WP/11/247

IMF Working Paper

An Assessment of Estimates of Term Structure Models for the United States

Ying He and Carlos Medeiros

INTERNATIONAL MONETARY FUND

 WP/11/247

IMF Working Paper

Monetary and Capital Markets Department

An Assessment of Estimates of Term Structure Models for the United States

Prepared by Ying He and Carlos Medeiros[1]

October 2011

Abstract

The paper assesses estimates of term structure models for the United States. To this end, this paper first describes the mathematics underlying two types of term structure models, namely the Nelson-Siegel and Cox, Ingersoll and Ross family of models, and the estimation techniques. It then presents estimations of some of specific models within these families of models—three-factor Nelson-Siegel Model, four-factor Svensson model, and preference-free, two-factor Cox, Ingersoll and Roll model—for the United States from 1972 to mid 2011. It subsequently provides an assessment of the estimations. It concludes that these estimations of the term structure models successfully capture the dynamics of the term structure in the United States.

JEL Classification Numbers: G12; E43; E44; E58

Keywords: Term structure models; term structure of interest rates; yield curves; yields on U.S. Treasury securities

Authors' E-Mail Addresses: yhe@imf.org and cmedeiros@imf.org.

[1]This paper has benefited from comments from Christopher Towe, Herman Kamil and Patricia Medeiros, and conversations with colleagues at the IMF.

Contents Page

I. INTRODUCTION

The term structure of interest rates in the United States has received considerable attention in recent years, particularly in the context of the global financial crisis that began in 2007. Even slight changes in the term structure have generated significant interest in light of their implications for financial asset prices, issuance of debt instruments, extension of credit, consumption, and investment. The efforts of the Federal Reserve to influence the term structure to address the financial crisis, including through the purchase of medium- and long-term fixed income instruments, have also helped keep changes in the term structure in the limelight.

Financial economists have developed a myriad of term structure models. Nawalha, Believa, and Soto (2007) provide a taxonomy of these models. They note that these models fit into either fundamental models or preference-free models. The fundamental models share a time-homogeneous short-rate process and an explicit specification of the market price of risk. These models value default-free zero coupon bonds using information related to investors' risk aversion and expected movements in interest rates. Such models include the well known Vasicek and Cox, Ingersoll, and Ross (CIR) models developed in 1977 and 1985, respectively, and the multifactor models in the affine class models, including the by now classic model developed by Dai and Singleton (2000) and quadratic class models. The preference-free models do not require explicit specifications of the market price of risk for valuing bonds. In other words, these models do not require knowledge of market participants' risk preferences. These models include the Nelson-Siegel family of models, including the Svensson model, and the preference-free version of the CIR model.

This paper assesses estimations of some term structure models. As Nawalha, Believa, and Soto (2007) note, many papers have undertaken assessments of estimations of term structure models. This paper focuses on an assessment of the estimations of some models within the Nelson-Siegel and Cox, Ingersoll, and Ross family of models—a three-factor Nelson-Siegel Model (NSM), a four-factor Svensson model, and a preference-free, two-factor CIR model— for the United States from January 1972 to June 2011. It estimates these models using the IMF's term structure software described in Gasha et al. (2010) that provides a common platform to assess these models. In carrying out this assessment, the paper tries to answer the question: *How well do the estimates of these models capture the dynamics of the observed term structure of interest rates of the United States?*

The paper is divided as follows. Section II provides a description of the mathematics underlying the Nelson-Siegel family of models, including the four-factor Svensson model, and the estimation methodology of these models. Section III offers a summary of the mathematics underlying the one- and two-factor CIR models and the estimation methodology of these models. Section IV presents an assessment of the estimates of a three-factor NSM, a four-factor Svensson model, and a preference-free, two-factor CIR model using data for the United States from 1972 to mid 2011. This section also offers a comparison of these estimations. Section V presents a conclusion.

II. Nelson-Siegel Models

As is well known, the term structure depicts a set of yields on U.S. Treasury securities of different maturities. The set of yields suggest the presence of a relationship among short-, medium- and long-term yields. This relationship does not appear stable over time, particularly because the term structure exhibits different shapes at different moments. Nevertheless, as Diebold and Li (2006) note, changes in the term structure follow certain patterns. They note that the Nelson-Siegel family of models captures these patterns, while reproducing the historical average shape of the term structure. These models also account for the existence of unobservable, or latent, factors and their associated factor loadings and key macroeconomic variables that underlie U.S. Treasury security yields.

A. Yield-Only Nelson-Siegel Model [2]

As Gasha et al. (2010) note, the NSM successfully fits the term structure of U.S. Treasury security yields, while capturing the dynamics of the term structure. The NSM provides a tractable framework to fit the term structure by approximating the forward rate curve by a constant plus a polynomial times an exponential decay term given by[3]

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

where $f_t(\tau)$ is the instantaneous forward rate. This yields a corresponding term structure

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

where $\beta_{1t}, \beta_{2t}, \beta_{3t}$ and λ_t are parameters and 1, $\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}\right)$ and $\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right)$ are their loadings. The parameter λ_t controls both the exponential decay rate and the maturity at which the loading on β_{3t} reaches its maximum. Even though the NSM appears to be static, Diebold and Li (2006) interpret the parameters β_{1t}, β_{2t} and β_{3t} as dynamic latent factors. They show that these parameters could represent the level, slope, and curvature factors, respectively,

[2] This subsection and subsection B follow closely Medeiros and Rodriguez (2011).

[3] 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 τ periods into the future and that matures τ^* 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.

particularly because their loadings are a constant, a decreasing function of τ, and a concave function of τ.[4]

As Gasha et al. (2010) stress, this framework:

- provides a parsimonious approximation of the term structure, since 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;

- generates a forward curve and term structure that start at the instantaneous rate $\beta_{1t} + \beta_{2t}$ and then level off at the finite infinite-maturity value of β_{1t}, which is constant;[5]

- makes it possible to interpret the three factors β_{1t}, β_{2t} and β_{3t} as long-, short- and medium-factors, respectively, in light of its three loadings $\left[1, \left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}\right) \text{ and } \left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right)\right]$;[6] and

- establishes that the time-series statistical properties of the three factors β_{1t}, β_{2t} and β_{3t} underlie the dynamic patterns of the term structure.

Diebold, Rudebusch, and Aruoba (2006) argue that the state-space representation provides a framework for analysis and estimation of dynamic models. As Gasha et al. (2010) explain, this representation provides a way of specifying a dynamic system, while making it possible to handle a wide range of time series models. It facilitates estimation, the extraction of latent term structure factors, and the testing of hypotheses about the dynamic interactions between the term structure and macroeconomic factors. The state-space representation is

---

[4]A heuristic interpretation of the factors along these lines is the following: (i) since yields at all maturities load identically on β_{1t}, an increase in β_{1t} increases all yields equally, changing the level of the yield curve; (ii) since short rates load more heavily on β_{2t}, an increase in β_{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 β_{3t}, an increase in β_{3t} will increase medium-term yields, which load more heavily on it, increasing the yield curve 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.

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

[6]To appreciate this interpretation, notice that the loading on β_{1t} is 1, which does not decay to zero in the limit; the loading on β_{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 β_{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 the Diebold and Li (2006) interpretation of the three factors as level, slope and curvature.

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

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

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

Alternatively, it is possible to express equations (II.4) and (II.5) in matrix form as

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

(II.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 (II.6), or the *transition equation*, specifies 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 μ is a 3x1 vector of coefficients, the transition matrix A is a 3x3 matrix of coefficients, and η_t is a white noise transition disturbance with a 3x3 non-diagonal covariance matrix Q.[7] Equation (II.7), or the *measurement equation*, is the specification of the term structure itself, and relates N observable yields to the three unobservable factors. The vector of yields Y_t contains N different maturities $Y_t = [y_t(\tau_1) \;\; \cdots \;\; y_t(\tau_N)]'$. The measurement matrix Λ is an Nx3 matrix whose columns are the loadings associated with the respective factors, and ε_t is a white noise *measurement* disturbance with an NxN 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 . Formally,

(II.8) $$\begin{pmatrix} \eta_t \\ \varepsilon_t \end{pmatrix} \backsim WN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} Q & 0 \\ 0 & H \end{pmatrix}\right]$$

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

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

In addition to facilitating computational tractability, these assumptions are essential to estimate both equations.

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

B. Yield-Macro Nelson-Siegel Model

As Diebold, Rudebusch, and Auroba (2006) note, recent latent factor models of the term structure make explicit use of macroeconomic factors.[8] These models use a state-space representation to incorporate macroeconomic factors in a latent factor model of the term structure to facilitate the analysis of the potential bidirectional feedback between the term structure and the economy. They enhance the state vector to include some key macroeconomic variables associated with economic activity, monetary stance, and inflation, specifically manufacturing capacity utilization (CU_t), the federal funds rate (FFR_t), and annual price inflation$(INFL_t)$. In so doing, they offer an insight into the underlying economic forces that drive the evolution of interest rates.

In this light, the state-space representation takes on the form

(II.9) $$F_t = \mu + AF_{t-1} + \eta_t$$

(II.10) $$Y_t = \Lambda F_t + \varepsilon_t$$

where $F_t = (\beta_{1t}\ \beta_{2t}\ \beta_{3t}\ CU_t\ FFR_t\ INFL_t)'$, and the dimensions of μ, A, and η_t are increased accordingly, to 6x1, 6x6 and 6x1, respectively. The matrix Λ 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 6x6, and the measurement disturbance covariance matrix *H* are non-diagonal and diagonal matrices, respectively,[9]

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

---

[8] Diebold, Piazzesi, and Rudebusch (2005) discuss the importance of a joint macro-finance modeling strategy to better understand the term structure of interest rates.

[9] Diebold, Rudebusch, and Aruoba (2006) stress that these macroeconomic variables represent the minimum set of fundamentals required to capture basic macroeconomic dynamics.

(II.12)

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

The estimations for yield-only and yield-macro NSMs are similar. In particular, they involve estimating the latent factors β_i, the coefficients in the transition matrix A, the mean state vector μ, the measurement coefficient matrix Λ, the transition disturbance covariance matrix Q, and the measurement disturbance covariance matrix H.

