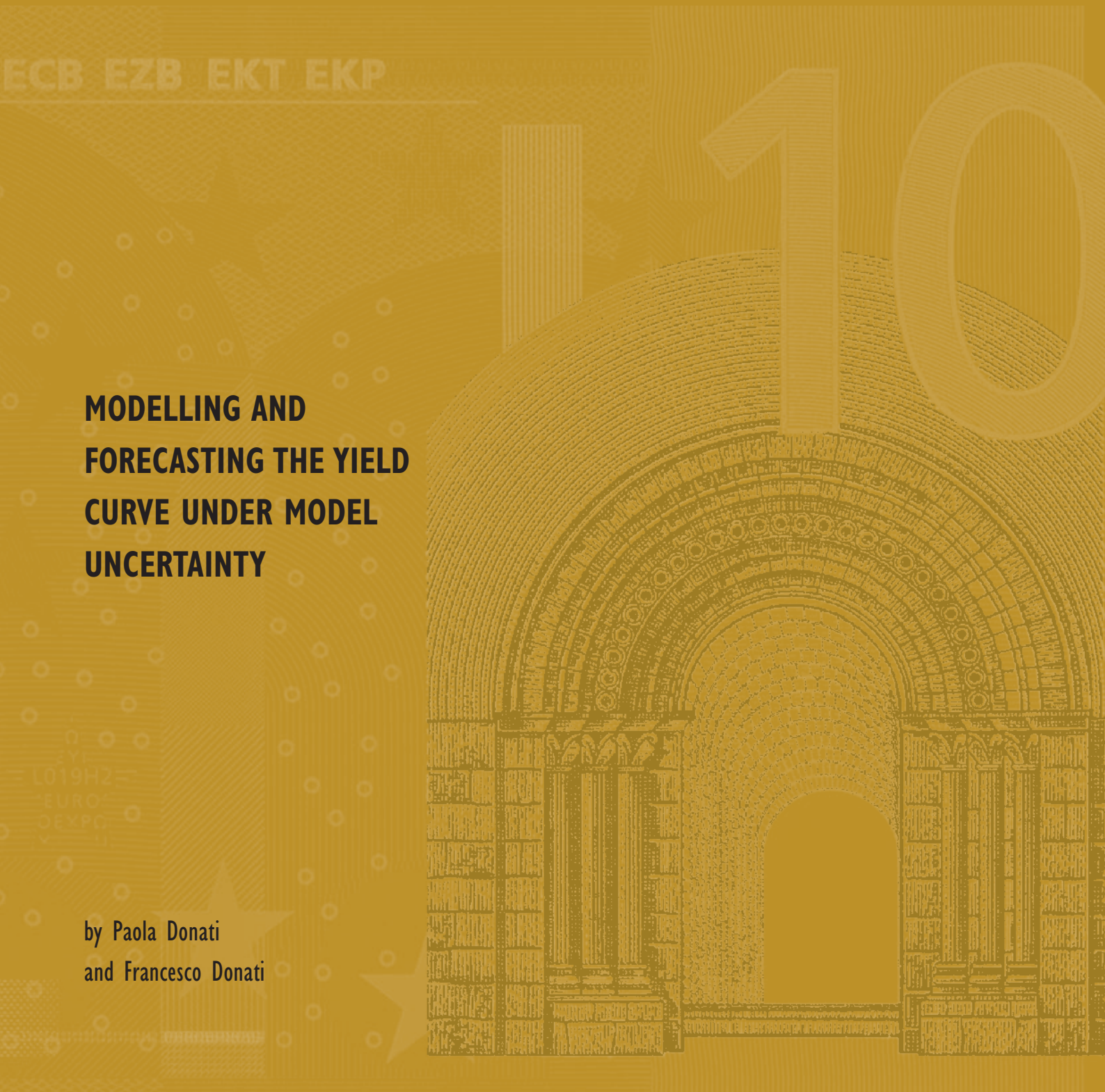
EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 917 / JULY 2008



MODELLING AND FORECASTING THE YIELD CURVE UNDER MODEL UNCERTAINTY

by Paola Donati

and Francesco Donati

EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 917 / JULY 2008

MODELLING AND FORECASTING THE YIELD CURVE UNDER MODEL UNCERTAINTY[1]

by Paola Donati[2] and Francesco Donati[3]

In 2008 all ECB publications feature a motif taken from the €10 banknote.

1 Most of this paper was written while Paola Donati was visiting the Federal Reserve Bank of New York which she thanks for its hospitality and for providing an excellent research environment. We would like to thank Andrew Ang, Geert Bekaert, Leonardo Bartolini, Matteo Ciccarelli, Robert Dekle, Roberto De Santis, Fabio Fornari, Arturo Estrella, Paolo Pesenti, Tony Rodrigues, Cedric Tille and an anonymous referee for helpful comments and discussions. We thank also the seminar participants at the Federal Reserve Bank of New York, the European Central Bank, the Euro Working Group on Financial Modelling and the 15th World Congress of the International Economic Association. We are grateful to the programmers of EicasLab for providing the software suite used in this project. The views expressed are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of New York or the European Central Bank.

2 European Central Bank, Kaiserstrasse 29, D-60311 Frankfurt am Main, Germany; phone +49-(0)69-1344-7885; e-mail: paola.donati@ecb.europa.eu

3 Politecnico of Torino, Department of Control and Computer Engineering, Corso Duca degli Abruzzi 24, I-10129 Torino, Italy; e-mail: francesco.donati@polito.it

Address
Kaiserstrasse 29
60311 Frankfurt am Main, Germany

Postal address
Postfach 16 03 19
60066 Frankfurt am Main, Germany

Telephone
+49 69 1344 0

Website
http://www.ecb.europa.eu

Fax
+49 69 1344 6000

ISSN 1561-0810 (print)
ISSN 1725-2806 (online)

CONTENTS

Abstract

This paper proposes a procedure to investigate the nature and persistence of the forces governing the yield curve and to use the extracted information for forecasting purposes. The latent factors of a model of the Nelson-Siegel type are directly linked to the maturity of the yields through the explicit description of the cross-sectional dynamics of the interest rates. The intertemporal dynamics of the factors is then modeled as driven by long-run forces giving rise to enduring effects, and by medium- and short-run forces producing transitory effects. These forces are reconstructed in real time with a dynamic filter whose embedded feedback control recursively corrects for model uncertainty, including additive and parameter uncertainty and possible equation misspecifications and approximations. This correction sensibly enhances the robustness of the estimates and the accuracy of the out-of-sample forecasts, both at short and long forecast horizons.

JEL classification: G1, E4, C5

Keywords: Yield curve; Model uncertainty; Frequency decomposition; Monetary policy

Non technical summary

The anomalous behavior the yield curve displayed in many countries in the recent past has attracted a great deal of attention on how to identify, and possibly predict, the factors driving the forward and the spot rates as a function of their maturity. Central banks responding to economic developments to maintain price and macroeconomic stability and investors taking positions in financial markets, to act effectively need to know, timely, about the nature and the persistence of the shocks driving the observed movements in the yields. In particular, given that the evolution of the forward rates, and then of the yield curve, is driven by a large number of economic forces acting simultaneously that may temporary overlap or offset each other, to avoid a blurred picture the effect produced by each type of shock should be tracked and forecasted individually.

For the purpose of this investigation, we propose to use a dynamic, parametric, yield curve model defined both in maturity-time (to explain the cross-sectional dynamics of yields) and in calendar-time (to explain the intertemporal dynamics yields). The cross-sectional dynamics of the rates is described by a variation of the Nelson-Siegel model (1987), that employs two different exponential decays terms and three latent factors, which retain the interpretation of level, slope and curvature of the yield curve common in the literature. At the same time, the latent factors can be directly associated with the maturity of the yields. In particular, a factor, labeled the *starting value*, aggregates the information contained in the short-end of the curve; another factor, labeled the *shape factor*, aggregates the information contained in middle-range maturities, while the third factor, labeled the *final value*, aggregates the information contained in the long-end of the yield curve. By partitioning the maturity spectrum the latent factors makes it easier to investigate the calendar-time dynamics of the yields.

The novelty of the approach proposed in this paper rests on the modeling of the calendar-time dynamics of the yields. The calendar-time dynamics of each latent factor is modeled as the response to three independent classes of forces: 1) long-run shocks giving rise to enduring effects that may persist up to infinity; 2) medium-run forces giving rise to transitory effects that wane within business-cycle horizons; 3) short-run forces giving rise to transitory and short-lived effects. The effects of these three types of forces are tracked by the frequency components composing the time series of each latent factor. Specifically, the long-run forces drive the low-frequency component of these time-functions, the business-cycle forces drive their medium-frequency component, and the short-lived forces govern their high-frequency component. Therefore, the low-, medium- and high-frequency component of the latent factors will reflect the effect of the long-, medium- and short-run forces driving the yields of short-, medium- and long-maturities.

The frequency components belong to three disjoint, pre-determined, frequency bandwidths so that the effect of each type of force on the correspondent frequency component can be modeled and investigated individually. The frequency components and the forces steering their dynamics are

estimated with a filter, i.e. an input-output state observer, that compares with the Kalman filter that is often used in the term structure literature, which through a feedback control ensures that each frequency component evolves within its pre-specified frequency bandwidth and that the sum of the three frequency components accurately reconstructs and out-of-sample forecasts the latent factor.

The advantages of using this state observer are: First, that it permits to investigate the behavior of the yield curve in the frequency domain starting from a model designed in the time domain, which makes the presentation and evaluation of the results straightforward. Second, that it permits to decompose the series of interest into the selected frequency components in real time, and in a way that ensures that the availability of new observations does not alter the pattern of the frequency components extracted in the past. This is important because to investigate the determinants of yield curve movements we need to monitor how the frequency components of the yields evolve with time. Thus, we need to avoid that when mapped into the frequency domain, which is time invariant by definition, the values that the frequency components take on at the beginning of the sample we consider gets averaged with the values they take on at the end of the sample, thereby mixing the past with the future. Third, the state observer permits to carry out the frequency decomposition also of nonstationary time series, without Fourier-transforming the data, thereby avoiding the potential information losses which may occur when the variables are de-trended. Fourth, it permits to extract all the frequency components at the same time. Finally, this state observer has been designed to take into account that the modeler may have limited information on the actual cause-and-effects relationships ruling the dynamics of the yields and on the statistical properties of the shocks and the employed state variables. We call this lack of knowledge model uncertainty, in line with the notion of uncertainty originally put forth by Knight (1921). By correcting for model uncertainty, this state observer produces robust estimates of the states and inputs of the models and a guaranteed performance also when the models are used to out-of-sample forecast the yields.

This paper investigates the behavior of the U.S. yield curve from January 1984 to December 2007. Our results show that the long-term rates have been essentially governed by long-run, slowly evolving forces, whereas the short-end of the yield curve has been largely driven by business cycles forces. Moreover, the contrast of the frequency components and of the shocks extracted from the variables modeling the rates shows that the long-run forces that have steered the underlying trend of the long-term rates display a rather different pattern than the long-run forces governing the trend of the short-term rates, thereby questioning the functional dependence between the long-term rates and the short-term rates. Finally, the adopted model specification in conjunction with the use of the state observer correcting for model uncertainty permits to obtain accurate out-of-sample forecast of the yields. These are better than the commonly used random walk benchmark both for short and long maturities, and at short and long forecast horizons, also during the 2004-2006 "conundrum" years, when the behavior of the long-end of the yield curve was difficult to understand.

