion-level vector, $\overrightarrow{\theta}^{\mathbb{Q}}$.

Given these quantities, and after estimating the market price of risk vector, $\overrightarrow{T} = \overrightarrow{q}_0 + \underline{\mathcal{R}}_t \overrightarrow{x}_t$, we can evolve in the $\mathbb{P}$ measure the state variables or the yields from 'today' to any horizon, τ, using the evolution equations

$$d\overrightarrow{x}_t = \underline{\mathcal{K}}^{\mathbb{P}} \left(\overrightarrow{\theta}^{\mathbb{P}} - \overrightarrow{x}_t \right) dt + \underline{S} \overrightarrow{dz}^{\mathbb{P}} \tag{79}$$

$$d\overrightarrow{y}_t = \underline{\mathcal{K}}_y^{\mathbb{P}} \left(\overrightarrow{\theta}_y^{\mathbb{P}} - \overrightarrow{y}_t \right) dt + \underline{\Omega S} \overrightarrow{dz}^{\mathbb{P}} \tag{80}$$

with

$$\underline{\mathcal{K}}_y^{\mathbb{P}} = \underline{\Omega} \left(\underline{\mathcal{K}}^{\mathbb{Q}} - \underline{S\mathcal{R}} \right) \Omega^T \tag{81}$$

and

$$\overrightarrow{\theta}_y^{\mathbb{P}} = \underline{\Omega} \left(\underline{\mathcal{K}}^{\mathbb{Q}} - \underline{S\mathcal{R}} \right)^{-1} \overrightarrow{\widetilde{\theta}}$$

$$\overrightarrow{\widetilde{\theta}} = \overrightarrow{\theta}^{\mathbb{Q}} + \left(\underline{I} - \left(\underline{\mathcal{K}}^{\mathbb{Q}} \right)^{-1} \underline{S\mathcal{R}} \right) \underline{\Omega}^T \overrightarrow{\widetilde{y}} + \left(\underline{\mathcal{K}}^{\mathbb{Q}} \right)^{-1} \underline{S} \overrightarrow{q}_0 \tag{82}$$

This $\mathbb{P}$-measure projection can be carried out exactly for any horizon, τ. In particular, it can be carried out for $\tau = \infty$. For this projection horizon, the expectations of the yields will be equal to their reversion levels, $\overrightarrow{\theta}_y^{\mathbb{P}}$. Therefore we have

$$\overrightarrow{\theta}_y^{\mathbb{P}} = \Omega \left(\underline{\mathcal{K}}^{\mathbb{Q}} - S\mathcal{R} \right)^{-1} \overrightarrow{\widetilde{\theta}} = E^{\mathbb{P}} \left[\overrightarrow{y}_{\infty} \right] \tag{83}$$

Note that these model quantities, which have been obtained after a $\mathbb{P}$-measure evolution, are a function of the vector $\overrightarrow{\widetilde{y}}$, which, in turn, depends on the vector $\overrightarrow{\theta}^{\mathbb{Q}}$.

Separately from this, we can also calculate from our historical record the long-term time average of yields. We denote these measured quantities by $\langle \overrightarrow{y} \rangle$.

As a last step, we can equate the measured, $\langle \overrightarrow{y} \rangle$, and the model, $E^{\mathbb{P}} \left[\overrightarrow{y}_{\infty} \right]$, quantities

$$\langle \overrightarrow{y} \rangle = E^{\mathbb{P}} \left[\overrightarrow{y}_{\infty} \right] = \underline{\Omega} \left(\underline{\mathcal{K}}^{\mathbb{Q}} - S\mathcal{R} \right)^{-1} \overrightarrow{\widetilde{\theta}} =$$

$$\underline{\Omega} \left(\underline{\mathcal{K}}^{\mathbb{Q}} - S\mathcal{R} \right)^{-1} \left(\overrightarrow{\theta}^{\mathbb{Q}} + \left(\underline{I} - \left(\underline{\mathcal{K}}^{\mathbb{Q}} \right)^{-1} \underline{S\mathcal{R}} \right) \underline{\Omega}^T \overrightarrow{\widetilde{y}} + \left(\underline{\mathcal{K}}^{\mathbb{Q}} \right)^{-1} \underline{S} \overrightarrow{q}_0 \right) \tag{84}$$

We can therefore determine the quantity, $\overrightarrow{\theta}^{\mathbb{Q}}$, that best achieves this match between the model-implied and the empirically-observed long-term yields. (Note that the reversion-speed vector, $\overrightarrow{\theta}^{\mathbb{Q}}$, also enters the expression for $\overrightarrow{\widetilde{y}}$. The dependence of the unknown vector, $\overrightarrow{\theta}^{\mathbb{Q}}$, remains linear, and can therefore be worked out exactly by matrix inversion without any numerical search, because $\overrightarrow{\widetilde{y}}$ depends on $\overrightarrow{\theta}^{\mathbb{Q}}$ linearly.)

We note in passing that, as long as the market price of risk, $\overrightarrow{T}_t = \overrightarrow{q}_0 + \underline{\mathcal{R}}\overrightarrow{x}_t$, is estimated consistently from the same data used to estimate the long-term averages, $\langle \overrightarrow{y} \rangle$, one can rest assured that, by using these reversion levels, the $\mathbb{P}$-measure future evolution of the factors, $\overrightarrow{x}$, will give rise to consistent excess returns.

7.3 Bounding the Values of the Scalar $\widetilde{y}_1$

The expression obtained above for the reversion-speed vector, $\overrightarrow{\theta}^{\mathbb{Q}}$, is parametrized by the arbitrary value of the first element, $\widetilde{y}_1$, of the vector, $\overrightarrow{\widetilde{y}}$. More precisely, for any choice of $\widetilde{y}_1$, a different $\mathbb{Q}$-measure reversion level vector, $\overrightarrow{\theta}^{\mathbb{Q}}$, will ensue. This last indeterminacy cannot be resolved by looking at time averages of yields. However, the value of $\widetilde{y}_1$ will affect extrapolated yields to which the model has not been calibrated, such as the consol yield (when available), or a yield longer than the maximum maturity yield in the set of N yields $\{\overrightarrow{y}\}$.

With this last piece of information the model is fully calibrated.

8 Results

We show in this section the result obtained by calibrating the model using the procedure described above to Treasury data provided by the Fed (Gurkaynak et al, 2006) for the period 1990-2014. Qualitatively similar results were obtained using a longer data sets (extending back to the late 1970s. However we used the shorter time window of 24 years in order to avail ourselves of information about the 20 year yields, useful for assessing the quality of the extrapolation. We used 5 and 10 years for the intermediate and long yield maturities, respectively.

