WP/10/258

# IMF Working Paper

# On the Estimation of Term Structure Models and An Application to the United States

*Giancarlo Gasha, Ying He, Carlos Medeiros, Marco Rodriguez, Jean Salvati, and Jiangbo Yi*

INTERNATIONAL MONETARY FUND

 WP/10/258

**IMF Working Paper**

Monetary and Capital Markets Department

**On the Estimation of Term Structure Models and An Application to the United States**

**Prepared by Giancarlo Gasha, Ying He, Carlos Medeiros, Marco Rodriguez, Jean Salvati, and Jiangbo Yi [1,2]**

November 2010

**Abstract**



This paper discusses the estimation of models of the term structure of interest rates. After reviewing the term structure models, specifically the Nelson-Siegel Model and Affine Term-Structure Model, this paper estimates the terms structure of Treasury bond yields for the United States with pre-crisis data. This paper uses a software developed by Fund staff for this purpose. This software makes it possible to estimate the term structure using at least nine models, while opening up the possibility of generating simulated paths of the term structure.

JEL Classification Numbers: G12; E43; E44; E58

Keywords: Term structure models of interest rates; interest rates; yields on bonds; United States

Authors' E-Mail Addresses: ggasha@imf.org; yhe@imf.org; cmedeiros@imf.org; mrodriguez@imf.org; jean.salvati@gmail.com; and jyi@imf.org.

---

[1] Giancarlo Gasha is on leave from the Fund, working as a Senior Advisor in the Ministry of Economy and Finance of Peru. Jean Salvati was a former Fund staff member. He now works for Fannie Mae.

[2] This paper has benefited from comments by Jorge Chan-Lau, Francesco Columba, Alin Mirestean, Richard Munclinger, Christopher Towe, and Liliana Schumacher.

Contents Page

Tables



Figures



Appendix Table

# I. INTRODUCTION

The modeling of the term structure of interest rates is a critical endeavor for investors, market analysts, and policymakers. The modeling of the term structure helps these economic agents not only understand the pricing and interest rate risk of particular financial instruments and investment portfolios, but also appreciate the potential impact of changes in interest rate policy on the yield curve. The modeling of the term structure therefore facilitates (i) the valuation of financial instruments, including credit derivatives; (ii) the simulation of interest rate scenarios; and (iii) the assessment of the impact of interest rate movements on the default probabilities of different financial instruments. Undoubtedly, the need to undertake such tasks has taken on added importance as a result of the sharp interest rate movements in the context of the global financial crisis.

The academic literature tends to focus on two models of the term structure, namely the Nelson-Siegel Models, or NSMs, and Affine-Term Structure models, or ATSMs (Diebold, Piazzesi, and Rudebush, 2005; Van Deventer, Imai, and Mesler, 2005; Baz and Chacko, 2004; and Boulder, 2001). Both types of models make use of stochastic processes and particular assumptions (Cochrane, 2001). For instance, term-structure models rely either on the stochastic process of a single factor, namely the short-term interest rate, or stochastic processes of multiple factors, such as the short-term interest rates and the yields of bonds of various maturities at any point in time. The models also depend on assumptions about the presence or lack of arbitrage to understand the evolution of the yields on bonds.

The models of the term structure attempt to replicate an observed yield curve. In particular, these models focus on ensuring that the models fit the data, while ensuring that the estimated rates are continuous and smooth (Nawalkha, Soto, and Believa, 2005). The NSMs tend to rely on at least three latent factors—interpreted as level, slope, and curvature—which are the parameters of a class of mathematical approximating functions. These models can also include observable macroeconomic variables, reflecting the importance of the joint behavior of the yield curve and macroeconomic variables for bond pricing, investment decisions, and public policy (Ang and Piazzesi, 2003). The ATSMs include some of the traditional term-structure models in the finance literature, including the general single-factor model, the Cox-Ingersoll-Ross (CIR) model, and the multi-factor model.

This paper discusses the estimation of models of the term structure of interest rates. In particular, this paper first reviews some of the main term structure models, specifically the NSM and ATSM models, and then estimates the United States' term structure of Treasury bond yields with data from 1972 to 2007. The paper uses a software developed by Fund staff for this purpose. This software makes it possible to model the term structure using at least nine models, while opening up the possibility of generating different paths of the term structure. This facilitates the computation of, among others, risk indicators such as Value-at-Risk (VaR) for managing the risk of investment holdings.

The paper focuses on two types of term-structure models:

- *The Nelson-Siegel Models, or NSMs.* These models postulate a particular form for the term structure of interest rates, and need not depend on the existence of arbitrage possibilities. These models consider both unobservable factors as well as observable macroeconomic factors.

- *The Affine Term-Structure Models, or ATSMs.* These models, which may depend on the absence of arbitrage opportunities, assume that the unobservable factors underlying the term structure follow stochastic processes.

The paper is organized as follows. Section II presents simple versions of both the NSMs and ATSMs, focusing on an extension of the CIR model. Section III illustrates the capabilities of the software developed by Fund staff through the estimation of the term structure of interest rates of the United States. Section IV discusses possible extensions of the software. Appendix I presents a detailed derivation of some of the main ATSM, Appendix II summarizes the estimation techniques of both NSMs and ATSMs, and Appendix III provides an overview of the capabilities of the software developed by Fund staff to estimate the term structure.

## II. Term Structure Models

### A. Background

The term structure of interest rates or yield curve can be depicted as a plot of a set of interest rates on bonds of different maturities. More than that, observations on the yield curve at different points in time suggest the presence of links among short-, medium- and long-term nominal bond rates. These links, though, do not appear stable through time, as statistical yield curves exhibit different shapes at different moments. However, such changes seem to follow systematic patters that economists have usefully summarized (Diebold and Li, 2006).

One key challenge facing modeling approaches to the yield curve is to provide a useful summary of information at any point in time, for a large number of traded nominal bonds, through a parsimonious model. Such a model should be able both to reproduce the historical stylized facts of the average shape of the yield curve and to forecast future interest rates. In this regard, most models of the yield curve are built on the assumption of the existence of only a few unobservable, or latent, factors and their associated factor loadings relating yields of different maturities to those factors underlying the pricing of tradable bonds (Litterman and Scheinkman, 1991; Balduzzi, Foresi, and Sundaran, 1996; Bliss, 1997a, b; and Dai and Singleton, 2000). The NSMs and ATSMs are two of the most popular classes of factors models used by academics, market participants, and central bank practitioners.

## B. Nelson-Siegel Models

### *The yield-only model*

The class of NSMs has proved satisfactory in fitting the yield curve and capturing its dynamics.[3] Nelson and Siegel (1987) initiated a modeling strategy that provides a powerful and tractable yield curve modeling framework in which the forward rate curve is fit at a given point in time by a class of mathematical approximating function. In particular, they offer a methodology to approximate the forward rate curve by a constant plus a polynomial times an exponential decay term given by[4]

$$f_t(\tau) = \beta_{1t} + \beta_{2t} e^{-\lambda_t \tau} + \beta_{3t}\lambda_t e^{-\lambda_t \tau}, \quad (1)$$

where $f_t(\tau)$ is the instantaneous forward rate.[5] The corresponding yield curve is given by[6]

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right) \quad (2)$$

The parameters of the derived yield curve model are $\beta_{1t}, \beta_{2t}, \beta_{3t}$ and $\lambda_t$, and their respective loadings are given by 1, $\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}\right)$ and $\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right)$. The parameter $\lambda_t$ controls both the exponential decay rate and the maturity at which the loading on $\beta_{3t}$ reaches its maximum.

Although the NSM is presented as a static model, Diebold and Li (2006) interpret $\beta_{1t}, \beta_{2t}$ and $\beta_{3t}$ as dynamic latent factors. They show that these factors can be construed as the level, slope, and curvature factors, respectively, since their loadings are, respectively, a constant, a decreasing function of $\tau$, and a concave function of $\tau$.[7]

---

[3] This modeling strategy has become very popular among market and central-bank practitioners (Bank of International Settlements, 2005).

[4] A forward rate $f_t(\tau, \tau^*)$ is the interest rate of a forward contract, set at time $t$, on an investment that is initiated $\tau$ periods into the future and that matures $\tau^*$ periods beyond the start date of the contract. The instantaneous forward rate $f_t(\tau)$ is obtained by letting the maturity of the contract go to zero.

[5] This curve could be written as $f_t(\tau) = \beta_{1t} + (\beta_{2t} + \beta_{3t}\lambda_t)e^{-\lambda_t \tau}$, which fits the description in the text.

[6] $y_t(\tau)$ denotes the continuously compounded zero-coupon nominal yield to maturity of a $\tau$-period discount bond. The relationship between the yield to maturity and the forward rate is given by

$y_t(\tau) = \frac{1}{\tau}\int_0^\tau f_t(u)du$, which states that the zero-coupon yield is an equally-weighted average of forward rates.

[7] A heuristic interpretation of the factors along these lines is the following: (i) since yields at all maturities load identically on $\beta_{1t}$, an increase in $\beta_{1t}$ increases all yields equally, changing the level of the yield curve; (ii) since short rates load more heavily on $\beta_{2t}$, an increase in $\beta_{2t}$ raises short yields more than long yields, thereby changing the slope of the yield curve; and (iii) since short rates and long rates load minimally on $\beta_{3t}$, an increase in $\beta_{3t}$ will increase medium-term yields, which load more heavily on it, increasing the yield curve

(continued)

The NSM is a popular model for the yield curve for a number of reasons:

- It provides a parsimonious approximation of the yield curve. In particular, the three loadings $\left[1, \left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}\right) \text{ and } \left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right)\right]$ give the model sufficient flexibility to reproduce a range of shapes of observed yield curves.

- It generates forward and yield curves that start at the instantaneous rate $\beta_{1t} + \beta_{2t}$ and then level off at the finite infinite-maturity value of $\beta_{1t}$, which is constant.[8]

- Its three loadings $\left[1, \left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}\right) \text{ and } \left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right)\right]$ allow the three factors $\beta_{1t}, \beta_{2t}$ and $\beta_{3t}$ to be interpreted as long-, short-, and medium-term factors, respectively.[9]

- The time-series statistical properties of the three factors $\beta_{1t}, \beta_{2t}$ and $\beta_{3t}$ underlie the dynamic patterns of the yield curve.

While the three-factor NSM is capable of replicating a variety of stylized facts of empirical yield curves including a diversity of yield curve shapes,[10] the model does exhibit difficulties in fitting the yield curve when yield data are dispersed, with multiple interior minima and maxima. Although this has led to extending the three-factor NSM model in various ways to increase its flexibility, there is a consensus that, for interest rate forecasting and dynamic analysis, the desirability of extensions of the NSM is not obvious.[11] In addition, such extensions may compound the complexity of the estimation problem, especially for the case of multi-country analysis (see section IV).

---

curvature. An additional implication of the NS model is that $y_t(0) = \beta_{1t} + \beta_{2t}$, i.e., the instantaneous yield depends on both the level and the slope factors.

[8] These values are obtained by taking the limits of $y_t(\tau)$ as $\tau$ goes to zero and to infinity, respectively.

[9] To appreciate this interpretation, notice that the loading on $\beta_{1t}$ is 1, which does not decay to zero in the limit; the loading on $\beta_{2t}$ is $\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}\right)$, which starts at 1 but decays quickly and monotonically to 0; the loading on $\beta_{3t}$ is $\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right)$, which starts at 0, increases, and then decays to 0. This coincides with Diebold and Li (2006) interpretation of the three factors as level, slope and curvature.

[10] See Section III C and Figure 2.

[11] See Diebold and Li (2006); and Diebold, Rudebusch, and Auroba (2006). However, more complex specifications have been implemented to obtain a close fit for the yield curve at a point in time, when one of the key objectives is to price yield-curve derivatives. For improving the fit at a particular point in time, Björk and Christensen (1999) add a fourth factor to their Nelson-Siegel specification; Bliss (1997) uses a three Nelson-Siegel specification, but adds an additional decay parameter; while Svensson (1994) adds a second curvature factor with its own separate decay parameter. See De Pooter (2007) for a description and analysis of these extensions of the NSM.

In general, the Nelson-Siegel specifications mentioned above can be placed in the context the following state-space representation[12]

(3) $$(F_t - \mu) = A(F_{t-1} - \mu) + \eta_t$$

(4) or $$F_t = \mu + AF_{t-1} + \eta_t$$

(5) $$y_t = \Lambda F_t + \varepsilon_t.$$

Equations (4) and (5) can be expressed as

(6) $$\begin{bmatrix}\beta_{1t}\\ \beta_{2t}\\ \beta_{3t}\end{bmatrix} = \begin{bmatrix}\mu_1\\ \mu_2\\ \mu_3\end{bmatrix} + \begin{bmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{23} & a_{33}\end{bmatrix}\begin{bmatrix}\beta_{1,t-1}\\ \beta_{2,t-1}\\ \beta_{3,t-1}\end{bmatrix} + \begin{bmatrix}\eta_{1t}\\ \eta_{2t}\\ \eta_{3t}\end{bmatrix}$$

(7) $$\begin{bmatrix}y_t(\tau_1)\\ \cdots\\ y_t(\tau_N)\end{bmatrix} = \begin{bmatrix}1 & \frac{1-e^{-\lambda_t\tau_1}}{\lambda_t\tau_1} & \frac{1-e^{-\lambda_t\tau_1}}{\lambda_t\tau_1} - e^{-\lambda_t\tau_1}\\ \cdots & \cdots & \cdots\\ 1 & \frac{1-e^{-\lambda_t\tau_N}}{\lambda_t\tau_N} & \frac{1-e^{-\lambda_t\tau_N}}{\lambda_t\tau_N} - e^{-\lambda_t\tau_N}\end{bmatrix}\begin{bmatrix}\beta_{1t}\\ \beta_{2t}\\ \beta_{3t}\end{bmatrix} + \begin{bmatrix}\epsilon_{1t}\\ \cdots\\ \epsilon_{Nt}\end{bmatrix}$$

Equation (6), called the *transition equation*, governs the dynamics of the state vector, which, for the three-factor NSM, is given by the unobservable vector $F_t = (\beta_{1t}\ \ \beta_{2t}\ \ \beta_{3t})'$. As in Diebold and Li (2006), it is assumed that these time-varying factors follow a vector autoregressive process of first order, VAR (1), where the mean state vector $\mu$ is a 3x1vector of coefficients, the transition matrix $A$ is a 3x3 matrix of coefficients, and $\eta_t$ is a white noise transition disturbance with a 3x3 non-diagonal covariance matrix $Q$.[13] Equation (7), called the *measurement equation*, is the specification of the yield curve itself, and relates $N$ observable yields to the three unobservable factors. The vector of yields $Y_t$, contains $N$ different maturities $Y_t = [y_t(\tau_1) \quad \cdots \quad y_t(\tau_N)]'$. The measurement matrix $\Lambda$ is an $N$x3 matrix whose columns are the loadings associated with the respective factors, and $\varepsilon_t$ is a white noise *measurement* disturbance with an $N$x$N$ diagonal covariance matrix $H$. It is assumed, mainly to facilitate computations, that both disturbances are orthogonal to each other and to the initial state, $F_0$.[14]

---

[12] The state-space representation is a way of specifying a dynamic system, which facilitates the handling of a wide range of time series models. In particular, the state-space representation facilitates estimation, the extraction of latent yield curve factors, and the testing of hypotheses about the dynamic interactions between the macroeconomy and the yield curve. See Hamilton (1994) and Harvey (1993).

[13] The VAR is expressed in terms of deviations from the mean since $F_t$ is a covariance-stationary vector process.

[14] Formally,

$$\begin{pmatrix}\eta_t\\ \varepsilon_t\end{pmatrix} \backsim WN\left[\begin{pmatrix}0\\ 0\end{pmatrix}, \begin{pmatrix}Q & 0\\ 0 & H\end{pmatrix}\right],$$

(continued)

***The yield curve with macro factors***

As pointed out earlier, movements in the yield curve can be captured by a framework in which yields are linear functions of a few dynamic latent factors. However, although factor models offer a relatively good description of the evolution of interest rates, they provide little insight into the nature of the underlying economic forces driving their movements. In an effort to understand such forces, recent latent factor models of the yield curve have started to incorporate explicitly macroeconomic factors.

Diebold, Rudebusch, and Auroba (2006) use a state-space representation to incorporate macroeconomic factors in a latent factor model of the yield curve to analyze the potential bidirectional feedback between the yield curve and the economy. Specifically, they complement the nonstructural nature of their yield curve representation with a simple nonstructural VAR representation of the macroeconomy to study the nature of the links between the factors driving the yield curve and macroeconomic fundamentals.

In terms of the state-space representation noted above, Diebold, Rudebusch, and Auroba (2006) enhance the state vector to include three key macroeconomic variables that stand for real activity, the stance of monetary policy, and inflation: manufacturing capacity utilization ($CU_t$), the federal funds rate ($FFR_t$), and annual price inflation($INFL_t$).

Explicitly, the state space model for yield-macro model is as follows.

(8) $$F_t = \mu + AF_{t-1} + \eta_t$$

(9) $$Y_t = \Lambda F_t + \varepsilon_t$$

where $F_t = (\beta_{1t}\ \beta_{2t}\ \beta_{3t}\ CU_t\ FFR_t\ INFL_t)'$, and the dimensions of $\mu$, A, and $\eta_t$ are increased accordingly, to 6x1, 6x6 and 6x1, respectively. The matrix $\Lambda$ now contains six columns, of which the three leftmost include the loadings on the three yield factors, and the three rightmost contain only zeroes, indicating that the yields still load only on the yield curve factors. The transition disturbance covariance matrix $Q$, with increased dimension to

---

$\mathbb{E}(F_0\, \eta_t') = 0,$

$\mathbb{E}(F_0\, \varepsilon_t') = 0.$

In addition to computational tractability, most of these assumptions are required to obtain optimality of the procedure used to estimate both equations.

6x6, and the measurement disturbance covariance matrix $H$ are, respectively, non-diagonal and diagonal matrices[15]

$$(10)\qquad \begin{bmatrix}\beta_{1t}\\ \beta_{2t}\\ \beta_{3t}\\ CU_t\\ FFR_t\\ INFL_t\end{bmatrix} = \begin{bmatrix}\mu_1\\ \mu_2\\ \mu_3\\ \mu_4\\ \mu_5\\ \mu_6\end{bmatrix} + \left[\begin{bmatrix}a_{11} & \cdots & a_{16}\\ \vdots & \ddots & \vdots\\ a_{61} & \cdots & a_{66}\end{bmatrix}\right] \begin{bmatrix}\beta_{1,t-1}\\ \beta_{2,t-1}\\ \beta_{3,t-1}\\ CU_{t-1}\\ FFR_{t-1}\\ INFL_{t-1}\end{bmatrix} + \begin{bmatrix}\eta_{1t}\\ \eta_{2t}\\ \eta_{3t}\\ \eta_{4t}\\ \eta_{5t}\\ \eta_{6t}\end{bmatrix}$$

$$(11)\qquad \begin{bmatrix}y_t(\tau_1)\\ \cdots\\ y_t(\tau_N)\end{bmatrix} = \begin{bmatrix}1 & \frac{1-e^{-\lambda_t\tau_1}}{\lambda_t\tau_1} & \frac{1-e^{-\lambda_t\tau_1}}{\lambda_t\tau_1} - e^{-\lambda_t\tau_1} & 0 & 0 & 0\\ \cdots & \cdots & \cdots & & & \\ 1 & \frac{1-e^{-\lambda_t\tau_N}}{\lambda_t\tau_N} & \frac{1-e^{-\lambda_t\tau_N}}{\lambda_t\tau_N} - e^{-\lambda_t\tau_N} & 0 & 0 & 0\end{bmatrix} \begin{bmatrix}\beta_{1t}\\ \beta_{2t}\\ \beta_{3t}\\ CU_t\\ FFR_t\\ INFL_t\end{bmatrix} + \begin{bmatrix}\epsilon_{1t}\\ \cdots\\ \epsilon_{Nt}\end{bmatrix}$$

After estimating the state-space model, Diebold, Rudebusch, and Auroba (2006) proceed to explore the dynamics of the yields-macro system using impulse response functions,[16] considering, in turn, four groups of impulse responses:

- Macro responses to macro shocks;
- Macro responses to yield curve shocks;
- Yield curve responses to macro shocks; and
- Yield curve responses to yield curve shocks.