The estimation of the decay parameter λ is key in this regard. The estimation of the decay parameter depends on an optimization algorithm that either minimizes the Root Mean Square Error (RMSE) of the measurement equation or maximizes the likelihood function estimated by the Kalman filter, which sequentially updates the linear projection of a state-space representation (Hamilton, 1994). This involves an iterative process of estimating and updating the measurement and transition equations until finding the optimal point. As a first step, for a given decay factor, it is necessary to determine the measurement coefficient matrix, and then run an OLS regression for each time t to obtain the latent factors β_t and the measurement errors ε_t. The resulting matrix of measurement errors serves to calculate the measurement disturbance covariance matrix H. In a second step, the latent factors are treated as dependent variables in a VAR(1), which makes it possible to estimate the transition matrix A, the mean state vector μ, and the transition errors v_t. These transition errors in turn open the way to compute the transition disturbance covariance matrix Q. In the case of the maximization of the likelihood function estimated by the Kalman filter, the well known Powell algorithm makes it possible to select the new values of the parameters associated with this process (see Gasha et al., 2010).

C. Four-Factor Svensson Model

Svensson (1994) extends the Nelson-Siegel yield-only model. In particular, he adds a fourth term, a second curvature or hump-shape, or $\beta_{4t}\left(\frac{1-e^{-\lambda_2\tau}}{\lambda_2\tau} - e^{-\lambda_2\tau}\right)$.

(II.13)

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau} - e^{-\lambda_1\tau}\right) + \beta_{4t}\left(\frac{1-e^{-\lambda_2\tau}}{\lambda_2\tau} - e^{-\lambda_2\tau}\right)$$

This fourth term has two additional parameters, namely β_{4t} and $\lambda_2\tau$ (the latter should be positive). Svensson argues that the addition of the fourth term increases the flexibility and fit of the Nelson-Siegel model. He further notes that the Nelson-Siegel model yields a satisfactory fit in most cases. However, he suggests that, when the term structure is complex, the extended model improves the fit considerably.

Svensson notes that the four-factor model could suffer from multicollinearity when the decay paremters—$\lambda_1\tau$ and $\lambda_2\tau$—have similar values. To overcome this difficulty, he suggests a slight adjustment to the second curvature factor as follows

(II.14)

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau} - e^{-\lambda_1\tau}\right) + \beta_{4t}\left(\frac{1-e^{-\lambda_2\tau}}{\lambda_2\tau} - e^{-2\lambda_2\tau}\right)$$

III. Cox, Ingersoll and Ross Models

The Cox, Ingersoll, and Ross (CIR) models are among the better known term structure models.[10] These models include one or more factors. In the context of the taxonomy of Nawalkha, Believa, and Soto (2007), such models could be either fundamental models or preference-free models. In this section, we derive a one-factor, fundamental CIR model. The difference between this model and a preference-free model is simply the knowledge of market risk. In this section, we summarize the description of Baz and Chacko (2004) of a single-factor model and a one-factor CIR model, explain briefly the difference between a one-factor CIR and a two-factor CIR models, and describe the estimation procedures of these models.

A. Single-Factor Model

As Baz and Chacko (2004) note, the short interest rate *r* follows an Itô process defined as

(III.1) $$dr = \mu(r,\tau)d\tau + \sigma(r,\tau)dW$$

where $\mu(r,\tau)$ is the drift term, $\sigma(r,\tau)$ is the diffusion term, *r* is the short time, and τ is time. A bond price *P* is a function of this short rate, time, and maturity, *T*

(III.2) $$P = P(r,\tau,T)$$

[10] See, for instance, Campbell, Low, and MacKinlay (1997) and Nawalkha, Believa, and Soto (2007).

Using Itô's lemma, it is possible to determine the bond return, and identify the drift α and the volatility b of this return

(III.3) $$\frac{dP}{P} = a(r,\tau,T)d\tau + b(r,\tau,T)dW$$

where a and b are

(III.4) $$a(r,\tau,T) \equiv \frac{\frac{\partial P}{\partial r}\mu(r,\tau) + \frac{\partial P}{\partial \tau} + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma^2(r,\tau)}{P}$$

and

(III.5) $$b(r,\tau,T) \equiv \frac{\frac{\partial P}{\partial r}\sigma(r,\tau)}{P}$$

The duration of D of a bond in a single-factor model is

(III.6) $$D \equiv -\frac{1}{P}\frac{\partial P}{\partial r}$$

and the volatility of a bond is equal to (minus) the duration multiplied by the volatility of the short rate

(III.7) $$b(r,\tau,T) = -\sigma(r,\tau)D$$

where b is negative for most bonds.

In this light, it is possible to consider two bonds—bond 1 and bond 2—priced at P_1 and P_2, maturing at T_1 and T_2, with drifts a_1 and a_2 and volatilities b_1 and b_2. A self-financing portfolio consists of

- bonds 1 worth V_1;
- bonds 2 worth V_2;

- an amount $(V_1 + V_2)$ borrowed at the riskless short rate *r;*

with no restriction on the sign of V_1 and V_2. This is a portfolio, π, with instantaneous change defined as

(III.8)
$$\begin{aligned} d\pi &= V_1 \frac{dV_1}{V_1} + V_2 \frac{dV_2}{V_2} - (V_1 + V_2) r d\tau \\ &= V_1(a_1 - r)d\tau + V_2(a_2 - r)d\tau + (V_1 b_1 + V_2 b_2)dW \end{aligned}$$

If the portfolio consists of

(III.9)
$$V_1 = -V_2 \frac{b_2}{b_1}$$

then the stochastic term dW disappears, and $d\pi$ is deterministic. Since the portfolio is riskless and self-financed, it can only earn an instantaneous rate of zero. Equation (III.8) becomes

(III.10)
$$-\frac{V_2 b_2}{b_1}(a_1 - r)d\tau + V_2(a_2 - r)d\tau = 0$$

or

(III.11)
$$\frac{a_1 - r}{b_1} = \frac{a_2 - r}{b_2}$$

Equation (III.11) shows that the expected bond return in excess of the risk-free rate per unit of volatility is the same for all interest-rate-sensitive securities, or λ. Since λ is the same for bonds of all maturities, it does not depend on T. However, it is possible to re-write this equation as

(III.12)
$$\frac{a_i - r}{b_i} = \lambda(r, \tau)$$

where α_i and b_i are the expected return and volatility of any bond *i,* respectively, and $\lambda(r,t)$ is the market price of the risk attached to bonds. As b_i is generally negative, a negative λ means that the market risk on a long bond is greater than market risk on a short bond and, therefore, the expected return on a long bond is greater than that on a short bond.

It is now possible to combine equation (III.12) with equations (III.4) and (III.5) to obtain a partial differential equation (PDE) that defines the pricing for any interest-rate-sensitive security

(III.13) $$\frac{\partial \mathrm{P}}{\partial r}\left[\mu(r,\tau)-\lambda(r,\tau)\sigma(r,\tau)\right]+\frac{\partial \mathrm{P}}{\partial \tau}+\frac{1}{2}\frac{\partial^2 \mathrm{P}}{\partial r^2}\sigma^2(r,\tau)=r\mathrm{P}$$

It is important to note that the payoff for a security described by this equation also depends on a boundary condition, which for a zero-coupon bond with a face value of 1 is

(III.14) $$\mathrm{P}(r,\mathrm{T},\mathrm{T})=1$$

The solution of (III.13) subject to boundary condition (III.14) is

(III.15) $$\mathrm{P}(\tau,\mathrm{T})=\widetilde{E}_\tau\left\{\exp\left[-\int_\tau^\mathrm{T} r(s)ds-\int_\tau^\mathrm{T}\lambda(r,s)dW(s)-\frac{1}{2}\int_\tau^\mathrm{T}\lambda^2(r,s)ds\right]\right\}$$

B. The Cox, Ingersoll, and Ross Model

As Baz and Chacko (2004) note, the CIR model defines a stochastic process for the short rate as

(III.16) $$dr=\mathrm{K}(\theta-r)d\tau+\sigma\sqrt{r}dW$$

where K is the speed of mean-reversion $(K>0)$, θ is the long-term target for r, $\mu(r,\tau)=\mathrm{K}(\theta-r)$, and $\sigma(r,\tau)=\sigma\sqrt{r}$. The CIR model uses a market price of risk defined as

(III.17) $$\lambda(r,\tau)=\frac{\lambda\sqrt{r}}{\sigma}$$

This makes it possible to re-write equation (III.13) as

(III.18) $$\frac{\partial \mathrm{P}}{\partial r}\left[\mathrm{K}(\theta-r)-\lambda r\right]+\frac{\partial \mathrm{P}}{\partial \tau}+\frac{1}{2}\sigma^2 r\frac{\partial^2 \mathrm{P}}{\partial r^2}=r\mathrm{P}$$

subject to $\mathrm{P}(r,\tau,\mathrm{T})=$ 1 for a zero-coupon bond paying 1 at maturity. As is common with term structure models, the solution for the price of a zero-coupon bond involves a guess in the form of

(III.19) $$\mathrm{P}(r,\tau,\mathrm{T})=\mathrm{A}(\tau)\exp\left[-\mathrm{B}(\tau)r\right]$$

This equation opens the way to solve the PDE in equation III.18 using two separable ordinary differential equations[11]

(III.20) $$A = \left\{ \frac{\eta \exp\left[\left(\frac{K + \lambda + \eta}{2}\right)\tau\right]}{\left(\frac{K + \lambda + \eta}{2}\right)\left[\exp(\eta\tau) - 1\right] + \eta} \right\}^{\xi}$$

and

(III.21) $$B = \frac{\exp(\eta\tau) - 1}{\left(\frac{\kappa + \lambda + \eta}{2}\right)\left[\exp(\eta\tau) - 1\right] + \eta}$$

with

$$\eta \equiv \sqrt{2\sigma^2 + (K + \lambda)^2}$$

and

$$\xi = \frac{2K\theta}{\sigma^2}$$

In this context, the zero-coupon rate is

(III.22) $$R(\tau, T) = \frac{-\log A(\tau) + rB(\tau)}{\tau}$$

As T goes to infinity, Baz and Chacko (2004) show that

(III.23) $$R(\tau, \infty) = \frac{2K\theta}{K + \lambda + \eta}$$

[11] Nawalkha, Believa, and Soto (2007) show a solution the PDE in equation using the Itô-Doblein formula and a Riccati equation.