After using covariance information from the period 1990 to 2014, the eigenvalues of the $\mathbb{Q}$-measure reversion-speed matrix, $\mathcal{K}^{\mathbb{Q}}$, turned out to be

$$\begin{bmatrix} \text{Eigenvalues of } \mathcal{K}^{\mathbb{Q}} \\ l_1 = 0.03660 \\ l_2 = 0.63040 \\ l_3 = 0.63036 \end{bmatrix} \tag{85}$$

The reversion-speed matrix in the $\mathbb{Q}$ measure was as follows

$$\mathcal{K}^{\mathbb{Q}} = \begin{bmatrix} -0.1880 & 1.0187 & -3.0499 \\ -0.2229 & 0.7731 & -4.2299 \\ 0.0242 & -0.0304 & 0.7123 \end{bmatrix} \tag{86}$$

As anticipated, note the presence of large and negative reversion speeds, also on the main diagonal.

With this reversion-speed matrix the model covariance matrices, and the difference between the model and the empirical covariance matrices were calculated. The results are shown in Fig (4) below:

Model COV

0.000124	0.000118	0.000112	0.000106	0.000101	0.000096	0.000092	0.000088	0.000085	0.000082
0.000118	0.000136	0.000140	0.000136	0.000130	0.000124	0.000118	0.000112	0.000107	0.000102
0.000112	0.000140	0.000148	0.000147	0.000142	0.000136	0.000130	0.000124	0.000119	0.000115
0.000106	0.000136	0.000147	0.000148	0.000145	0.000141	0.000136	0.000131	0.000126	0.000122
0.000101	0.000130	0.000142	0.000145	**0.000144**	0.000141	0.000138	0.000134	0.000130	0.000127
0.000096	0.000124	0.000136	0.000141	0.000141	0.000140	0.000138	0.000135	0.000132	0.000130
0.000092	0.000118	0.000130	0.000136	0.000138	0.000138	0.000137	0.000135	0.000133	0.000131
0.000088	0.000112	0.000124	0.000131	0.000134	0.000135	0.000135	0.000134	0.000133	0.000132
0.000085	0.000107	0.000119	0.000126	0.000130	0.000132	0.000133	0.000133	0.000133	0.000132
0.000082	0.000102	0.000115	0.000122	0.000127	0.000130	0.000131	0.000132	0.000132	**0.000131**

Mkt COV

0.000124	0.000121	0.000113	0.000107	0.000101	0.000096	0.000091	0.000088	0.000085	0.000082
0.000121	0.000140	0.000141	0.000135	0.000128	0.000122	0.000116	0.000112	0.000108	0.000105
0.000113	0.000141	0.000148	0.000146	0.000141	0.000135	0.000129	0.000124	0.000120	0.000116
0.000107	0.000135	0.000146	0.000147	0.000145	0.000140	0.000136	0.000131	0.000127	0.000123
0.000101	0.000128	0.000141	0.000145	**0.000144**	0.000142	0.000138	0.000134	0.000131	0.000127
0.000096	0.000122	0.000135	0.000140	0.000142	0.000141	0.000138	0.000136	0.000133	0.000129
0.000091	0.000116	0.000129	0.000136	0.000138	0.000138	0.000137	0.000136	0.000133	0.000131
0.000088	0.000112	0.000124	0.000131	0.000134	0.000136	0.000136	0.000135	0.000133	0.000131
0.000085	0.000108	0.000120	0.000127	0.000131	0.000133	0.000133	0.000133	0.000133	0.000132
0.000082	0.000105	0.000116	0.000123	0.000127	0.000129	0.000131	0.000131	0.000132	**0.000131**

Error

0.000000	-0.000002	-0.000001	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
-0.000002	-0.000004	-0.000001	0.000001	0.000002	0.000002	0.000001	0.000000	-0.000001	-0.000003
-0.000001	-0.000001	0.000000	0.000001	0.000001	0.000001	0.000001	0.000000	-0.000001	-0.000002
0.000000	0.000001	0.000001	0.000001	0.000001	0.000000	0.000000	0.000000	0.000000	-0.000001
0.000000	0.000002	0.000001	0.000001	**0.000000**	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000002	0.000001	0.000000	0.000000	-0.000001	-0.000001	0.000000	0.000000	0.000000
0.000000	0.000001	0.000001	0.000000	0.000000	-0.000001	-0.000001	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	-0.000001	-0.000001	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	-0.000003	-0.000002	-0.000001	0.000000	0.000000	0.000000	0.000000	0.000000	**0.000000**

Figure 4: Model and market covariance matrices, and the error (market - model). The rows and columns correspond to maturities from 1 to 10 years.

The quality of the fit is evident. To assess the ability of the model to predict quantities it has not been calibrated to, the empirical and model yield volatilities out to 20 years are shown in Fig (5). We stress that the volatilities beyond 10 years are extrapolated by the model.Fig (6) shows the observed and fitted yield

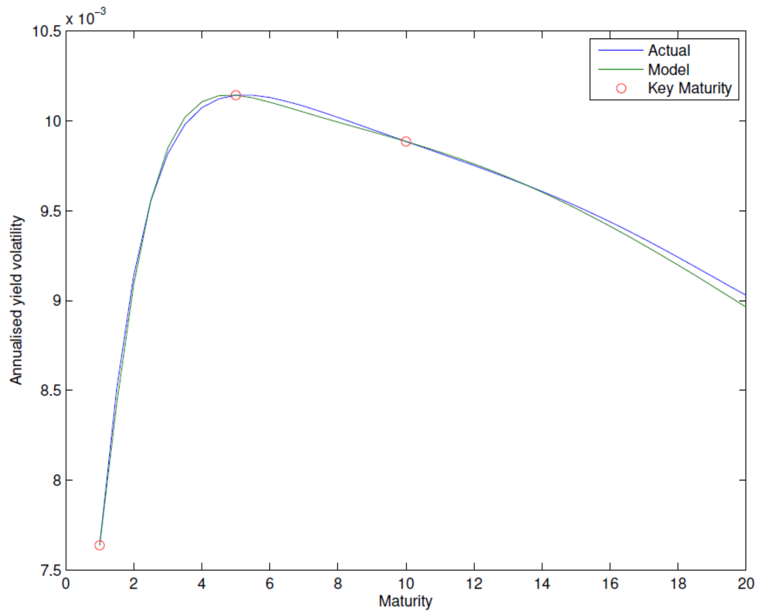

Figure 5: Model and empirical yield volatilities. Volatilities beyond 10 years are extrapolated.

curves on randomly selected dates between 1990 and 2014. The same model parameters were used for all the fits, and only the state variables were changed. They yields for the key maturities are, of course, perfectly recovered. It is interesting to note, however, that the intermediate and the extrapolated yields (ie, the yield beyond 10 years) are also well recovered.