In addition, Diebold, Rudebusch, and Auroba (2006) study the nature of macro and yield curve interactions by examining macroeconomic and yield curve variance decompositions,[17]

---

[15] Diebold, Rudebusch, and Auroba (2006) note that these macroeconomic variables are considered to be the minimum set of fundamentals required to capture basic macroeconomic dynamics. See, also, Rudebusch and Svensson (1999).

[16] To produce impulse responses from their model, Diebold, Rudebusch, and Auroba (2006) identify the covariances given by the off-diagonal elements of the $Q$ matrix by ordering the variables as follows: $\beta_{1t}$, $\beta_{2t}$, $\beta_{3t}$, $CU_t$, $INFL_t$, $FFR_t$. This follows from the fact that they use beginning-of-period yield data and end-of-period macro data. However, they point out that their results are robust to alternative identification strategies.

[17] Diebold, Rudebusch, and Auroba (2006) explore both, variance decomposition for yields and for macroeconomic variables at different time horizons. For yields, they contrast the yields-only model with the yields-macro model. For macroeconomic variables, they contrast the yields-macro model with a macro-only model, which is a simple first-order VAR for $CU_t$, $INFL_t$, $FFR_t$.

and the results from formal statistical tests.[18] Their results indicate that, although bidirectional causality is likely to be present, the effects of macroeconomic factors on future yield curves seem relatively more important than those of the yield curve factors on future macroeconomic developments. However, market yields do still contain relevant predictive information about the stance of monetary policy.[19]

### C. Estimation Approaches for the Nelson-Siegel Models

Using the state-space representation as the general framework for estimating the latent factors and parameters of the different Nelson-Siegel specifications, two general estimation approaches can be identified. In particular, depending on whether the two equations are estimated separately or jointly, they are a two-step approach or a one-step approach. In addition, for each of them the decay parameters are either pre-specified or estimated.

The two-step approach is exemplified by Diebold and Li (2006), who face the problem of estimating the parameters $\beta_{1t}, \beta_{2t}, \beta_{3t}$ and $\lambda_t$ for a Nelson-Siegel model without macro factors. Following a practice initiated by Nelson and Siegel (1987), they first fix $\lambda_t$ at a prespecified value for all *t*, i.e., $\lambda_t = \lambda \; \forall \; t$, and then use ordinary least squares to estimate the factors, for each month *t*.[20] In general, for the case of one prespecified decay parameter, this step generates time series of estimated values for each of the K factors: $\{\beta_{i,t}\}_{t=1}^{T}$ for *i*=1, 2,..., K. In the second step, the transition equation is estimated, assuming that *A* and *Q* are diagonal matrices. The strategy of fixing $\lambda_t$ at a prespecified value in the first step greatly simplifies the estimation procedure; otherwise, it would be necessary to use nonlinear least squares for each month *t*.[21]

---

[18] These formal tests consist in tests about restrictions on the *A* and *Q* matrices. Specifically, by partitioning *A* and *Q* into four 3x3 blocks, as: $A=\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$ and $Q=\begin{pmatrix} Q_1 & Q_2 \\ Q_2^T & Q_3 \end{pmatrix}$,
Diebold, Rudebusch, and Auroba (2006) test whether $A_2 = 0, A_1 = 0$, and $Q_2 = 0$ (i.e., there is no interaction between yields and macro factors), $A_2 = 0$ (i.e., there is no interaction from macro to yields), and $A_3 = 0$, and $Q_2 = 0$ (i.e., there is no interaction from yields to macro).

[19] It is worth noting that although Diebold, Rudebusch, and Auroba (2006) do not impose no-arbitrage restrictions, they argue that even if no-arbitrage restrictions hold for the data, they will, at least, be roughly captured by the fitted yield curves, particularly because they are flexible approximations to the data.

[20] The main role played by $\lambda$ is to determine the maturity $\tau$ at which the loading on the curvature factor, $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right)$, is at its maximum. In Diebold and Li (2006), the value of $\lambda$ that maximizes the curvature loading at exactly 30 months is $\lambda$= 0.0609.

[21] The case of more than one decay parameter cannot be handled by pre-specifying their values, although it is still assumed that their values are constant over time.

The one-step approach, in which all parameters are estimated simultaneously, is illustrated by Diebold, Rudebusch, and Auroba (2006) who estimate a Nelson-Siegel-type model with three factors and one prespecified decay parameter with and without macroeconomic variables. They place their Nelson-Siegel-type model in state-space form, which allows the application of Kalman filter techniques.[22] This, in turn, provides maximum-likelihood estimates and optimal filtered and smoothed estimates of the underlying factors. This approach is considered superior to the two-step approach since it produces correct inference via standard theory.[23] One drawback of the one-step approach, however, is that the number of parameters to estimate is considerable. For example, the yield-only model in Diebold, Rudebusch, and Auroba (2006) has 36 parameters, while their yield-macro model has 81 parameters that must be estimated by numerical optimization.

In both the yield-only and yield-macro Nelson-Siegel models, we need to estimate latent factors $F_t$ as well as coefficients in the transition matrix $A$, the mean state vector $\mu$, the measurement matrix $\Lambda$, the transition disturbance covariance matrix $Q$, and the measurement disturbance covariance matrix $H$. Depending on whether the transition equation and measurement equation are estimated separated or jointly they are a one-step approach or a two-step approach.

In our implementation, we also follow Diebold and Li (2006) in fixing the decay parameter over the time. However, we do not set the decay parameter at the prespecified value; instead, we choose this value based on optimizing estimation performance. In the two-step approach, we optimize the decay parameter based on the root mean square error (RMSE) of the measurement equation. For a given decay parameter, the measurement matrix is known. We can then consider the measurement equation as a cross-sectional model, and run an ordinary least square for each time epoch $t$ to obtain the latent factors $F_t$ and the measurement error $\epsilon_t$. From the measurement errors over time, we can calculate the measurement disturbance covariance matrix $H$ and the RMSE of the measurement equation. The $\lambda$,, $F_t$ and $H$ associated with the lowest RMSE are our estimated parameters. After we obtain the latent factors, we can consider the transition equation as a VAR(1) model and run an ordinary least square to get the transition matrix $A$, the mean state vector $\mu$ , and the transition error $\eta_t$ . The transition error makes it possible to compute the transition disturbance covariance matrix $Q$.

In our implementation of the one-step approach, we optimize the decay parameter $\lambda$, the mean state vector $\mu$, the measurement matrix $\Lambda$, the transition disturbance covariance

[22] The Kalman filter is applied to models placed on a state space form, and provides algorithms for prediction and smoothing. In a Gaussian model, the Kalman filter supplies the ways of constructing the likelihood function by the prediction error decomposition. See Harvey (1993).

[23] In other words, the estimates exhibit better large-sample properties, including the asymptotic distributions from which inferences are made.

matrix $Q$ and the measurement disturbance covariance matrix $H$ to maximize the log-likelihood of the state-space system that is given as follows

(12) $$L = \sum_t \left\{-\frac{1}{2}[ln(|S_t|) + v_t' S_t^{-1} v_t]\right\}$$

where the predicted error covariance matrix, $S_t$, and predicted error, $v_t$, are calculated using Kalman filtering. In our implementation, we use the following Kalman filtering iteration equations

(13) $$F_{t|t-1} = \mu + AF_{t-1|t-1}$$

$$\Sigma_{t|t-1} = A\Sigma_{t-1|t-1}A' + Q$$

$$v_t = Y_t - \Lambda F_{t|t-1}$$

$$S_t = \Lambda\Sigma_{t|t-1} + H,$$

$$K_t = \Sigma_{t|t-1}\Lambda' S_t^{-1}$$

$$F_{t|t} = F_{t|t-1} + K_t v_t$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t\Lambda\Sigma_{t|t-1}$$

We employ unconditional mean and unconditional variance of the latent factor vector $F_t$ for the initial $F_{0|0}$ and $\Sigma_{0|0}$. In this context, the latent factor vector $F_t$ is also estimated.

## D. Affine Term-Structure Models

***Background***

As mentioned above, the class of ATSMs, used mainly by finance academics and market participants, is another type of factor models linking the dynamics of the term structure of interest rates to the dynamics of a few unobserved, or state, variables that impinge upon the yields. The key of these dynamics is the instantaneous interest rate, $r_t$.[24] In general, an ATSM

---

[24] If $y_{t,t+\tau}$ denotes the interest at time $t$ on a loan contract between $t$ and $t+\tau$, the instantaneous interest rate is defined as $y_t = \lim_{n\to 0} y_{t,t+\tau}$. As such, there is no empirical counterpart to this concept; it is a theoretical construct that facilitates the application of the methods of continuous-time stochastic processes in the modeling of the term structure. The instantaneous short-term rate, and the zero-coupon bonds are the building blocks for the ATSMs.

begins with the specification of the instantaneous interest rate as a linear combination of a set of state variables, followed by a description of the evolution of the factor processes, typically as stochastic differential equations (SDEs).[25] These equations relate changes in the factor processes to a component changing with time itself (the well know trend or *drift* term) plus a stochastic component (the equally well known variance or *diffusion* term) whose changes depend on a linear combination of the stochastic processes driving the state variables (usually standard scalar Wiener processes defined on the same probability space).[26]

Once the stochastic structure of the state variables is specified, the price of the pure discount bond with maturity, i.e., $P(\tau)$, is postulated to be a function of the maturity itself and of the underlying risk factors. The dynamics, i.e., $dP$, require an appropriate application of techniques from stochastic calculus to the price function.[27] Appendix I summarizes the assumptions underlying the associated stochastic equations and derives one-single factor model, the CIR, and multi-factor models.

At this point, the generic term for the change in the price of a bond, $dP$, is used in obtaining an expression for the relative change in the value of a riskless self-financing portfolio.[28] This portfolio includes a long position in a pure discount bond with instantaneous maturity and as many short positions, each with a different maturity, in discount bonds as risk factors may exist.[29] The problem is then to select the weights on the portfolio so as to eliminate the underlying sources of risk, which requires that the portfolio earn the risk-free rate to ensure the absence of arbitrage. This process yields a relatively complex partial differential equation (PDE), which can be solved analytically since the term structure model is affine.[30]

---

[25] An SDE is an equation of the form: $dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t$,
where $W_t$is a scalar Wiener process. Heuristically, it indicates that the differential change in the variable $X_t$, i.e., $dX_t$, is made up of a non-stochastic component, or *drift* term, i.e., $\mu(X_t, t)dt$ , and a stochastic component, or *diffusion* term, i.e., $\sigma(X_t, t)dW_t$.

[26] A probability space is a triplet, $(\Omega, \mathcal{F}, \mathbb{P})$, where $\Omega$ is the set of all possible outcomes, $\mathcal{F}$ specifies the set of all events (subsets of $\Omega$), to which probability numbers will be assigned, and $\mathbb{P}$ is a probability measure operating on $\mathcal{F}$. A standard scalar Wiener process, or standard Brownian motion, $W_i(t)$, is a stochastic process having continuous sample paths, stationary independent increments, and $W_i(t)$has normal distribution with mean zero and variance $t$, i.e., $W_i(t) \sim N(0.t)$. The subscript $i$ is an index indicating the number of state variables for a particular ATSM.

[27] In particular, application of the Itô's formula (see Duffie (2001)).

[28] A self-financing portfolio is a portfolio whose value changes due to a profit or loss in the investment.

[29] Assuming that there are $n$ state variables, or risk factors, determining the instantaneous interest rate, the model requires $n + 1$ bonds to construct a riskless portfolio.

[30] A function $f: \mathbb{R}^n \longrightarrow \mathbb{R}$, is said to be *affine* if there exists $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}^n$ such that $f(x)= \alpha + \beta^T x$, for all $x \in \mathbb{R}^n$.

The result of the steps just described previously is an analytical relationship between the price of a zero-coupon bond of maturity $\tau$, at any point in time, and the values of the risk factors.

***Extension of the Cox-Ingersoll-Ross model***

A number of authors, including Duffie (2001), Duffie and Kan (1996, 1994) and Chaplin and Sharp (1993) generalize a very popular one-factor model of the term structure of interest rates, developed by Cox, Ingersoll and Ross (1985), to include several factors in an affine structure. This multi-factor generalization[31] starts by assuming that the instantaneous short-term interest rate, $y_t$, is s linear combination of $n$ independent state variables, or factors, denoted by $z_1(t), .., z_n(t)$, i.e.,

(14) $$y_t = \sum_{i=1}^{n} z_i(t).$$

Each of the state variables is assumed to follow a square-root process whose differential dynamics is given by [32]

(15) $$dz_i(t) = \kappa_i\big(\theta_i - z_i(t)\big)dt + \sigma_i\sqrt{z_i(t)}dW_i(t),$$

where $\kappa_i\big(\theta_i - z_i(t)\big)dt$ and $\sigma_i\sqrt{z_i(t)}dW_i(t)$, for $i = 1, \dots, n$, are, respectively, the *drift* and the *diffusion* terms of the process, and the $W_i(t)$'s are independent scalar Wiener processes defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.[33]

Given these specifications for the fundamental determinants of the instantaneous interest rate, an ATSM postulates that the price process of a pure discount bond is a function of the term to maturity, $\tau$, and the set of $n$ state variables $\big(z_1(t), .., z_n(t)\big)$. Specifically,

(16) $$P(\tau) = P(\tau, z_1, .., z_n)$$

The differential dynamics of the price process, $dP$, is obtained by applying Itô's theorem to equation (14), which produces an expression that includes factors associated with $dt$ and a linear combination of factors associated with $dW_i(t)$, for $i = 1, \dots, n$. This differential dynamics expression, $dP$, is then used in computing the return on a self-financing portfolio

[31] As indicated in the discussion of the class of Nelson-Siegel models, a key feature desired for term structure models is to provide specifications with enough flexibility to reproduce a range of shapes of observed yield curves. A one-factor specification seems very limited for this purpose.

[32] This assumption will ensure that the factors and the interest rate are nonnegative at all times almost surely, i.e., with probability one.

[33] The *drift* term is a mean-reverting factor with parameters $\kappa_i$ and $\theta_i$, while the *diffusion* term has volatility parameter $\sigma_i$.

composed of a long position in a pure discount bond with maturity $m_0$, and $n$ short short positions in pure discount bonds with maturities $m_1, \ldots, m_n$, where $m_0 \neq m_1 \neq \cdots \neq m_n$. Denoting the value of this portfolio by $V$, its rate of return is given by

(17) $$\frac{dV}{V} = \frac{dP(t,m_0)}{P(t,m_0)} - \sum_{k=1}^{n} u_k \frac{dP(t,m_k)}{P(t,m_k)}$$

The next step is then to select the weights on the portfolio $(u_1, \ldots, u_n)$ so that the $n$ underlying sources of risk are eliminated, which will require that the portfolio earns the risk-free rate to ensure the absence of arbitrage. This process will produce a PDE, which can be solved analytically using the affine specification for the zero-coupon bond with maturity $\tau$ given by

(18) $$P(\tau, z_1, .., z_n) = e^{\sum_{i=1}^{n}(A_i(\tau) - B_i(\tau) z_i)}$$

The solution to the $n$-factor CIR PDE involves obtaining closed-form expressions for the functions $A_i(\tau)$ and $B_i(\tau)$ given by

(19) $$A_i(\tau) = ln\left(\frac{2\gamma_i e^{\frac{(\gamma_i+\kappa_i+\lambda_i)\tau}{2}}}{(\gamma_i+\kappa_i+\lambda_i)(e^{-\gamma_i\tau}-1)+2\gamma_i}\right)^{\frac{2\kappa_i\theta_i}{\sigma_i^2}}$$

(20) $$B_i(\tau) = \frac{2(e^{-\gamma_i\tau}-1)}{(\gamma_i+\kappa_i+\lambda_i)(e^{-\gamma_i\tau}-1)+2\gamma_i}$$

Where $$\gamma_i = \sqrt{(\kappa_i + \lambda_i)^2 + 2\sigma_i^2}$$

and $\lambda_i$ is the market price of risk.[34]

Equations (16)-(20) provide the expressions linking the price of a zero-coupon bond of maturity $\tau$ to the risk factors. The functions $A_i(\tau)$ and $B_i(\tau)$ depend on the maturity $\tau$ and on the parameters of the model, and represent the loadings of the bond price on the state variable $i$, for $i = 1, \ldots, n$. In this regard, an ATSM can be conceived as a procedure for computing the zero-coupon yield of a given term to maturity, knowing the value of the state variables.

[34] The market price of risk, loosely speaking, represents the standardized excess return, over the risk-free rate, for holding a pure discount bond.

# III. An Application to the United States

## A. Background

This section identifies, catalogues and explains key stylized facts of the U.S. yield curve for the period 1972 to 2007. In particular, it uncovers a set of factors driving the dynamic evolution of the term structure of U.S. Treasury yields, and explores their links to the evolution of macroeconomic fundamentals. The section proceeds first by describing the nature of the data used in the analysis and then by identifying a set of stylized facts from the available yield data. Subsequently a series of Nelson-Siegel models are estimated to explain the stylized facts of the U.S. yield curve, as well as the dynamic interactions between the yield curve and a set of macroeconomic fundamentals.[35]

## B. Data

### *Yield data*

The empirical analysis in this paper uses U.S. Treasury monthly data on zero-coupon bond yields provided obtained from Bloomberg. The yields are annualized zero-coupon bond continuously compounded nominal yields. They are monthly observations on yields for U.S. Treasury bonds between January 1972 and December 2007, i.e., 432 months, and contain 3,888 monthly observations of yields for 9 maturities: 3, 6, 12, 24, 36, 48, 60, 84 and 120 months.[36]

### *Macroeconomic data*

Monthly data on three macroeconomic variables are used to study the potential bidirectional feedback from the yield curve to the economy. These variables, for the period January 1972 to December 2007, are: (i) the inflation variable, the annual percentage change in the monthly price deflator for personal consumption expenditures; (ii) the real economic activity relative to potential, manufacturing capacity utilization; and (iii) the monetary policy instrument, the monthly average federal funds rate. These variables are widely viewed to be the minimum set of macroeconomic factors able to capture basic macroeconomic dynamics.[37]

---

[35] The analysis of this section follows closely the approaches in Diebold and Li (2006) and Diebold, Rudebusch, and Auroba (2006).