Introduction

The anomalous behavior the yield curve displayed in many countries in the recent past has attracted a great deal of attention on how to identify, and possibly predict, the factors driving the forward and the spot rates as a function of their maturity. Central banks responding to economic developments to maintain price and macroeconomic stability and investors taking positions in financial markets, to act effectively need to know, timely, about the nature and the persistence of the shocks driving the observed movements in the yields. In particular, given that the evolution of the forward rates, and then of the yield curve, is driven by a large number of economic forces acting simultaneously that may temporary overlap or offset each other, to avoid a blurred picture the effect produced by each type of shock should be tracked and forecasted individually.

This paper proposes an approach to investigate systematically, and in real time, the nature of the forces driving the calendar-time dynamics of the forward and the spot rates and shows how to use the related information to enhance the accuracy of the out-of-sample forecasts of the yields, also at long forecast horizons.

To explain the maturity-time dynamics of the interest rates we use a variation of the model proposed by Nelson and Siegel (1987). This choice is motivated by the widespread popularity of models of the Nelson-Siegel type among practitioners and central banks. Such popularity is due to the fact that in addition to being parsimonious and easy to estimate, this type of model fits the data well and it may produce good out-of-sample forecasts, as shown by Diebold and Li (2006). Until recently, however, the Nelson-Siegel specification faced the criticism of not enforcing by construction the absence of arbitrage opportunities, as emphasized by Bjork and Christensen (1999), because it does not ensure that the intertemporal evolution of the yields is consistent with the shape that the yield curve takes on at each point in calendar-time. Christensen, Diebold and Rudebusch (2007) obviate such pitfall by adding to the dynamic Nelson-Siegel specification a mathematical correction, which ensures the fulfillment of the no-arbitrage constraints, but also increases the computational complexity. The results obtained by Coroneo, Nyholm and Vidova-Koleva (2008), however, suggest that the simplicity of the Nelson-Siegel model may be preserved, because its empirical performance is not statistically different from that of a three-factor no-arbitrage model, at least for U.S. data. The findings of these authors corroborate the hypothesis put forth by Diebold, Rudebusch and Aruoba (2006) that when a model is flexible enough to provide a satisfactory fit of the yields, it is likely to satisfy the no-arbitrage restrictions to the extent that they are satisfied by the data. This paper rests on this assumption.

To offer the possibility of selecting the model specification which better fits the data of interest, this paper considers a class of cross-sectional dynamic parametric models which generalizes the Nelson-Siegel formulation. Observing that the functions solving for differential or difference equations possess the fitting smoothness and flexibility requirements recommended by Vasicek and Fong

(1982), giving rise at the same time to a parsimonious model, Nelson and Siegel suggest to use the solution to a second-order differential equation with real and *equal* roots to model the forward rate curve. Instead of starting from a pre-determined solution, in this paper the maturity-time dynamics of the forward rates is explicitly modeled with a system of differential equations. When (optimally) fitted to the data, the solution to these equations may, or may not, give rise to the Nelson-Siegel solution. Therefore, if a three factor model is selected, the possibility that this solution includes real and *unequal* roots or *complex-conjugate* pairs is not excluded ex-ante. Indeed, when considering U.S. zero-coupon bond yields we obtain that a three-latent factor model with two unequal roots, and therefore two different exponential decay terms, fits the data better than the standard Nelson-Siegel yield curve formulation, especially over the last few years.

The system to model the cross-sectional dynamics of the forward rates (and thus of the yields) is designed in way that its three latent factors, when combined, can still be interpreted as the level, slope and curvature of the yield curve in the sense of Litterman and Scheinkam (1991), but they may be given also an interpretation with a more straightforward economic meaning. Specifically, one factor labeled *the starting value* aggregates the information contained in the short-end of the curve; a second factor labeled *the shape factor* aggregates the information contained in middle-range maturities, whereas the third factor, labeled the *final value*, aggregates the information contained in the long-end of the yield curve. We exploit the property of the latent factors to be directly associated with the maturity of the yields to gain insights on the nature and persistence of the forces driving the observed intertemporal movements in the yields.

The model of the calendar-time dynamics of the latent factors is the main contribution of this paper. At each point in calendar-time, we divide the frequency domain in which are defined the forces steering the intertemporal movements of the latent factors into three disjoint, pre-determined, frequency bandwidths. As a result, these forces are classified into: 1) *long-run shocks* giving rise to enduring effects that may persist up to infinity; 2) *medium-run forces* giving rise to transitory effects that wane within business-cycle horizons; 3) *short-run forces* giving rise to transitory and short-lived effects. To track the effects of these otherwise unidentified three classes of forces on the rates, we decompose the calendar-time series of each latent factor into three frequency components lying within the same three frequency bandwidths. Since the frequency bandwidths are disjoint, the effect of each type of force on the correspondent frequency component can be modeled and investigated individually.

We proceed as follows. We explicitly model the evolution of each frequency component of each latent factor as the output of a strictly-causal, linear, dynamic system whose parameters are optimally identified. The exogenous causes, or inputs, controlling the dynamics of these systems are the same forces lying within the three pre-specified frequency bandwidths. Thus, the long-run forces drive the low-frequency component of the latent factors, the business-cycle forces drive their medium-frequency component, and the short-lived forces govern their high-frequency component.

The inputs steering the dynamics of the frequency components are recursively estimated with a dynamic filter, which acts as a band-pass filter. At the same time, the filter estimates the states of the dynamic systems. We use a filter that compares with the Kalman filter, although it has less stringent requirements on the statistical properties of the input process, which may or may not be stationary, and on the knowledge that the modeler is expected to have on the statistical properties of these and the state processes. Actually, it is undertaken that the modeler has limited information on the properties of the shocks and the states, in line with the notion of uncertainty originally put forth by Knight (1921). This implies that it is also undertaken that the model used to describe the dynamics of the yields is a simplification, and possibly a misspecification, of the true data generating process. We call this lack of knowledge, which encompasses additive and parameter uncertainty, *model uncertainty*. The filter, namely the input-output state observer, which is a linear, time-invariant, dynamic system, in recursively estimating the states and the inputs, corrects also for the degradation in the model performance due to model uncertainty. This is achieved by exogenously setting the eigenvalues (or, equivalently, by placing the poles) characterizing its dynamics. Since the eigenvalues of the gain matrix of the filter have a frequency-domain interpretation, by placing its poles we are able to control also the spectral decomposition. In addition to being fairly straightforward, this technique explicitly permits to integrate the information the state observer extracts from the data with other sources of information as the experience of the modeler.

This approach to perform the frequency decomposition of the latent factors has a number of advantages. To start with, 1) it allows to investigate the behavior of the frequency components in the frequency domain although it works in the time domain. Importantly, it allows to filter the data in *real time*, which means that the decomposition of a time function at time t is performed without requiring the knowledge of the values that the time function will take on at time $t+1 > t$ and without altering the outcomes of the decomposition already performed at time $t-1 < t$. This is important because to investigate the determinants of yield curve movements we need to monitor how the frequency components of the yields evolve with time. Thus, we need to avoid that when mapped into the frequency domain, which is time-invariant by definition, the values that the frequency components take on at the beginning of the sample considered gets averaged with the values they take on at the end of the sample, thereby mixing the past with the future. Moreover, 2) the state observer permits to decompose also nonstationary time series, thereby avoiding the potential information losses which may occur when the variables are de-trended, and it does imply the Fourier-transformation of the data, thereby making the interpretation of the results straightforward. In addition, 3) it permits to extract all the selected frequency components at the same time. Finally, 4) through its feedback reaction, the state observer ensures that the frequency components of each latent factor keep orthogonal to each other, that each frequency component evolves within its pre-specified frequency bandwidth, and that the sum of the three frequency components accurately reconstructs, and out-of-sample forecast, each latent factor. The out-of-sample forecasts of the

latent factors computed by the state observers when plugged into the yield curve model give rise to the out-of-sample forecasts of the yields.

By explicitly modeling the cause-and-effect relationship between the yield curve and the shocks driving its dynamics the proposed approach generalizes the approach adopted by Estrella and Mishkin (1997) and Evans and Marshall (2007) who investigate in the time domain the link between a few selected shocks and the pattern followed in response by a few selected yields through the estimation of the implied impulse response functions, and the approach of Sargent (1968), Pippenger (1974), Brick and Thompson (1978), Assenmacher-Wesche and Gerlach (2008) who perform the spectral decomposition of selected interest rates series to investigate the present periodicities and how rates of different maturities, and on different assets, are related to each other and to monetary policy.

In this paper we investigate U.S. zero-coupon bond yields from January 1984 to December 2007. Our results show that the long-term rates have been essentially governed by long-run, slowly evolving forces, whereas the short-end of the yield curve has been largely driven by business cycles forces. Indeed, the long-end of the yield curve displays little volatile medium and high-frequency components, thereby appearing less sensitive to short-lived forces, and more anchored. Moreover, the contrast of the frequency components and of the shocks extracted from the variables modeling the rates shows that the long-run forces that have steered the underlying trend of the long-term rates display a rather different pattern than the long-run forces governing the trend of the short-term rates, which supports the literature questioning the functional dependence between the long-term rates and the short-term rates put forth by the Expectations Hypothesis.

Finally, the adopted model specification produces fairly accurate out-of-sample forecast of the yields, also thanks to the fact that the state observer corrects for model uncertainty when estimating the dynamic systems employed, thereby avoiding that the accumulation of computational and modeling inaccuracies undermines the model performance. The out-of-sample forecasts obtained are better than the commonly used random walk benchmark both for short and long maturities, and at short (1 to 6 month) and long (12 to 24 month) forecast horizons, also during the June 2004 to June 2006 "conundrum" years, when the behavior of the long-end of the yield curve was difficult to understand, and the outbreak of the 2007 financial turmoil. Whereas the random walk benchmark has been traditionally difficult to beat in the term-structure literature, even by explicitly arbitrage-free models as shown by Duffee (2002), using different models and different methodological approaches Diebold and Li (2006) and Chua, Foster, Ramaswamy and Stine (2007) have also obtained forecasts more accurate than the random walk.

The paper is organized as follows. After illustrating the notation and the data employed, Section 4 introduces the class of dynamic parametric yield curve models and presents the results obtained from the selected cross-sectional model. Section 5 introduces the class of calendar-time dynamic models and presents the frequency decompositions obtained when applying the model to the esti-

mated latent factors and the out-of-sample forecasts of the yields. Section 6 draws the conclusions.

1 Notation

For any given point in calendar-time $t \in \mathbb{N}$, the spot rate $y(m,t)$ on a zero-coupon bond providing a unit redemption payment at time $t+m$, is defined as

$$P(m,t) = e^{-y(m,t)m}$$

where $P(m,t)$ is the price of the bond. We refer to spot rates as a function of their maturity-time as the (spot) yield curve or the term structure of interest rates. By construction, the spot rates average instantaneous forward rates over the term of maturity of the zero-coupon bonds to exclude arbitrage opportunities,

$$y(m,t) = \frac{1}{m}\int_0^m fw(u,t)du. \tag{1}$$

where the instantaneous forward rate $fw(m,t)$ is the marginal rate of return from the reinvestment of a zero-coupon bond in a $(m+1)-$ maturity zero-coupon bond.