After empirical estimation of the regression matrix of excess returns we were able to estimate the $\mathbb{P}$-measure reversion levels, and the trace of the matrix. See Equations (65) and (66). To show the robustness of the procedure, we present the estimates obtained for three different subsections of the data. Observe that all the parameters remain reasonably stable, with the possible exception of the reversion level of the first factor in the first half of the sample. This is probably due to the quasi-unit-root nature of the first principal component.

$$\begin{bmatrix} \text{Data sample} & \theta_1^{\mathbb{P}} & \theta_2^{\mathbb{P}} & \theta_3^{\mathbb{P}} & \text{Trace}\left(\mathcal{K}^{\mathbb{P}}\right) \\ \text{1990-2002} & 0.0958 & 0.0090 & -0.0073 & 3.78 \\ \text{2002-2014} & 0.0367 & 0.0115 & -0.0057 & 3.62 \\ \text{1990-2014} & 0.0380 & 0.0132 & -0.0055 & 3.11 \end{bmatrix} \tag{87}$$

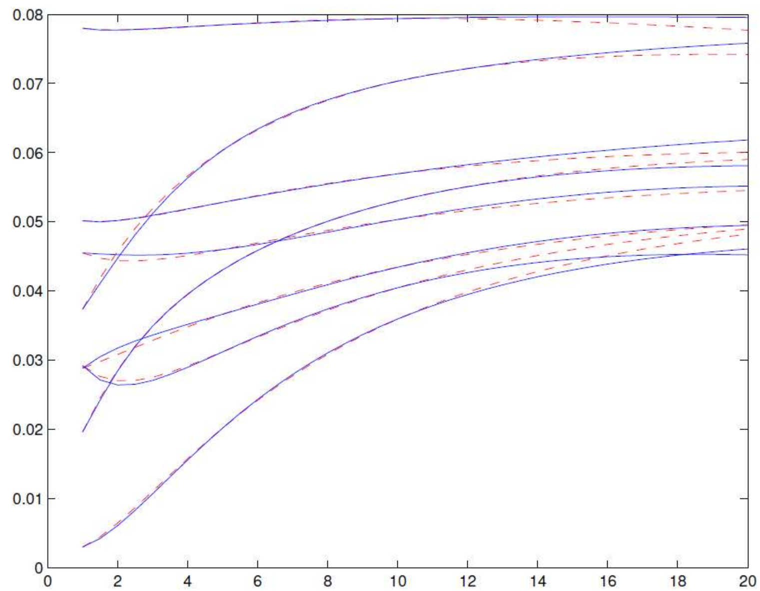

Figure 6: Fitted and empirical yiled curves for random dates between 1990 and 2014. Note that the same parameters are used throughout, and that yield beyond 10 years are extrapolated.

Once the term premia have been estimated, we can calculate the deterministic evolution of the reference yields under both measures from 'today's' yield curve (30-Jan-2014). This is shown in Fig (**??**).

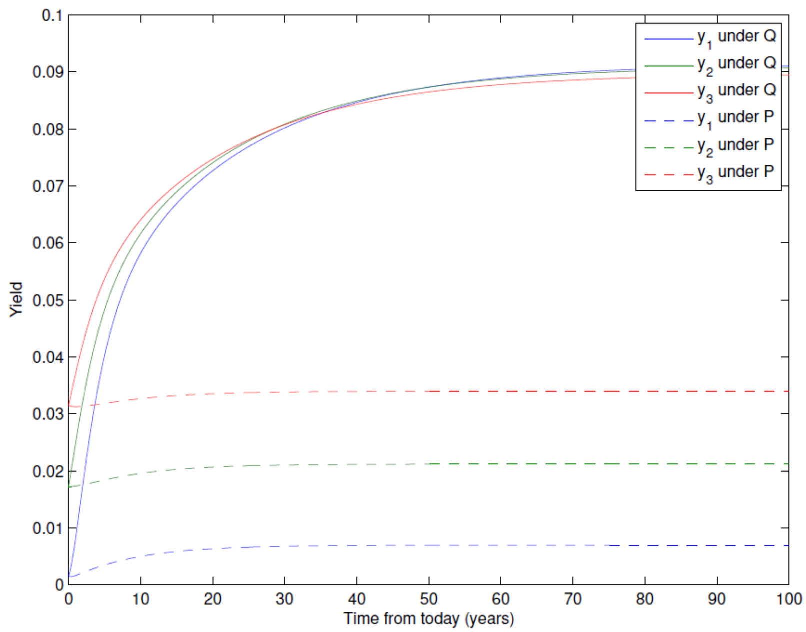

Figure 7: The deterministic evolution of the reference yields under $\mathbb{Q}$ (solid lines) and under $\mathbb{P}$ (dashed lines). The yield curve is flat if the three lines (which correspond to three different maturities) are superimposed.

Note that under the risk-neutral measure the asymptotic yield curve is almost exactly flat. This should indeed be the case: apart from convexity effects (whose magnitude is estimated below), all factors will be at their reversion levels, and in the $\mathbb{Q}$ measure no term premia can steepen the yield curve. As a consequence, apart from convexity terms investigated below, under $\mathbb{Q}$ the curve will evolve deterministically to a flat shape.

The magnitude of the convexity term is shown in Fig (8), which shows yields and value of convexity for different eigenvalues of the mean-reversion speed matrix $\mathcal{K}$. Solid lines represent the yield curves, and dashed lines represent the yield curves without the convexity term. The colours correspond to different eigenvalues of $\mathcal{K}$: blue - $(0.02; 0.2; 0.5)$, green - $(0.03; 0.3; 0.6)$, red - $(0.04; 0.4; 0.7)$.

We see that it is possible to obtain very similar fits on the traded portion of the yield curve with different eigenvalues of $\mathcal{K}$, but the consequences for extrapolation are very different. We emphasize, however that, despite the similar quality of the fit for the yields, *the fit to the covariance matrix produced by the three sets of eigenvalues was very different*. This stresses again the importance of making use of the full covariance matrix information in the fitting phase.

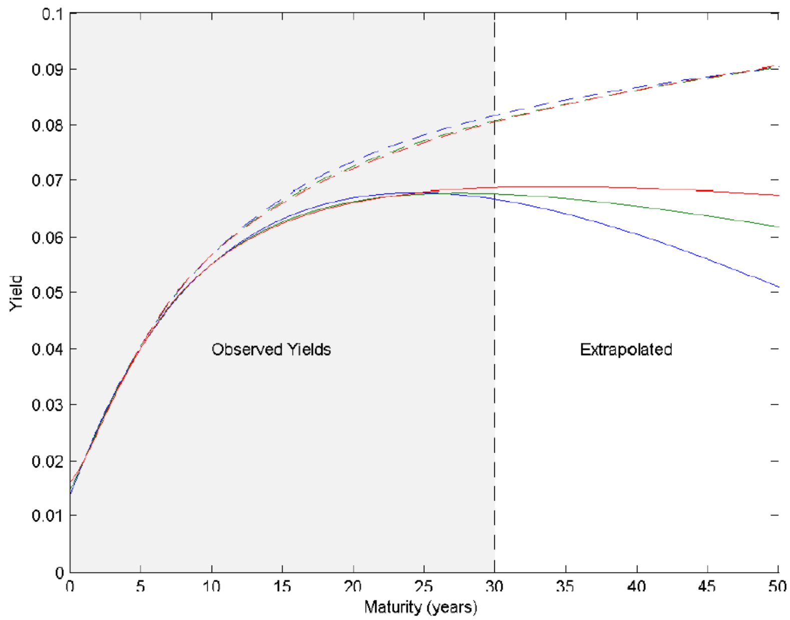

Figure 8: Yields and value of convexity for different eigenvalues of the mean-reversion speed matrix $\mathcal{K}$. Solid lines represent the yield curves, and dashed lines represent the yield curves without the convexity term. The colours correspond to different eigenvalues of $\mathcal{K}$: blue - (0:02; 0:2; 0:5), green - (0:03; 0:3; 0:6), red - (0:04; 0:4; 0:7).