The term structure of zero-coupon rates is upward sloping when $r < R(\tau,\infty)$, has a hump when $R(\tau,\infty) < r < \frac{\mathrm{K}\theta}{\mathrm{K}+\lambda}$, and is downward sloping when $r > \frac{\mathrm{K}\theta}{\mathrm{K}+\lambda}$.

As Shreve (2004) argues, an extension of the one-factor CIR model is important to capture better the dynamics of the term structure of interest rates. While the one-factor model tends to capture shifts in the term structure, it does not explain changes in the slope or the curvature of the term structure. A two-factor CIR model addresses at least in part this shortcoming. As Gasha at al. (2010) show, it is straightforward to extend the one-factor CIR model to a two-factor CIR model. A two-factor CIR is simply the sum of two state variables given as

(III.24) $$r(t) = \beta_1(t) + \beta_2(t)$$

where $\beta_1(t)$ and $\beta_2(t)$ are the state variables. The one-factor CIR model depends on the stochastic process of the short interest rate, or *r*, described by the stochastic differential equation (III.16) to obtain a general solution, but, as Gasha et al. (2010), Nawalkha, Believa, and Soto (2007) and Shreve (2004) show, the two-factor CIR depends on two stochastic differential equations.

A general state-space representation allows us to solve the CIR model. This state-space representation is virtually identical to the state-space representation of the Nelson-Siegel family of models. Using a slightly different nomenclature than the one presented above, the state-space representation for the CIR model is given as

(III.25) $$\begin{aligned} \beta_t &= \mu + \Phi\beta_{t-1} + v_t \\ R_t &= \rho + \Lambda\beta_t + \varepsilon_t \end{aligned}$$

where

(III.26) $$\mu = \begin{bmatrix} \theta_1(1-e^{-K_1\Delta_t}) \\ \vdots \\ \theta_K(1-e^{-K_K\Delta_t}) \end{bmatrix}$$

(III.27) $$\Phi = \begin{bmatrix} e^{-K_1\Delta_t} & & \\ & \ddots & \\ & & e^{-K_k\Delta_t} \end{bmatrix}$$

(III.28)

$$\rho = \begin{bmatrix} \sum_{k=1}^{K} -\frac{\ln A_k(\tau_1)}{\tau_1} \\ \vdots \\ \sum_{k=1}^{K} -\frac{\ln A_k(\tau_N)}{\tau_N} \end{bmatrix}$$

and

(III.29)

$$\Lambda = \begin{bmatrix} B_1(\tau_1) & \cdots & B_K(\tau_1) \\ \vdots & \ddots & \vdots \\ B_1(\tau_N) & \cdots & B_K(\tau_N) \end{bmatrix}$$

and Δ_t represents the length of the time interval in the discrete sample. The equations for the state variables follow directly from a Chi, or χ^2, distribution. The variances of the state or transition errors and observation (i.e., measurement) errors are Q_t and H, where Q_t is a diagonal matrix with conditional expectation given by zero and conditional variance of the state variable with diagonals given by

(III.30)

$$\begin{bmatrix} \sigma_1^2\left(\frac{1-e^{\kappa_1\Delta_t}}{\kappa_1}\right)\left(\frac{1}{2}\theta_1\left(1-e^{\kappa_1\Delta_t}\right)+e^{\kappa_1\Delta_t}\beta_{1,t-1}\right) & & \\ & \ddots & \\ & & \sigma_K^2\left(\frac{1-e^{\kappa_K\Delta_t}}{\kappa_K}\right)\left(\frac{1}{2}\theta_K\left(1-e^{\kappa_K\Delta_t}\right)+e^{\kappa_K\Delta_t}\beta_{K,t-1}\right) \end{bmatrix}$$

IV. Estimations of the Term Structure Models

After briefly summarizing the data used in this paper, this section presents estimations of the term structure of U.S. Treasury security yields using the Nelson-Siegel family of models and the CIR model.

A. Data

This paper uses yields of U.S. Treasury securities and macroeconomic variables. The yields are annualized zero-coupon bond nominal yields continuously compounded. The yields, obtained from Bloomberg, are monthly observations of U.S. Treasury securities of 9 maturities—3, 6, 12, 24, 36, 48, 60, 84 and 120 months—for the period of 1972:1 to 2011:6. The macroeconomic variables include (i) the inflation variable, or the annual percentage change in the monthly price deflator for personal consumption expenditures; (ii) the real economic activity relative to potential, or manufacturing capacity utilization; and (iii) the monetary policy instrument, or the monthly average federal funds rate.

B. Nelson-Siegel Models

As Gasha et al. (2010) note, the term structure, including during the global financial crisis, exhibits the following characteristics:

- The term structure is on average upward sloping and concave.
- The term structure takes on a variety of shapes through time, including upward sloping, downward sloping, humped, and inverted humped.
- The term structure has shifted downward noticeably in the context of the Federal Reserve policy to lower the fed funds rate to nearly zero in recent years.
- The level of the term structure is more persistent that the slope and the curvature, as evidenced by the smaller variation of the level relative to its mean than the variation of the slope and curvature relative to their means. The slope in turn is more persistent than the curvature.

The estimate of the three-factor, yield-only NSM using the state-space representation explains well the term structure of U.S. Treasury security yields for the period 1972:1–2011:6. This is consistent with previous findings of estimates of the term structure of interest rates for the United States (Medeiros and Rodriguez, 2011). Figure 1 shows the estimate of the term structure of U.S. Treasury security yields, and the estimation residuals, or the difference between the estimated and observed term structures. The estimation residuals tend to be small, and, for long periods of time, are close to zero. The goodness-of-fit test, measured by the Chi-square test statistic, attests how well the estimated term structure fits the observed term structure

(Table 1).[12] Also, reflecting the goodness of this fit, Figure 3 shows that the estimated term structure of the yields on the one-year and five-year U.S. Treasury securities and observed yields on these U.S. Treasury securities nearly overlap. These results hold for the other seven maturities as well. As further evidence of the goodness of fit, Figure 4 shows the close fit of the estimated term structure curve with respect to particular yields on the observed term structure. Figure 5 displays the estimated three factors for the period under analysis.[13] Table 2 summarizes the goodness-of-fit results of the also promising estimates of the yield-macro NSM.[14]

Impulse response functions and the variance decomposition provide additional metrics to assess the term structure. The impulse response functions based on the yield-only NSM show that shocks of the level and slope have a statistically significant impact on the level of the term structure (Figure 6). Shocks to the slope also have a statistically significant impact on the slope of the term structure. The variance decomposition shows that innovations or shocks to the three factors impact the variance of U.S Treasury security yields for a 60-month period (Table 3). The variance decomposition shows that shocks to the slope account for the largest proportion of the variance of U.S. Treasury security yields. However, shocks to the slope account for a declining proportion of the variance as the maturity of the U.S. Treasury securities increases. Shocks to the level account for a smallest proportion of the variance of the yields of U.S. Treasury securities with maturity of one year or less, but explain an increasing proportion of the variance of yields of U.S. Treasury securities with maturity beyond one year. Shocks to the curvature explain only a small proportion of the variance of the yields of U.S. Treasury security across all maturities.

A Monte Carlo simulation further complements the analysis of the term structure. This simulation makes it possible to obtain a distribution of future yields that are consistent with the historical dynamics of the term structure. This requires first using one-step-ahead factors in the transition equation (equation II.6), and then utilizing these results to determine the term structure using the measurement equation (equation II.7). Specifically, to determine the one-step-ahead factors, we generate a K number of standard normal random variables that are independent and identically distributed, and derive the transition disturbance random errors, with a covariance matrix Q, by multiplying the vector of standard normal variables with the Cholesky decomposition of Q.[15] For fixed initial values of the factors, we first use these errors to calculate new values for the factors using the transition equation, and then utilize these new

---

[12]The Chi-square statistic is defined as the square of the difference between the observed term structure and the estimated term structure divided by the variance of the observed yields. The null hypothesis states that the observed term structure is the same as the estimated term structure.

[13]Note that the slope of the term structure is actually -β_{2t}, i.e., the program estimates it as the negative of the slope.

[14] Medeiros and Rodriguez (2011) explain the estimates of the yield-macro NSM for the United States.

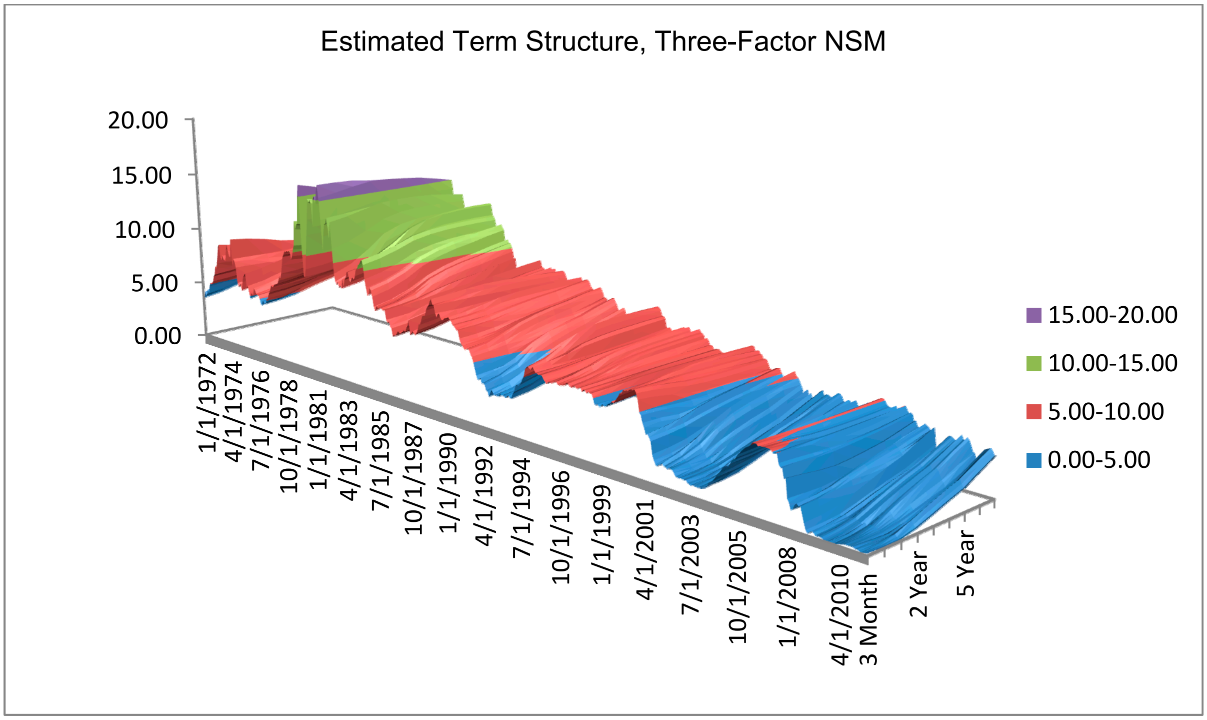
[15] K refers to the number of latent term structure factors.