2 The Data

We consider month-end U.S. zero-coupon bond yields of maturities from 6 months to 120 months, regularly spaced at 6-month maturity intervals (20 maturities $m = m_1, \ldots, m_{20}$, with $m_1 = 6$ and $m_{20} = 120$ months) covering the period from 31 January 1984 to 31 December 2007 (denoted $t = 1, \ldots, 288$ monthly observations). The yields are extracted from the market prices of U.S. Treasury coupon bonds by applying the smoothing splines method of Fisher, Nychka and Zervos (1995) and are collected daily by the Bank for International Settlements.

Table 1 shows that over the period examined, the average U.S. yield curve is upward sloped, short-maturity yields are more volatile than the long-maturity ones and that, whereas yields at all maturities are potentially non-stationary, their autocorrelation is more persistent for long-maturity yields. Noticeably, with the increasing of the maturity, the yield distributions progressively depart from the Gaussian distribution because the skewness of the yield distributions progressively departs from zero, thereby becoming increasingly asymmetric, and the kurtosis progressively departs from three, so that the distribution display increasingly thick tails. These are not properties specific to the data we consider; also Ang and Piazzesi (2003) remark that the data they use, i.e. the Fama CRSP zero-coupon bond yields, fail to meet Gaussian distribution assumptions over the period they examine.

3 The Class of Cross-Sectional Dynamic Parametric Models

We consider a class of cross-sectional dynamic models with a state-space representation designed in continuous (maturity-) time in order to avoid discontinuities in the estimate and forecast of the term structure (but the model can be easily formulated also in discrete time). The models are linear and time-invariant, and are formulated in a canonical form that guarantees simplicity and computation tractability. They are defined by the following properties:

1. For any point in calendar-time t, the $m-$maturity instantaneous forward rate $fw(m,t)$ is modeled as the output of a linear, time-invariant, canonical, n^{th}- order dynamic system and a reconstruction error $e(m,t)$,

$$\dot{\mathbf{x}}(m,t) = \mathbf{A}\mathbf{x}(m,t) \quad (2)$$

$$fw(m,t) = \mathbf{C}\mathbf{x}(m,t) + e(m,t) \quad (3)$$

in which state equation (2) consists of n first-order ordinary differential equations where $\mathbf{x}(m,t) = \left[\ \dots \ \ x_i(m,t) \ \ \dots \ \right]'$ is a $n-$dimensional vector of state variables, $\dot{\mathbf{x}}(m,t) \equiv d\mathbf{x}(m,t)/dm$ is a $n-$dimensional vector of state derivatives with respect to maturity-time, and $\mathbf{A}$ is a $(n \times n) -$dimensional, real, time-invariant, matrix. $\mathbf{A}$ is completely defined by its n eigenvalues. Given the property of the nominal forward rates to be only positive and to possibly display more than one local maximum (or minimum), and to ensure the stability of the system, the eigenvalues of $\mathbf{A}$ are negative or null real-valued, or they occur in complex-conjugate pairs with negative or null real parts. Given the property of the nominal forward rates to stabilize asymptotically, one eigenvalue of $\mathbf{A}$ is always null. The non-null eigenvalues are identified optimally when fitting the model to the data. According to the output equation (3), the $m-$maturity instantaneous forward rate $fw(m,t)$ is determined by the product between the $(1 \times n) -$dimensional, real, time-invariant, matrix $\mathbf{C}$ and the state vector $\mathbf{x}(m,t)$. Given that the modeler is assumed to have limited information on the process generating the forward rates, in order to take into account the related possible model approximations and misspecifications, a reconstruction error $e(m,t)$ is added to the reconstruction of each forward rate $fw(m,t)$. No constraint is imposed on the statistical properties of $e(m,t)$.

2. Given that the model state variables $x_i(m,t)$ are not directly accessible to measurement, they need to be estimated. To ensure that the knowledge of the forward rates $fw(m,t)$ of eq. (3) suffices to uniquely estimate $\mathbf{x}(m,t)$, we select a canonical representation for the model ensuring that the $(n \times n) -$dimensional matrix $O(m,t) = \int_0^m \mathrm{e}^{\mathbf{A}'u}\mathbf{C}'\mathbf{C}\mathrm{e}^{\mathbf{A}u}du$ is nonsingular.

3. By plugging the unique solution $\mathbf{x}(m,t) = \mathrm{e}^{\mathbf{A}m}\mathbf{x}(0,t)$ of eq. (2) into eq. (3) we obtain the

instantaneous forward rate curve

$$fw(m,t) = \mathbf{C}\, \mathrm{e}^{\mathbf{A}m}\, \mathbf{x}(0,t) + e(m,t). \tag{4}$$

4. By plugging eq. (4) into eq. (1) we obtain the (spot) yield curve or term structure of interest rates

$$y(m,t) = \frac{1}{m}\, \mathbf{C}\mathbf{A}^{-1}\, \mathrm{e}^{\mathbf{A}m}\, \mathbf{x}\,(0,t) + \boldsymbol{\varepsilon}(m,t) \tag{5}$$

where $\varepsilon(m,t)$ are the reconstruction error associated to the spot rate $y(m,t)$. The errors $e(m,t)$ and $\varepsilon(m,t)$ are linked by the same relation that links the forward to the spot rates.

5. When the eigenvalues of the matrix $\mathbf{A}$ are single, real and unequal, the forward rate curve of eq. (4) can also be written in the explicit form

$$fw(m,t) = \sum_{i=1}^{n} \beta_i(m,t)\, \mathrm{e}^{\lambda_i m} + e(m,t) \tag{6}$$

and the yield curve of eq. (5) can be rewritten as

$$y(m,t) = \sum_{i=1}^{n} \beta_i(m,t) \left(\frac{\mathrm{e}^{\lambda_i m} - 1}{\lambda_i m} \right) + \varepsilon(m,t) \tag{7}$$

where λ_i are the eigenvalues of $\mathbf{A}$ and the coefficients β_i linearly combining the exponential functions depend on the initial states in a form determined by the structures of the matrices $\mathbf{A}$ and $\mathbf{C}$.

Note that eqs. (4) and (5) generate a family of forward rate curves $fw\,(\mathbf{A},\mathbf{C};m,t)$ and yield curves $y\,(\mathbf{A},\mathbf{C};m,t)$ that are fully parametrized by the pair $\{\mathbf{A},\mathbf{C}\}$. This means that when the initial states $\mathbf{x}(0,t)$ are known, the triplet $\{\mathbf{A},\mathbf{C},\mathbf{x}(0,t)\}$ fully characterizes all the maturity spectrum of the forward and the spot rates. Therefore, the pair $\{\mathbf{A},\mathbf{C}\}$ may be interpreted as reflecting time-invariant characteristics of the economy. Instead, the initial states may be interpreted as aggregating all the information that, at each point in calendar-time t, market participants use to price the bonds included in the time-t term structure. If at a subsequent point in time $t+1$, no matter how close to time t, other information becomes available, the values taken on by the initial states change thereby giving rise to another term structure.

3.1 Selection and Estimation of the Cross-Sectional Dynamic Model

Using different model assumptions and methodological approaches, Nelson and Siegel (1987), Litterman and Scheinkman (1991), de Jong (2000), Duffee (2002), Lamoureux and Witte (2002), among others, show that three variables may satisfactorily explain most of the movements in U.S. Treasury

yields (and prices). Against this background, to reconstruct the spot rates over the twenty-four years we consider we initially select a cross-sectional dynamic model of the $3^{rd}-$ order and we appraise its fitting performance against an overall *fitting accuracy target* defined by a root mean squared error smaller than 5.0 basis points. We find that this model largely meets our accuracy objective by reconstructing the $5,760$ spot rates we consider with root mean squared error of less than 3.5 basis points. Thus, we do not investigate the performance of more complex models.

In particular, we explain the evolution of the forward rates with the following linear, continuous-time, time-invariant, dynamic system employing three state variables,

$$\underbrace{\begin{bmatrix} \dot{x_0}(m,t) \\ \dot{x_1}(m,t) \\ \dot{x_2}(m,t) \end{bmatrix}}_{\dot{\mathbf{x}}(m,t)} = \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ a & -a & 0 \\ 0 & b & -b \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_0(m,t) \\ x_1(m,t) \\ x_2(m,t) \end{bmatrix}}_{\mathbf{x}(m,t)} \tag{8}$$

$$fw(m,t) = x_2(m,t) + e(m,t). \tag{9}$$

Output equation (9) states that for any point in calendar-time t, the forward rate of maturity m is given by the sum of the state $x_2(m,t)$ and the reconstruction error $e(m,t)$. In turn, $x_2(m,t)$ depends on $x_0(m,t)$ and $x_1(m,t)$ and on the eigenvalues of the state transition matrix $\mathbf{A}$ of state equation (8). The matrix $\mathbf{A}$ is lower-triangular and contains two time-invariant parameters a and b, which are the system eigenvalues along with the first null element in its principal diagonal. Therefore, a and b are also the inverse of the time constants that characterize the impulse response of the system: $\tau_1 = 1/a$ and $\tau_2 = 1/b$. The values of a and b will be optimally identified.

By eq. (6), the solution to the system of eqs. (8) and (9) is the forward rate curve

$$fw(m,t) = \mathbf{f}(m)\boldsymbol{\beta}(0,t) + \mathbf{e}(m,t) \tag{10}$$

where the instantaneous forward rates $fw(m,t)$ of maturities $m \in \Re$ observed at calendar-time t are defined by the vector of initial states $\boldsymbol{\beta}(0,t) \equiv \begin{bmatrix} \beta_1(0,t) & \beta_2(0,t) & \beta_3(0,t) \end{bmatrix}'$, with $\beta_1(0,t) \equiv x_0(0,t)$, $\beta_2(0,t) \equiv x_2(0,t) - x_0(0,t)$ and $\beta_3(0,t) \equiv x_1(0,t) - x_0(0,t)$; the vector of factor loadings $\mathbf{f}(m) = \begin{bmatrix} 1 & fa(m) & fb(m) \end{bmatrix}$ with $1 = \mathrm{e}^{-\infty\, m}$, $fa(m) = \mathrm{e}^{-b\, m}$ and $fb(m) = \frac{b}{b-a}\left(\mathrm{e}^{-a\, m} - \mathrm{e}^{-b\, m}\right)$; and the vector $\mathbf{e}(m,t)$ of reconstruction errors.

By eq. (5), eq. (10) gives rise to the yield curve,

$$y(m,t) \;\; = \;\; \mathbf{g}(m)\boldsymbol{\beta}(0,t) + \boldsymbol{\varepsilon}(m,t) \tag{11}$$

where the instantaneous spot rates $y(m,t)$ of maturities $m \in \Re$ observed at calendar-time t are defined by the same vector of initial states of eq. (10), the vector of factor loadings $\mathbf{g}(m) =$

$\begin{bmatrix} 1 & ya(m) & yb(m) \end{bmatrix}$ with $1 = \mathrm{e}^{-\infty\, m}$, $ya(m) = \frac{1-\mathrm{e}^{-b\, m}}{b\, m}$ and $yb(m) = \frac{b}{b-a}\left(\frac{1-e^{-a\, m}}{a\, m} - \frac{1-e^{-b\, m}}{b\, m}\right)$, and the vector $\boldsymbol{\varepsilon}(m,t)$ of spot rate reconstruction errors.

Note that the forward rate curve and the yield curve of eqs. (10) and (11) are fully characterized by the pair $(a,\ b)$ and the initial states $x_0(0,t)$, $x_1(0,t)$ and $x_2(0,t)$.