Finally we show in Fig (9) the time series of the 10-year yield observed in the market ('yield under $\mathbb{Q}$'), and the yield that would have been observed if investors had the same expectations, but were risk-neutral ('yield under $\mathbb{P}$'). In the same figure we also show the term premium (red line) which is the difference between the two yields. The average risk premium for the last 24 years averages around 3%, which compares well with our empirical estimates of unconditional excess returns for the 10 year maturity, shown in Fig (**??**).

Finally, we show in Figs (10) and (11) the model and empirical asymptotic autocorrelation for the three principal components. Empirically, it is well-known that the first principal component should be far more persistent than the second and the third (see, eg, the discussion in Diebold and Rudebush (2013).) This is not borne out by our model, which also gets the overall speed of decrease of the autocorrelation significantly wrong.

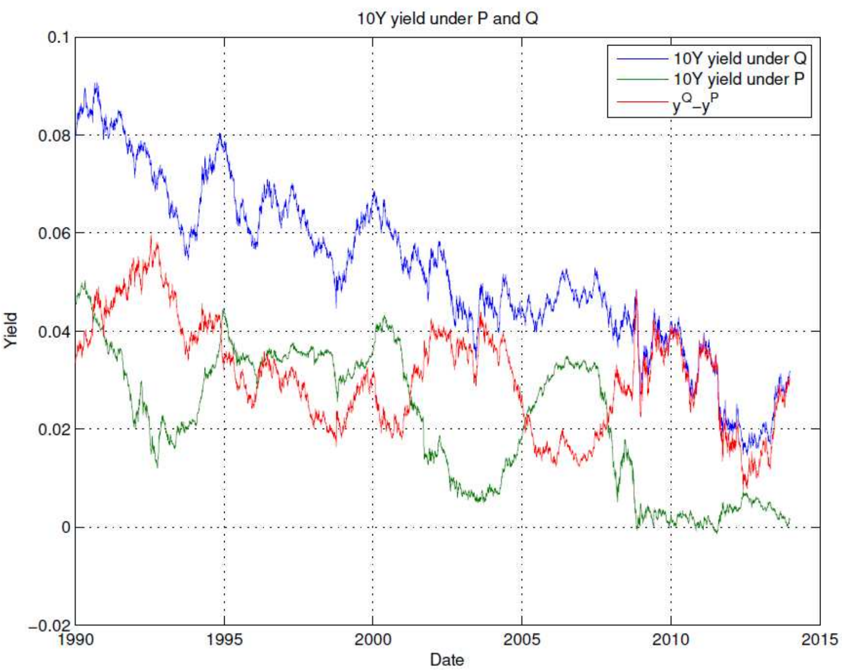

Figure 9: Observed 10-year zero coupon yield (blue), the '$\mathbb{P}$'10-year zero-coupon yield (green) and the term premium (red).

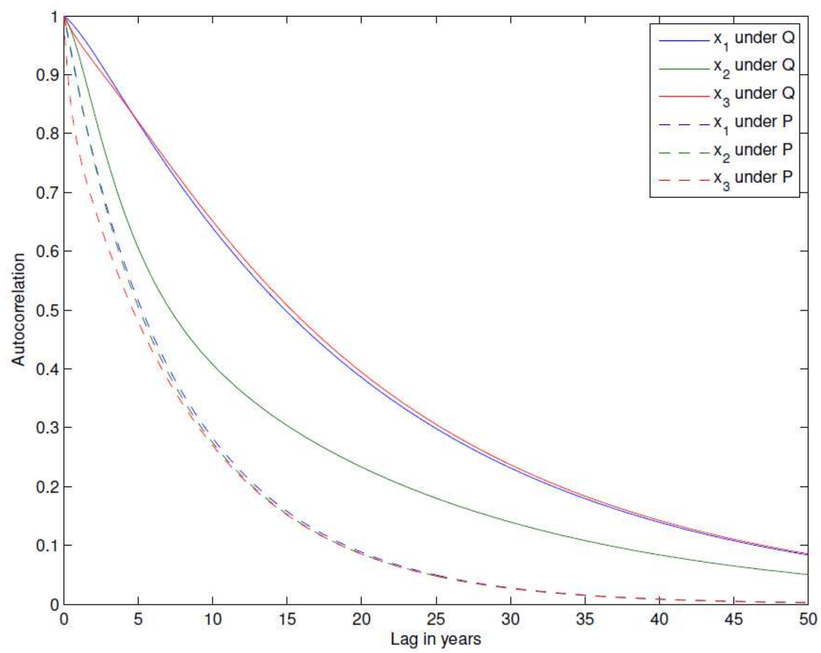

Figure 10: Asymptotic autocorrelation of principal components under $\mathbb{Q}$ (solid lines) and under $\mathbb{P}$ (dashed lines).

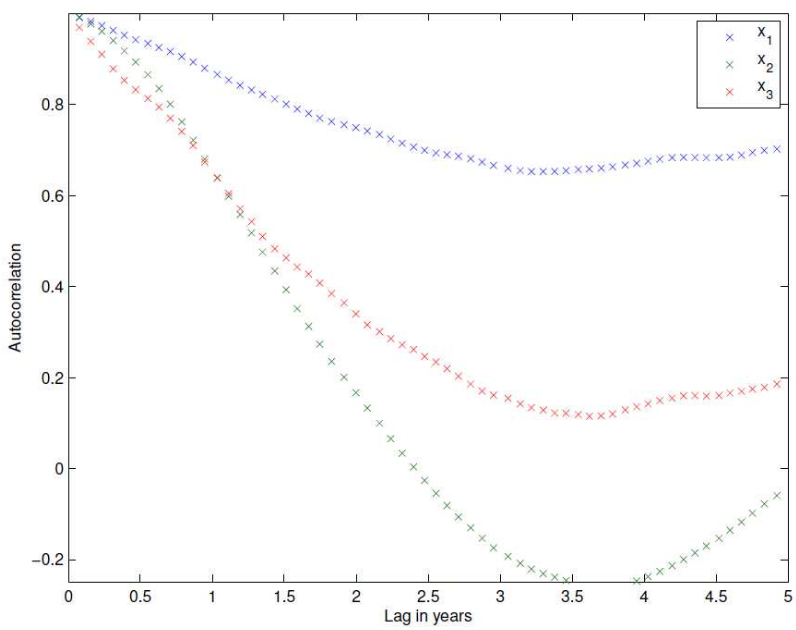

Figure 11: Estimated empirical asymptotic autocorrelation of principal components. Observe that empirically the first PC features a much higher autocorrelation than what the model produces. The model is not able to capture this feature.