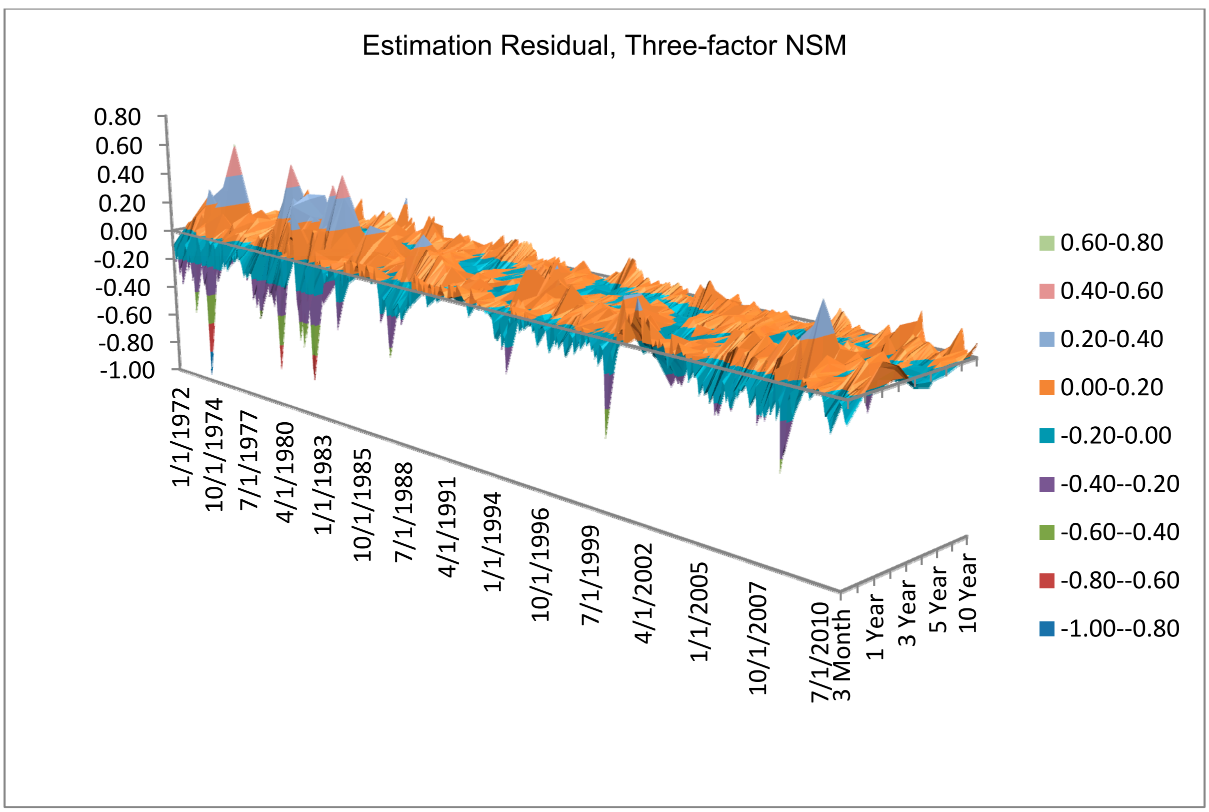
values to determine the term structure using the measurement equation. The process to determine the multiple-step-ahead factors depends on the number of time steps or periods, S, which yields a final simulated value for the term structure at the end of these periods. We repeat this process a number of times, or m, to obtain model statistics, or implied distribution, of future yields. By way of illustration, Figure 7 shows some five-year paths of the Monte Carlo simulation for the one-year and five-year U.S. Treasury securities repeated 1000 times. This simulation yields some statistics of the estimated term structure, including mean, variance, skewness and kurtosis (Table 5).

The estimation of the four-factor Svensson model also explains well the term structure of interest rates of the United States over 1972:1–2011:6. In particular, the four factors of the Svensson model—level, slope, and two curvatures—capture well the many changes in the shape of the term structure. As in the case of the three-factor, yield-only NSM, the estimation residuals are generally small (Figure 2). The Chi-square test statistic confirms that the estimation of the term structure using the Svensson model fits well the observed term structure (Table 2). This test also shows that the estimation of the four-factor Svensson model appears to fit somewhat better the observed term structure of the United States than the estimation of the three-factor NSM. However, given that the estimation of the four-factor Svensson model is only marginally better than the estimation of the three-factor NSM, it may well be appropriate to rely on one model. The estimations of the term structure using the four-factor Svensson model for one-year and five-year U.S. Treasury securities essentially lie on the observed term structure for these maturities (Figure 3). These results hold not only for these two maturities, but also for the other seven maturities used in this paper. Figure 4 shows the good fit of the estimated term structure curve with respect to particular points in the observed term structure, and Figure 5 displays the estimates of the four factors over the estimation period.

As noted above, the variance decomposition provides an additional metric to assess the term structure. The variance decomposition shows how shocks to the four factors of the Svensson model impact the variance of U.S Treasury security yields for a 60-month period (Table 4). As in the case of the three-factor NSM, shocks to the slope represent the largest proportion of the variance of U.S. Treasury security yields. However, shocks to the slope account for a declining proportion of the variance of these yields as the maturity of the U.S. Treasury securities increases. Shocks in the level account for a small proportion of the variance of U.S. Treasury security yields, particularly in the case of U.S. Treasury securities with maturity of one year or less. Shocks of the level account for an increasing share of the variance of U.S. Treasury security yields as the maturity of these instruments increases. Shocks to the curvature factors account for a small proportion of the variance of U.S. Treasury security yields, even less so as the maturity of the U.S. Treasury securities increase.

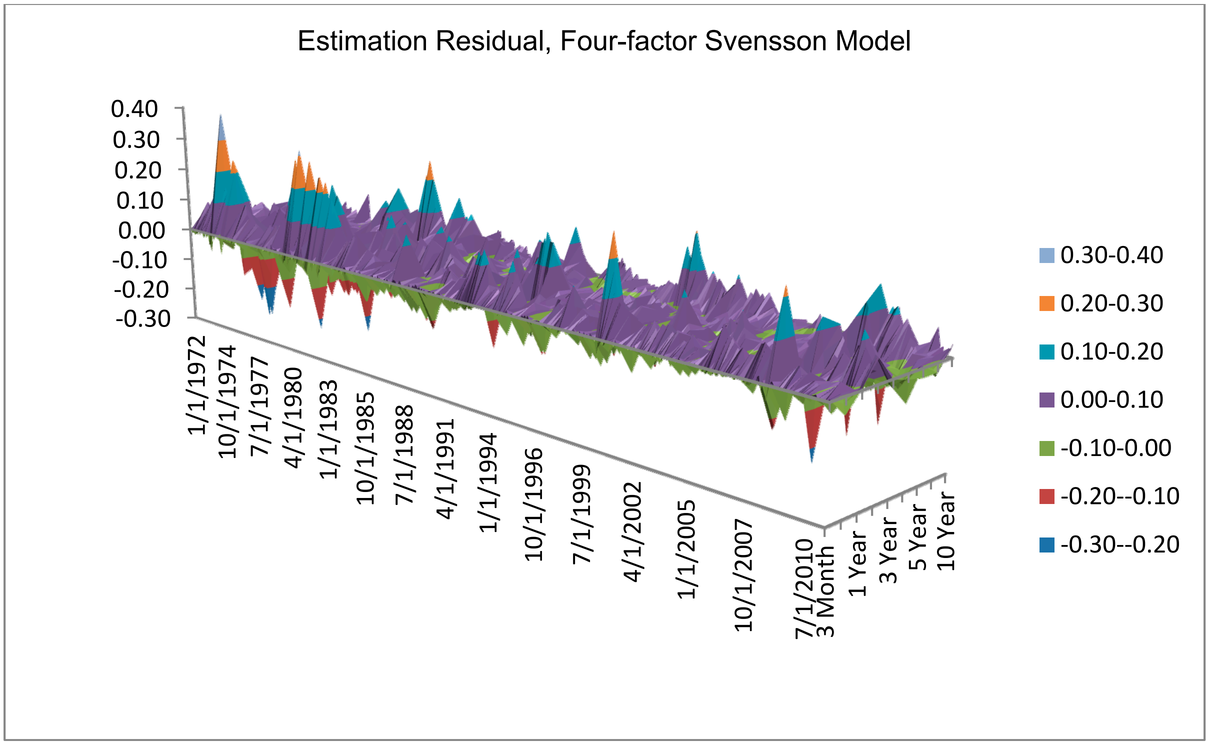
Figure 1. Estimated Term Structure and Estimation Residual