3.1.1 The Starting Value, the Final Value and the Shape Factor Interpretation

The initial states $x_0(0,t)$, $x_1(0,t)$ and $x_2(0,t)$ of the cross-sectional dynamic model of eqs. (8) and (9), henceforth denoted the *latent factors*, are directly associated with the maturity of the forward rates and the yields. In particular:

- $x_0(0,t)$, which is maturity-time invariant by construction, corresponds to the *asymptotic final value* which both the forward rate curve and the yield curve take on at their longest maturity $m = \infty$. Hence: $f(\infty,t) = y(\infty,t) = x_0(\infty,t) = x_0(0,t)$. As a result, the calendar-time series $x_0(0,t)$ aggregate the information contained in the *long-end of the yield curve.*

- $x_2(0,t)$ defines the *starting value* which both the forward rate curve and the yield curve take on at their shortest maturity $m = 0$, that is: $f(0,t) = y(0,t) = x_2(0,t)$. Then, the calendar-time series $x_2(0,t)$ aggregate the information contained in the *short-end of the yield curve.* As $m \to \infty$, the state variable $x_2(m,t)$ tends asymptotically towards $x_0(m,t)$ with a dynamic law characterized by the time constant $\tau_2 = 1/b$. Hence, $x_2(\infty,t) = x_0(\infty,t)$.

- Finally, $x_1(0,t)$ links the starting value to the final value of the forward rate curve and the yield curve. In particular, by bridging the two ends of the yield curve, the calendar-time series $x_1(0,t)$ aggregate the information contained in the *middle-range maturities of the yield curve.* Note that thanks to its bridging role, the initial state $x_1(0,t)$ defines the *shape* of both forward rate curve and the yield curve. For example, if $x_1(0,t)$ is smaller than both $x_0(0,t)$ and $x_2(0,t)$ the yield curve exhibits an inverted hump, whereas if its level is included between the initial and the final values, $x_2(0,t) < x_1(0,t) < x_0(0,t)$, the yield curve is positively and monotonically sloped. As $m \to \infty$, also the state variable $x_1(m,t)$ tends asymptotically towards $x_0(\infty,t)$ according to an exponential law specified by the function $\mathrm{e}^{-a\, m}$. Thus, $x_0(\infty,t) = x_2(\infty,t) = x_1(\infty,t)$.

3.1.2 The Level, Slope and Curvature Interpretation

The three latent factors $x_0(0,t)$, $x_1(0,t)$ and $x_2(0,t)$ can also be seen as the level, slope and curvature of the yield curve by applying the interpretation proposed by Diebold and Li (2006) and Diebold, Rudebush, Aruoba, (2006). Consider the transformations of the latent factors into the factors β and the factor loadings $\mathbf{f}(m)$ and $\mathbf{g}(m)$ introduced in eqs. (10) and (11). To compare with the literature, disregard their maturity-time dimension. In this case:

– The loading on $\beta_1(t)$ is the constant function 1 which rules the long-term component of both the forward rate curve and the yield curve given that it does not decay to zero in the limit. Given that $\beta_1(t)$ equals the asymptotic final value of the forward rate curve and the yield curve, the factor $\beta_1(t)$ is the long-term factor of the curves governing their *level*.

– The loadings on $\beta_2(t)$ rule the short-term component of the forward rate curve and the yield curve because $fa(m)$ and $ya(m)$ take on value 1 when $m = 0$ and steadily decay thereafter reaching zero when $m = \infty$. Being defined by the difference between the starting value and the asymptotic final value of the forward rate curve and the yield curve, the factor $\beta_2(t)$ is the short-term factor of the curves governing their *slope*.

– Finally, the loading on $\beta_3(t)$ governs the medium-term component of the forward rate curve and the yield curve because $fb(m)$ and $yb(m)$ are zero-valued when $m = 0$ and when $m = \infty$, so they cannot rule their short- or long-term component. Being defined by the difference between the shape factor of the forward rate curve and the yield curve and their (constant) level, the factor $\beta_3(t)$ is the medium-term factor of the curves governing their *curvature*.

3.1.3 The Link with the Nelson-Siegel model

If the two parameters a and b, which are the non-zero eigenvalues of the cross-sectional dynamic model of eqs. (8) and (9) explaining the evolution of the forward rates in maturity-time, tend in the limit to take on the same value, the two factor loadings, or basis functions, $fa(m)$ and $fb(m)$ introduced in eq. (10) become

$$fa^{NS}(m) = \lim_{b \to a} e^{-bm} = e^{-am}$$

and

$$fb^{NS}(m) = \lim_{b \to a} \frac{b}{b-a}(e^{-am} - e^{-bm}) = ame^{-am}.$$

Together with the constant function 1, the functions $fa^{NS}(m)$ and $fb^{NS}(m)$ span the space of forward rate curve

$$fw(m,t) = \beta_1(t) + \beta_2(t)\, \mathrm{e}^{-a\,m} + \beta_3(t)\, a\, m\, \mathrm{e}^{-a\,m}$$

proposed by Nelson and Siegel (1987).

Therefore, the Nelson-Siegel forward rate curve belongs to the class of cross-sectional dynamic parametric models presented in this paper. Several central banks estimate the term structure of their domestic interest rates using the Nelson-Siegel model (for a comprehensive overview see BIS (2005)). Clearly, if the forward rates and the spot rates are generated with a Nelson-Siegel model, when the three-factor model of eqs. (8) and (9) is optimally fitted to the data, it gives rise to the Nelson-Siegel specification of the forward and spot rate curves with the parameters taking on the same value, $a = b$.

3.2 Estimation of the Cross-Sectional Model

To identify the parameters a and b and estimate the latent factors $x_0(0,t)$, $x_1(0,t)$, and $x_2(0,t)$ of the cross-sectional dynamic model of eqs. (8) and (9), we iteratively minimize, employing the numerical algorithm based on the conjugate-gradient method contained in the software suite EicasLab, the quadratic loss functional

$$\min_{a,b} Fspe = \sum_{m_i=1}^{20} \sum_{t=1}^{288} [\varepsilon(m,t)]^2 \tag{12}$$

where $\varepsilon(m,t)$ are the spot rate reconstruction errors of eq. (11). We begin with an initial guess for the pair (a,b) and at the end of each iteration the Householder transformations yields the ordinary least squares estimates of the initial state values, which when plugged into eq. (11), provide the estimate of the yields.

We find that $Fspe$ is minimized by the unequal eigenvalues, $a = 0.070$ and $b = 0.039$ which correspond to the time constants, $\tau_1 = 14.257$ months and $\tau_2 = 25.874$ months. When $Fspe$ is at its minimum, the global root mean squared error, $RMSE = \sqrt{Fspe/(288 * 20)} = 3.46$ basis points, which largely meets our ex-ante fitting accuracy target of a $RMSE < 5.0$ basis points.

The model performs satisfactorily. It is flexible and, on average, it is able to reconstruct accurately with a contained systematic bias the maturity spectrum of the yields considered. The mean and the mean average error ($MAE = \frac{1}{288} \sum_{t=1}^{288} |\varepsilon(m,t)|$) of the fitting errors have low values for all maturities, and the yield curve average MAE edges below 2.5 basis points, as shown in the first and fifth columns of Table 2. The model has also a reliable reconstruction ability as indicated by the small values of the standard deviation, the minimum and maximum of the fitting errors reported in the second, third and fourth columns of Table 2, and the contained RMSE reported in the sixth column. In particular, the model performs well in reconstructing the relatively more volatile middle-range maturities of 12 to 60 months, as, on average, their fitting errors give rise to a RMSE of less than 2.9 basis points. The average RMSE for all of the twenty maturities is 3.33 basis points.

However, the fitting errors at all maturities exhibit a high, positive, first-order autocorrelation of about 0.77, which decreases to zero, on average, after 17 to 18 monthly displacements to stabilize within the range $[0.05 \div -0.28]$ thereafter. The global fitting error computed for the entire sample of $5,760$ yields displays a first-order autocorrelation of 0.81, which decreases to zero after 18 displacements and stabilizes at about -0.08 thereafter. Despite such autocorrelation, the fitting performance of the model leaves no room for any further reconstruction improvement. This becomes evident when we measure the explanatory power of the model with the norm ℓ^2. Specifically, the $5,760$ yields in our sample measure $\|y\| = \sqrt{\sum_{m_i=0}^{20} \sum_{t=1}^{288} [y(m_i,t)]^2} = 491.268$, their reconstructed counterparts measure $\|\hat{y}\| = 491.261$ and the fitting errors measure $\|\hat{\varepsilon}\| = 2.652$. Note that:

1. the model reconstructs the yields optimally, because $\frac{\sqrt{\|\hat{\varepsilon}\|^2 + \|\hat{y}\|^2}}{\|y\|} = 99.999\%$. By the Pythagorean

theorem, the fitting errors $\hat{\varepsilon}(m,t)$ are orthogonal to the reconstructed yields $\hat{y}(m,t)$ when $\|y\| = \sqrt{\|\hat{\varepsilon}\|^2 + \|\hat{y}\|^2}$.

2. the model explains as much as $\|\hat{y}\| / \|y\| = 99.998\%$ of the actual yields. Hence, what remains unexplained is far too small to justify the employment of another state variable without overfitting the data, which would translate into parameter instability, and thus in a too large sensitivity of the parameters of the model to small computational numerical errors.

3.2.1 The Estimated Latent Factors

All the three latent factors $x_0(0,t)$, $x_1(0,t)$, and $x_2(0,t)$ display very high persistence and are potentially nonstationary.

As indicated in Table 3, they all exhibit a very high first-order autocorrelation and for all of them the autocorrelation function declines as the length of the displacements increases, but slowly, especially $x_0(0,t)$ modeling the long-end of the forward rate curve and the yield curve. The Phillips-Perron test confirms that all the three initial states may have a unit root. As expected, $x_1(0,t)$ modeling mid-range maturities is the most volatile of the three factors whereas the final value $x_0(0,t)$ is the least volatile. Finally, the mean values of the three initial states reproduce the average upward-sloping U.S. yield curve of the period under consideration, with the starting value $x_2(0,t)$ displaying the lowest mean level and the final value $x_0(0,t)$ the highest.

In what follows we examine the behavior of the three latent factors in conjunction with the behavior of the slope, curvature and the level factors introduced in section 3.1.2.

The top panels of Figure 1, show the *short-term factors* together with two empirical counterparts. The 99.1% correlation between the starting value and the federal funds rate against which it is contrasted, supports our interpretation of $x_2(0,t)$ as modeling the short-end of the forward rate and yield curves. The equally very high 99.8% correlation which $\beta_2(t)$ exhibits with the slope-transformation of the yield curve defined as $[y(6,t) - y(120,t)]$ supports our interpretation of $\beta_2(t)$ as a slope factor.

The panels in the center of Figure 1, show the *medium-term factors*. The shape factor $x_1(0,t)$ displays a 97.8% correlation with a medium-maturity transformation of the yield curve $[y(24,t) + y(36,t) + y(60,t)]/3$ thereby supporting our hypothesis that it aggregates the information contained in the middle-range maturities of the yield curve. Also the latent factor $\beta_3(t)$ supports our interpretation of governing the curvature of the yield curve given that it exhibits a 98.5% correlation with the curvature-transformation defined as $[2 * y(24,t) - y(6,t) - y(120,t)]$.

Finally, the panels at the bottom of Figure 1 plot the *long-term factors*. The final value $x_0(0,t)$ displays a 97.0% correlation with $y(120,t)$, which is the longest-maturity yield we consider, thereby supporting our hypothesis that $x_0(0,t)$ aggregates the information contained in the long-end of the yield curve. Also $\beta_1(t)$ supports our interpretation as level factor given that it exhibits a 84.8%

correlation with the level-transformation of the yield curve defined as $[y(6,t)+y(24,t)+y(120,t)]/3$.