9 Conclusions

We have presented the theoretical results for the affine evolution of mean-reverting principal component models, we have shown how the model can be effectively calibrated using both $\mathbb{Q}$- and $\mathbb{P}$-measure information.

We have stressed that the no-arbitrage constraints imposed by using a pre-specified model, and the choice of principal components as the mapping between the model state variables and the observable yields affect deeply the model evolution. In particular, we have shown that the reversion speed-matrix must contain negative and 'large' entries. This contributes strongly to the model's dynamic richness and 'complexity'.

Once the calibration has been carried out, we have presented the model behaviour.

With the exception of the degree of persistence of the principal components, the behaviour of the model seems to be robust, believable and interpretable.

The cross-measure results of course depend strongly on the estimation of the state dependence of the market price of risk (the excess returns). For illustrative purposes, we have carried out in this study a rather simple-minded, regression-based estimate. There is plenty of room for more careful analysis here (and the Cochrane tent-shaped return-predicting factor could be accommodated, if one so wished, in the modelling framework). We are planning to present the results of our empirical study of excess returns for Treasuries as a separate piece of work.

Acknowledgement 3 *It is a pleasure to acknowledge useful suggestions provided by Dr Naik Vasant and Dr Andrei Lyashenko.*

References

Ahn D, Dittmar R F, Gallant A R, (2002), *Quadratic Term Structure Models: Theory and Evidence*, Review of Financial Studies, 15, 243-288

Andersen T G, Lund J, (1997), *Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate*, Journal of Econometrics, 77, 343-377

Ang A, Piazzesi M, (2003), *A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables*, Journal of Monetary Economics, vol 50, 745-87

Black F, Derman E, Toy W, (1990), *A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options*, Financial Analyst Journal, 46, 33-39

Chan K C, Karoly A, Logstaff F A, Sanders A B, (1992), *An Empirical Comparison of Alternative Models of the Short-Term Interest Rate*, Journal of Finance, 47, 1209-1227

Cheridito P, Filipovic D, Kimmel R L, (2010), *A Note on the Dai-Singleton Canonical Representation of Affine Term Structure Models*, Mathematical Finance, Vol 2, No 3, (July), 509-519

Christensen J H E, Diebold F X, Rudebush G D, (2011), *The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models*, Journal of Econometrics, 164, 4-20

Cochrane J H, Piazzesi M, (2005), *Bond Risk Premia*, American Economic Review, Vol 95, no 1, 138-160

Cochrane J H, Piazzesi M, (2008), *Decomposing the Yield Curve*, University of Chicago and NBR working paper, March 2008

Cox J, Ingersoll J, Ross S, (1985a), *An Intertemporal General Equilibrium Model of Asset Prices, Econometrica*, 53, 363-384

Cox J, Ingersoll J, Ross S, (1985b), *A Theory of the Term Structure of Interest Rates, Econometrica*, 53, 385-408

D'Amico S, Kim D H, Min Wei, (2010), *Tips from TIPS: The Informational Content of Treasury Inflation Protected Security Prices*, working paper, Finance and Economics Discussion Series, Federal Reserve Board, Washington, DC

Dai Q, Singleton K J, (2000), *Specification Analysis of Affine Term Structure Models*, Journal of Finance, Vol LV, No 4, 1943-1978

Diebold, F X, Rudebush G D, (2013), *Yield Curve Modelling and Forecasting – The Dynamic Nelson-Siegel Approach*, Princeton University Press, Princeton, NJ, and Oxford, UK

Duffee, G R, (2002), *Term Premia and Interest Rate Forecasts in Affine Models*, Journal of Finance, Vol LVII, No 1, 405-443

Fama E, Bliss R, (1987), *The Information in Long-Maturity Forward Rates,* American Economic Review, 77, 680-692

Fama E, French K, (1989), *Business Conditions and Expected Returns on Stocks and Bonds*, Journal of Financial Economics, 25, 23-49

Fama E, French K, (1993), *Common Risk Factors in the Returns on Stocks and Bonds*, Journal of Financial Economics, 33, 3-56

Gurkaynak R S, Sack B, Wright J H, (2066), *The US Treasury Yield Curve: 1961 to the Present*, working paper no 2006-28, Federal Reserve Board, Washington, DC, Division of Research & Statistics and Monetary Affairs

Heath D, Jarrow R A, Morton A, (1987), *Bond Pricing and the Term Structure of Interest Rates: A New Methodlogy*, Working paper, Cornell University

Hull, J, White, A, (1990), *Pricing Interest Rate Derivatives Securities*, The Review of Financial Studies, 3, 573-592

Hellerstein R, (2011), *Global Bond Risk Premiums*, Federal Reserve Bank of New York Staff Report no 499, June 2011

Joslin, S, Anh Le, Singleton, K J, (2013), *Why Gaussian Macro-Finance Models Are (Nearly) Unconstrained Factor-VARs*, Journal of Financial Economics,

Joslin, S, Priebsch, M, Singleton, K J, (2014), *Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks*, Journal of Finance, forthcoming and working paper, Stanford University, Graduate School of Business

Joslin, S, Singleton, K J, Zhu H, (2011), *A New Perspective on Gaussian Dynamic Term Structure Models*, The Review of Financial Studies, v 24, 926-970

Litterman, R, Scheinkman, J, (1991), *Common Factors Affecting Bond Returns*, Journal of Fixed Income, 1, 54-61

Kim D H, (2007), *Challenges in Macro-Finance Modelling*, BIS working paper no 240, December 2007

Mayer J, Khairy K, Howard J, (2010), *Drawing an Elephant with Four Complex Parameters*, American Journal of Physics, 78, (6), June, 648-649

Nawalka S K, Rebonato R, (2011), *What Interest Models To Use? Buy Side Versus Sell Side, Journal of Investment Management*, Vol 9, No 3, 5-18

Nelson C R, Siegel A F, (1987), *Parsimonious Modeling of Yield Curves*, Journal of Business, 60, 473-489

Piazzesi, M, (2010), *Affine Term Structure Models*, in *Handbook of Financial Econometrics*, Yacin Ait-Sahalia, editor, vol 1, *Tools and Techniques*, North Holland

Saroka I, (2014), *Affine Principal-Component-Based Term Structure Models*, Oxford University, April 2014, MSc thesis in Mathematical Finance.