[36] Diebold, Rudebusch, and Auroba (2006) examine U.S. Treasury yields with maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108 1nd 120 months for the period January 1972 through December 2000, using the unsmoothed Fama and Bliss (1987) approach, as described in Diebold and Li (2006). As shown below, the results of this section are similar to those in Diebold and Li (2006) and Diebold, Rudebusch, and Auroba (2006).

[37] See Rudebusch and Svensson (1999).

## C. The U.S. Yield Curve[38]

### *Stylized facts*

U.S. Treasury yields exhibit a sizable inter temporal variation during the period of analysis (see Figure 2). To summarize the yield information at any point in time for the nominal bonds that are traded, we follow the principle that, since only a small number of sources of systematic risk underlie the pricing of financial assets, almost all price information can be extracted with a few constructed factors.[39] In the context of our modeling approach, we assume that three factors—level, slope and curvature—are enough to summarize the essential features of the term structure at any given point in time, as well as its evolution through time.[40] Figure 2 shows that, for the period 1972:1-2007:12, the U.S. yield curve exhibits sizable inter temporal variation in its level, and, although the variation in the slope and curvature is less marked, it is nonetheless evident.

A set of stylized facts characterizing the U.S. yield curve can be extracted for the period of analysis. Table 1 presents descriptive statistics for the yields at different maturities, and for the yield curve *empirical* level, slope and curvature factors. The last three columns include sample autocorrelations at displacements of 1, 12, and 30 months. Based on these results and a detailed look at the yield data for the period, we can identify the following stylized facts, whose replication should be the test for any potential model of the U.S. yield curve:

- The average yield curve is upward sloping and concave.
- The yield curve assumes a variety of shapes through time, including upward sloping, downward sloping, humped, and inverted humped.[41]
- Yield dynamics are persistent, while spread dynamics are less persistent.[42]
- The short end of the yield curve is more volatile than the long end.[43]

---

[38] As noted in the Introduction, the main goal of this section is to illustrate the capabilities of the MCM-TGS software. For this reason, the period of analysis chosen (1972:1-2007:12) leaves out the recent global financial crisis which is the object of an ongoing research project in the Monetary and Capital Markets Department of the Fund.

[39] See Diebold, Piazzesi and Rudebusch (2005) and Litterman and Scheinkman (1991).

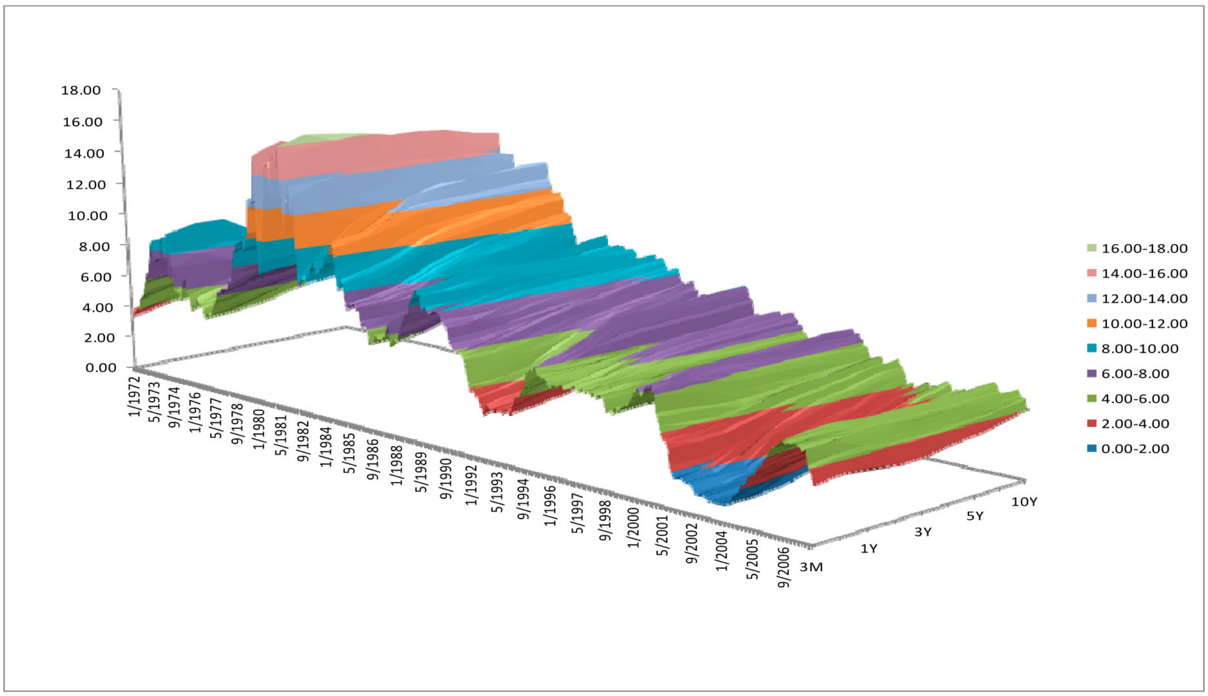
[40] In constructing the *empirical* factors, we define the level as the 4-year yield, the slope as the difference between the one-month and the 4-year yields, and the curvature as twice the one-year yield 1minus the sum of the one-month and four-year yields. See Diebold and Li (2006).

[41] The *empirical* slope takes negative values in 73 of the 432 months.

[42] Persistent yield dynamics are associated with the strong persistence of the level, and less persistent spread dynamics are related to weaker persistence of the slope.

- The level of the yield curve is highly persistent, but exhibits small variation relative to its mean. In fact the level is more persistent than any single yield.
- The slope is less persistent than any single yield but highly variable relative to its mean.
- The curvature is the least persistent of all factors and displays the largest variability relative to its mean.

**Figure 1. Observed Yield Curves**




Source: Fund staff estimates.

***A yield-only model for the United States***

A three-factor NSM model fits well the series of cross sections of U.S. monthly Treasury yields. In particular, fitting equation (2) to the U.S. yield data provides estimates of the three factors and the decay parameter in the three-factor NS model—$\beta_{1t}, \beta_{2t},\ \beta_{3t}$ and $\lambda$.[44] Analysis

[43] Volatility at the short end of the curve results from the added volatility of the slope and the level, while the long end volatility is influenced only by the volatility of the level.

[44] This paper uses software developed by MCM and TGS to estimate the parameters of NSM models (see Appendices II and III). In the context of the empirical analysis of this section, the procedure is used to estimate 432 yield curves, one for each month.

of the residuals from the estimation procedure, shown in Table 2, indicates that the three-factor NSM model fits well the U.S. yield data during the period of analysis.[45] Figure 2 illustrates the ability of the three-factor NSM model to capture a variety of shapes that the U.S. yield curve assumes through time. Using the estimation results, Figure 3 shows that the implied average fitted curve and the average actual yield curve are very close, reinforcing the assessment of the overall good fit provided by the model, and matching some of the stylized facts of the U.S. yield curve.

Table 1. United States: Yield-Only Model Descriptive Statistics, Yield Curves

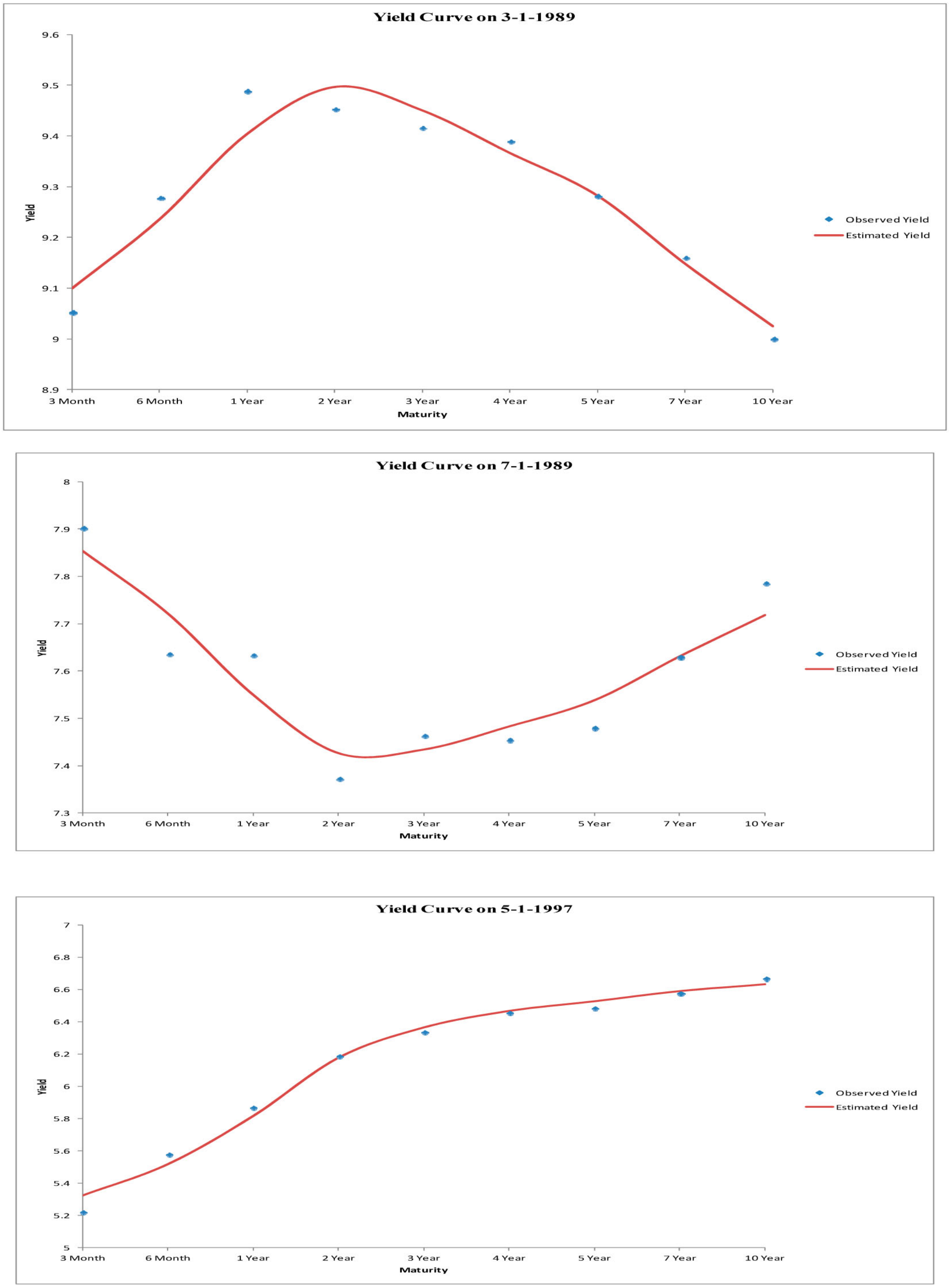
| Maturity (Months) | Mean | Std. Dev. | Minimum | Maximum | ρ(1) | ρ(12) | ρ(30) |
|---|---|---|---|---|---|---|---|
| 3 | 6.101 | 2.969 | 0.860 | 15.999 | 0.981 | 0.777 | 0.454 |
| 6 | 6.326 | 3.010 | 0.976 | 16.511 | 0.982 | 0.791 | 0.486 |
| 12 | 6.536 | 2.987 | 1.027 | 16.345 | 0.983 | 0.809 | 0.530 |
| 24 | 6.799 | 2.872 | 1.251 | 16.145 | 0.985 | 0.829 | 0.599 |
| 36 | 6.974 | 2.772 | 1.570 | 15.825 | 0.986 | 0.842 | 0.635 |
| 48 | 7.119 | 2.696 | 1.966 | 15.847 | 0.987 | 0.849 | 0.660 |
| 60 | 7.219 | 2.651 | 2.272 | 15.696 | 0.988 | 0.855 | 0.680 |
| 84 | 7.398 | 2.564 | 2.855 | 15.283 | 0.990 | 0.866 | 0.705 |
| 120 (level) | 7.514 | 2.484 | 3.372 | 15.065 | 0.991 | 0.873 | 0.728 |
| Slope | 1.413 | 1.447 | -3.223 | 4.140 | 0.941 | 0.419 | -0.154 |
| Curvature | -0.016 | 0.782 | -2.062 | 3.012 | 0.838 | 0.332 | 0.094 |

Source: Fund staff estimates.

[45] The residual sample autocorrelations indicate that pricing errors are somewhat persistent, reflecting possible persistent tax and liquidity effects. Also, the estimated means and standard deviations of the residuals, expressed in basis points, show that the mean error is negligible at all maturities and that the average standard deviation for the relevant middle range of maturities from 6 to 60 months is very small—about 6.5 basis points.

## Figure 2. Selected Fitted (Model Based) Yield Curves

The plots include fitted yield curves for selected dates, together with actual yields, for the United States



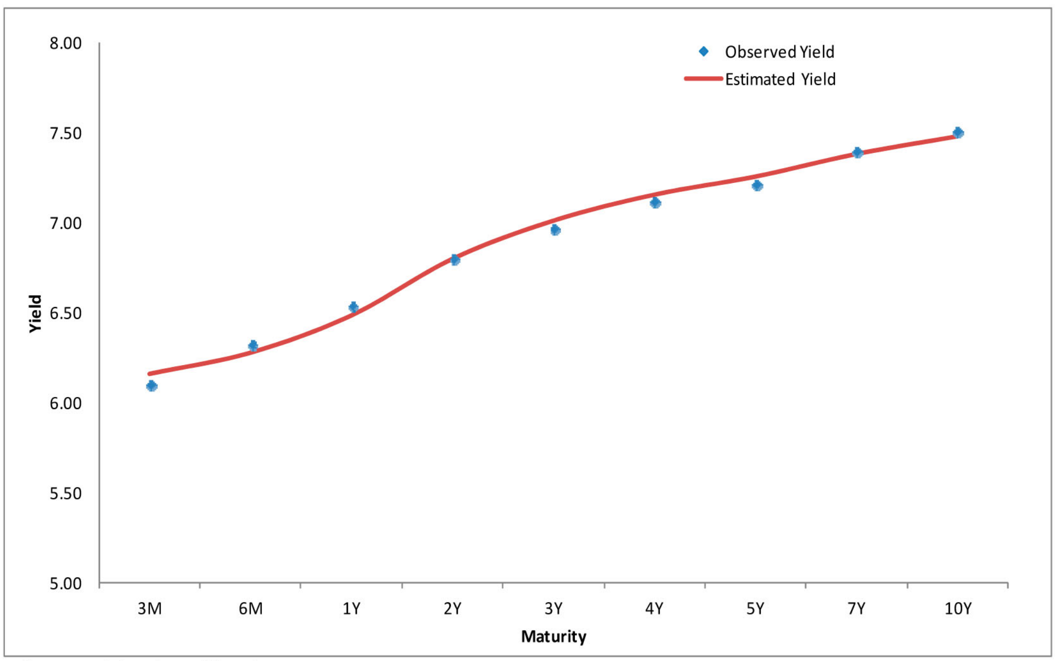

Source: Fund staff estimates.

Table 2. United States: Yield-Only Model Descriptive Statistics, Yield Curve Residuals

| Maturity (Months) | Mean | Std. Dev. | Minimum | Maximum | ρ(1) | ρ(12) | ρ(30) |
|---|---|---|---|---|---|---|---|
| 3 | -0.059 | 0.092 | -0.628 | 0.226 | 0.655 | 0.274 | 0.108 |
| 6 | 0.046 | 0.078 | -0.223 | 0.472 | 0.616 | 0.282 | 0.120 |
| 12 | 0.050 | 0.102 | -0.248 | 0.527 | 0.671 | 0.244 | 0.003 |
| 24 | 0.005 | 0.058 | -0.220 | 0.206 | 0.654 | 0.145 | -0.089 |
| 36 | -0.031 | 0.049 | -0.245 | 0.123 | 0.590 | 0.234 | 0.086 |
| 48 | -0.030 | 0.055 | -0.243 | 0.130 | 0.757 | 0.145 | 0.070 |
| 60 | -0.032 | 0.050 | -0.220 | 0.166 | 0.666 | 0.159 | -0.127 |
| 84 | 0.019 | 0.047 | -0.101 | 0.258 | 0.666 | 0.220 | 0.074 |
| 120 | 0.032 | 0.069 | -0.159 | 0.311 | 0.739 | 0.356 | 0.032 |

Source: Fund staff estimates.

Figure 3. Observed and Estimated Average Yield Curve




Source: Fund staff estimates.

The evolution of the estimated factors, $\{\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}\}$, characterizes the yield curve dynamics and reproduces several stylized facts of the U.S. yield curve.

- Table 3 shows descriptive statistics for the estimated factors, which suggest that the first factor is the most persistent and the least volatile relative to its mean, and that the second factor is more persistent and less volatile relative to its mean than the third.[46] Figure 4 displays the three estimated factors of the model for a comparative assessment, and Figure 5 plots each of the factors together with their respective empirical proxies. The level factor, which is positive and fluctuates around 7.7 percent, is highly persistent, and the slope and curvature are less persistent and take on both positive and negative values.[47] The plots in Figure 5 corroborate the claim that the three factors of the model match up the level, slope and curvature. The correlations between the estimated factors and their empirical proxies are $\rho(\hat{\beta}_{1t}, l_t)$ = 0.99, $\rho(\hat{\beta}_{2t}, s_t)$ = 0.99, and $\rho(\hat{\beta}_{3t}, c_t)$ = 0.97, where $\{l_t, s_t, c_t\}$ stands for the empirical level, slope and curvature of the yield curve, as defined above. In sum, the level, slope, and curvature factors provide a good representation of the yield curve.

- As noted earlier, the evolution of the yield curve factors is assumed to follow a VAR of order 1. Table 4 presents the estimates of the coefficient matrix $A$. They show highly persistent own dynamics of $\hat{\beta}_{1t}, \hat{\beta}_{2t}$, and $\hat{\beta}_{3t}$, with estimated own-lag coefficients of 1.00, 0.92 and 0.84, respectively.[48] Cross-factor dynamics appear significant.

- In addition to the strong persistence of the individual factors, results from the estimated VAR suggest that, during the period of analysis, the level influences positively the slope and curvature, the slope influences negatively the level and curvature, and the curvature affects negatively the level and slope. These results suggest complex dynamic interactions among the yields at different maturities induced by the underlying forces driving the factors' dynamics. An economically meaningful interpretation of these results, however, would require a framework that relates changes in yield curve factors to macroeconomic fundamentals.

---

[46] Since the long rates load heavily on the level factor, this result matches the fact that the long end of the curve is more persistent than the short end of the curve. On the other hand, since the short end of the curve loads on both the level and the slope they are more volatile than the long end of the curve.

[47] As defined earlier, the slope factor in the estimated equation, i.e., $\hat{\beta}_{2t}$, corresponds to the negative of the slope as traditionally defined, i.e., long minus short yields.