4 The Class of Intertemporal Dynamic Parametric Models

The movements observed in the intertemporal evolution of the yields are driven by a broad set of exogenous forces, or shocks. In order to gain insights, in real time, on the nature of such forces, this paper proposes to aggregate and classify them in function of the persistence of their effect on the rates into: 1) *long-run forces* giving rise to enduring effects that may persist up to infinity; 2) *medium-run forces* giving rise to transitory effects waning within business-cycle horizons; 3) *short-run forces* giving rise to transitory and short-lived effects. At each point t in calendar-time, the whole frequency domain in which the forces moving the rates are defined, is then divided into three disjoint, pre-determined, frequency bandwidths.

To track the effects of these three types of forces on the evolution of the forward rates and the yields, we decompose the calendar-time series of each latent factor used in the cross-sectional dynamic model into three frequency components lying in the same three frequency bandwidths. Since the frequency bandwidths are disjoint, the effect of each type of force on the correspondent frequency component can be then modeled and investigated individually. The intertemporal time model has been designed in discrete time because macroeconomic developments, which are behind interest rate calendar-time dynamics, are measured at possibly large discrete time intervals.

In what follows, first we provide an overview on how we relate the forces acting in the time domain to the frequency decomposition of the latent factors, then we illustrate the models and the dynamic filter that we employ to extract, explain and predict the frequency components. Finally, we describe how we identify the models of the frequency components and how we perform the out-of-sample forecasts.

4.1 From the Time Domain to the Frequency Domain

Consider the time series of the latent factors, which in what follows we generally denote $z(t)$ for simplicity. Their signal, or discrete-time function $z(t)$ is defined in the frequency range $[0 \div f_{max}]$, where $f_{max} = \frac{1}{2T}$ and T is the discrete-time sampling step. Following the procedure proposed by Donati (1971), we partition the frequency range into a set of *finite frequency resolution intervals*: a low-frequency domain $[0 \div f_{lf}]$, a medium-frequency domain $[f_{lf} \div f_{mf}]$, a high-frequency domain $[f_{mf} \div f_{hf}]$ and a residual-frequency domain $[f_{hf} \div f_{max}]$. Then, we associate to each of these four finite frequency domains four *finite time resolution intervals* T_{lf}, T_{mf}, T_{hf} and T_{max}, respectively, selected in such a way that the product of the frequency and the time intervals is strictly greater than one: $T_{lf} \cdot f_{lf} \gg 1$, $T_{mf} \cdot (f_{mf} - f_{lf}) \gg 1, T_{hf} \cdot (f_{hf} - f_{mf}) \gg 1$ and $T_{max} \cdot (f_{max} - f_{hf}) \gg 1$. The signal power is then simultaneously decomposed with finite resolutions both in the frequency domain and in the time domain (see Appendix I for details).

As a result, the function $z(t)$ is partitioned into a *low-frequency component* $z_{lf}(t)$, a *medium-frequency component* $z_{mf}(t)$, a *high-frequency component* $z_{hf}(t)$, and a residual frequency component $we(t)$,

$$z(t) = z_{lf}(t) + z_{mf}(t) + z_{hf}(t) + we(t). \tag{13}$$

With a reasonable approximation, the "locally averaged" power of the four frequency components on the right-hand side of eq. (13) corresponds to the partitions into the four frequency domains $[0 \div f_{lf}]$, $[f_{lf} \div f_{mf}]$, $[f_{mf} \div f_{hf}]$ and $[f_{hf} \div f_{fmax}]$ of the "locally averaged" power of the signal $z(t)$. The "local averaging" is carried out over the time resolution intervals T_{lf}, T_{mf}, T_{hf} and T_{max}.

Thanks to the properties above, when time resolution intervals larger than T_{lf}, T_{mf}, T_{hf} and T_{max} are considered, the four frequency components partitioning $z(t)$ are orthogonal to each other. The orthogonality property is especially relevant for the purpose of this study because it permits to treat each frequency component of $z(t)$ independently both when processing the data and when assessing the results. The residual frequency component, $we(t)$, being essentially noise, will not be examined.

4.2 The Dynamic Models of the Frequency Components

To explain and predict the evolution of the frequency components $z_{lf}(t)$, $z_{mf}(t)$ and $z_{hf}(t)$ in which we decompose $z(t)$, we select three linear, time-invariant, dynamic systems employing the minimum number of variables and parameters guaranteeing an acceptable level of model performance in out-of-sample forecasting the spot rates. The models are expressed in a canonical state space representation ensuring also computation simplicity and tractability. They explain the cause-and effect relationship between a single exogenous cause, e.g. the long-, or the medium- or the short-run forces, and the correspondent response of a single variable, e.g. the class of low-, or the medium- or the high-frequency component of the time function $z(t)$. The exogenous causes, which are the model inputs, are estimated within the three pre-specified low-, medium and high-frequency bandwidths $[0 \div f_{lf}]$, $[f_{lf} \div f_{mf}]$ and $[f_{mf} \div f_{hf}]$. The model parameters are optimally identified within the same three frequency domains. As a result, the model outputs, that is the frequency components of $z(t)$, evolve within the same domains.

Denote M_j the linear, time-invariant, discrete-time, strictly causal, system of the $2^{nd}-$order modeling the generic frequency component $z_j(t)$ of $z(t)$,

$$\underbrace{\begin{bmatrix} q_{j,1}(t+1) \\ q_{j,2}(t+1) \end{bmatrix}}_{\mathbf{q}_j(t+1)} = \underbrace{\begin{bmatrix} 1-s_j & -r_j \\ 1 & 1 \end{bmatrix}}_{\mathbf{P}_j} \underbrace{\begin{bmatrix} q_{j,1}(t) \\ q_{j,2}(t) \end{bmatrix}}_{\mathbf{q}_j(t)} + \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\mathbf{B}_{Mj}} \underbrace{u_j(t)}_{u_j(t)} \tag{14}$$

$$z_j(t+1) \quad = q_{j,2}(t+1) \qquad\qquad j = lf,\, mf,\, hf \tag{15}$$

where $\mathbf{q}_j$ is the two-dimensional vector of the system state variables, u_j is the system single unknown input with unknown statistical properties, and $(s_j,\, r_j)$ are the system parameters. The latter, are in one-to-one correspondence with the eigenvalues of the matrix $\mathbf{P}_j$ and will be optimally identified within the frequency bandwidth of interest by minimizing the out-of-sample prediction errors of the spot rates, as illustrated in section 4.4.

The model described by eqs. (14) and (15) is *strictly causal* because the effects follow in time the causes that produced it: at time $t+1$, the effect − i.e. the system output $z_j(t+1)$ − depends solely on the external causes − i.e. the single input $u_j(k)$ for $k = t,\, t-1,\, t-2,\, \ldots, 0$ − that acted upon the system M_j before time $t+1$.

Finally, the response of model M_j can be decomposed into the response determined solely by its initial states $\mathbf{q}_j(0)$, as if $u_j(t)$ were identically zero for all t (i.e. the *zero-input response* or *free response*), and the response determined exclusively by its input $u_j(t)$ as if the initial states $\mathbf{q}_j(0)$ were zero (i.e. the *zero-state response* or *forced response*).

4.3 The Dynamic Filter

The scope of the calendar-time dynamic models M_j is to identify the forces moving the yields and use this information for forecasting purposes. These forces, namely the input functions $u_{lf}(t)$, $u_{mf}(t)$ and $u_{hf}(t)$ of eq. (14), are unknown to the modeler and need to be estimated. Also the six state variables $\mathbf{q}_{lf}(t)$, $\mathbf{q}_{mf}(t)$ and $\mathbf{q}_{hf}(t)$ of the systems M_{lf}, M_{mf} and M_{hf} are not directly accessible to measurement and need therefore to be estimated. To reconstruct both the inputs $u_j(t)$ and the state variables $\mathbf{q}_j(t)$ for *j=lf, mf, hf,* we use a dynamic filter, namely an *input-output state observer*, which extracts the required information from the time series $z(t)$.

Following the method introduced by Luenberger (1964, 1966, 1971), we obtain the estimates of $\mathbf{q}_j(t)$ as the output of the input-output state observer, which is another dynamic system with the same form as the original system M_j. Thus, it also consists of a linear, time-invariant, strictly causal, discrete-time model, but it has the function $z(t)$ as input. Luenberger show it is possible to design state observers having the property that the estimation error, defined by the difference between the state of the actual system and the state of the observer, can be made go to zero as fast as one may desire (see, e.g., B. Friedland (1986)). The observer acts through a feedback control system. This means that it recursively computes the inputs $u_{lf}(t)$, $u_{mf}(t)$ and $u_{hf}(t)$, which are imposed to lie within their pre-defined frequency domains $[0 \div f_{lf}]$, $[f_{lf} \div f_{mf}]$ and $[f_{mf} \div f_{hf}]$, ensuring that the outputs of the systems M_j, for each point in calendar-time t, when added up do not significantly differs from the the actual value of $z(t)$, i.e. $z_{lf}(t) + z_{mf}(t) + z_{hf}(t) \cong z(t)$. Therefore the observer controls the dynamics of the systems M_j towards the achievement of a guaranteed performance defined in terms of a targeted level of accuracy in out-of-sample forecasting the yields. In doing so,

it also filters out the possible measurement errors contained in $z(t)$ and corrects for the deterioration in the performance of the models M_j caused by model approximations and misspecifications (see Appendix II for details).

We select this filtering approach because of its numerous advantages: 1) it allows to investigate the behavior of the frequency components in the frequency domain although it works in the time domain. This is relevant because the data we consider need not be Fourier-transformable and this approach does not require the Fourier transformations of the signals for their behavior to be examined in the frequency domain. Moreover, 2) this approach does not require any pre-filtering, or transformation, of the data because it foregoes the stationarity requirement. Importantly, 3) it permits to filter the data in *real time*. This means that the decomposition of a time function at time t is performed without requiring the knowledge of the values that the function will take on at time $t+1 > t$ and without altering the outcomes of the decomposition already performed at time $t-1 < t$. This is an important property for this study because to investigate the determinants of yield movements we need to monitor in a consistent fashion how the frequency components of the function $z(t)$ evolve with time. In particular, we want to avoid that when mapped into the frequency domain the values that the frequency components take on at the beginning of the sample gets averaged with the values they take on at the end of the sample, thereby mixing the past with the future. Finally, 4) this approach permits to extract all the frequency components of $z(t)$ jointly and in way that minimizes the information loss when switching from a frequency bandwidth to the other. In fact, given that the sum of the three frequency components reconstructs the actual pattern of $z(t)$, the oscillations whose period is neither significantly lower nor significantly higher than the selected frequency cuts end up with lying in one of two neighboring frequency domains.