Ungari S, Turc J, (2012), *Macro-Financial Model*, working paper SocGen Cross Asset Quant Research, 19 October 2012

Vasicek O, (1977), *An Equilibrium Characterization of the Term Structure*, Journal of Financial Economics, 5, 177-188

10 Appendix A – Exponential of Orthogonal or Inverse Matrices

Let U be an $[n \times n]$ orthogonal matrix. Let A be an arbitrary $[n \times n]$ matrix. Consider the expression

$$c = Ue^A U^T \tag{88}$$

Expand the exponent to obtain

$$c = U\left[I + A + \frac{1}{2}A^2 + ...\right]U^T \tag{89}$$

Expand the epxression

$$c = UIU^T + UAU^T + \frac{1}{2}UA^2U^T + ... = \tag{90}$$

$$I + UAU^T + \frac{1}{2}UA^2U^T + ... \tag{91}$$

where the last line follows because of the orthogonality of U. Consider now the exponential $d = e^{UAU^T}$:

$$d = e^{UAU^T} = I + UAU^T + \frac{1}{2}\left(UAU^T\right)^2 + ... =$$

$$I + UAU^T + \frac{1}{2}UAU^TUAU^T + ... =$$

$$I + UAU^T + \frac{1}{2}UAAU^T + ...$$

$$I + UAU^T + \frac{1}{2}UA^2U^T + ... \tag{92}$$

Repeating for higher-order terms and comparing Equations (92) and (91), we can conclude

$$Ue^AU^T = e^{UAU^T} \tag{93}$$

The same proof applies to non-orthogonal matrices, as long as matrix inversion replaces transposition:

$$Ze^AZ^{-1} = e^{ZAZ^{-1}} \tag{94}$$

11 Appendix II – Solving the ODE for $B(\tau)$

Let the solution for the bond price in a multidimensional OU process be given by

$$P_t^\tau = e^{A(\tau) + \vec{B}(\tau)^T \vec{x}_t} \tag{95}$$

with

$$B_{ik}(\tau) = \frac{\partial P_t^{\tau_k}}{\partial x_i} \tag{96}$$

(Note the positive sign of the exponent and the transpose.) Given the process

$$d\overrightarrow{x}_t = \underline{\mathcal{K}}\left(\overrightarrow{\theta} - \overrightarrow{x}\right)dt + \underline{S}d\overrightarrow{z}$$

the ODE for $B(\tau)$ will be:

$$\begin{aligned} \frac{d\overrightarrow{B}(\tau)}{d\tau} &= -\underline{\kappa}\overrightarrow{B}(\tau) - \Omega^T\overrightarrow{e}_1 && (97) \\ \overrightarrow{B}(0) &= \overrightarrow{0} && (98) \end{aligned}$$

with

$$\underline{\kappa}^T = \underline{\mathcal{K}}$$

This is a inhomogeneous ODE. We solve it by i) finding the solution to the homogeneous ode; ii) finding a particular solution; iii) joining the two and iv) satisfying the boundary condition.

An aside: We have the *transpose* of $\underline{\mathcal{K}}$ in Equation (97) because the expression for P_t^τ is expressed as a function of $\overrightarrow{B}(\tau)^T$. Therefore we really have

$$\frac{d\overrightarrow{B}^T(\tau)}{d\tau} = -\underline{\mathcal{K}}\overrightarrow{B}^T(\tau) - \overrightarrow{e}_1^T\Omega \tag{99}$$

and therefore

$$\begin{aligned} \frac{d\overrightarrow{B}^T(\tau)}{d\tau} &= \left(-\overrightarrow{B}^T\underline{\mathcal{K}}(\tau) - \overrightarrow{e}_1^T\Omega\right)^T = \\ \frac{d\overrightarrow{B}(\tau)}{d\tau} &= -\underline{\kappa}\overrightarrow{B}(\tau) - \Omega^T\overrightarrow{e}_1 && (100) \end{aligned}$$

11.1 The Homogeneous ODE

The homogeneous ODE has the form

$$\frac{d\overrightarrow{B}_{\text{hom}}(\tau)}{d\tau} + \kappa\overrightarrow{B}_{\text{hom}}(\tau) = \overrightarrow{0} \tag{101}$$

By inspection, an obvious candidate solution is

$$\overrightarrow{B}_{\text{hom}}(\tau) = e^{-\kappa\tau}\overrightarrow{H} \tag{102}$$

Indeed

$$\frac{d\overrightarrow{B}_{\text{hom}}(\tau)}{d\tau} = -\underline{\kappa}e^{-\kappa\tau}\overrightarrow{H} \tag{103}$$

and therefore

$$\frac{d\overrightarrow{B}_{\text{hom}}(\tau)}{d\tau} + \underline{\kappa}\overrightarrow{B}_{\text{hom}}(\tau) = -\underline{\kappa}e^{-\kappa\tau}\overrightarrow{H} + \underline{\kappa}e^{-\kappa\tau}\overrightarrow{H} = 0 \tag{104}$$

11.2 The Particular Solution

We now have to find any solution to the inhomogeneous ODE. Let's try

$$\overrightarrow{B}(\tau) = \overrightarrow{C} \tag{105}$$

for some constant vector, $\overrightarrow{C}$. Then we have

$$\frac{d\overrightarrow{B}(\tau)}{d\tau} = \frac{d\overrightarrow{C}}{d\tau} = \overrightarrow{0} \tag{106}$$

and therefore

$$\overrightarrow{0} = -\kappa\overrightarrow{C} - \Omega^T\overrightarrow{e}_1 \tag{107}$$

$$\overrightarrow{C} = -\kappa^{-1}\Omega^T\overrightarrow{e}_1 \tag{108}$$

11.3 The Full Solution and the Initial Condition

We have found

$$\overrightarrow{B}(\tau) = e^{-\kappa\tau}\overrightarrow{H} + \overrightarrow{C} = e^{-\kappa\tau}\overrightarrow{H} - \kappa^{-1}\Omega^T\overrightarrow{e}_1 \tag{109}$$

The initial condition, $\overrightarrow{B}(0) = \overrightarrow{0}$, imposes that

$$\overrightarrow{B}(0) = e^{-\kappa 0}\overrightarrow{H} - \kappa^{-1}\Omega^T\overrightarrow{e}_1 = \overrightarrow{0} \tag{110}$$

and therefore

$$\overrightarrow{H} = \kappa^{-1}\Omega^T\overrightarrow{e}_1 \tag{111}$$

Therefore we have

$$\begin{aligned}\overrightarrow{B}(\tau) = e^{-\kappa\tau}\kappa^{-1}\Omega^T\overrightarrow{e}_1 - \kappa^{-1}\Omega^T\overrightarrow{e}_1 = \\ \left(e^{-\kappa\tau} - I\right)\kappa^{-1}\Omega^T\overrightarrow{e}_1\end{aligned} \tag{112}$$

11.4 The Integral Expression

Consider the integral

$$Int = \int_0^\tau e^{-\kappa(\tau-s)}ds = \int_0^\tau e^{-\kappa(\tau-s)}ds \tag{113}$$

This has solution

$$-\left[e^{-\kappa(\tau-s)}\right]_0^\tau \kappa^{-1} = \left[e^{-\kappa(\tau-s)}\right]_\tau^0 \kappa^{-1} = \left(e^{-\kappa\tau} - I\right)\kappa^{-1}\Omega^T\overrightarrow{e}_1 \tag{114}$$

Therefore the solution (112) can be equivalently expressed as

$$\overrightarrow{B}(\tau) = -\int_0^\tau e^{-\kappa(\tau-s)}\Omega^T\overrightarrow{e}_1 ds \tag{115}$$