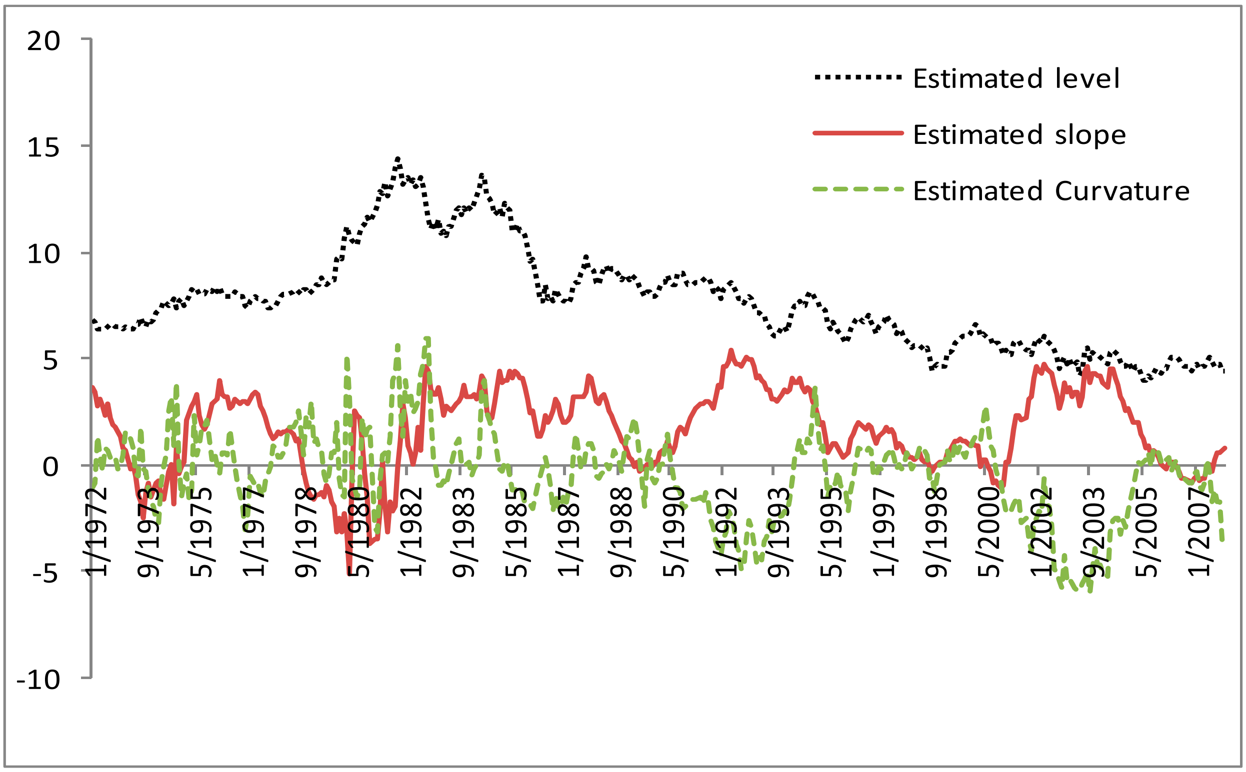
[48] Although the own-lag coefficient of the level factor is slightly greater than 1, stationarity is assured since all the eigenvalues of the matrix of the estimates have modulus less than 1.

Table 3. United States: Yield-Only Model Descriptive Statistics, Estimated Factors

| Factor | Mean | Std. Dev. | Minimum | Maximum | ρ(1) | ρ(12) | ρ(30) |
|---|---|---|---|---|---|---|---|
| $\hat{\beta}_1$ | 7.723 | 2.377 | 3.934 | 14.399 | 0.992 | 0.879 | 0.734 |
| $\hat{\beta}_2$ | 1.693 | 1.903 | -5.136 | 5.370 | 0.949 | 0.456 | -0.134 |
| $\hat{\beta}_3$ | -0.472 | 2.040 | -5.953 | 6.007 | 0.884 | 0.439 | 0.062 |

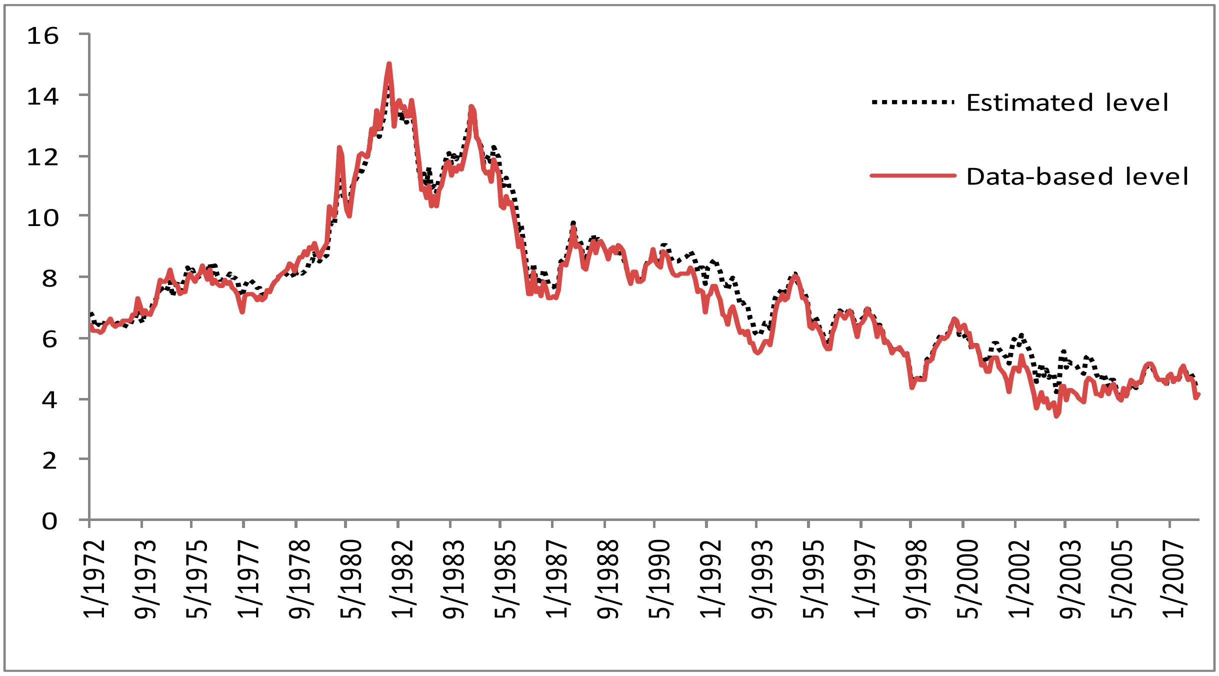
Source: Fund staff estimates.

Figure 4. Estimates of the Level, Slope, and Curvature in the Yields-Only Model




Source: Fund staff estimates.

Figure 5. Model-Based vs. Data-Based Level, Slope, and Curvature




Source: Fund staff estimates.

Figure 5. Model-Based vs. Data-Based Level, Slope and Curvature (continued)

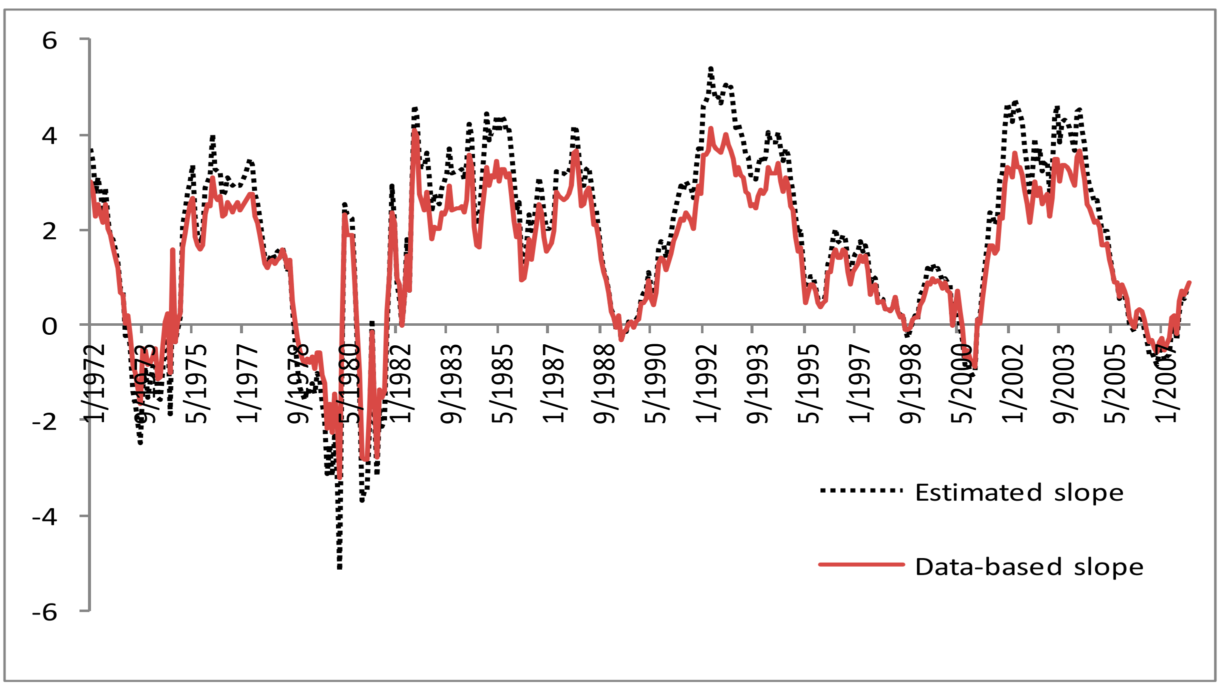


Figure 5. Model-Based vs. Data-Based Level, Slope and Curvature (concluded)

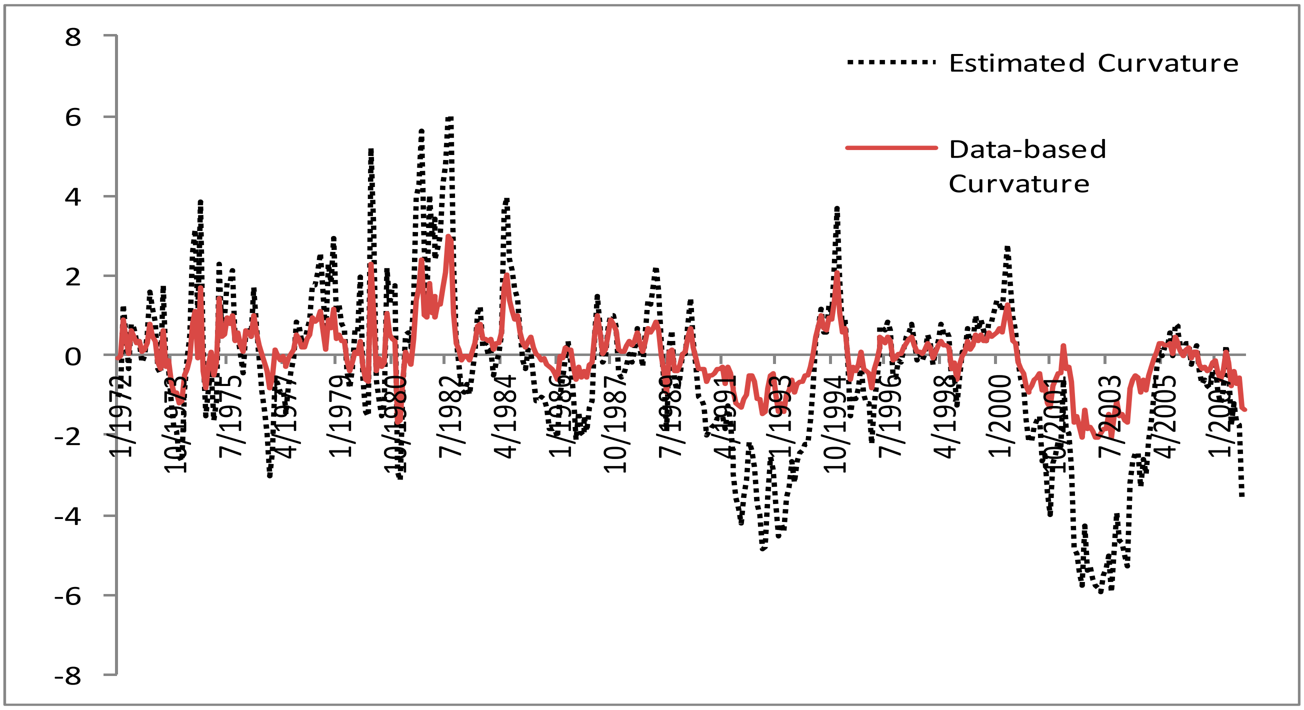



Table 4. United States: Yield-Only Model VAR Parameter Estimates
(Standard errors in parentheses and t-statistics in brackets)

| | Level | Slope | Curvature |
|---|---|---|---|
| Level (-1) | **1.00** | **0.04** | **0.06** |
| | (0.00668) | (0.01378) | (0.02200) |
| | [150.416] | [2.78315] | [2.68736] |
| Slope (-1) | **-0.03** | **0.92** | **-0.05** |
| | (0.00807) | (0.01665) | (0.02657) |
| | [-4.31019] | [55.1756] | [-2.05823] |
| Curvature (-1) | **-0.03** | **-0.07** | **0.84** |
| | (0.00838) | (0.01719) | (0.02759) |
| | [-3.02939] | [-3.97735] | [30.3649] |

Source: Fund staff estimates.

***A yield-macro model for the United States***

Yield curve factors appear to be related to macroeconomic variables. Figures 6–8 plot the estimated level, slope and curvature factors from the previous section—$\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$, and $\hat{\beta}_{3t}$, together with likely related macroeconomic variables. Figure 6 shows the estimated yield-only model level, $\hat{\beta}_{1t}$, and a measure of inflation (the 12-month percent change in the deflator for personal consumption expenditures), whose correlation, 0.52, appears to identify a link between the level of the yield curve and inflationary expectations, as suggested by the Fisher hypothesis. Similarly, Figure 7 displays the estimated slope factor, $\hat{\beta}_{2t}$ , and an indicator of macroeconomic activity (demeaned capacity utilization), whose correlation, -0.48, suggests that the yield curve slope is highly connected to the cyclical dynamics of the economy. With regard to the curvature, as this section will show, there is no reliable macroeconomic links to $\beta_{2t}$.

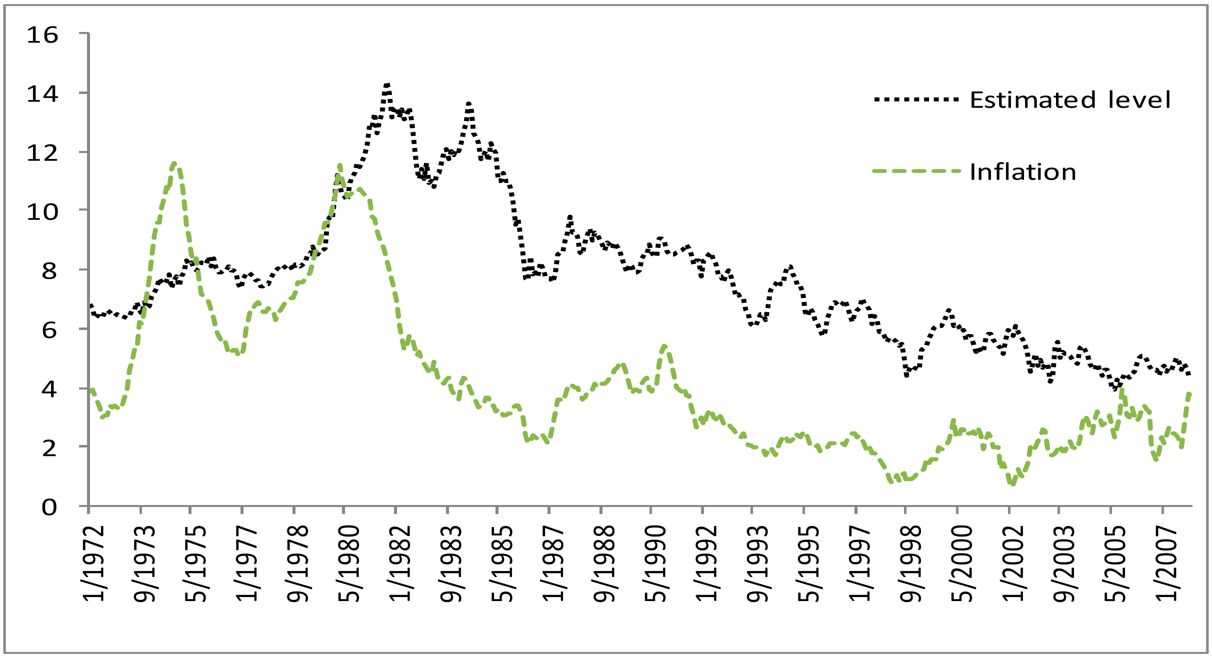
An expanded NSM model with macroeconomic variables could help explain macro-financial interactions and fits U.S. Treasury yield data well. As mentioned earlier, extending the NSM model to include three macroeconomic factors—manufacturing capacity utilization $(CU_t)$, the federal funds rate, $(FFR_t)$, and annual price inflation, $(INF_t)$—to the set of state variables under the assumption that the yields load only on the yield curve factors would provide a framework for studying the potential bilateral feedback between the yield curve and the macroeconomy. The time series of estimates of the level, slope and curvature factors in the yield-macro model are very similar to those obtained in the yield-only model.[49] Table 5 shows the descriptive statistics of the estimated factors in the yield-macro model, whose values and statistical properties are very similar to those of the yield-only model presented in Table 3. In addition, Table 6 displays the means and standard deviations of the measurements errors from the yield-macro model, which are also very similar to those of the yield-only model shown in Table 2. Specifically, the mean errors and the standard deviations are very small, suggesting a very good fit of the yield-macro model to the U.S. Treasury yield data.[50]

---

[49] The MCM Term Structure Software includes an additional subroutine that extends the econometric procedures used to estimate the yield-only model to estimate the yield factors of a model that include macroeconomic variables. See Appendix II.

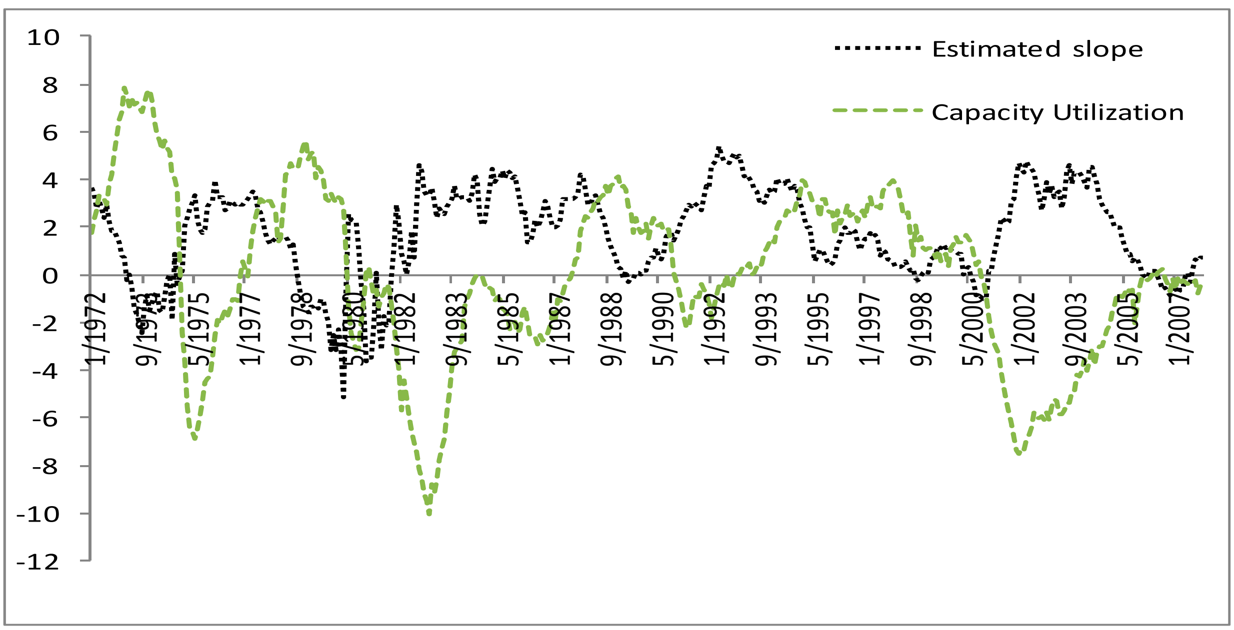
[50] As in the yield-only model, the estimated means and standard deviations of the residuals, expressed in basis points, show that the mean error is negligible at all maturities and that the average standard deviation for the relevant middle range of maturities from 6 to 60 months is very small—about 6.4 basis points.