We design the interactions between the models M_j and the observer within a unique model, encompassing the three models M_j. In what follows, we will refer to this overall model as the model of the dynamic filter. Given that the observer has the same form as the original process, overall, the dynamic filter has a dimension as big as the double of the order of the systems M_j. This means that the dynamic filter employs four state variables $\left[\begin{array}{cccc} q_{j,1} & q_{j,2} & q_{j,3} & q_{j,4} \end{array} \right] = \mathbf{q}_j$ to model and estimate the generic frequency component $z_j(t)$ of $z(t)$. Since each model M_j is of the 2^{nd} order and we use three frequency components to reconstruct the time function $z(t)$, the state observer is of the 12^{th} order, that is it works with 12 state variables,

$$\mathbf{q}(t+1) = \mathbf{H}\mathbf{q}(t) + \mathbf{B}z(t) \tag{16}$$

$$\mathbf{y}(t+1) = \mathbf{G}\mathbf{q}(t+1) \tag{17}$$

As shown by state eq. (16), at calendar-time t the state observer receives as input the time function $z(t)$ and through the $12-$dimensional vector of state variables $\mathbf{q}(t)$, the real, time-invariant, (12×12) - dimensional matrix $\mathbf{H}$, and the (12×1) - dimensional matrix $\mathbf{B}$ it computes the values

taken on by the state vector $\mathbf{q}(t+1)$. As shown by output eq. (17) through the real, time-invariant, (7×12) - dimensional matrix $\mathbf{G}$ the state vector $\mathbf{q}(t+1)$ is turned into the output vector $\mathbf{y}(t+1)$ for $\mathbf{y} \equiv \left[\begin{array}{cccccc} u_{lf} & u_{mf} & u_{hf} & z_{lf} & z_{mf} & z_{hf} & \hat{z} \end{array} \right]$. Thus, the output of the dynamic filter includes the one-step ahead forecast carried out at time t of the three inputs, the three frequency components and the time function z.

At time $t+1$, the forecast $\hat{z}(t+1)$ is contrasted with the actual $z(t+1)$ and the input-output state observer reacts to their difference by computing the inputs $u_{lf}(t+2)$, $u_{mf}(t+2)$ and $u_{hf}(t+2)$ which maintain the systems M_{lf}, M_{mf} and M_{hf} on track in the subsequent period. As explained in more detail in Appendix II, the parameters of the gain matrix $\mathbf{H}$ governing the dynamics of the filter are selected to make the one-step-ahead, out-of-sample predictions, as accurate as possible.

4.4 The Out-of-Sample Forecasting Algorithm

The forecasts of the yields are obtained by plugging the out-of-sample predictions of the latent factors $x_i(0,t)$, $i = 1, \ldots, n$ into the yield curve of eq. (5). In turn, the predictions of the initial states are obtained from the dynamic filter illustrated in the preceding section.

In keeping with the notation introduced in the last three sections, starting from the generic discrete-time series $z(t)$ known up to time t, the dynamic filter of eqs. (16) and (17) yields the out-of-sample forecast $\hat{z}(t+1)$. In particular, $\hat{z}(t+1) = \mathbf{Dq}(t+1)$ where $\mathbf{D}$ is the sub-matrix formed by seventh row of matrix $\mathbf{G}$ of eq. (17) (see Appendix II for details).

Note that at time t, when forecasting $\hat{z}(t+1)$, the inputs $u_{lf}(t+1)$, $u_{mf}(t+1)$ and $u_{hf}(t+1)$ determining the actual value $z(t+1)$ are still unknown. Therefore, $\hat{z}(t+1)$ stems from the zero-input response of the systems M_{lf}, M_{mf} and M_{hf}. As a result, the out-of-sample forecast error denoted $fe(t,1) = z(t+1) - \hat{z}(t+1)$, is due to the effect of both the possible model approximations included in the zero-input responses of the dynamic systems and the effect of the missing inputs.

In the subsequent period, once they are disclosed, we compute the effect of the $(t+1)$-inputs on the systems. We call this latter effect *innovation*. If the dynamic filter works properly, the innovations are the main, and possibly the only, reason of the out-of-sample forecast errors.

To produce out-of-sample forecasts of z for longer horizons, we proceed recursively and use the dynamic filter outputs, i.e. its one-step-ahead predictions, as its next-step inputs. For example, $\hat{z}(t+1)$ is used as the dynamic filter input of time $t+1$ for it to produce the forecast $\hat{z}(t+2)$, and so on. The out-of-sample forecasts performed at time t for the horizon $\tau \in [1,h]$ are then obtained from equation $\hat{z}(t+\tau) = [\mathbf{H} + \mathbf{BD}]^{\tau-1}\mathbf{q}(t+1)$ where $\mathbf{H}$ and $\mathbf{B}$ are the matrices of eq. (16).

4.5 The Identification Algorithm

The cause-and-effect relationship between the long-, medium- and short-run forces acting on the yields and the pattern followed in response by the frequency components of each latent factor $x_i(0,t)$,

$i = 1, \ldots, n$ is modeled with the three strictly-causal dynamic systems M_j for $j = lf, mf, hf$, introduced in section 4.2. In this section, we illustrate how to estimate the parameters $(s_j,\ r_j)$ of the systems M_j.

The procedure we adopt is not standard for three reasons: 1) because the system inputs u_j are unknown; 2) because we want to identify the parameters that minimize the out-of-sample forecast errors $fe(t, \tau)$ so that the identification of $(s_j,\ r_j)$ is not performed in sample, but on the out-of-sample forecasts of a sub-set of data; 3) because at the same time we want also to optimally estimate the initial conditions $\mathbf{q}(0)$ of the dynamic filter, which are also unknown. Although the dynamic filter ensures that the temporary disturbances generated by an inaccurate knowledge of the initial conditions are progressively attenuated to zero, by assigning as accurate initial conditions as possible, the attenuation occurs faster. The related efficiency gains can be appreciated especially for the system M_{lf} modeling the slowly evolving low-frequency component z_{lf}. In fact, the error $\mathbf{eq}(0) = \mathbf{q}^*(0) - \mathbf{q}(0)$, where $\mathbf{q}^*(0)$ are the unknown, accurate, initial conditions and $\mathbf{q}(0)$ is an arbitrary guess of their values, evolves according to the law $\mathbf{eq}(t) = \mathbf{H}^t\mathbf{eq}(0)$ set by state equation (11) of the dynamic filter. Given that all of the eigenvalues of $\mathbf{H}$ eq. (16) are real, positive, scalars smaller than one by construction, the error $\mathbf{eq}(0)$ tends asymptotically to zero as calendar-time t goes to infinity. However, the closer to one are the eigenvalues of the sub-matrix of $\mathbf{H}$ ruling the dynamics of each frequency component, the larger is the number of steps that the dynamic filter needs to correct for $\mathbf{eq}_{lf}(0)$.

We partition the data set. The data running from t_1 to t_2 are used both to identify $(s_j,\ r_j)$ and $\mathbf{q}(0)$, and the data running from t_2 to the end of the sample are used to evaluate forecast accuracy. We proceed as follows. Starting from some guessed initial conditions $\mathbf{q}(0)$, we reconstruct the states $\mathbf{q}(t)$ employing the dynamic filter and, at each point in calendar-time t, we compute the out-of-sample forecasts $\hat{z}(t + \tau)$ for $\tau \in [1, h]$. Then, we identify the parameters $(s_j,\ r_j)$ and the initial conditions $\mathbf{q}(0)$ by minimizing the quadratic loss functional Ft

$$\min_{s_j, r_j, \mathbf{q}(0)} Ft = \sum_{t=t_1}^{t_2-1} \sum_{\tau=1}^{h} [fe(t, \tau) w(\tau)]^2 \tag{18}$$

where $w(\tau)$ is a negative exponential function that assigns to the forecast errors a weight decreasing with the lengthening of the prediction horizon τ.

In this study, we minimize the functional Ft employing the numerical algorithm based on the conjugate-gradient method contained in the software suite EicasLab, whose embedded tests check for the local uniqueness of the minimum.

4.6 The Estimation of the Intertemporal Dynamic Model

In this section, we report the results from the estimation of the intertemporal dynamic model introduced in section 4. We begin by presenting the three disjoint low-, medium- and high-frequency bandwidths selected for the purpose of this study to decompose the time series of the latent factors $x_0(0,t)$, $x_1(0,t)$ and $x_2(0,t)$. In section 4.6.2, we describe the dynamic systems modeling the frequency components and in section 4.6.3 we examine their properties and those of their driving forces. In section 4.4 we conclude by presenting the out-of-sample forecasts of the yields.

4.6.1 The Selected Frequency Bandwidths and the Poles of the Filter

Before performing the spectral decomposition of the latent factors, we exogenously assign a value to the eigenvalues, or poles, of the gain matrix **H** governing the dynamics of the input-output state observer. The poles have a time and a frequency domain interpretation:

- *Frequency domain interpretation.* Given the sampling period T, which in our case corresponds to 1 month, the pole p is related to the angular frequency ω, measured in rad/month, by the equation $p = e^{-\omega T}$. If the frequency is measured in cycles/month, it will be denoted f. The values of the frequency, either measured in rad/month or in cycles/month, determined by the poles, set the upper bound of the pass-band frequency of the filter. Recall that the oscillation period, defined as $\frac{1}{f} = \frac{2\pi}{\omega}$, is the inverse of the frequency.
- *Time domain interpretation.* The inverse of the angular frequency ω is the time constant, denoted τ, which characterizes the impulse response of the state observer.

Now we proceed with the description of the selected frequency bandwidths.

The *long-run shocks*, which give rise to enduring effects that may persist up to infinity, are extracted from the time series of each latent factor by means of a low pass-band filters of the $4^{th}-$order embedded in the input-output state observer. We set its four poles are all equal to 0.985. This means that the frequency bandwidth within which the dynamic system M_{lf} responds to the long-run shocks u_{lf} by producing as output the low-frequency components z_{lf} is $[0 \div 0.015]$ rad/month. The poles fix the limit below which the system M_{lf} does not respond to the action of the exogenous shocks placed upon it. In particular, the oscillations with a period longer than 415.5 months pass, whereas the oscillations with a period significantly lower than that − that are therefore more frequent − do not pass but get filtered by the state observer associated to the system M_{mf}. The oscillations with a period in between are attenuated by $(1/\sqrt{2} * 1/\sqrt{2} * 1/\sqrt{2} * 1/\sqrt{2}) = 0.25$. This also means that we define long-run shocks those shocks whose effects broadly take at least 5.5 years to decay significantly. Hence, the filter smooths the original time series by averaging its instantaneous values over time resolution intervals larger than $66.2 * 2 * \pi = 415.5$ months.

The *medium-run shocks*, whose effects wane within what in the context of this study we denote business-cycle horizons, are extracted from the residual time series obtained by removing the low-frequency component from the original ones with another low pass-band filters of the $4^{th}-$order embedded in the input-output state observer whose four poles are all equal to 0.6. This means that the dynamic system M_{mf} works within the frequency range of $[0.015 \div 0.51]$ rad/month. The medium-frequency component of each initial state is then characterized by a period of at least 12.29 months, but significantly lower than 415.5 months. Equivalently, the effect of the medium-run shocks broadly take at least 1.96 months, but less than 5.5 years, to decay significantly. This also means that each medium-frequency component is extracted by smoothing over time resolution intervals larger than $1.96 * 2 * \pi = 12.29$ months.

The *short-run shocks*, whose effects abate very quickly, are extracted from the time series obtained after removing the low- and medium-frequency components from the original ones, with a third low passband filter of the $4^{th}-$order whose four poles are all equal to 0.2. This means that the dynamic system M_{hf} works within the frequency bandwidth of $[0.51 \div 1.6]$ rad/month. The high-frequency component of each initial state is then characterized by a period of at least 3.90 months, but significantly lower than a year. Equivalently, the effects of the short-lived shocks broadly take at least 0.6 months, but less than 1.96 months, to decay significantly. This also means that each high-frequency component is extracted by smoothing over time resolution intervals larger than $0.6 * 2 * \pi = 3.90$ months.