11.5 The Expression for $\overrightarrow{B}^T$

The transpose of Equation (112) becomes

$$\overrightarrow{B}^T = \left[\left(e^{-\kappa\tau} - I\right)\kappa^{-1}\Omega^T\overrightarrow{e}_1\right]^T =$$

$$\overrightarrow{e}_1^T\Omega\left(\kappa^{-1}\right)^T\left[\left(e^{-\kappa\tau}\right)^T - I\right] \tag{116}$$

The transpose of Equation (118) becomes

$$\overrightarrow{B}^T = -\left[\int_0^\tau e^{-\kappa(\tau-s)}\Omega^T\overrightarrow{e}_1 ds\right]^T = -\int_0^\tau \left[e^{-\kappa(\tau-s)}\Omega^T\overrightarrow{e}_1\right]^T ds =$$

$$-\int_0^\tau \overrightarrow{e}_1^T\Omega\left[e^{-\kappa(\tau-s)}\right]^T ds \tag{117}$$

12 Appendix III – Obtaining an Expression for $B(\tau)$

We obtained that

$$\overrightarrow{B}(\tau) = -\int_0^\tau e^{-\underline{\kappa}(\tau-s)}\Omega^T\overrightarrow{e}_1 ds \tag{118}$$

Now, orthogonalize the matrix $\underline{\kappa}$:

$$\underline{\kappa} = \underline{a}\Lambda_\kappa\underline{a}^{-1} \tag{119}$$

This gives

$$\begin{aligned}
\overrightarrow{B}(\tau) &= -\int_0^\tau e^{-\underline{\kappa}(\tau-s)}\Omega^T\overrightarrow{e}_1 ds = \\
&= -\int_0^\tau e^{-\underline{a}\Lambda_\kappa\underline{a}^{-1}(\tau-s)}\Omega^T\overrightarrow{e}_1 ds = \\
&= -\underline{a}\left[\int_0^\tau e^{-\underline{\Lambda}_\kappa(\tau-s)}ds\right]\underline{a}^{-1}\Omega^T\overrightarrow{e}_1 = \\
&\equiv \underline{a}\,\overrightarrow{\xi}
\end{aligned} \tag{120}$$

with

$$\overrightarrow{\xi} \equiv \left[\int_0^\tau e^{-\underline{\Lambda}_\kappa(\tau-s)}ds\right]\underline{a}^{-1}\Omega^T\overrightarrow{e}_1 \tag{121}$$

Now,

$$\int_0^\tau e^{-\underline{\Lambda}_\kappa(\tau-s)}ds =$$

$$= \begin{bmatrix} \frac{1-e^{-l_1\tau}}{l_1} & 0 & 0 & ... & & 0 \\ 0 & \frac{1-e^{-l_2\tau}}{l_2} & 0 & ... & ... & 0 \\ 0 & 0 & \frac{1-e^{-l_3\tau}}{l_3} & ... & ... & 0 \\ 0 & 0 & & ... & & 0 \\ 0 & 0 & & & \frac{1-e^{-l_{N-1}\tau}}{l_{N-1}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1-e^{-l_N\tau}}{l_N} \end{bmatrix} \equiv$$

$$= \underline{D}(\tau) = \text{diag}\left[\frac{1-e^{-l_{ii}\tau}}{l_{ii}}\right], \quad i = 1, N \tag{122}$$

and, finally,

$$\overrightarrow{B}(\tau) = \underline{a}\,\overrightarrow{\xi} = \underline{aD}(\tau)\,\underline{a}^{-1}\Omega^T\overrightarrow{e}_1 \tag{123}$$

13 Appendix IV – Obtaining an Expression for $A(\tau)$

We obtained that

$$\overrightarrow{B}(\tau) = \underline{a}\,\overrightarrow{\xi} = \underline{aD}(\tau)\,\underline{a}^{-1}\Omega^T\overrightarrow{e}_1$$

with $\underline{\kappa}$:

$$\underline{\kappa} = \underline{a}\Lambda_\kappa\underline{a}^{-1} \tag{124}$$

and

$$\overrightarrow{\xi} \equiv \underline{D}(\tau)\,\underline{a}^{-1}\Omega^T\overrightarrow{e}_1 \tag{125}$$

$$\underline{D}(\tau) = \text{diag}\left[\frac{1-e^{-l_{ii}\tau}}{l_{ii}}\right], \quad i = 1, N \tag{126}$$

We have to solve

$$\frac{dA}{d\tau} = -\widetilde{y}_1 + \overrightarrow{B}^T\underline{K}\overrightarrow{\theta} + \frac{1}{2}\overrightarrow{B}^T SS^T\overrightarrow{B} \tag{127}$$

Therefore

$$A(\tau) = Int_1 + Int_2 + Int_3 \tag{128}$$

with

$$Int_1 = -\int_0^\tau \widetilde{y}_1 ds = -\widetilde{y}_1\tau \tag{129}$$

$$Int_2 = \left[\int_0^\tau \overrightarrow{B}^T(\tau)\,ds\right]\underline{K}\,\overrightarrow{\theta} \tag{130}$$

$$Int_3 = \frac{1}{2}\int_0^\tau \overrightarrow{B}^T(\tau)\,SS^T\overrightarrow{B}(\tau)\,ds \tag{131}$$

13.1 Evaluation of Int_3

We have

$$Int_3 = \int_0^\tau \overrightarrow{B}^T(\tau)\, SS^T \overrightarrow{B}(\tau)\, ds =$$

$$\int_0^\tau \left(\underline{a}\,\overrightarrow{\xi}\right)^T SS^T \left(\underline{a}\,\overrightarrow{\xi}\right) ds = \int_0^\tau \overrightarrow{\xi}^T \underline{a}^T \underline{SS}^T \underline{a}\,\overrightarrow{\xi}\, ds =$$

$$\int_0^\tau \overrightarrow{\xi}^T \underline{Q}\,\overrightarrow{\xi}\, ds \tag{132}$$

with

$$\underline{Q} = \underline{a}^T \underline{SS}^T \underline{a} \tag{133}$$

To lighten notation define

$$\overrightarrow{f} = \underline{a}^{-1}\Omega^T \overrightarrow{e}_1 \tag{134}$$

then

$$\overrightarrow{\xi} = -\underline{D}\,\overrightarrow{f} \tag{135}$$

and therefore

$$Int_3 = \int_0^\tau \overrightarrow{\xi}^T \underline{Q}\,\overrightarrow{\xi}\, ds =$$

$$\overrightarrow{f}^T \left[\int_0^\tau D^T \underline{Q}\underline{D} ds\right] \overrightarrow{f} \tag{136}$$

14 Appendix V – Derivation of the Variance of an N-d O-U Process

Let

$$d\overrightarrow{x}_t = \underline{\mathcal{K}}\left(\overrightarrow{\theta} - \overrightarrow{x}_t\right) dt + \underline{S}\overrightarrow{dz} \tag{137}$$