Figure 6. Yield-only Model Level Factor and Inflation




Source: Fund staff estimates.

Figure 7. Yield-only Model Slope Factor and Capacity Utilization




Source: Fund staff estimates.

The NSM yield-macro model is able to capture the degree and the nature of the dynamic interactions between the economy and the yield curve. Table 7 presents the estimates of the parameters of the yield-macro model, which contains the key macroeconomic and yield curve interactions. Individually, all the diagonal elements are significant, while 12 out of 30 off-diagonal elements appear insignificant. In particular the 3x3 lower left block of the matrix of coefficients, showing the influence of macroeconomic factors on the yield curve factors, contains five insignificant coefficients, while the 3x3 upper right block of the matrix of coefficients, showing the influence of yield curve factors on macroeconomic factors, contains three insignificant coefficients.[51] Although results from the VAR estimates appear to show limited bilateral feedback between the yield curve and macroeconomic variables, a more thorough analysis of the relation between yield movements and shocks in macro variables, and vice versa, would use the impulse responses implied by the estimated VAR.

Table 5. United States: Yield-Macro Model Descriptive Statistics, Estimated Factors

| Factor | Mean | Std. Dev. | Minimum | Maximum | ρ(1) | ρ(12) | ρ(30) |
|---|---|---|---|---|---|---|---|
| $\hat{\beta}_1$ | 7.706 | 2.396 | 3.921 | 14.503 | 0.992 | 0.877 | 0.733 |
| $\hat{\beta}_2$ | 1.674 | 1.875 | -5.023 | 5.279 | 0.946 | 0.442 | -0.134 |
| $\hat{\beta}_3$ | -0.601 | 2.125 | -6.259 | 5.756 | 0.890 | 0.465 | 0.050 |

Source: Fund staff estimates.

Table 6. United States: Summary Statistics for Measurement Errors of Yields

| Maturity | Yield-Macro Model | |
|---|---|---|
| | Mean | Std. Dev. |
| 3 | -0.059 | 0.084 |
| 6 | 0.048 | 0.078 |
| 12 | 0.052 | 0.095 |
| 24 | 0.002 | 0.058 |
| 36 | -0.035 | 0.050 |
| 48 | -0.033 | 0.055 |
| 60 | -0.034 | 0.050 |
| 84 | 0.021 | 0.047 |
| 120 | 0.038 | 0.072 |

Source: Fund staff estimates.

[51] Specifically, using the partitioning of matrix A, as was introduced in Section II B, A=$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$, $A_2$ shows the influence of yield curve factors on macroeconomic factors, while $A_3$ shows the influence of macroeconomic factors on yield curve factors.

Table 7. United Sates: Yield-Macro Model Parameter Estimates
(Standard errors in parenthesis and t-statistics in brackets)

| | Level | Slope | Curvature | CU | INF | FFR |
|---|---|---|---|---|---|---|
| Level (-1) | **0.97** | -0.09 | **0.21** | **0.29** | **0.56** | -0.04 |
| | (0.02476) | (0.04906) | (0.08044) | (0.04287) | (0.04298) | (0.02178) |
| | [ 39.2045] | [-1.78343] | [ 2.64645] | [ 6.84723] | [ 12.9222] | [-1.78560] |
| Slope (-1) | 0.00 | **1.00** | **-0.18** | **-0.21** | **-0.52** | 0.03 |
| | (0.02398) | (0.04753) | (0.07793) | (0.04153) | (0.04164) | (0.0211) |
| | [-0.17922] | [ 20.9513] | [-2.32893] | [-5.04640] | [-12.5521] | [ 1.22017] |
| Curvature (-1) | **-0.03** | **-0.06** | **0.83** | 0.02 | **0.05** | **-0.02** |
| | (0.00872) | (0.01728) | (0.02834) | (0.0151) | (0.01514) | (0.00767) |
| | [-2.91824] | [-3.75630] | [ 29.4552] | [ 1.26505] | [ 3.27703] | [-2.02061] |
| CU (-1) | 0.00 | **-0.03** | 0.03 | **1.01** | **0.03** | **0.02** |
| | (0.00471) | (0.00934) | (0.01532) | (0.00816) | (0.00819) | (0.00415) |
| | [ 0.00117] | [-3.27940] | [ 1.82799] | [ 123.177] | [ 3.28826] | [ 5.83939] |
| INF (-1) | 0.02 | **0.12** | **-0.13** | **-0.26** | **0.48** | 0.03 |
| | (0.02085) | (0.04132) | (0.06775) | (0.03611) | (0.0362) | (0.01835) |
| | [ 1.06520] | [ 2.87320] | [-1.98930] | [-7.24906] | [ 13.3776] | [ 1.62117] |
| FFR (-1) | 0.01 | **-0.03** | 0.02 | -0.01 | **0.04** | **0.99** |
| | (0.00767) | (0.01519) | (0.02491) | (0.01328) | (0.01331) | (0.00675) |
| | [ 1.84976] | [-2.18756] | [ 0.66352] | [-0.70491] | [ 3.22852] | [ 147.270] |

Source: Fund staff estimates.

Results from the impulse response functions of the yield-macro model reveal complex and subtle dynamic interactions between macroeconomic variables and yield curve factors. Figure 8 displays the impulse response functions of the complete yield-macro system.[52] Four groups of impulse responses are considered: (i) responses of the macroeconomic variables to macroeconomic shocks; (ii) responses of the macroeconomic variables to yield curve shocks; (iii) responses of the yield curve to macroeconomic shocks; and (iv) responses of the yield curve to yield curve shocks.

- *Responses of the macroeconomic variables to macroeconomic shocks.* These are similar to those obtained in standard small macro models.[53] With the exception of the $FFR$, the macroeconomic variables show significant persistence.[54] Also, an increase in the $FFR$

[52] Producing impulse responses from the VAR model requires to assume a particular ordering of the variables. The order of the variables used in this paper is similar to that followed by Diebold, Rudebusch and Aroba (2006), i.e., $\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{1t}, CU_t, FFR_t$and $INF_t$. The term structure factors enter prior to the macroeconomic variables since they are dated at the beginning of the period. The results obtained are robust to a different ordering of the macroeconomic variables.

[53] See Diebold, Rudebusch and Aroba (2006) and the papers cited within.

[54] This result for the $FFR$ is not found in different Diebold, Rudebusch and Aroba (2006). However the period studied in Diebold et al. goes from 1972:01 to 2000:12.

lowers $CU$ over the following years. The $FFR$ rises with $CU$ and with $INF$ in a manner consistent with an estimated monetary policy reaction function. Lastly, $INF$ increases with $CU$, and has a negative delayed response to the $FFR$.

- *Responses of the macroeconomic variables to yield curve shocks.* The macroeconomic variables exhibit unimportant, and mostly insignificant, responses to shocks in the curvature factor. However, an increase in the slope produces an important decline in the $FFR$, signaling a close relation between the slope factor and the monetary policy instrument.[55] Finally, an increase in the level raises $CU$, $FFR$, and $INF$, where the latter result confirms the link between inflation and the level factor highlighted in Figure 6 in which the level factor is perceived as the bond market's expectation of long-run inflation. In this regard, an increase in the level reduces the ex-ante real rate of interest, $FFR - \beta_1$, followed by an increase in $CU$. However, during the period of analysis, the Fed accommodated only a fraction of the expected rise in inflation by increasing the nominal funds rate, reducing $CU$, and limiting the increase in $INF$ to a fraction of the initial shock to the level. establish

- *Responses of the yield curve to macroeconomic shocks*. Shocks to macroeconomic variables do have little, and mostly insignificant, influence on the curvature factor. On the other hand, a positive shock to the $FFR$ quickly increases the slope factor, making the yield curve steeper (or less negatively sloped).[56] Positive shocks to $CU$, and to a less degree inflation, produce an inverse and more delayed response in the slope, suggesting a bond market anticipation of monetary policy tightening. In addition, shocks to the macroeconomic variables influence the level of the term structure. In particular, shocks to $INF$ and $CU$ appear to generate a prolonged increase in the level factor, suggesting that long-term inflation expectations may not be firmly anchored. Finally, a positive shock to $FFR$ induces a small temporary increase in the level factor, suggesting that a surprise tightening may be signaling a Fed's concern with inflationary pressures in the economy, which may heighten inflationary expectations and an increase in the level factor.

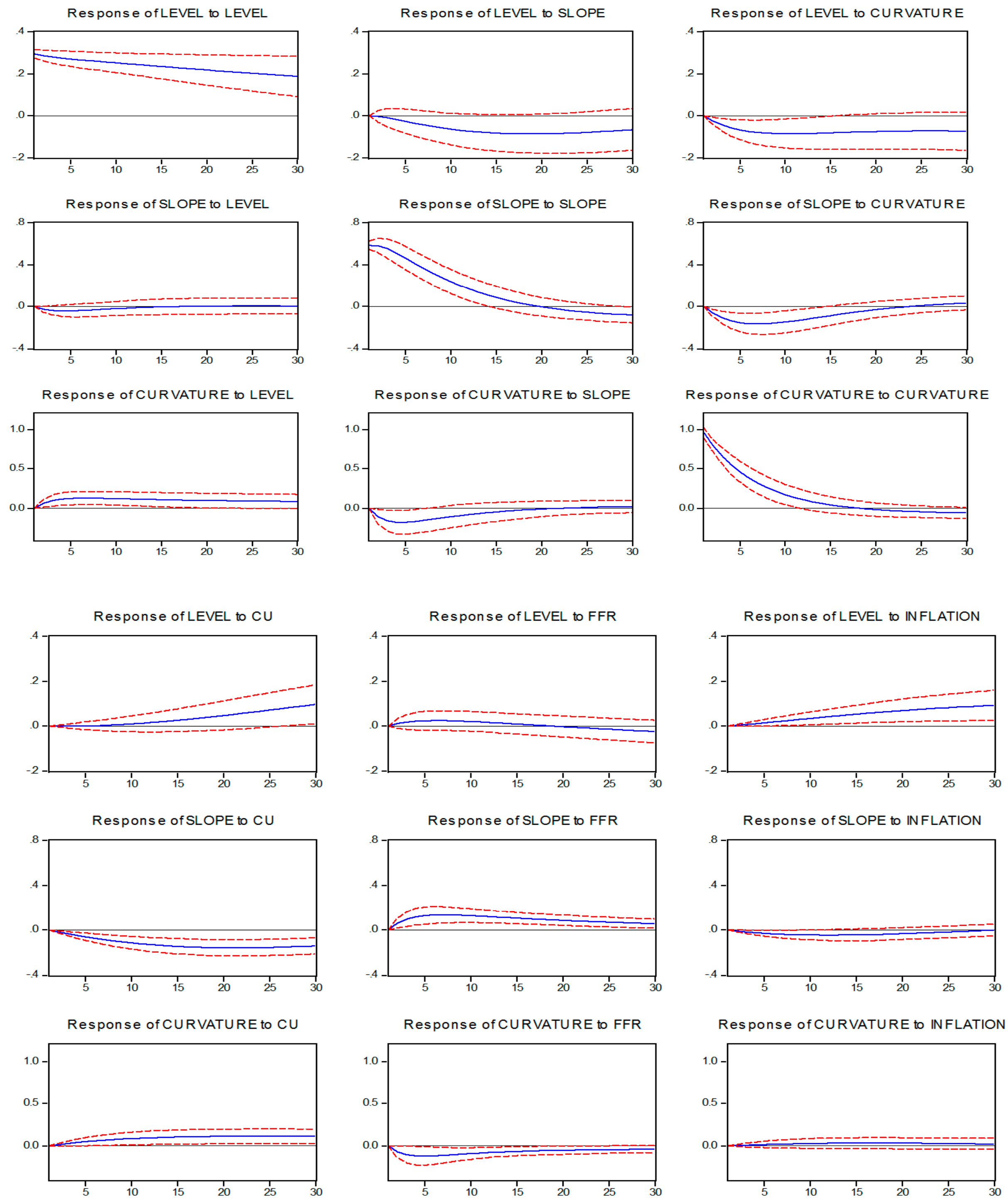
  *Responses of the yield curve to yield curve shocks*. The three yield curve factors display significant persistence. Also, a shock to the level factor, interpreted as higher inflation expectations, will increase the slope, which is associated with a lowering of the short end of the curve relative to the long end, and a loosening of monetary policy. Finally, a surprise increase in the slope factor reduces the level factor, suggesting a shift in the total curve downwards, but with a relatively stronger decline in the short end of the curve.

---

[55] As explained in Diebold, Rudebusch and Aroba (2006), there are two interpretations of this link: Either the Fed may be reacting to yields in setting the funds rate, or the yields are reacting to macroeconomic information in anticipation of the Fed decisions. This last possibility occurs when the Fed has been able to set up a predictable policy reaction to macroeconomic information.

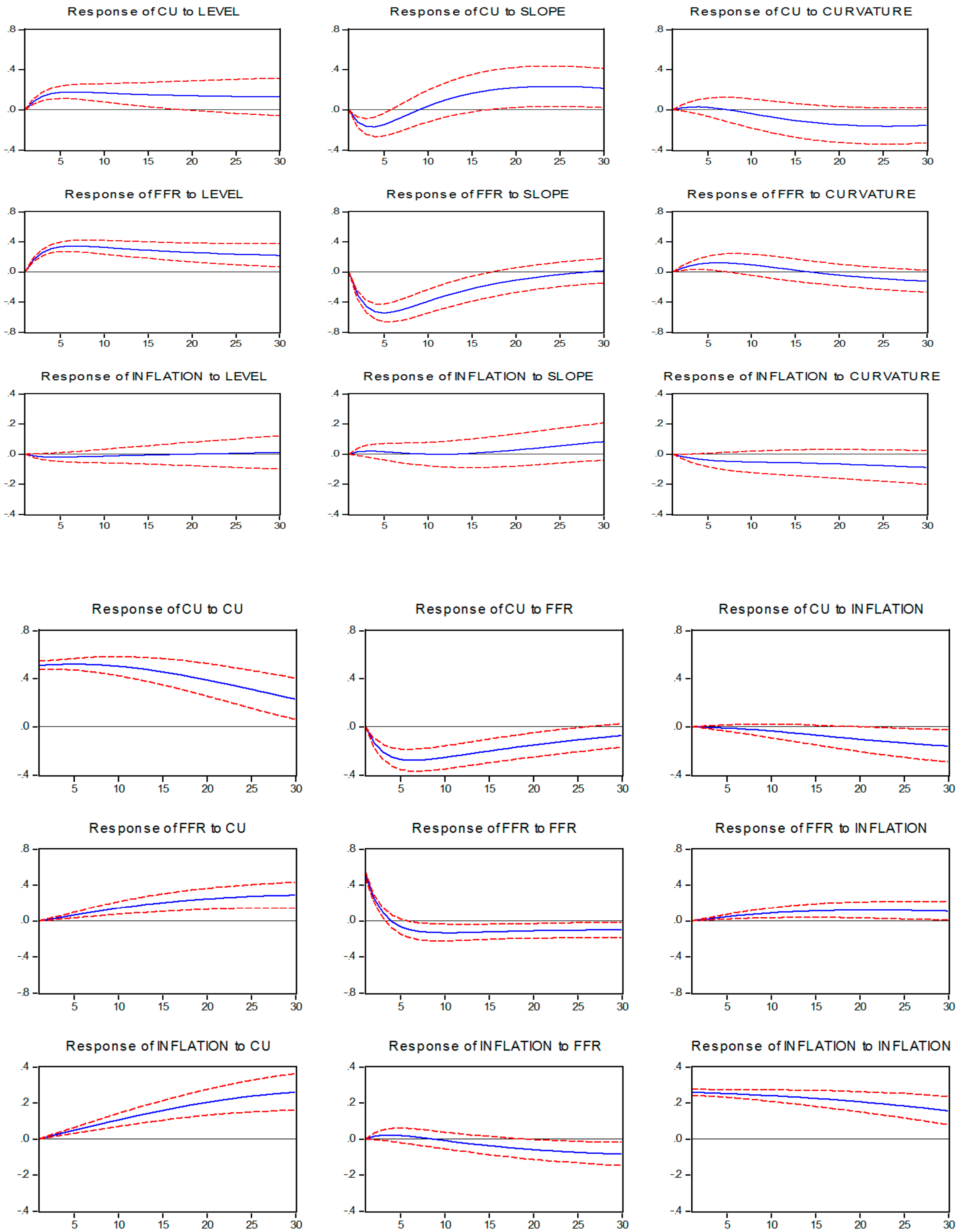
[56] This result differs from Diebold, Rudebusch and Aroba (2006) in which a positive shock to the $FFR$ almost immediately makes the yield curve less positively sloped (or more negatively sloped).

## Figure 8. Impulse-Response Functions




Source: Fund staff estimates.

## **Figure 8. Impulse-Response Functions** (concluded)



# IV. POSSIBLE EXTENSIONS

This section explores two possible extensions of the software**,** namely a multi-country and global yield curve analysis, and the introduction of no-arbitrage restrictions.

## A. Global yield curve dynamics and interactions

The Nelson-Siegel-type models surveyed in the previous sections specify and estimate a single country's yield curve in isolation, relating domestic yields to domestic yield factors and to macroeconomic factors. Recently, Diebold, Li, and Yue (2008) generalized the Nelson-Siegel approach to study the degree to which domestic yield dynamics are driven by the dynamics of both global and country specific factors. In particular, they construct a hierarchical dynamic factor model for a set of several countries' yield curves, in which country yields may depend on country factors, and country factors may depend on global factors.[57] This extension of the NSM framework permits to assess the existence of commonality in country-factor dynamics, and makes it possible to examine the extent to which the extracted global factors reflect developments in key macroeconomic variables during the sample period.

Diebold, Li, and Yue (2008) are able to extract global factors and country–specific factors, showing that global factors do in fact exist and are economically relevant, accounting for a significant fraction of variation in country bond yields. They also find evidence that global yield factors may be linked to global macroeconomic fundamentals (inflation and real activity), and that this appear more important for the period of larger global financial integration.