The residual time series, which is obtained by removing the low- , medium- and high-frequency components from the original time series, lies in the residual frequency range $[1.6 \div 3.14]$ rad/month. We neglect such residual frequency component because it is too noisy for the purpose of this study.

4.6.2 The Dynamic Models of the Frequency Components

To explain the evolution of the frequency components of the latent $x_0(0,t)$, $x_1(0,t)$ and $x_2(0,t)$ plotted in Figure 1, and to carry out their out-of-sample forecasts, we model each frequency component as the output of the $2^{nd}-$order dynamic systems $M_{i,j}$ for $i = 0, 1, 2$ and j=*lf, mf, hf* defined by eqs. (14) and (15). Each dynamic system $M_{i,j}$ is subject to a single shock u_{ij} and it is fully defined by the parameters $(s_{i,j},\ r_{i,j})$.

Since we have three latent factors and each of them is decomposed into three frequency components, the overall model of their frequency components has $(3 \times 3 \times 2) = 18$ parameters. Of these, three parameters are *a priori* imposed to be zero, namely $r_{i,lf} = 0$, to ensure that the effect produced by the long-run forces on the low-frequency components may persist up to infinity. The values of the other 15 parameters are optimally identified from the time series of each latent factor over the period from 31 January 1990 (equal to t_1) to 31 December 2003 (equal to t_2). Recall that the parameter estimates are carried out on the out-of-sample forecasts produced by the dynamic filter. The reason why we skip the first predictions (our sample starts on 31 January 1984) is that the

errors in the initial values, which we assign arbitrarily, of the state variables employed by $4^{th}-$order low pass-band filters have a significant transient effect on forecast accuracy. By skipping the first 72 outputs of the filters, possible errors in the initial state values $\mathbf{q}_{i,mf}(0)$ and $\mathbf{q}_{i,hf}(0)$ attributed to the high- and medium-frequency low pass-band filters do not affect the out-of-sample forecasts anymore.

We optimally estimate the initial conditions $\mathbf{q}_{i,lf}(0)$ of the four state variables used to model the low-frequency components $x_{i,lf}(0,t)$ along with the model parameters. The simplest approach to the problem may seem the following. First, decompose the original time series of the latent factors into their frequency components. Second, by considering a single frequency component at the time, identify the parameters $(s_{i,j},\ r_{i,z})$ $i = 1, 2, 3$, j=*lf, mf, hf* and z=*mf, hf*, of the respective dynamic system $M_{i,j}$, then reconstruct the unknown input u_{ij} and estimate the dynamic system's state variables $\mathbf{q}_{i,j}$. Instead, we decompose the time series and carry out the state and input estimations in a single step by employing in parallel, for each considered time series, the set of three low pass-band filters. These, given the frequency bandwidths stated in Table 4 and the parameters $(s_{i,j},\ r_{i,j})$ of the dynamic systems $M_{i,j}$, simultaneously decompose the time series into their frequency components. For each frequency component, they reconstruct the unknown input u_{ij} and they estimate the state variables $\mathbf{q}_{i,j}$ of the respective dynamic system $M_{i,j}$.

As explained in section 4.5, we carry out the identification of the parameters $(s_{i,j},\ r_{i,z})$ by minimizing the out-of-sample forecast errors of the yields. The minimization of the out-of-sample forecast errors is carried out by introducing the cost functional,

$$\min_{s_{i,j},r_{i,z}} Ft = \sum_{t=t_1}^{t_2-1}\sum_{\tau=1}^{h}[fe(t,\tau)w(\tau)]^2 \quad i=0,1,2 \quad j=lf,\ mf,\ hf \quad z=mf,\ hf$$

where $fe(t,\tau)$ is the out-of-sample forecast error for the prediction carried out at calendar-time t of the value taken on by the variable of interest at time $t+\tau$; $w(\tau) = \mathrm{e}^{-\frac{\tau}{T}}$ is the weight associated to the out-of-sample forecast errors $fe(t,\tau)$; we set $T = 300$ months in order to attach to the forecast errors a weight only slightly decreasing with the forecast horizon; we consider a 24-month forecast horizon $\tau \in [1,\ 24]$ months so that $h = 24$. The estimated values of the parameters are reported in Table 5.

4.6.3 Analysis of the Frequency Components and their Driving Forces

In this section we examine the low-, medium- and high-frequency components in which we decompose the calendar-time functions of the latent factors, and the long-run shocks driving their dynamics.

Given that the initial outputs of the dynamic filter are affected by the errors made when guessing the initial conditions of the state variables modeling the medium- and the high-frequency components, as discussed in section 4.5, here we skip those initial estimates and we examine the calendar-

time period from 31 October 1985 to 31 December 2007. The frequency components of the three latent factors are plotted in Figure 2.

The low-frequency components (LF henceforth) of all the three latent factors exhibit a downward trend (see panels (a), (b) and (c) of Figure 2). Over the 267 months we examine in this section, the LFs all decreased by about 5 percentage points: the LF of the starting value $x_{2,lf}(0,t)$ aggregating the information contained in the shortest-end of the forward rate curve and the yield curve decreased by 4.98 percentage points, the LF of shape factor $x_{1,lf}(0,t)$ aggregating the information contained in middle-range maturities rates declined by 4.48 percentage points and the LF of the final value $x_{0,lf}(0,t)$ aggregating the information contained in the longest-end of the curves declined by 5.00 percentage points.

Although the fact that they exhibit a similar pattern might induce to believe that the LF of the three latent factors have responded to closely related long-run forces, the analysis of the long-run shocks $u_{i,lf}(t)$ for $i = 0, 1, 2$ suggests otherwise. Whereas the long-run forces driving the LF of the starting value appear closely related to the economic cycle fluctuations the Federal Reserve has reacted to in setting the federal funds rate target (see panel (a) of Figure 3), the long-run forces driving the LF of the mid-maturity rates exhibit much less volatility, especially from the end of the 1980s onwards (see panel (b) of Figure 3). Finally, the long-run forces driving final value seem to respond little to business-cycle developments given that they display a fairly different pattern than the long-run shocks moving $x_{2,lf}(t)$: not only are they much less volatile, but they follow a much slower dynamics characterized by a single long trough and a single peak in the twenty-two years we consider.

The medium-frequency components (MF henceforth) of the latent factors display dissimilar behaviors suggesting that yields of different maturities are differently affected by the economic forces exerting their effects at business cycle frequencies (see panels (d), (e) and (f) of Figure 2). To start with, the MF differ in variability. The MF of the *s*tarting value is the most volatile, with a standard deviation of 1.87 percentage points, followed by the MF of the shape factor, whose standard deviation is 0.64 percentage points, and by the little volatile MF of the final value, which displays a standard deviations of only 0.36 percentage points. The MF of the long-term rates is also very little correlated, by -3.38%, with the MF of the short-term rates, whereas the latter is more correlated, by 36.95%, with the MF of mid-maturity rates.

Finally, the high-frequency component (HF henceforth) of the shape factor is the most volatile, with a standard deviation of 1.03 percentage points, confirming that short-lived economic forces affect mostly middle-range rates, while the HF of the final value is the least volatile, with a standard deviation of 0.33 percentage points. The short-end of the yield curve appears affected by short-run shocks more heavily than the long-end, but less than middle-range maturities, displaying a standard deviation of 0.64 percentage points.

Given that this paper takes a yield-only approach, we leave a deeper investigation of these

findings and of their economic implications to future fruitful research.

4.6.4 The Forecasts of the Yields

At each at time t included in the interval from 31 December 2003 up to 31 December 2007, we forecast out-of-sample the three latent factors, then we substitute the forecasted $\hat{x}_2(0,t,\tau)$, $\hat{x}_1(0,t,\tau)$, $\hat{x}_0(0,t,\tau)$ for $\tau = 1,\ldots,24$ months into the yield curve $y(m,t)$ of eq. (11) and obtain the predicted values $\hat{y}(m,t,\tau)$ of the yields. At the end of each forecast exercise, before carrying out the next prediction, we update the estimate of the state variables $\mathbf{q}_{i,lf}(t)$, $\mathbf{q}_{i,mf}(t)$ and $\mathbf{q}_{i,hf}(t)$ of the systems $M_{i,lf}$, $M_{i,mf}$ and $M_{i,hf}$ for $i = 0,1,2$, whereas their parameters remain unchanged. The obtained results are presented in Table 6.

To evaluate the performance of our model in forecasting out-of-sample we use as benchmark the random walk model (RW), which foresees that the m-maturity yield at time t is the forecast of the m-maturity yield at time $t+\tau$, with $\tau = 1,\ldots,24$ months. Moreover, the forecast accuracy is measured with the root mean squared forecast error. Although we forecast all of the twenty yield maturities included in our data set, Table 6 reports the statistics for the out-of-sample forecasts of yields with maturities 6 months, and 1, 3, 5, 10 years at forecasting horizons of 1, 6, 12 and 24 months.

The yield curve model we developed used in conjunction with the state observer to estimate the state of the dynamic models, outperforms the random walk in forecasting out-of-sample the yields of all maturities and at all the forecasting horizons, as shown in the last three columns Table 6, with a margin that increases with the lengthening of the forecast horizon. Within the same forecast horizon, the yield curve model predicts better than the random walk in particular the short- and the long-term yields.

The one-month ahead prediction errors have small mean values and low levels of the first-order autocorrelation (and are also stationary), as shown in the top panel of Table 6. These results confirms that the input-output state observer works properly and that the innovations are the primary and possibly the only, unavoidable, reason of the out-of-sample forecast errors.

These results indicate that the performance of the yield curve model used in conjunction with the dynamic filter to forecast the yields is satisfactory. The robustness of the methodology is corroborated also by the fact that we predict out-of-sample the yields during years, especially 2004 and 2005, in which the behavior of the long-end of the yield curve was difficult to understand and forecast, leading the then-Fed Chairman Alan Greenspan to speak of a "conundrum". The findings of this study indicate that although anomalous, the behavior of the long-term rates during such period has remained in the realm of a model, whose performance is corrected for model uncertainty.

5 Conclusion

This paper proposes a class of dynamic, parametric, models for the term structure of interest rates in which the calendar-time dynamics of the forward rates, and thus of the spot rates, is described both in the time domain and in the frequency domain, which permits to gain a better understanding on the determinants of the movements observed in the yields. To estimate the selected model, to perform the required frequency decompositions and to forecast the yields out of sample we introduce a recursive procedure controlling for the possible misspecifications, approximations and data measurement errors that may undermine the empirical performance of the model. As a result, we obtain out-of-sample forecasts that are more accurate than the random walk benchmark also at short forecast horizons, and that remain fairly accurate also at long forecast horizons.

References

[1] Ang, A., and M. Piazzesi. 2003. A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables. *Journal of Monetary Economics* 50:745-87.

[1] Assenmacher-Wesche K., and S. Gerlach. 2008. The Term Structure of Interest Rates across Frequencies. Working Paper.

[2] Bijork, T., and B. J. Christensen. 1999. Interest rate Dynamics and Consistent Forward Rate Curves. *Mathematical Finance*. 9:323-348.

[3] BIS. 2005. Zero-Coupon Yield Curves. Technical Documentation. BIS Paper 25.

[4] Brick, J. R., and H. E. Thompson. 1978. Time Series Analysis of Interest Rates: Some Additional Evidence. *Journal of Finance* 33:93-103.