Then

$$P\left(x_t|x_0\right) = \mathcal{N}\left(\overrightarrow{\mu}_t, \underline{\Sigma}_t\right) \tag{138}$$

with

$$\overrightarrow{\mu}_T = e^{-\underline{\mathcal{K}}t} + \left(\underline{I} - e^{-\underline{\mathcal{K}}t}\right)\overrightarrow{\theta} \tag{139}$$

We derive the expression for Σ_t. We drop the subscript t in what follows. We assume that the reversion speed matrix can be orthgonalized:

$$\underline{\mathcal{K}} = \left(\underline{G}^{-1}\right)^T \underline{\Lambda}\underline{G}^T \tag{140}$$

The i, j-th element of Σ_t is given by $[\Sigma]_{ij}$:

$$[\Sigma]_{ij} = \left[\int_0^t e^{-\underline{\mathcal{K}}(t-u)} \underline{SS}^T \left(e^{-\underline{\mathcal{K}}(t-u)}\right)^T du\right]_{ij} \tag{141}$$

Then we have

$$[\Sigma]_{ij} = \left[\left(G^{-1}\right)^T \int\limits_0^t e^{-\underline{\mathcal{K}}(t-u)} \underline{H} e^{-\underline{\mathcal{K}}(t-u)} du G^{-1} \right]_{ij} \tag{142}$$

where we have used the up-and-down theorem, and

$$\underline{H} = \underline{G}^T \underline{S} \underline{S}^T \underline{G} \tag{143}$$

Then we have

$$[\Sigma]_{ij} = \left[\left(G^{-1}\right)^T \left(\int\limits_0^t \left[e^{-(\lambda_i+\lambda_j)u} h_{ij} \right] du \right) G^{-1} \right] \tag{144}$$

and therefore

$$\underline{\Sigma} = \left(\underline{G}^{-1}\right)^T \underline{M} \underline{G}^{-1} \tag{145}$$

with

$$\underline{M} = h_{ij} \frac{1 - e^{-(\lambda_i+\lambda_j)u}}{\lambda_i + \lambda_j} \tag{146}$$

15 Appendix VI – Proof of the Constraints on the Reversion-Speed Matrix $\mathcal{K}^Q$

We have defined

$$\underline{\beta} = \begin{bmatrix} \frac{\left(\overrightarrow{B}_{\tau_1}\right)^T}{\tau_1} \\ \frac{\left(\overrightarrow{B}_{\tau_2}\right)^T}{\tau_2} \\ \ldots \\ \frac{\left(\overrightarrow{B}_{\tau_N}\right)^T}{\tau_N} \end{bmatrix} \tag{147}$$

and we have obtained that

$$\overrightarrow{B}(\tau) = -\int_0^\tau e^{-\underline{\mathcal{K}}(\tau-s)} \Omega^T \overrightarrow{e}_1 ds \tag{148}$$

$$\underline{\Omega}^T = \underline{\beta} \tag{149}$$

and that

$$\underline{\beta\beta}^T = \underline{I} \tag{150}$$

Therefore

$$\underline{I} = \underline{\Omega\Omega}^T = \underline{\Omega\beta} = \underline{\Omega} diag\left[\frac{1}{\tau}\right] \underline{B} =$$

$$\left(\frac{1}{\tau_1} \int_0^{\tau_1} e^{-\underline{\Omega}\mathcal{K}(\tau_1-s)\underline{\Omega}^T} ds \overrightarrow{e}_1, ..., \frac{1}{\tau_N} \int_0^{\tau_N} e^{-\underline{\Omega}\mathcal{K}(\tau_N-s)\underline{\Omega}^T} ds \overrightarrow{e}_1 \right) \tag{151}$$

Define

$$C = \underline{\Omega}\mathcal{K}\underline{\Omega}^T \tag{152}$$

Orthogonalize C:

$$C = T\underline{\Lambda}T^{-1} \tag{153}$$

and define

$$U = T^{-1} \tag{154}$$

Then we have

$$\underline{I} =$$

$$\left(\frac{1}{\tau_1}\int_0^{\tau_1} Te^{-\underline{\Lambda}(\tau_1-s)}T^{-1}ds\overrightarrow{e}_1, ..., \frac{1}{\tau_N}\int_0^{\tau_N} Te^{-\underline{\Lambda}(\tau_N-s)}T^{-1}ds\overrightarrow{e}_1\right) =$$

$$\left(Tdiag\left[\frac{1-e^{-\lambda^i\tau_1}}{\lambda^i\tau_1}\right]T^{-1}\overrightarrow{e}_1, ..., Tdiag\left[\frac{1-e^{-\lambda^i\tau_N}}{\lambda^i\tau_N}\right]T^{-1}\overrightarrow{e}_1\right) \tag{155}$$

Multiplying both sides by T^{-1} one gets

$$U =$$

$$= \left(diag\left[\frac{1-e^{-\lambda_C^i\tau_1}}{\lambda^i\tau_1}\right]U\overrightarrow{e}_1, ..., diag\left[\frac{1-e^{-\lambda_C^i\tau_N}}{\lambda^i\tau_N}\right]U\overrightarrow{e}_1\right) \tag{156}$$

Therefore

$$U = \begin{bmatrix} u_{11}f_{11} & u_{11}f_{12} & ... & u_{11}f_{1N} \\ u_{21}f_{21} & u_{21}f_{22} & ... & u_{N1}f_{2N} \\ ... & ... & ... & ... \\ u_{N1}f_{N1} & u_{N1}f_{N2} & & u_{N1}f_{NN} \end{bmatrix} \tag{157}$$

with

$$f_{ij} = \frac{1-e^{-\lambda_C^i\tau_j}}{\lambda^i\tau_j} \tag{158}$$

$$f_{11} = f_{21} = f_{31} = 1 \tag{159}$$

One can therefore write

$$U = DF \tag{160}$$

$$T = F^{-1}D^{-1} \tag{161}$$

with

$$D = diag\,[u_{11}, u_{21}, ..., u_{N1}]$$

and u_{i1}^{-1} are chosen in such a way that the first, second, ..., nth column of F are normalized. Therefore the matrix F consists of indepedent vectors in its columns, and u_{i1}^{-1} normalize it to a basis.

Putting the pieces together one gets

$$C = \underline{\Omega}\mathcal{K}\underline{\Omega}^T \tag{162}$$

and therefore

$$\mathcal{K} = \underline{\Omega}^T C \underline{\Omega} = \underline{\Omega}^T T \underline{\Lambda} T^{-1} \underline{\Omega} =$$
$$\underline{\Omega}^T F^{-1} \underline{\Lambda} F \underline{\Omega} \tag{163}$$

We have written the orthogonalization of $\underline{\mathcal{K}}$ as

$$\underline{\mathcal{K}} = \underline{a} \Lambda_{\mathcal{K}} \underline{a}^{-1} \tag{164}$$

Therefore, equating terms we have

$$\underline{a} = \underline{\Omega}^T F^{-1} \tag{165}$$

and

$$\underline{\Lambda}_{\mathcal{K}} = \underline{\Lambda} \tag{166}$$