As in Diebold, Rudebusch, and Auroba (2006), Diebold, Li, and Yue (2008) use a one-step approach to estimate their global yield curve factor model by exploiting its state-space structure for both parameter estimation and factor structure. However, due to the large number of parameters to estimate in multi-country contexts,[58] they apply a Bayesian approach in which they use Markov Chain Monte Carlo methods to perform a posterior analysis of the model. Undoubtedly, extension of the software to include these methods would contribute to enhance the analytical and empirical tools to understand the workings of the global bond market, as well as to explore the extent to which variation in individual countries' yield curves come from global or idiosyncratic sources, with profound implications for the countries' macroeconomic policies. In the same vein, it would provide a

---

[57] Diebold, Li, and Yue (2008) use a monthly dataset of government bond yields for Germany, Japan, the US and the UK from 1985:9 to 2005:8.

[58] Diebold, Li, and Yue (2008) also estimate a total of 257 parameters in each of their separate global level and slope models.

tool to facilitate the monitoring and assessment of economic and financial interconnections, and to increase the understanding of international policy spillovers.[59]

## B. No-Arbitrage Restrictions

In addition to the Nelson-Siegel-type models reviewed above, a second approach to construct bond yield factors and factor loadings is the no-arbitrage dynamic latent-factor model, which is widely used in the finance literature. Since, for the case of advanced economies, bond trading occurs in deep well-organized markets, the restriction ruling out remaining riskless arbitrage opportunities across maturities and over time has been central to the literature on the empirical analysis of bond pricing. Unfortunately, as indicated by Duffee (2002), these types of models do not perform well empirically, in particular with respect to out-of-sample forecasting.[60] Dynamic versions of Nelson-Siegel-type model, despite their good empirical performance, do not impose the theoretical restriction of absence of arbitrage.

Recently, Christensen, Diebold, and Rudebusch (2007) have integrated the Nelson-Siegel model with the absence of arbitrage by constructing an affine arbitrage-free model that maintains the Nelson-Siegel factor structure for the yield curve, and that exhibits superior empirical forecasting performance. Clearly, extension of the software to include the arbitrage-free Nelson-Siegel would open the way to use the latest state-of–the-art analytical and empirical tools for studying the structure and dynamics of countries' yield curves.

# V. Conclusions

This paper discusses the estimation of models of the term structure of interest rates. In this context, this paper first summarizes some of the main models of term structure, namely the Nelson-Siegel models and the Affine Term-Structure models, perhaps the most widely used models by market participants and central bank officials. The paper then presents estimations of the terms structure of the U.S. Treasury bond yields from 1972 to 2007. In line with the findings in the literature, the paper concludes that:

- The U.S. yield curve for the period 1972:1-2007:12 exhibits large variation across all its maturities, which, in turn, can be characterized in terms of the variation of its three key factors, namely the level, slope and curvature. These factors both link the yields at different maturities at any given moment and restrict their dynamic evolution in a systematic way from which stylized facts can be identified.

---

[59] See Lipsky (2010).

[60] An additional problem, as reported by Christensen, Diebold, and Rudebusch (2007, 2008) is that the estimation of these models is adversely affected by the existence of several model likelihood maxima that fit the data similarly, but have different implications for economic behavior.

- An estimated three-factor Nelson-Siegel model for the yield curve reproduces well the stylized facts of the U.S. yield curve for the period of analysis, and provides a framework for assessing its dynamic evolution. However, a meaningful interpretation of the yield curve dynamics requires exploring the links of yield curve factors with macroeconomic variables. This can be achieved by expanding the Nelson-Siegel methodology to include macroeconomic factors in addition to the yield curve ones.

- An estimated six-factor Nelson-Siegel model fits the U.S. yield curve well and provides information on the nature of macro-financial linkages for the period of analysis. The analysis of the dynamic interactions between macroeconomic and yield curve factors is widened by the information provided by the impulse response functions of the yield-macro model. One important conclusion of the analysis is that market yields contain relevant predictive information about the Fed's policy rate.

The paper uses a software developed by Fund staff to estimate the term structure of Treasury bond yields for the United States. This software makes it possible to estimate at least nine term structure models of interest rates, focusing particularly on Nelson-Siegel models and Affine-Term Structure models. These models make use of state-of-the-art solution techniques. Even though it uses C# to solve the term structure models, the software relies on a friendly Excel interface.

The paper proposes possible extensions of the software. In particular, the paper argues that a promising extension of the software is to include global factors that would make it possible to understand better the dynamics of the domestic yield curve. This extension would open the way to identify common factors that are important in the determination of the domestic yield curve across different countries. The paper also notes that an extension of the Nelson-Siegel model to include the absence of arbitrage could enrich the structure and dynamics of the countries' yield curves.

## References

Ang, A. and M. Piazzesi, 2003, "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics,* 50, 745–787.

Balduzzi, P., S. R. Das, S. Foresi, and R. Sundaran, 1996, "A Simple Approach to Three-Factor Affine Term Structure Models," *The Journal of Fixed Income,* 6, 14–31.

Bank of International Settlements (BIS), 2005, "Zero-Coupon Yield Curves: Technical Documentation," (Basle: Bank for International Settlements).

Baz, J., and G. Chacko, 2004, *Financial Derivatives: Pricing, Applications and Mathematics* (Cambridge, United Kingdom: Cambridge University Press).

Bjork, T. and B. Christensen, 1999, "Interest Rate Dynamics and Consistent Forward Rate Curves," *Mathematical Finance*, 9, 323–348.

Bliss, R. R., 1997a, "Movements in the Term Structure of Interest Rates," *Economic Review*, 82, 16–33. (Atlanta: Federal Reserve Bank of Atlanta).

Bliss, R. R., 1997b, "Testing Term Structure Estimation Methods," *Advances in Futures and Options Research*, 9, 197–231.

Bolder, D. J., 2001, "Affine Term-Structure Models: Theory and Implementation," Working Paper 2001-15 (Ottawa: Bank of Canada).

Bolder, D. and D. Strzelecki, 1999, "Yield Curve Modeling at the Bank of Canada," Bank of Canada Technical Report No. 84 (Ottawa: Bank of Canada).

Campbell, J. Y., A. W. Low and A. C. MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton, New Jersey: Princeton University Press).

Chaplin, G. and K. Sharp, 1993, "Analytic Solutions for Bond and Bond Options Under Correlated Processes," Institute of Insurance and Pension Research, Research Report 93-16, University of Waterloo.

Chen R. R. and L. Scott, 1993, "Maximum Likelihood Estimation for a Multi-Factor Equilibrium Model of the Term Structure of Interest Rates," *The Journal of Fixed Income,* 4, 14–31.

Christensen J., F. Diebold and D. Rudebusch, 2008, "An Arbitrage-Free Generalized Nelson-Siegel Term Structure Model," *Econometrics Journal*, 1, 1-31.

Christensen J., F. Diebold and D. Rudebusch, 2007, "The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Model," (San Francisco: Federal Reserve Bank of San Francisco).

Cochrane, J. H., 2001, *Asset Pricing* (Princeton, New Jersey: Princeton University Press).

Cox, J. C., J. E. Ingersoll and S. A. Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica,* 53, 385-407.

Dai, Q. and K. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *The Journal of Finance,* 55, 1943-1978.

De Pooter M., 2007, "Examining the Nelson-Siegel Class of Term-Structure Models," Tinbergen Institute Discussion Paper.

Diebold, F. X., C. Li and V.Z. Yue, 2008, "Global Yield Curve Dynamics and Interactions: A Dynamic Nelson-Siegel Approach," *Journal of Econometrics*, 146, 351–363.

Diebold, F. X. and C. Li, 2006, "Forecasting the Term Structure of Government Bond Yields," *Journal of Econometrics*, 130, 337–364.

Diebold, F. X., G. D. Rudebusch and B. Auroba, 2006, "The Macroeconomy and the Yield Curve: a Dynamic Latent Factor Approach," *Journal of Econometrics*, 131, 309–338.

Diebold, F. X., M. Piazzesi and G. D. Rudebusch, 2005, "Modeling Bond Yields in Finance and Macroeconomics," *American Economic Review Papers and Proceedings*, 95, 415–420.

Duffee, G., 2002, "Term Premia and Interest Rate Forecasts in Affine Models," *The Journal of Finance*, 57, 405–443.

Duffie, D., 2001, *Dynamic Asset Pricing Theory*, Third Edition (Princeton, New Jersey: Princeton University Press).

Duffie, D., and R. Kan, 1996, "A Yield-Factor of Interest Rates," *Mathematical Finance,* 6, 379-404.

Duffie, D., and R. Kan, 1994, "Multi-Factor Term Structure Models," *The Royal Society,* 347 577-586.

Fama, E. and R. Bliss, 1987, "The Information in Long-Maturity Forward Rate, *American Economic Review*, 77, 680–692.

Filipovic, D. and E. Mayerhofer, 2009, "Affine Diffusion Processes: Theory and Applications," in *Advanced Financial Modelling,* edited by H. Albrecher, W. J. Runggaldier and W. Schachermayer (Berlin: Walter de Gruyter).

Hamilton, J., 1994, *Time Series Analysis* (Princeton, New Jersey: Princeton University Press).

Harvey, A.C., 1981, *Time Series Models*, Second Edition (Cambridge, Massachusetts: MIT Press, 1993).

Kahl, C., M. Gunther and T. Rossberg (2004), "Structure Preserving Stochastic Integration Schemes in Interest Rate Derivative Modeling," Bergische Universität Wuppertal, BUW-AMNA 04/10.

Lipsky, J., 2010, "Reconsidering the International Monetary System," Remarks at the Federal Reserve Bank of Kansas City symposium on "Macroeconomic Challenges: The Decade Ahead," Jackson Hole, Wyoming, August 26-28.

Litterman, R. and J. Scheinkman, 1991, "Common Factors Affecting Bond Returns," *The Journal of Fixed Income*, 1, 54–61.

Nawalkha, S. K., G. M. Soto and N. A. Believa, 2005, *Interest Rate Risk Modeling* (Hoboken, New Jersey: John Wiley & Sons, Inc.).

Nelson, C. R. and A. F. Siegel, 1987, "Parsimonious Modeling of Yield Curves," *Journal of Financial and Quantitative Analysis*, 60, 473–489.

Powell, M. J. D., 1964 "An Efficient Method of Finding the Minimum of a Function of Several Variables Without Calculating Derivatives," *Computer Journal,* 7, 152-162.

Rebonato, R., 1996, *Interest-Rate Option Models* (Hoboken, New Jersey: John Wiley & Sons, Inc.).

Rudebusch, G. and T. Wu, 2003, "A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy," Federal Reserve Bank of San Francisco Working Paper 03-17 (San Francisco: Federal Reserve Bank of San Francisco).

Rudebusch, G. and L.E.O. Svensson, 1999, "Policy Rules for Inflation Targeting," in *Monetary Policy Rules*, edited by J. B. Taylor (Chicago: University of Chicago Press).

Shreve, S., 2004, *Stochastic Calculus for Finance II: Continuous-Time Models* (New York:

Springer Finance).

Siegel, A. F. and C. R. Nelson, 1988, "Long-Term Behavior of Yield Curves," *Journal of Business*, 23, 105–110.

Svensson, L. E. O., 1994, "Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994," NBER Working Paper Series, 4871.

Van Deventer, D. R., K. Imai, and M. Mesler, 2005, *Advanced Financial Risk Management* (Singapore: John Wiley & Sons (Asia) Pte Ltd).

# Appendix I. Term-Structure Modeling Using the MCM Term Structure Software

After providing a general overview of term-structure modeling under Absence of Arbitrage (AOA), this appendix summarizes the mathematics behind the models for the term structure models, particularly the CIR, included in the MCM Term Structure Software (MCMTS). The summary in this section uses a somewhat different notation than in Section II.

## I. The Short Rate

Assume that there exists an instantaneous riskless interest rate. This rate is called the "short rate", and its value at time $t$ is noted $r_t$. More formally, let $r_{t,t+\tau}$ be the riskless interest rate at time $t$ for a loan with maturity date $t+\tau$. The short rate at time $t$ is defined as

(1) $$r_t = \lim_{\tau \to 0} r_{t,t+\tau}$$

This definition makes it clear that the short rate is not comparable to short-term rates quoted in real-world markets. The short rate is the riskless interest rate for an infinitesimal time to maturity. Since there is no instantaneous interest rate in the real world, the short rate is not observable. Therefore, for practical applications, it must be either estimated or replaced by a proxy.

## II. Equivalent Martingale Measure

Let *P* be the probability measure that represents uncertainty in the "real world." *P* is often called the market measure.

Assume that there exists a probability measure *Q* with the following properties

1. *Q* is equivalent to *P* in the sense that, for any event *E*,
$$P(E) > 0 \Leftrightarrow Q(E) > 0$$ [61]

   In other words, *P* and *Q* assign zero probability to the same events.
2. Discounted security prices are martingales under *Q*.

A measure *Q* with these two properties is called an *equivalent martingale measure*. In a continuous-time setting, the existence of such a measure guarantees the absence of

[61] Formally, uncertainty is represented by a probability space $\Omega, F, P$ where $F$ is a tribe on $\Omega$. A measure $Q$ is said to be equivalent to the measure $P$ if, for any element $E$ of the tribe $F$, $P(E) = 0 \Leftrightarrow Q(E) = 0$.

arbitrage.[62] This result is sometimes called the "First Fundamental Theorem of Asset Pricing."

The martingale property of discounted security prices under $Q$ can be formulated more precisely. Let $S_t$ be the price at time $t$ of a security that does not pay any dividend or coupon,[63] let $r_t$ be the short rate at time $t$, and define the *discount factor process* $D_t$ as

(2) $$D_t = \exp\left( -\int_{u=0}^{t} r_u \cdot du \right)$$

The discounted security price process is defined as $D_t \cdot S_t$. Under the equivalent martingale measure $Q, D_t \cdot S_t$ is a martingale. Therefore, for any $\tau \geq 0$

(3) $$E_t^Q \left( D_{t+\tau} \cdot S_{t+\tau} \right) = D_t \cdot S_t$$

Using the definition of $D_t$, this equality can be rewritten as

(4) $$E_t^Q \left[ \exp\left( -\int_{u=0}^{t+\tau} r_u \cdot du \right) \cdot S_{t+\tau} \right] = \exp\left( -\int_{u=0}^{t} r_u \cdot du \right) \cdot S_t$$

Therefore

(5) $$E_t^Q \left[ \exp\left( -\int_{u=t}^{t+\tau} r_u \cdot du \right) \cdot S_{t+\tau} \right] = S_t$$

The fact that $D_t \cdot S_t$ is a martingale can also be expressed as

(6) $$E_t^Q \left[ d\left( D_t \cdot S_t \right) \right] = 0$$

When $r_t$ and $S_t$ follow Ito processes, this equation can be rewritten as

(7) $$E_t^Q \left( -r_t \cdot D_t \cdot S_t \cdot dt + D_t \cdot dS_t \right) = 0 \Leftrightarrow E_t^Q \left[ D_t \cdot \left( -r_t \cdot S_t \cdot dt + dS_t \right) \right] = 0$$

---

[62] See, for example, Duffie (2001), Chapter 6, or Shreve (2004), Chapter 5. In a finite-dimensional setting, the existence of an equivalent martingale measure is equivalent to the absence of arbitrage. In an infinite-dimensional setting, the existence of an equivalent martingale measure is equivalent to the absence of "approximate arbitrage." See Duffie (2001), Chapter 6, for details.

[63] The analysis can easily be extended to securities that pay dividends or coupons.

Since $D_t > 0$, we finally obtain

(8) $$E_t^Q\left(dS_t\right) = r_t \cdot S_t \cdot dt$$

Therefore, under the equivalent martingale measure *Q*, the instantaneous expected rate of return on any security is equal to the short rate. This is the reason why *Q* is often called the "risk-neutral" measure: expected returns under *Q* are the same as in an artificial risk-neutral world.

Equations (5) and (6) are general pricing equations under AOA; they apply to all types of securities (with slight modifications for securities that pay dividends or coupons). Looking forward, these equations will be applied to zero-coupon bonds in order to determine the term-sctructure of interest rates.

In the remainder of this appendix, $W^P$ represents a Brownian motion under $P$, and $W^Q$ represents a Brownian motion under $Q$. The increments of $W^P$ and $W^Q$ over an infinitesimal interval of time $dt$ are noted $dW^P$ and $dW^Q$, respectively. $dW^P$ and $dW^Q$ are Normal random variables with mean 0 and variance $dt$.

## III. EXISTENCE OF AN EQUIVALENT MARTINGALE MEASURE

In the previous section, the existence of an equivalent martingale measure was assumed. The objective of this section is to construct an equivalent martingale measure. That is, given the market measure *P*, we wish to find a probability measure *Q* with the following properties:

1. *Q* is equivalent to *P*;
2. Discounted security prices are martingales under *Q*.

For simplicity, this section makes the following assumptions:

1. There is a fixed horizon *T*;
2. Uncertainty is generated by a single one-dimensional Brownian motion $W^P$.

Consider a stochastic process $\theta$ such that

(9) $$\int_{t=0}^{T} \theta_t^2 \cdot dt < \infty \text{ a.s.}$$
$$E^P\left[\exp\left(\frac{1}{2} \int_{t=0}^{T} \theta_t^2 \cdot dt\right)\right] < \infty$$

It can be shown that the following process is a martingale under $P$

(10)
$$\xi_t = \exp\left( -\int_{u=0}^{t} \theta_u \cdot dW_u^P - \frac{1}{2} \int_{u=0}^{t} \theta_u^2 \cdot du \right)$$

In addition, the probability measure $Q$ defined by the following equation is equivalent to $P$ [64]

(11)
$$\frac{dQ}{dP} = \xi_T$$

$dQ/dP$ is the Radon-Nikodym derivative of $Q$ with respect to $P$.

Furthermore, the following result holds:

Girsanov's Theorem: Consider the process $W_t^Q$ defined as

(12)
$$W_t^Q = W_t^P + \int_{u=0}^{t} \theta_u \cdot du$$

When the process $\theta$ satisfies this process, the following process is a Brownian motion (and hence a martingale) under $Q$

(13)
$$W_t^Q = W_t^P + \int_{u=0}^{t} \theta_u \cdot du$$

Consider a security price $S$ that follows a geometric Brownian motion under $P$

(14)
$$dS_t = \mu_S \cdot S_t \cdot dt + \sigma_S \cdot S_t \cdot dW_t^P$$

As shown in the previous section, the discounted security price process under $P$ is given by

(15)
$$\begin{aligned} d\left(D_t \cdot S_t\right) &= -r_t \cdot D_t \cdot S_t \cdot dt + D_t \cdot dS_t \\ &= -r_t \cdot D_t \cdot S_t \cdot dt + D_t \cdot \left(\mu_S \cdot S_t \cdot dt + \sigma_S \cdot S_t \cdot dW_t^P\right) \\ &= \left(\mu_S - r_t\right) D_t \cdot S_t \cdot dt + \sigma_S \cdot D_t \cdot S_t \cdot dW_t^P \end{aligned}$$

From Girsanov's Theorem, this equation can be rewritten as

---

[64] See, for example, Duffie (2001), Appendix D.