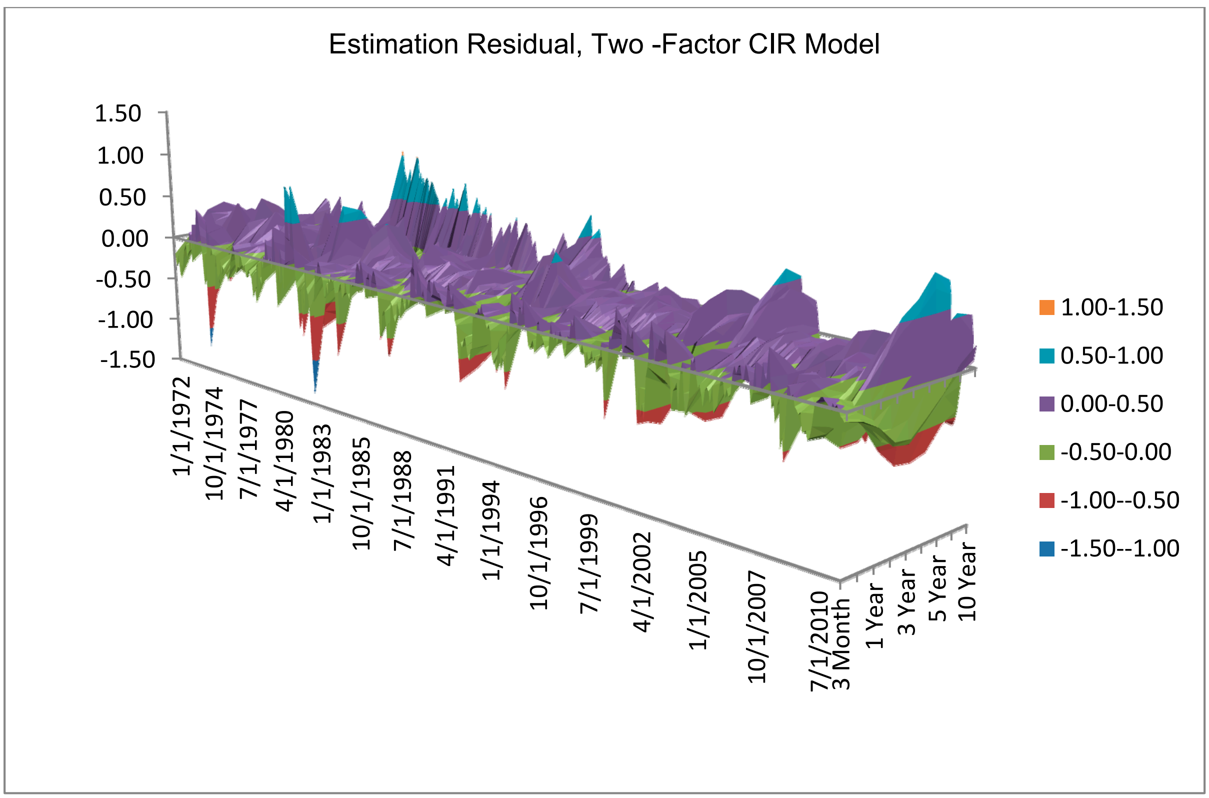




Source: Fund staff estimates.

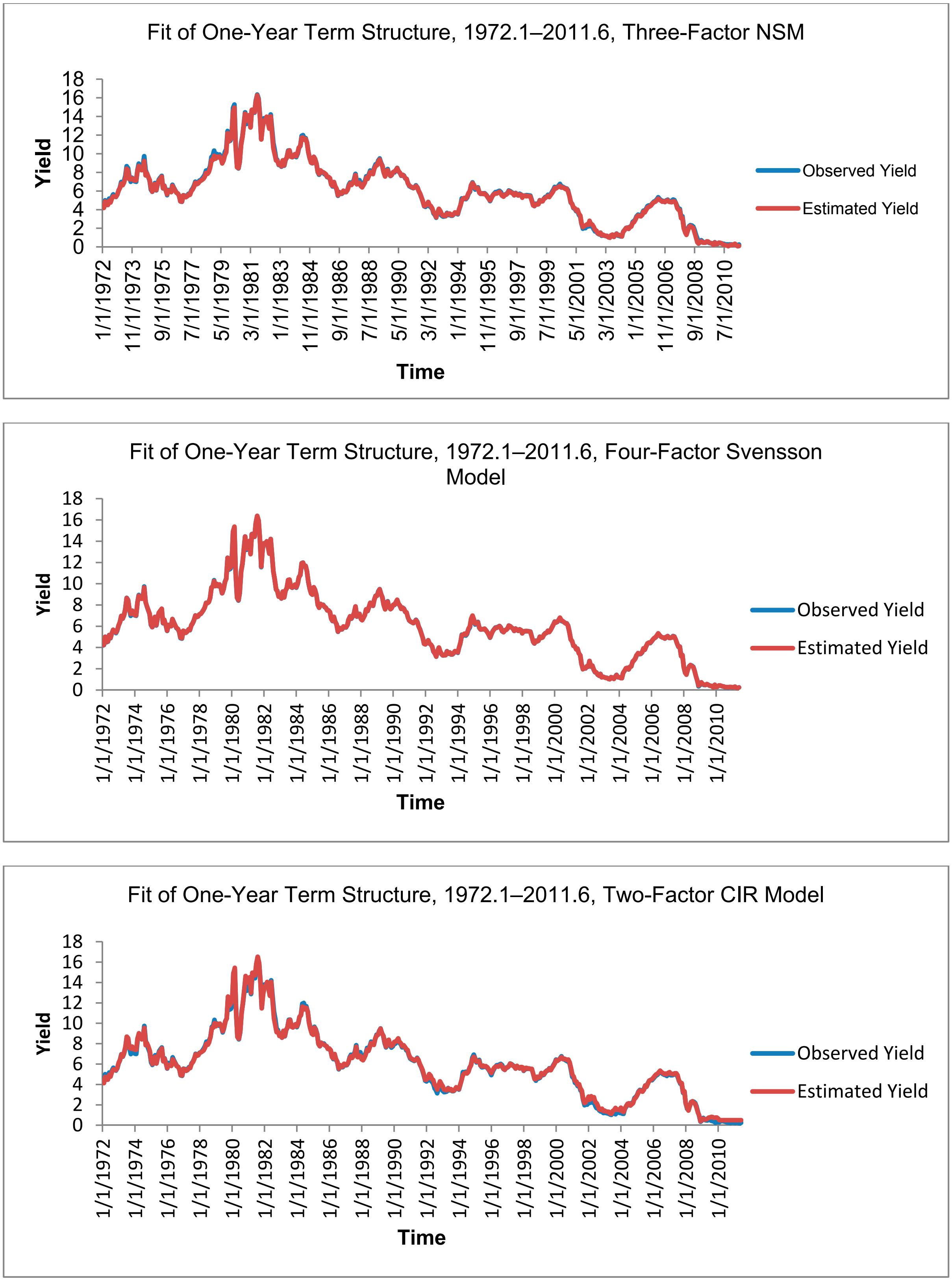
Figure 2. Estimation Residuals





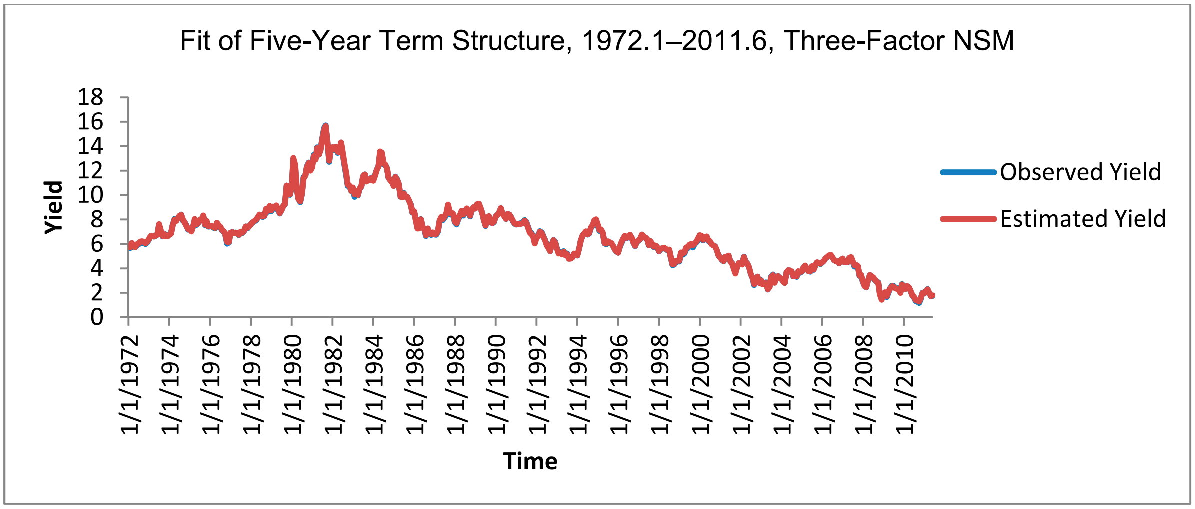
Source: Fund staff estimates.

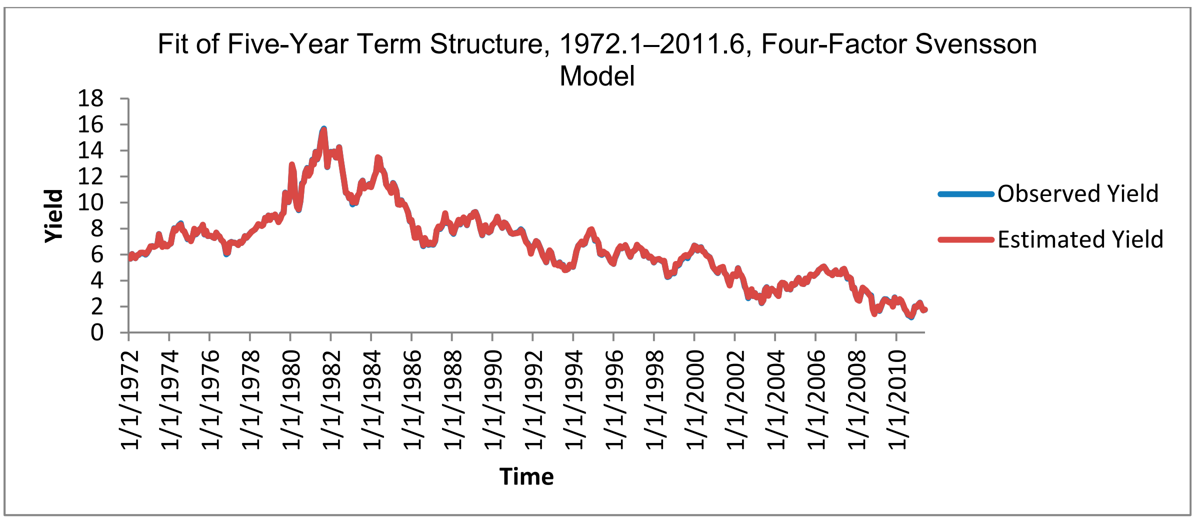
Figure 3. Performance Evaluation of Models