(16)
$$d\left(D_t\cdot S_t\right)=\left(\mu_S-r_t\right)\cdot D_t\cdot S_t\cdot dt+\sigma_S\cdot D_t\cdot S_t\cdot\left(dW_t^Q-\theta_t\cdot dt\right)$$
$$=\left(\mu_S-r_t-\sigma_S\theta_t\right)\cdot D_t\cdot S_t\cdot dt+\sigma_S\cdot D_t\cdot S_t\cdot dW_t^Q$$

This is the discounted security price process under *Q*. We are looking for a probability measure *Q* such that $D\cdot S$ is a martingale under *Q*. In other words, we are looking for a probability measure *Q* such that the drift of $D\cdot S$ is zero under *Q*. This leads to the following condition

(17)
$$\left(\mu_S-r_t-\sigma_S\cdot\theta_t\right)\cdot D_t\cdot S_t\cdot dt=0$$

This condition must be satisfied by any security price process. Therefore, the process $\theta$ must satisfy the following condition for any security price process *S*

(18)
$$\mu_S-r_t-\sigma_S\cdot\theta_t=0\Leftrightarrow\frac{\mu_S-r_t}{\sigma_S}=\theta_t$$

The process $\theta$ defined in this way is called the *market price of risk*.

To summarize, consider the probability measure *Q* defined by

(19)
$$\frac{dQ}{dP}=\xi_T=\exp\left(-\int_{t=0}^{T}\theta_t\cdot dW_t^P-\frac{1}{2}\cdot\int_{t=0}^{t}\theta_t^2\cdot dt\right)$$

where $\theta$ is the market price of risk process. This probability measure is equivalent to *P*, and discounted security prices are martingales under *Q*.

## IV. Equivalent Martingale Measures and Monte-Carlo Simulations

It must be emphasized that the equivalent martingale measure *Q* does not represent real-world uncertainly. *Q* is an artificial probability measure that has no relevance in the real world. Put differently, information revelation in the real world is represented by the filtration of the state space generated by the Brownian motion $W^P$; in contrast, the filtration of the state space generated by the Brownian motion $W^Q$ represents information revelation in an artificial risk-neutral world. For these reasons, performing Monte-Carlo simulations under *Q* is not always meaningful.

As shown by equation (5), arbitrage-free security prices can be represented as expectations under *Q*. In many cases, these expectations must be estimated by Monte-Carlo simulation because they do not have analytical representations. When the purpose of Monte-Carlo simulations is to compute expectations under *Q*, obviously, performing the simulations under

$Q$ is perfectly legitimate. In other words, Monte-Carlo simulations under $Q$ are appropriate for pricing purposes.

This, however, is not the case when the purpose of the simulations is to compute the VaR of a portfolio. As mentioned above, the martingale measure $Q$ and the market measure $P$ are only equivalent in the sense that they assign 0 probability to the same events. In general, the distribution of portfolio values under $Q$ is not the same as the distribution of portfolio values under $P$. Since a VaR is just a quantile of the distribution of portfolio values, it is not the same under $Q$ as under $P$. In this case, the relevant probability measure is the market measure $P$; the martingale measure $Q$ does not provide any information about the distribution of portfolio values in the "real-world." Therefore, when the purpose of Monte-Carlo simulations is to compute a VaR, the simulations should be performed under the objective market measure $P$.

## V. Single-Factor Term Structure Models

Let $B(t,\tau)$ be the price at time $t$ of a zero-coupon bond that pays one unit of currency at time $T = t+\tau$, and let $x_t$ be a scalar stochastic process. A one-factor term-structure model is a function $b$ such that

(20) $$B(t,\tau) = b(x_t, t, t+\tau, \beta) + e(t,\tau), \quad \tau \geq 0$$

where $\beta$ is a vector of parameters to be estimated, $x_t$ is the single factor that is assumed to drive the entire term-structure, and $e(t,\tau)$ is a pricing error. If the model is "true", then $e(t,\tau) = 0$ for any $t$ and any $\tau$. In the remainder of this section, pricing errors are assumed to be null.

This section shows that, when there exists an equivalent martingale measure, the function $b(x_t, t, T, \beta)$ is the solution of a particular stochastic partial differential equation (PDE).

Let $W^Q$ represent a Brownian motion under $Q$. Assume that the process of $x$ under $Q$ is the following Ito process

(21) $$dx_t = \mu_{x,Q} dt + \sigma_{x,Q} dW_t^Q$$

Assuming that $b$ is a smooth, twice-differentiable function of $x$ and a differentiable function of $t$, Ito's lemma implies

(22) $$db\left(x_t,t,T,\beta\right)=\left[\frac{\partial b}{\partial t}+\frac{\partial b}{\partial x_t}\cdot\mu_{x,Q}+\frac{1}{2}\cdot\sigma_{x,Q}^2\cdot\frac{\partial^2 b}{\partial x_t^2}\right]\cdot dt+\frac{\partial b}{\partial x_t}\cdot\sigma_{x,Q}\cdot dW_t^Q$$

Under *Q*, the instantaneous expected rate of return on any security is equal to $r_t$. This means in particular that the drift of $db\left(x_t,t,T,\beta\right)$ must be equal to $r_t b\left(x_t,t,T,\beta\right)$. Therefore

(23) $$\frac{\partial b}{\partial t}+\frac{\partial b}{\partial x}\cdot\mu_{x,Q}+\frac{1}{2}\cdot\sigma_{x,Q}^2\cdot\frac{\partial^2 b}{\partial x^2}=r_t\cdot b\left(x_t,t,T,\beta\right)$$

Single-factor models of the term structure generally assume that the single factor that drives the term structure is the short rate

(24) $$x_t=r_t \qquad \forall t$$

This is true, for example, in the Cox-Ingersoll-Ross (CIR) models. Equation (23) then becomes

(25) $$\frac{\partial b}{\partial t}+\frac{\partial b}{\partial r}\cdot\mu_{r,Q}+\frac{1}{2}\cdot\sigma_{r,Q}^2\cdot\frac{\partial^2 b}{\partial r^2}=r_t\cdot b\left(x_t,t,T,\beta\right)$$

This equation is a stochastic partial differential equation (PDE) known as the Backward Kolmogorov equation. It is also often referred to as the "no-arbitrage" pricing equation.[65] The solution $b$ to this equation is the term structure function under AOA.

The no-arbitrage pricing equation is subject to the following boundary condition

(26) $$b\left(r_t,T,T,\beta\right)=1$$

A general solution to the no-arbitrage pricing equation is provided by the Feynman-Kac formula

(27) $$b\left(r_t,t,T,\beta\right)=E_t^Q\left[\exp\left(-\int_{u=t}^{T} r_u d_u\right)\right]$$

---

[65] See Rebonato (1996), Chapter 7, Section 1, for a derivation of this equation without using the concept of equivalent martingale measure. While we focus here in the prices of zero-coupon bonds, we note that this equation is valid for the price of any security whose cash flows can be expressed as a function of an Ito process.

Note that this is just equation (5) applied to a security that pays one unit of currency at time *T*: in this case, $t+\tau=T$ and $S_{t+\tau}=1$.

Depending on the process of $r$, there may or may not be an analytical solution to equation (25) and a closed-form expression for the conditional expectation on the right-hand side of equation (27). When the short rate process is a simple mean-reverting process or a mean-reverting square root process (as in the CIR model), there is a closed-form expression for zero-coupon bond prices.

To summarize, single-factor arbitrage-free models of the term structure are derived as follows:

1. The existence of an equivalent martingale measure *Q* is assumed;
2. The existence of *Q* guarantees AOA;
3. The single factor that drives the entire term structure is assumed to follow an Ito process under *Q*;
4. Ito's lemma is used to derive an expression for the drift of a zero-coupon bond price process under *Q*;
5. Since *Q* is an equivalent martingale measure, any security's instantaneous expected rate of return under *Q* must be equal to the short rate; this provides a second expression for the drift of a zero-coupon bond price process under *Q*;
6. Combining the two expressions for the drift of a zero-coupon bond price process under *Q* yields the no-arbitrage PDE;
7. The solution to the no-arbitrage PDE is the term structure function under AOA.

## VI. The Cox-Ingersoll-Ross Model

The Cox-Ingersoll-Ross (CIR) model of the term structure assumes that the single factor is the short rate, and that it follows a mean-reverting square root process.

More precisely, the short rate process under *P* is assumed to behave as

(28)
$$dr_t=\kappa\left(\theta-r_t\right)dt+\sigma\sqrt{r_t}\,dW_t^P$$

where $\kappa$ is the mean reversion rate, $\theta$ is the long-term mean of the short rate, and $\sigma$ is a diffusion (volatility) parameter.

Suppose that the short rate at time $0$ is $r_0>0$. The short rate at time $t$ is then given by

(29)
$$r_t=r_0+\int_{u=0}^{t}\kappa\left(\theta-r_u\right)du+\int_{u=0}^{t}\sigma\sqrt{r_u}\,dW_u^P$$

When $\kappa > 0$ and $\theta > 0$, the short rate is non-negative. When $\kappa \cdot \theta \geq \sigma^2/2$, the short rate is strictly positive at any time.

From Girsanov's theorem, the corresponding process of the short rate under the martingale measure $Q$ is

(30) $$dr_t = \kappa\left(\theta - r_t - \lambda\sigma\sqrt{r_t}\right)dt + \sigma\sqrt{r_t}dW_t^Q$$

where $\lambda$ is the market price of risk.

In the CIR model, the vector of parameters that need to be estimated is $\beta = (\kappa, \theta, \sigma, \lambda)$. Under the assumption that the short rate follows a mean-reverting square root process, the no-arbitrage PDE can be solved analytically, and $b(r_t, t, t+\tau, \beta)$ is given by[66]

(31) $$b(r_t, t, t+\tau, \beta) = \exp\left[A(\tau,\beta) - C(\tau,\beta) \cdot r_t\right]$$

where

(32)
$$C(\tau,\beta) = \frac{2\left(e^{\gamma\tau} - 1\right)}{(\gamma + \kappa + \lambda)\left(e^{\gamma\tau} - 1\right) + 2\gamma}$$
$$A(\tau,\beta) = \ln\left[\frac{2e^{\frac{(\gamma+\kappa+\lambda)\tau}{2}}}{(\gamma + \kappa + \lambda)\left(e^{\gamma\tau} - 1\right) + 2\gamma}\right]^{\frac{2\kappa\theta}{\sigma^2}}$$
$$\gamma = \sqrt{(\kappa+\lambda)^2 + 2\sigma^2}$$

Note that, in the CIR model, $b(r_t, t, t+\tau, \beta)$ actually does not depend on $t$, so that it can be rewritten $b(r_t, \tau, \beta)$.

The continuously compounded yield to maturity of the zero-coupon bond with maturity $t+\tau$ is $y(t,\tau,\beta)$, and serves as the solution of the following equation

(33) $$\exp\left[-y(t,\tau,\beta) \cdot \tau\right] = b(r_t, t, t+\tau, \beta)$$

---

[66] See, for example, Bolder (2001) or Rebonato (1996).

Therefore

(34) $$y(r_t,\tau,\beta)=\frac{1}{\tau}\left[C(\tau,\beta)\cdot r_t - A(\tau,\beta)\right]$$

This is the continuously compounded yield to maturity at time *t* of a zero-coupon bond maturing at $t+\tau$, according to the single-factor CIR model. Note that $y(r_t,\tau,\beta)$ is an affine function of $r_t$; this makes the CIR model a member of the affine term-structure model family.

## VII. Affine Multi-Factor Models

Assume that zero-coupon bond prices are now functions of two factors $x_1$ and $x_2$. For example, the price at time *t* of a zero-coupon bond that pays 1 at $T=t+\tau,\ \tau>0$ is now written as

(35) $$b(x_{1,t},x_{2,t},t,t+\tau)=b(x_{1,t},x_{2,t},t,T)$$

In the rest of this section, $b_t$ will often be used as an abbreviation for $b(x_{1,t},x_{2,t},t,t+\tau)$.

As before, assume that there exists an equivalent Martingale measure *Q*. The factors $x_1$ and $x_2$ are assumed to follow Ito processes

(36) $$\begin{aligned} dx_{1,t} &= \mu_{1,t}dt+\sigma_{1,t}dW_{1,t}^{Q} \\ dx_{2,t} &= \mu_{2,t}dt+\sigma_{2,t}dW_{2,t}^{Q} \end{aligned}$$

where $W_1^Q$ and $W_2^Q$ are independent Brownian motions under $Q$.

From Ito's Lemma, if $b(\ )$ is twice-continuously-differentiable in $x_1$ and $x_2$, then the bond price satisfies the following stochastic differential equation

(37)
$$\begin{aligned} db(x_{1,t},x_{2,t},t,t+\tau) &= \left[\frac{\partial b}{\partial t}+\frac{\partial b}{\partial x_1}\cdot\mu_{1,t}+\frac{\partial b}{\partial x_2}\cdot\mu_{2,t}+\frac{1}{2}\cdot\sigma_{1,t}^2\cdot\frac{\partial^2 b}{\partial x_1^2}+\frac{1}{2}\cdot\sigma_{2,t}^2\cdot\frac{\partial^2 b}{\partial x_2^2}\right]\cdot dt \\ &\quad +\frac{\partial b}{\partial x_1}\cdot\sigma_{1,t}\cdot dW_{1,t}^{Q}+\frac{\partial b}{\partial x_2}\cdot\sigma_{2,t}\cdot dW_{2,t}^{Q} \end{aligned}$$

The process for the discounted bond price is then given by

(38)

$$
\begin{aligned}
d\left(D_t \cdot b_t\right) &= -r_t \cdot D_t \cdot b_t \cdot dt + D_t \cdot db_t \\
&= D_t \cdot \left[ -r_t \cdot b_t + \frac{\partial b}{\partial t} + \frac{\partial b}{\partial x_1} \cdot \mu_{1,t} + \frac{\partial b}{\partial x_2} \cdot \mu_{2,t} + \frac{1}{2} \cdot \sigma_{1,t}^2 \cdot \frac{\partial^2 b}{\partial x_1^2} + \frac{1}{2} \cdot \sigma_{2,t}^2 \cdot \frac{\partial^2 b}{\partial x_2^2} \right] \cdot dt \\
&\quad + D_t \cdot \left[ \frac{\partial b}{\partial x_1} \cdot \sigma_{1,t} \cdot dW_{1,t}^Q + \frac{\partial b}{\partial x_2} \cdot \sigma_{2,t} \cdot dW_{2,t}^Q \right]
\end{aligned}
$$

Since $Q$ is an equivalent martingale measure, by definition, the discounted bond price $D_t \cdot b_t$ is a martingale under $Q$. This means that the coefficient of $dt$ in the expression $d\left(D_t \cdot b_t\right)$ must be zero. Since $D_t > 0 \ \forall t$, the following must be true

(39)
$$
-r_t \cdot b_t + \frac{\partial b_t}{\partial t} + \frac{\partial b_t}{\partial x_{1,t}} \cdot \mu_1 + \frac{\partial b_t}{\partial x_{2,t}} \cdot \mu_2 + \frac{1}{2} \cdot \sigma_1^2 \cdot \frac{\partial^2 b_t}{\partial x_{1,t}^2} + \frac{1}{2} \cdot \sigma_2^2 \cdot \frac{\partial^2 b_t}{\partial x_{2,t}^2} = 0
$$

This is the no-arbitrage PDE for bond prices in the two-factor case, under the assumption that uncertainty is generated by two independent Brownian motions.

Suppose now that the processes for the factors $x_1$ and $x_2$ are affine diffusions.[67] In other words, assume that $\mu_1$, $\mu_2$, $\sigma_1^2$ and $\sigma_2^2$ are affine functions of the factors

(40)
$$
\begin{aligned}
\mu_1\left(t, x_{1,t}\right) &= \varepsilon_1 + \delta_1 \cdot x_{1,t} \\
\mu_2\left(t, x_{2,t}\right) &= \varepsilon_2 + \delta_2 \cdot x_{2,t} \\
\sigma_1^2\left(t, x_{1,t}\right) &= \omega_1 + \lambda_1 \cdot x_{1,t} \\
\sigma_2^2\left(t, x_{2,t}\right) &= \omega_2 + \lambda_2 \cdot x_{2,t}
\end{aligned}
$$

Also assume that the short rate is an affine function of the factors:

(41)
$$
r_t = \gamma_0 + \gamma_1 \cdot x_{1,t} + \gamma_2 \cdot x_{2,t}
$$

Under these assumptions, we can conjecture that all zero-coupon yields are affine functions of the factors. In other words, we can conclude that bond prices have the exponential affine form

---

[67] See Filipovic and Mayerhofer (2009) for more details about affine diffusion processes and their applications.

$$b\left(x_{1,t},x_{2,t},t,t+\tau\right)=\exp\left[\alpha(\tau)+\beta_1(\tau)\cdot x_{1,t}+\beta_2(\tau)\cdot x_{2,t}\right]$$
$$=\exp\left[\alpha(T-t)+\beta_1(T-t)\cdot x_{1,t}+\beta_2(T-t)\cdot x_{2,t}\right] \quad (42)$$

where $\alpha(\ )$, $\beta_1(\ )$ and $\beta_2(\ )$ are deterministic functions of the time-to-maturity.