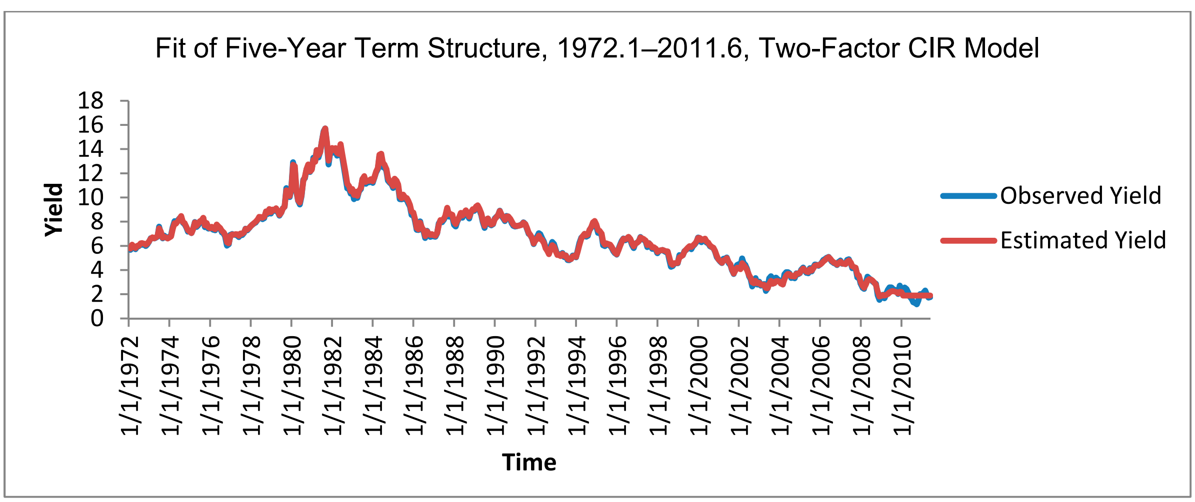


Source: Fund staff estimates.

Figure 3. Performance Evaluation of Models (Continued)



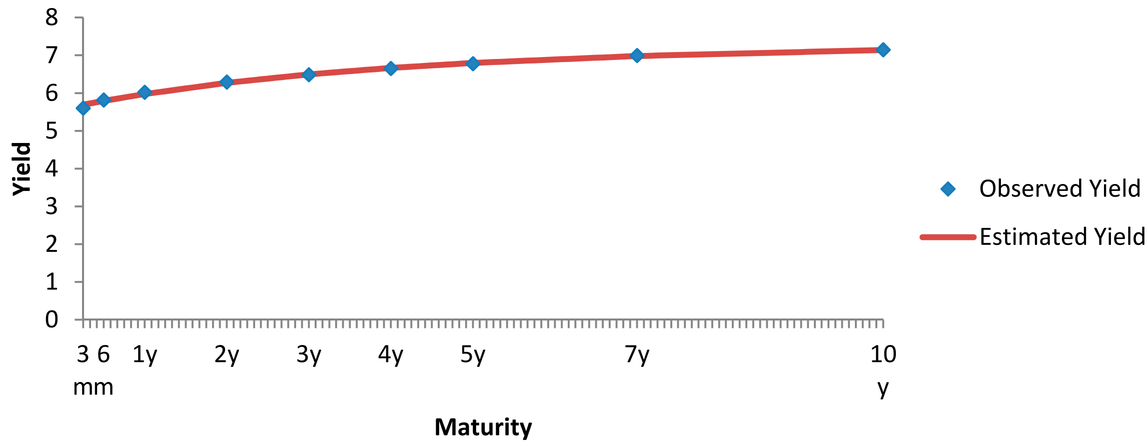




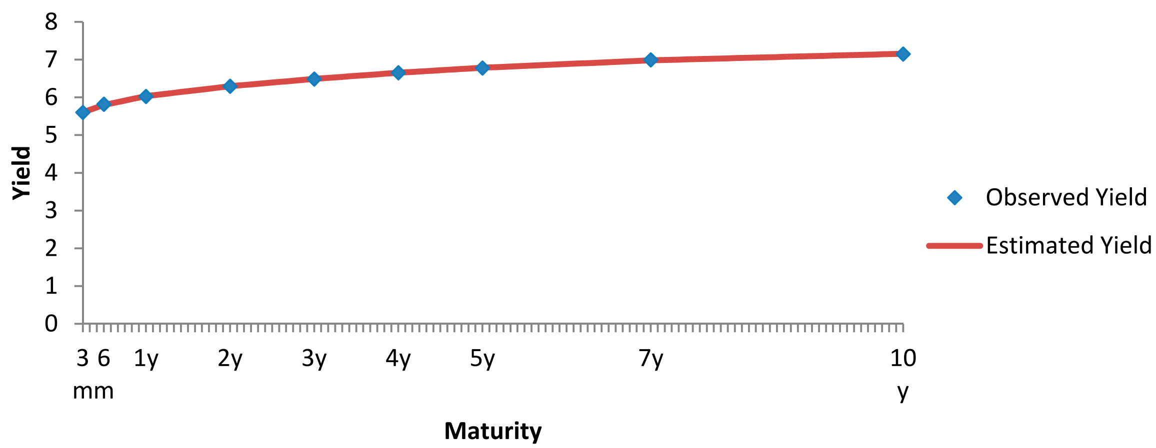
Source: Fund staff estimates.

Figure 4. Observed and Estimated Average Yield Curve

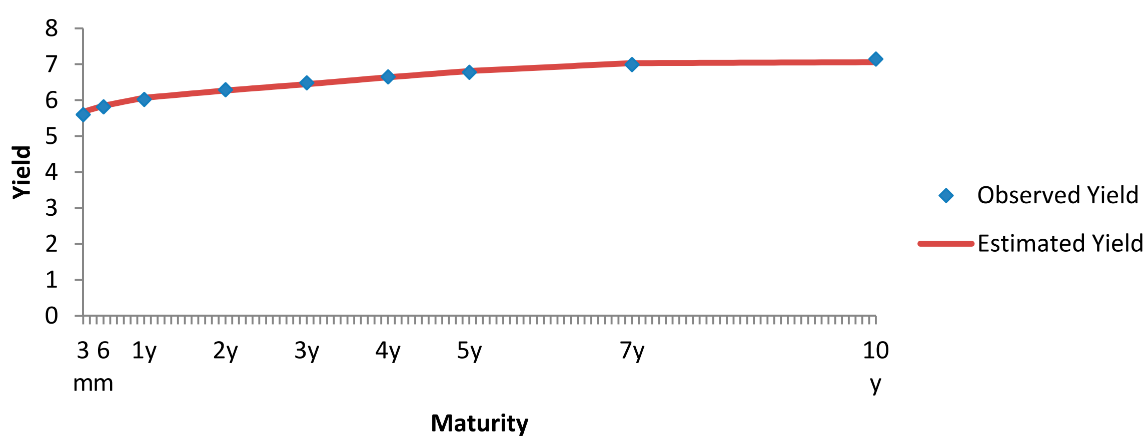
Average Term Structure Fitting, 1972.1–2011.6, Three-Factor NSM