When zero-coupon bond prices have the exponential affine form, PDE reduces to three ordinary differential equations. To see this, compute the derivatives of $b(\ )$ in equation (39)

(43)

$$\frac{\partial b_t}{\partial t}=-\left[\frac{\partial\alpha(T-t)}{\partial t}+\frac{\partial\beta_1(T-t)}{\partial t}\cdot x_{1,t}+\frac{\partial\beta_2(T-t)}{\partial t}\cdot x_{2,t}\right]\cdot b_t$$
$$=\left[\frac{\partial\alpha(\tau)}{\partial\tau}+\frac{\partial\beta_1(\tau)}{\partial\tau}\cdot x_{1,t}+\frac{\partial\beta_2(\tau)}{\partial\tau}\cdot x_{2,t}\right]\cdot b_t$$
$$\frac{\partial b_t}{\partial x_{i,t}}=\beta_i(\tau)\cdot b_t \qquad i=1,2$$
$$\frac{\partial^2 b_t}{\partial x_{i,t}^2}=\left[\beta_i(\tau)\right]^2\cdot b_t \qquad i=1,2$$

After replacing $r_t$, $\mu_1$, $\mu_2$, $\sigma_1^2$, $\sigma_2^2$ and the derivatives of $b_t$ with their expressions, equation (39) becomes

$$\begin{aligned}0=&-\left(\gamma_0+\gamma_1\cdot x_{1,t}+\gamma_2\cdot x_{2,t}\right)\\&+\left[\frac{\partial\alpha(\tau)}{\partial\tau}+\frac{\partial\beta_1(\tau)}{\partial\tau}\cdot x_{1,t}+\frac{\partial\beta_2(\tau)}{\partial\tau}x_{2,t}\right]\cdot b_t\\&+\beta_1(\tau)\cdot\left(\varepsilon_1+\delta_1\cdot x_{1,t}\right)\cdot b_t+\beta_2(\tau)\cdot\left(\varepsilon_2+\delta_2\cdot x_{2,t}\right)\cdot b_t\\&+\frac{1}{2}\cdot\left[\beta_1(\tau)\right]^2\cdot\left(\omega_1+\lambda_1\cdot x_{1,t}\right)\cdot b_t+\frac{1}{2}\cdot\left[\beta_2(\tau)\right]^2\cdot\left(\omega_2+\lambda_2\cdot x_{2,t}\right)\cdot b_t\end{aligned} \quad (44)$$

Since $b_t>0$, after re-arranging we obtain

(45)

$$\begin{aligned}0=&-\gamma_0+\frac{\partial\alpha(\tau)}{\partial\tau}+\beta_1(\tau)\cdot\varepsilon_1+\beta_2(\tau)\cdot\varepsilon_2+\frac{1}{2}\cdot\left[\beta_1(\tau)\right]^2\cdot\omega_1+\frac{1}{2}\cdot\left[\beta_2(\tau)\right]^2\cdot\omega_2\\&+x_{1,t}\cdot\left[-\gamma_1+\frac{\partial\beta_1(\tau)}{\partial\tau}+\beta_1(\tau)\cdot\delta_1+\frac{1}{2}\cdot\left[\beta_1(\tau)\right]^2\cdot\omega_1\right]\\&+x_{2,t}\cdot\left[-\gamma_2+\frac{\partial\beta_2(\tau)}{\partial\tau}+\beta_2(\tau)\cdot\delta_2+\frac{1}{2}\cdot\left[\beta_2(\tau)\right]^2\cdot\omega_2\right]\end{aligned}$$

Since this equation is true for any value of $x_{1,t}$ and $x_{2,t}$, we finally have

(46)

$$\frac{\partial \beta_1(\tau)}{\partial \tau} = \gamma_1 - \delta_1 \cdot \beta_1(\tau) - \frac{1}{2} \cdot \omega_1 \cdot \left[\beta_1(\tau)\right]^2$$
$$\frac{\partial \beta_2(\tau)}{\partial \tau} = \gamma_2 - \delta_2 \cdot \beta_2(\tau) - \frac{1}{2} \cdot \omega_2 \cdot \left[\beta_2(\tau)\right]^2$$
$$\frac{\partial \alpha(\tau)}{\partial \tau} = \gamma_0 - \varepsilon_1 \cdot \beta_1(\tau) - \varepsilon_2 \cdot \beta_2(\tau) - \frac{1}{2} \cdot \omega_1 \cdot \left[\beta_1(\tau)\right]^2 - \frac{1}{2} \cdot \omega_2 \cdot \left[\beta_2(\tau)\right]^2$$

Therefore, when the factors follow affine diffusions, the no-arbitrage PDE reduces to three ordinary differential equations. The equations for $\beta_1$ and $\beta_2$ are known as Riccati equations. The boundary conditions for these differential equations are

(47)

$$\alpha(0) = 0$$
$$\beta_i(0) = 0 \quad i = 1,2$$

These boundary conditions guarantee that $b\left(x_{1,T}, x_{2,T}, T, T\right) = 1$.

The two-factor CIR model is the special case where each factor follows a mean-reverting square-root diffusion process. That is

(48)

$$\mu_i\left(t, x_{i,t}\right) = \kappa_i \cdot \left(\theta_i - x_{i,t}\right) \quad i = 1,2$$
$$\sigma_i^2\left(t, x_{i,t}\right) = x_{i,t} \Leftrightarrow \sigma_i\left(t, x_{i,t}\right) = \sqrt{x_{i,t}} \quad i = 1,2$$
$$\kappa_i > 0 \quad i = 1,2$$
$$\theta_i > 0 \quad i = 1,2$$

## APPENDIX II. ESTIMATION TECHNIQUES

### I. IMPLEMENTATION OF NELSON-SIEGEL MODELS

**Implementation overview**

For the implementation of the Nelson-Siegel family of models, the key issue is the definition of the decay parameters. These parameters can be either fixed or variable during the time period of the calculation. Fixed decay parameters provide a stable state space framework, which is crucial to predict yield curve dynamics. The estimation of fixed decay parameters requires joint optimization in both cross-section and time series dimensions. This optimization is mathematically challenging.

MCMTS uses a one-stage solution with fixed decay parameters to estimate the Nelson-Siegel models. The main procedure relies on the Powell's method to maximize the log-likelihood value. This method iterates an arbitrary number of times until the tolerance constraint are met.[68] In each iteration, several optimization steps are computed sequentially.

- Cross-Sectional Optimization:
  - Tool: OLS
  - Input: Cross-sectional yield data
  - Output: Time series of factors

- Time Series Optimization:
  - Tool: Auto-regression or Vector auto-regression
  - Input: Time series of factors
  - Output: Parameters of state space model

- Joint Analysis:
  - Tool: Kalman Filter
  - Input: Initial time series factors, and state space model parameters
  - Output: Adjusted time series factors, and likelihood value

The covariance matrix of measurement errors is computed by OLS in the context of the cross-sectional optimization, and the covariance matrix of state innovation errors is computed by an AR(1) or VAR(1) during the time series optimization. This makes it unnecessary to estimate the values of these covariance matrices, which can significantly reduce the total number of parameters to estimate and makes the overall estimation process faster and easier to converge. The yield curve fitting results are also good. By way of example, when estimating the Nelson-Siegel models, the Root Mean Square Error (RMSE) is normally less than 5 basis points.

[68] See Powel (1964).

**Algorithm description**

The Powell algorithm—for example for the three-factor model—starts with an initial value for $\lambda$. In each iteration, the algorithm first tries a new value of $\lambda$, and then computes the log-likelihood value based on the given $\lambda$. In detail

(1) For a given $\lambda$, for each maturity $\tau$, compute $x_{\tau,i}$ for the $j_{th}$ factor

$$x_1(\tau)=1,\ x_2(\tau)=\frac{1-\exp(-\tau/\lambda)}{(\tau/\lambda)},\ x_3(\tau)=\frac{1-\exp(-\tau/\lambda)}{(\tau/\lambda)}-\exp(-\tau/\lambda)$$

and the measurement coefficient matrix

$$X=\begin{bmatrix} x_1(\tau_1) & x_2(\tau_1) & x_3(\tau_1) \\ \ldots & \ldots & \ldots \\ x_1(\tau_N) & x_2(\tau_N) & x_3(\tau_N) \end{bmatrix}$$

(2) For each time period $t \in [1...T]$, run an OLS regression on

$$Y_t = \rho + X\beta_t + \varepsilon_t$$

where $Y_t$ denotes the nominal interest rates, $\rho = 0$ and $\beta_t = [\beta_{1,t} \quad \beta_{2,t} \quad \beta_{3,t}]'$

The outputs are the time series of factors $\beta_t$ and measurement errors $\varepsilon_t$. Then, compute covariance matrix of measurement error, *H*, using time series of $\varepsilon_t$.

(3) Run an auto-regression, AR(1) or VAR(1), on the time series of factors $\beta_t$

$$\beta_t = \mu + \Phi\beta_{t-1} + \nu_t$$

The outputs include the state constant vector $\mu$, state transition matrix $\Phi$, and time series of state innovation errors $\nu_t$. Then, compute the covariance matrix of state innovation errors, *Q*, using the time series of $\nu_t$.

(4) Now, we have all the parameters for the state space model

$$Y_t = \rho + X\beta_t + \varepsilon_t$$
$$\beta_t = \mu + \Phi\beta_{t-1} + \nu_t$$

with

$$\begin{bmatrix} \varepsilon_t \\ \nu_t \end{bmatrix} \sim N\left( \begin{bmatrix} 0_{N\times 1} \\ 0_{K\times 1} \end{bmatrix}, \begin{bmatrix} H & 0 \\ 0 & Q \end{bmatrix} \right)$$

A Kalman Filter can be used to calculate adjusted time series of factors. The log-likelihood value $L$ is also be calculated by

$$L = \sum_{t=1}^{T} -\frac{1}{2}\left[\ln\left(\left|f_{t|t-1}\right| + \eta'_{t|t-1} f_{t|t-1}^{-1} \eta_{t|t-1}\right)\right]$$

(5) The end of this iteration.

## II. Implementation of Cox-Ingersoll-Ross Models

**Implementation overview**

MCMTS also uses a one-stage solution based on the Powell method to estimate CIR models. In each iteration, this solution takes the following optimization steps

- Cross-Sectional Optimization:
  - Tool: OLS
  - Input: Cross-Sectional yield data
  - Output: Time series of factors

- Compute the parameters of a state space model

- Joint Analysis:
  - Tool: Kalman Filter
  - Input: Initial time series factors, and state space model parameters
  - Output: Adjusted time series factors, and likelihood value

Again, the covariance matrix of measurement errors is computed by OLS during cross-sectional optimization, and the covariance matrix of state innovation errors is dynamically calculated at each time step of the Kalman Filter. Therefore, it is unnecessary to estimate the values of these covariance matrices.

**Algorithm description**

The algorithm starts with an initial value of the parameter vector *P,* and $P = [\kappa_1, \theta_1, \sigma_1, \lambda_1; \kappa_2, \theta_2, \sigma_2, \lambda_2]'$. In each iteration of the Powell method, the algorithm first tries a new value of *P*, and then computes the log-likelihood value based on the given *P*. In detail, for example, for a two-factor CIR

- The factors are independent with each other. The state transition matrix $\Phi$ and state error covariance $Q$ are therefore diagonal.
- The state variables follow non-central $\chi^2$ distribution. The estimator here is a quasi maximum likelihood estimator.

(1) For the given parameter vector *P*, for each maturity $\tau_i$, compute the measurement coefficient $x_j(\tau_i)$ for the $j_{th}$ factor, and measurement constant $\rho_i$

$$\rho_i = -\sum_{j=1}^{K} \frac{\ln a_j(\tau_i)}{\tau},\ x_j(\tau_i) = \frac{\ln b_j(\tau_i)}{\tau}$$

with

$$\rho = [\rho_1 \quad \ldots \quad \rho_N]',\ X = \begin{bmatrix} x_1(\tau_1) & \ldots & x_K(\tau_1) \\ \ldots & \ldots & \ldots \\ x_1(\tau_N) & \ldots & x_K(\tau_N) \end{bmatrix}$$

(2) For each time period $t \in [1 \ldots T]$, run an OLS regression of

$$\tilde{Y}_t = X\beta_t + \varepsilon_t$$

here, $\tilde{Y}_t = Y_t - \rho$ and $\beta_t = [\beta_{1,t} \quad \ldots \quad \beta_{K,t}]'$

The outputs are the time series of factors $\beta_t$ and measurement errors $\varepsilon_t$. Then, compute *H* using the time series of $\varepsilon_t$.

(3) For the given parameter vector *P*, compute the state equation

$$\beta_t = \mu + \Phi\beta_{t-1} + \nu_t$$

with

$$\mu_j = \theta_j(1 - \exp(-\kappa_j \Delta t))$$

$$Q_{j,j}(t) = \mathrm{Var}(\nu_j(t)) = \frac{\sigma_j^2}{\kappa_j}[1 - \exp(-\kappa_j \Delta t)]\left[\frac{1}{2}\theta_j(1 - \exp(-\kappa_j \Delta t)) + \exp(-\kappa_j \Delta t)\beta_{j,t-1}\right]$$

The outputs include the state constant vector $\mu$, state transition matrix $\Phi$, and time series of state innovation errors $v_t$. Note that $Q_{j,j}$ is a linear function of $\beta_{j,t-1}$, for a given parameter values.

(4) Now, we have all the parameters for the state space model

$$Y_t = \rho + X\beta_t + \varepsilon_t$$
$$\beta_t = \mu + \Phi\beta_{t-1} + v_t$$

and

$$\begin{bmatrix} \varepsilon_t \\ v_t \end{bmatrix} \sim N\left( \begin{bmatrix} 0_{N\times 1} \\ 0_{K\times 1} \end{bmatrix}, \begin{bmatrix} H & 0 \\ 0 & Q \end{bmatrix} \right)$$

A Kalman Filter can be used to calculate adjusted time series of factors. Since $Q_{j,j}$ depends on $\beta_{j,t-1}$, it is dynamically updated at each time period *t* in the Kalman Filter. The log-likelihood value *L* is calculated by

$$L = \sum_{t=1}^{T} -\frac{1}{2}\left[\ln\left(\left|f_{t|t-1}\right| + \eta'_{t|t-1} f_{t|t-1}^{-1} \eta_{t|t-1}\right)\right]$$

(5) The end of this iteration.

## III. Estimation Analysis

Appendix Table 1 shows the number of parameters and usual number of iterations required to achieve convergence depending on the model.

**Appendix Table 1. Estimation Analysis for Nelson-Siegel and CIR models**

| Model | Number of Parameters | Parameters | Number of Iterations before Convergence |
|---|---|---|---|
| Nelson-Siegel model with one decay parameter | 1 | C | <10 |
| Nelson-Siegel model with two decay parameters | 2 | $\lambda_1, \lambda_2$ | <50 |
| One-factor Cox-Ingersoll-Ross model | 4 | $\kappa, \theta, \sigma, \lambda$ | <200 |
| Two-factor Cox-Ingersoll-Ross model | 8 | $\kappa_1, \theta_1, \sigma_1, \lambda_1$ $\kappa_2, \theta_2, \sigma_2, \lambda_2$ | <2,000 |

Source: Fund staff estimates.

## IV. MONTE CARLO SIMULATIONS

The Monte Carlo simulations are computed based on the state-space framework. The dynamic evolutions of the factors is driven by the AR(1) or VAR(1) processes. The Nelson-Siegel models and Cox-Ingersoll-Ross models use two different approaches to simulate the innovation errors.

For the Nelson-Siegel models, *K* (or the number of factors) independent and identically-distributed standard normal random variables are created. Then, a joint normal distribution, with covariance matrix *Q* can be obtained, by multiplying the vector of *K* standard normal random variables with the Cholesky decomposition of *Q*. Defining

$$\omega \sim N(0, I), \text{ and } C = \text{CholeskyDecompose}(Q)$$
$$\Rightarrow q = C \times \omega, \text{ such that } q \sim N(0, Q)$$

At each time step, a draw of random variable *q* is created. This draw is used to drive the evolution of the state space model. Therefore, the time series of the simulated factors and interest rates are calculated by the state space model.

For the CIR models, a Euler scheme to "discretize" the square root diffusion process is used. The factor dynamics can be obtained by

$$\beta_{j,t+1} = \beta_{j,t} + \kappa_j (\theta_j - y_j)\Delta t + \sigma_j \sqrt{\beta_{j,t}} \Delta z_j, \quad \Delta z_j \sim N(0, \Delta t)$$

The simulated interest rates can be calculated from the measurement equation of the state space model.

The user can set the number of simulated paths *N,* and the number of steps *S*. The simulator first "reads" these numbers, and then starts the simulation process. After generating all the simulated paths, MCMTS calculates the summary statistics of the interest rates for all paths at the last time period, including mean, median, standard deviation and percentiles. The software also draws a histogram of the distribution in an Excel chart. In addition, the simulated paths at each time period are displayed.

## APPENDIX III. DESCRIPTION OF THE MCM TERM STRUCTURE SOFTWARE

**Basics**

The MCM Term Structure Software (MCMTS) explores a number of solutions for models of the term structure of interest rates. To this end and by way of introduction, the MCMTS makes it possible to perform the following operations:

- Create data sets, including of yield curves and macroeconomic variables.
- Specify time horizons.
- Select term structure models from a variety of model families.
- Estimate model parameters, and evaluate estimation performance.
- Run Monte Carlo simulations.
- Estimate impulse responses and variance decomposition using a Macro Factor Model.

MCMTS relies on a user-friendly Excel "add-in" which guides the user to explore different families of models.[69] The main page contains five menu items:

- *Select Model*
- *Model Estimation*
- *Performance Evaluation*
    - ➢ *Cross Sectional Yield Fitting*
    - ➢ *Time Series Yield Fitting*
    - ➢ *Factors Dynamic Evolution*
- *Monte Carlo Simulation*
- *Model Test*
    - ➢ *Impulse Response*
    - ➢ *Variance Decomposition*

---

[69] The algorithms of the program have been developed in C#.

- *Chi-square Test of Overall Fit*

**Quick tour**

**Creating Data Sets.** As a first step in using the MCMTS, the user needs to create two data sets: the first with information on yields, and the second with macroeconomic variables. Data on yield curves are essential to run the term structure models. Data on macroeconomic variables are critical only for the macro-factor models.

- The "Yield Curve" sheet contains yields of zero-coupon bonds for all different maturities, across a specific time period.
- The "Macro Factors" sheet contains time series data for various macroeconomic variables.

**Selecting a Model.** The second step in using the MCMTS is to specify a time horizon, and to select a model within the various alternatives.

- There are three alternatives for frequency values: daily, weekly or monthly.
- The "Simulation Horizon" for simulating the models can be set in line with the following options:
  - For daily time interval, one year after the end date
  - For weekly time interval, three years after the end date
  - For monthly time interval, ten years after the end date
- MCMTS has three families of models of interest rate models: Nelson-Siegel Models, Cox-Ingersoll-Ross Models, Macro-Factor Models.

**Estimating the Models.** Depending on the family of models chosen, MCMTS offers different options for estimating the models. The Nelson-Siegel and the Macro-Factor models, share the same estimation options, while the Cox-Ingersoll-Ross models have a separate set of options.

- For the Nelson-Siegel models:
  - MCMTS offers five alternatives: (i) a three-factor base model; (ii) the Bjork and Christensen four-factor model; (iii) the Bliss three-factor model; (iv) the Svensson four-factor model; and (v) the Adjusted Svensson four-factor model.

- There are two options for defining the "Factor Relationships": (i) Independent Factors, AR(1); or (ii) Correlated Factors, VAR(1).[70]
- There are three "Optimization Methods": (i) No Optimization; (ii) Minimize Measurement RMSE; and (iii) Maximum Likelihood with Kalman Filter.[71]

- For the Cox-Ingersoll-Ross models, MCMTS offers two options: (i) a one-factor model; and (ii) a two-factor model.[72]
- The output of the estimation is summarized in two charts: "Estimated Factor Loads" which displays the factor loadings for each maturity, and "Estimated Factors" which displays the dynamic evolution of each factor over time.

**Evaluating the Performance.** After estimating the parameters of particular model, MCMTS makes it possible to evaluate the fit of a yield curve.

- Cross Sectional Yield Fitting: selecting a date within the date range of the model estimation, MCMTS compares observed and estimated yields.

- Time Series Yield Fitting: selecting a maturity, MCMTS compares time series of observed and estimated yields for the given maturity.

- MCMTS also offers a Chi-square test for evaluating the overall performance of the estimation.

**Monte Carlo Simulations.** MCMTS also allows the user to perform Monte Carlo simulations based on the estimated models. After setting the number of simulations and paths to be displayed, MCMTS presents the distribution and summary statistics of simulated yields

**Impulse Responses and Variance Decomposition.** One of the most interesting features of MCMTS is that it allows the estimation of the contribution of each factor and macroeconomic variable to the variance of each yield maturity, and the impulse response of the factors to shocks in the macroeconomic variables (and vice versa).

MCMTS allows the user to choose the shocks and the responding variables, as well as the time horizon of the exercise.

---

[70] As described in Section II, a Nelson-Siegel model can explain the dynamic evolutions of three or four factors. The relationship among those factors can be modeled as independent or correlated. The second option is recommended for most cases.

[71] The Maximum Likelihood with Kalman Filter is undoubtedly the most sophisticated solution. Nevertheless, it is worth clarifying that there are multiple layers of optimizations implemented in the software. The optimization options offered in the interactive menus apply only for the highest level of optimization. Even under the "No Optimization" alternative, optimizations of lower levels are computed.

[72] A two-factor model can achieve a better performance in yield curve fitting, but it also takes a longer time to estimate the parameters of the model.