Average Term Structure Fitting, 1972.1–2011.6, Four-Factor Svensson Model

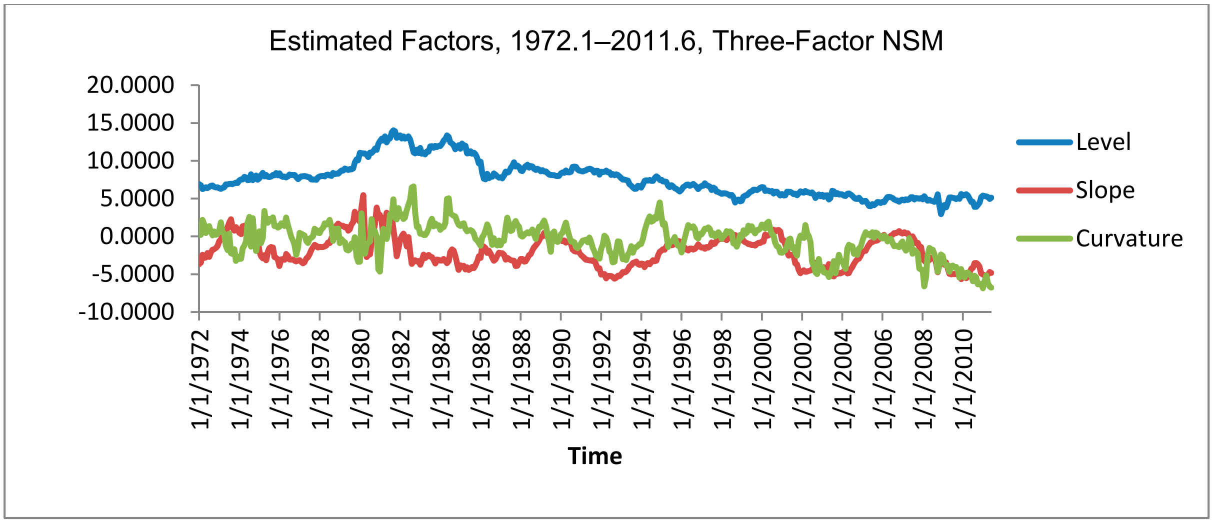


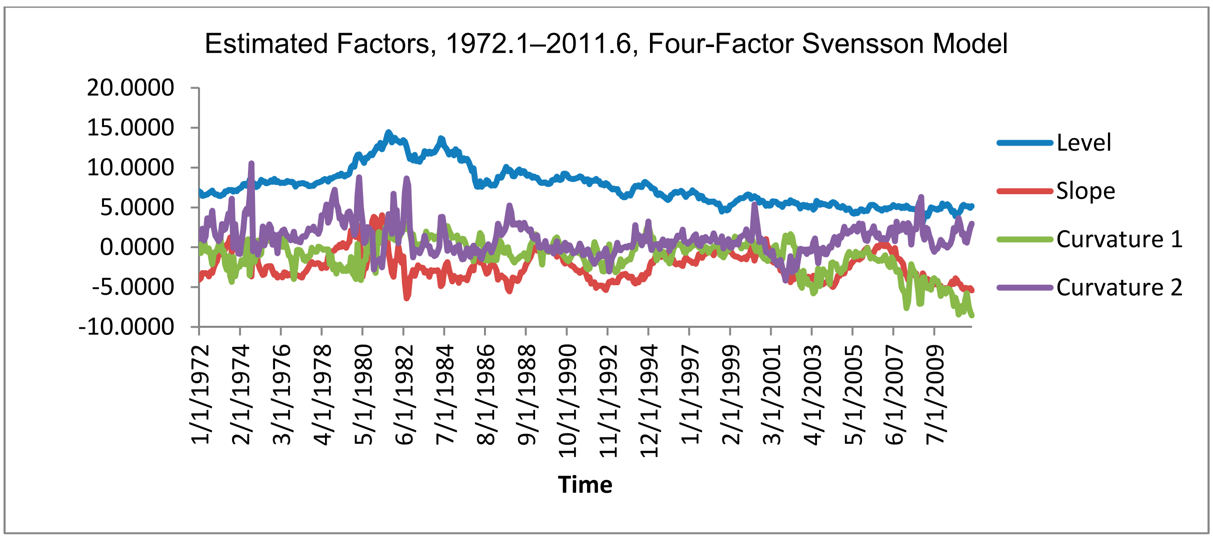
Average Term Structure Fitting, 1972.1–2011.6, Two-Factor CIR Model

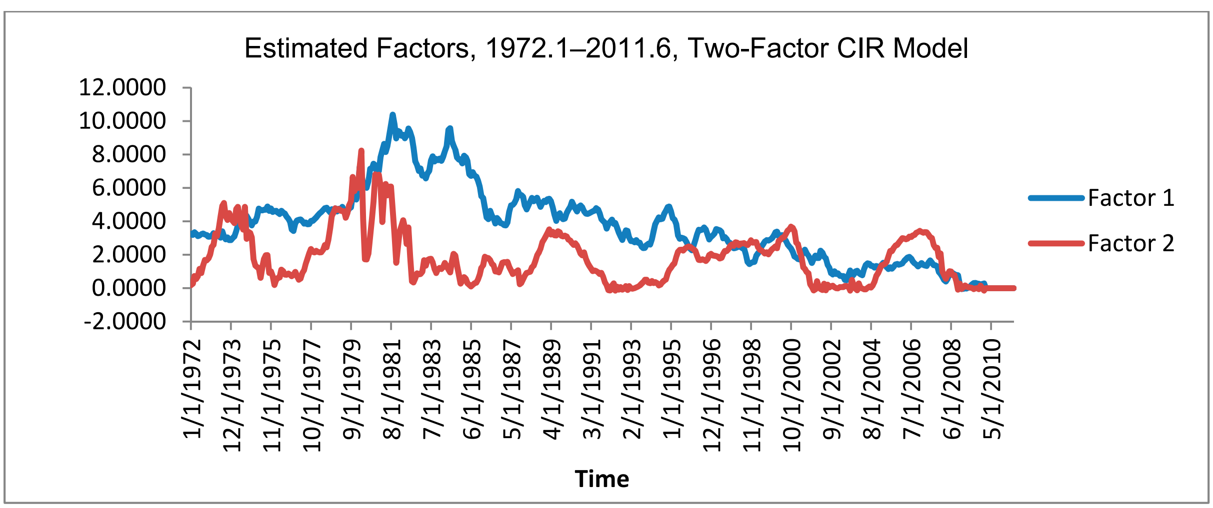


Source: Fund staff estimates.

Figure 5. Estimation Factors







Source: Fund staff estimates.

Table 1. Goodness of Fit

Chi-square Test of Fit, Three-Factor NSM, 1972.1–2011.6

	Value
SSE	34.2018
Chi-square	6.9892
DF	4263

Chi-square Test of Fit, Four-Factor Svensson Model, 1972.1–2011.6

	Value
SSE	9.2137
Chi-square	2.0924
DF	4262

Chi-square Test of Fit, Two-Factor CIR Model, 1972.1–2011.6

	Value
SSE	184.9339
Chi-square	49.3508
DF	4264

Source: Fund staff estimates.

Table 2. Goodness of Fit of the Yield- Macro NSM

Chi-square Test of Fit, Three-Factor NSM, 1972.1–2011.6

	Value
SSE	34.2669
Chi-square	6.4113
DF	4260

Source: Fund staff estimates.

Table 3. Variance Decomposition, Three-Factor NSM

	Level	Slope	Curvature
3 Month	0.1067	0.7463	0.1471
6 Month	0.1225	0.7232	0.1543
1 Year	0.1522	0.6810	0.1669
2 Year	0.2039	0.6166	0.1795
3 Year	0.2477	0.5737	0.1787
4 Year	0.2854	0.5440	0.1706
5 Year	0.3179	0.5223	0.1598
7 Year	0.3696	0.4919	0.1386
10 Year	0.4213	0.4628	0.1159

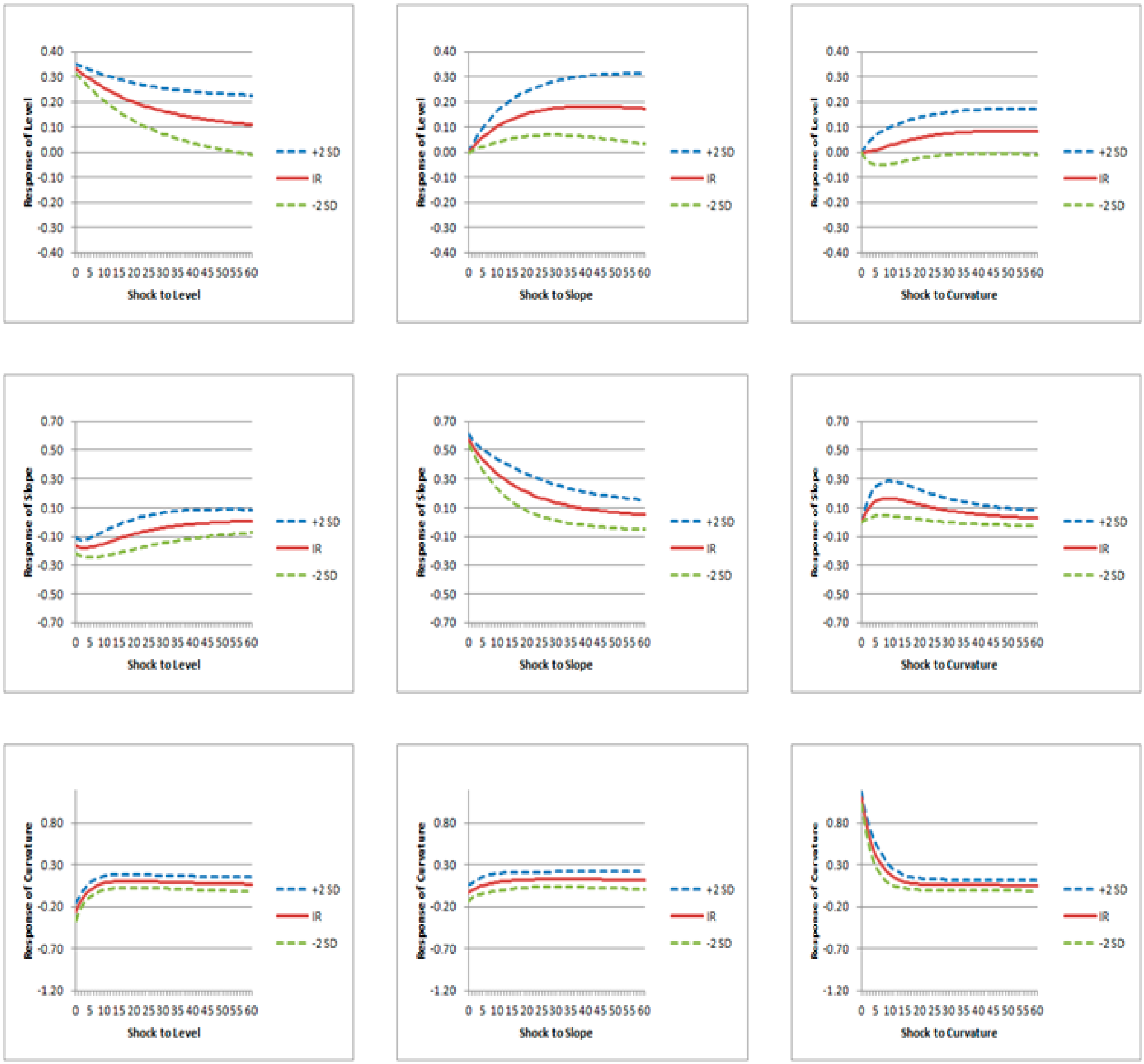
Source: Fund staff estimates.

Table 4. Variance Decomposition, Four-Factor Svensson Model

	Level	Slope	Curvature 1	Curvature 2
3 Month	0.1167	0.6212	0.1395	0.1225
6 Month	0.1317	0.5942	0.1438	0.1303
1 Year	0.1587	0.5656	0.1527	0.1230
2 Year	0.2049	0.5422	0.1667	0.0862
3 Year	0.2420	0.5279	0.1697	0.0604
4 Year	0.2721	0.5172	0.1649	0.0459
5 Year	0.2966	0.5089	0.1570	0.0375
7 Year	0.3331	0.4965	0.1411	0.0293
10 Year	0.3671	0.4837	0.1245	0.0248

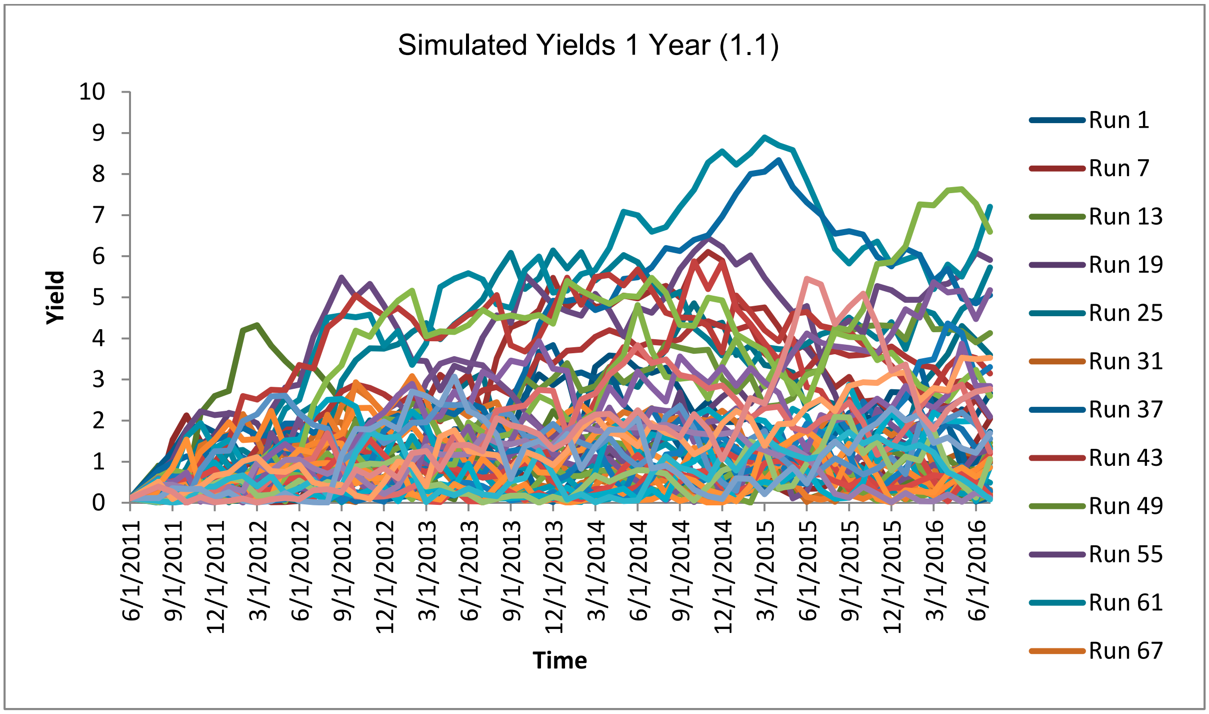
Source: Fund staff estimates.

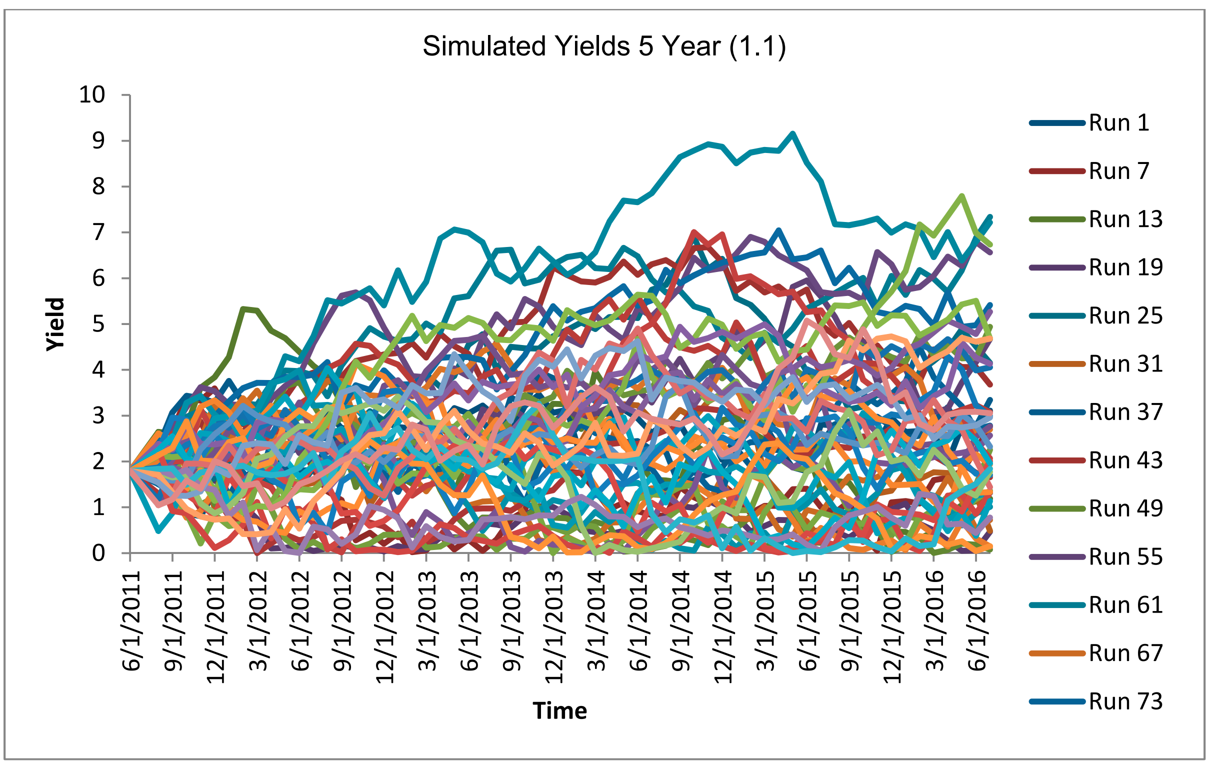
Figure 6. Impulse Functions Based on Yield-Only NSM



Source: Fund staff estimates.

Figure 7. Simulated Yields Based on Yield-Only NSM





Source: Fund staff estimates.

Table 5. Simulation Statistics

	1 Year	5 Year
Mean	2.5390	3.1707
Median	2.0113	2.9964
Std Deviation	2.0150	2.1031
Skewness	0.9250	0.4552
Kurtosis	0.0286	-0.5214
Min	0.0820	0.0015
Max	8.5674	9.1877

Source: Fund staff estimates.

C. Cox, Ingersoll and Ross (CIR) Models

The estimation of the preference-free, two-factor CIR model using a state-space representation explains well the term structure of U.S. Treasury security yields.[16] This estimation captures the many and significant changes in the shape of the observed term structure over 1972:1–2011:6. The estimation residuals are small (Figure 2), and the goodness-of-fit test measured by the Chi-square confirms that the estimated term structure fits well the observed term structure (Table 2). The time series fit of the two-factor CIR model also accounts well for the observed time series for the yields of the one-year and five-year U.S. Treasury securities (Figure 3). This result holds for the other maturities of the U.S. Treasury securities included in this paper as well.

Even though the estimation of the preference-free, two-factor CIR model does not appear to be as good as the estimates of the three-factor NSM and four-factor Svensson models, a comparison of these estimates requires caution. As Figures 1 and 2 show, the dispersion of the estimation residuals of the CIR model are larger than the dispersion of the estimation residuals of the three-factor NSM and four-factor Svensson models. As the Chi-square test statistics show, the goodness-of-fit of the estimation of the CIR model is not as good as the goodness-of fit of the estimations of the three-factor NSM and four-factor Svensson models. However, the differences in the estimation of these models reflect the fact the three models have a different number of factors, and a comparison is not that straightforward.

[16] The estimation of the preference-free, one-factor CIR model yields a poor result, and is not reported in this paper.

V. Conclusions

This paper assesses estimates of term structure models for the United States. In this context, it first describes the mathematics underlying both the Nelson-Siegel and Cox, Ingersoll and Ross family of models and estimation methodologies. It then presents estimations of some of these models within these families of models—a three-factor, yield-only Nelson-Siegel model, a four-factor Svensson model, and a preference-free, two-factor CIR model—for the United States from 1972 to mid 2011. It subsequently assesses these estimations.

The paper finds that the estimations of the term structure models presented in this paper capture well the dynamics of the term structure in the United States. The estimations of the three-factor NSM, four-factor Svensson model, and preference-free, two-factor CIR model capture well the dynamics of the term structure over 1972:1-2011:6. These estimations encapsulate the changes in expectations of short-term future interest rates, while confirming that the yield-factors of the term structure of interest rates—level, slope and curvatures—provide a good representation of the term structure. Such estimations provide support for the notion that these latent factors help explain the dynamics of the term structure.

REFERENCES

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

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

Dai, Q., and K.J. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *The Journal of Finance*, 55, 1443–78.

Diebold, F. X., M. Piazzesi, and G. D. Rudebusch, 2005, "Modeling Bond Yields in Finance and Macroeconomics," *American Economic Review Papers and Proceedings*, 95, 415–20.

___________, and C. Li, 2006, "Forecasting the Term Structure of Government Bond Yields," *Journal of Econometrics*, 130, 337–64.

___________, G. D. Rudebusch, and B. Aruoba, 2006, "The Macroeconomy and the Yield Curve: a Dynamic Latent Factor Approach," *Journal of Econometrics*, 131, 309–38.

Gasha, G., Y. He, C. Medeiros, M. Rodriguez, J. Salvati, and J. Yi, 2010, "On the Estimation of Term Structure Models and An Application to the United States," Working Paper 10/258, November (Washington DC: International Monetary Fund).

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

Medeiros, C., and M. Rodríguez, 2011, "The Dynamics of the Term Structure of Interest Rates in the United States in Light of the Financial Crisis of 2007–10," Working Paper 11/84, April (Washington DC: International Monetary Fund).

Nawalkha, S.K., N.A. Believa, and G.M. Soto, 2007, *Dynamic Term Structure Modeling* (Hoboken: John Wiley & Sons, Inc.).

Shreve, S.E., 2004, *Stochastic Calculus for Finance II, Continuous-Time Models* (New York: Springer Science + Business Media, Inc.).

Svensson, L.E.O., 1994, "Estimating and Interpreting Forward Interest Rates: Sweden 1992–1994," Working Paper No. 4871, September (Cambridge, Massachusetts: National Bureau of Economic Research).