[5] Carlucci, D., and F. Donati. 1975. Control of Norm Uncertain Systems. *IEEE Transactions on Automatic Control*. 792-95.

[6] Chen, C.-T. 1999. *Linear System Theory and Design*. Oxford University Press, New York.

[7] Christensen, J., F. Diebold, and G. Rudebusch. 2007. The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models. Working Paper 2007-20. Federal Reserve Bank of San Francisco.

[8] Coroneo, L., K. Nyholm and R. Vidova-Koleva. 2008. How Arbitrage Free is the Nelson-Siegel Model? Working Paper 874. European Central Bank.

[9] Diebold, F. X., and C. Li. 2006. Forecasting the term structure of government bond yields. *Journal of Econometrics*. 130:337-364.

[10] Diebold, F. X., G. D. Rudebusch, and S. B. Aruoba. 2006. The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach. *Journal of Econometrics*. 131:309-38.

[11] Donati, F. 1971. Finite-Time Averaged Power Spectra. *IEEE Transaction on Information Theory.* 17:7-16.

[12] Donati, F., and M. Vallauri. 1984. Guaranteed Control of "Almost-Linear" Plants. *IEEE Transactions on Automatic Control.* 29:34-41.

[13] Duffee, G. 2002. Term Premia and Interest Rate Forecasts in Affine Models. *Journal of Finance.* 57:405-43.

[14] Estrella, A., and F. S. Mishkin. 1997. The Predictive Power of the Term Structure of Interest Rates in Europe and United States: Implications for the European Central Bank. *European Economic Review.* 41:1375-1401.

[15] Evans, C. L., and D. A. Marshall. 2007. Economic Determinants of the Nominal Treasury Yield Curve. *Journal of Monetary Economics.* 54:1986-2003.

[16] Fisher, M., D. Nychka and D. Zervos. 1995. Fitting the Term Structure of Interest Rates with Smoothing Splines. Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series. 95-1.

[17] Friedland, B. 1986. Control System Design. An Introduction to State Space Methods. McGraw-Hill, Inc.

[18] Lamoureux C. G., and D. Witte. 2002. Empirical Analysis of the Yield Curve: The Information in the Data Viewed through the Window of Cox, Ingersoll, and Ross. *Journal of Finance.* 57:1479-1520.

[19] Litterman, R., and J. A. Scheinkman. 1991. Common Factors Affecting Bond Returns. *Journal of Fixed Income.* 1:54-61.

[20] Luenberger, D. G. 1964. Observing the State of a Linear System. *IEEE Transactions on Military Electronics.* 8:74-80.

[21] Luenberger, D. G. 1966. Observers for Multivariable Systems. *IEEE Transactions on Automatic Control.* 11:190-97.

[22] Luenberger, D. G. 1971. An Introduction to Observers. *IEEE Transactions on Automatic Control.* 16:596-602.

[23] Knight, F. H. 1921. *Risk, Uncertainty and Profit.* Boston / New York: Houghton Mifflin.

[24] Nelson, C. R., and A. F. Siegel. 1987. Parsimonious Modeling of Yield Curve. *Journal of Business.* 60:473-489.

[25] Pippenger, J. E. 1974. A Time Series Analysis of Post-Accord Interest Rates: A Comment. *Journal of Finance.* 29:1320-25.

[26] Vasicek, O. A., and H. G. Fong. 1982. Term Structure Modeling Using Exponential Splines. *Journal of Finance.* 37:177-188.

Appendix I

Consider the discrete-time function $z(t)$ whose pattern we want to investigate in the frequency domain. The function $z(t)$ is a finite power sampled time function satisfying the condition $\lim_{N\to\infty}\frac{1}{2N}\sum_{t=-N}^{N}[z(t)]^2 = Pz < \infty$ where Pz is the mean power of the signal $z(t)$. If the data sampling unit considered is one month, as in this paper, the upper limit of the frequency domain is $f_{max} = 0.5\, cycles/month = \pi\, rad/month$. Given that the mean power Pz is finite, the function $z(t)$ is not Fourier transformable. In this case, if $z(t)$ follows an ergodic stationary process, Pz is distributed in the frequency domain with a power spectrum $\Phi(f)$, such that: $Pz = \int_0^{fmax}\Phi(f)\,df$. If $z(t)$ follows a nonstationary processes, as many economic time series, the signal power is likely to be time-varying. In this case, as shown by Donati (1971), it is possible to define a *class* of time-varying power spectral functions $\varphi(f_i,t)$, where f_i, with $i \in (1, Nf)$, denotes the frequency values belonging to a finite set of Nf elements, with the following properties:

- $pz(t) = \sum_{i=1}^{Nf}\varphi(f_i,t)$ is a "*locally time averaged*" instantaneous power obtained by a suitable smoothing of the signal instantaneous power $[z(t)]^2$, such that: $Pz = \lim_{N\to\infty}\frac{1}{2N}\sum_{t=-N}^{N} pz(t)$;
- $\bar{\Phi}(f_i) = \lim_{N\to\infty}\frac{1}{2N}\sum_{t=-N}^{N}\varphi(f_i,t)$, is a "*locally frequency averaged*" power spectral value such that $Pz = \sum_{i=1}^{Nf}\bar{\Phi}(f_i)$. If the signal $z(t)$ is the realization of an ergodic stationary stochastic process, the power spectrum $\bar{\Phi}(f_i)$ corresponds to a "*locally frequency averaging*" of the stochastic process power spectrum $\Phi(f)$.

The elements $\varphi(f_i,t)$ of the time-varying power spectrum class are related to the criteria selected to perform the local averaging in the time and frequency domains. While different averaging criteria may be adopted, they should meet the following general rules:

1. A weighted averaging approach must be applied, with the weighting function defined in such a way that the averaged value may be attributed (even roughly) to a *finite interval,* whose amplitude is denoted T when referring to the *time interval,* and Δf when referring to the *frequency interval*. The intervals of amplitudes T and Δf define the *finite resolution* of the performed averages. As a result, in the time domain, two values $\varphi(f_i,t_1)$ and $\varphi(f_i,t_2)$ cannot differ significantly if the time instant difference $\|t_2 - t_1\|$ is not significantly larger than T. Similarly, in the frequency domain, two values $\varphi(f_1,t)$ and $\varphi(f_2,t)$ cannot be significantly different if $\|f_2 - f_1\|$ is not significantly larger than Δf.
2. The time-varying power spectral decomposition is possible only if adopting finite resolutions T and Δf such that $T\cdot\Delta f \gg 1$.
3. If we adopt the greatest time resolution $T = 1\, month$, which is equal to the sampling unit, the required frequency resolution is $\Delta f = fmax$. Therefore no spectral decomposition is possible.
4. If we adopt the greatest frequency resolution, which in this paper is $\Delta f = 1/288\ cycles/month$ since our data extend to 288 monthly observations, then the required time resolution coincides with all the time interval (running from 1984:01 to 2007:12) and no time-varying power spectrum may be considered, but only the power spectrum of the power averaged over all the available time series data.

Having clarified the above conditions, for the purpose of the study presented here we decide to opt for a good resolution in the time domain and to accept a low resolution in the frequency domain. As a result, we decompose the time series $z(t)$ in only four spectral components: a low-, a medium-, a high-frequency component and a residual decomposition error lying within a residual very high-frequency domain, which we do not investigate.

Appendix II

In this appendix we explain how, starting from the actual time function $z(t)$, we reconstruct the inputs $u_j(t)$ and estimate the state variables $q_{j,1}(t)$, $q_{j,2}(t)$ of the dynamic systems M_j, $j = lf, mf, hf$, introduced in sections 4.2 and 4.3. From the theory of dynamic systems with a state space representation we know (see, e.g., Chen (1999)) that given the system outputs $z_j(t)$, $j = lf, mf, hf$, of systems expressed in the canonical form of eqs. (14) and (15) it is possible to estimate the values taken by all the states $\mathbf{q}_j(t)$ and to reconstruct all the inputs $u_j(t)$, $j = lf, mf, hf$ by means of another dynamic system, namely a state observer. As explained in section 4.3, we consider an input-output state observer consisting of a linear, discrete-time, strictly causal, time-invariant, system of the 12^{th} order.

Its state equation has the following representation:

$$\underbrace{\begin{bmatrix} q_{lf,1}(t+1) \\ q_{lf,2}(t+1) \\ q_{lf,3}(t+1) \\ q_{lf,4}(t+1) \\ q_{mf,1}(t+1) \\ q_{mf,2}(t+1) \\ q_{mf,3}(t+1) \\ q_{mf,4}(t+1) \\ q_{hf,1}(t+1) \\ q_{hf,2}(t+1) \\ q_{hf,3}(t+1) \\ q_{hf,4}(t+1) \end{bmatrix}}_{(12\times 1)} = \underbrace{\begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 & \mathbf{H}_3 \end{bmatrix}}_{(12\times 12)} \underbrace{\begin{bmatrix} q_{lf,1}(t) \\ q_{lf,2}(t) \\ q_{lf,3}(t) \\ q_{lf,4}(t) \\ q_{mf,1}(t) \\ q_{mf,2}(t) \\ q_{mf,3}(t) \\ q_{mf,4}(t) \\ q_{hf,1}(t) \\ q_{hf,2}(t) \\ q_{hf,3}(t) \\ q_{hf,4}(t) \end{bmatrix}}_{(12\times 1)} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ d_{lf} \\ c_{lf} \\ 0 \\ 0 \\ d_{mf} \\ c_{mf} \\ 0 \\ 0 \\ d_{hf} \\ c_{hf} \end{bmatrix}}_{(12\times 1)} z(t) \tag{19}$$

$$\mathbf{q}(t+1) = \mathbf{H}\,\mathbf{q}(t) + \mathbf{B}\,z(t)$$

with

$$\underset{(12\times 4)}{\mathbf{H}_1} = \begin{bmatrix} 1-a_{lf} & -b_{lf} & h_{lf} & k_{lf} \\ 1 & 1 & 0 & 0 \\ 0 & -d_{lf} & 1-r_{lf} & -s_{lf} \\ 0 & -c_{lf} & 1 & 1-c_{lf} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -d_{mf} & 0 & 0 \\ 0 & -c_{mf} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -d_{hf} & 0 & 0 \\ 0 & -c_{hf} & 0 & 0 \end{bmatrix} \quad \underset{(12\times 4)}{\mathbf{H}_2} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1-a_{mf} & -b_{mf} & h_{mf} & k_{mf} \\ 1 & 1 & 0 & 0 \\ 0 & -d_{mf} & 1-r_{mf} & -s_{mf} \\ 0 & -c_{mf} & 1 & 1-c_{mf} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -d_{hf} & 0 & 0 \\ 0 & -c_{hf} & 0 & 0 \end{bmatrix}$$

and

$$\underset{(12\times 4)}{\mathbf{H}_3} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1-a_{hf} & -b_{hf} & h_{hf} & k_{hf} \\ 1 & 1 & 0 & 0 \\ 0 & -d_{hf} & 1-r_{hf} & -s_{hf} \\ 0 & -c_{hf} & 1 & 1-c_{hf} \end{bmatrix}$$

with $r_j = a_j - h_j$ and $s_j = b_j + d_j - h_j$ for $j = lf, mf, hf$.

Its output equation has the following representation:

$$\underbrace{\begin{bmatrix} u_{lf}(t+1) \\ u_{mf}(t+1) \\ u_{hf}(t+1) \\ z_{lf}(t+1) \\ z_{mf}(t+1) \\ z_{hf}(t+1) \\ \hat{z}(t+1) \end{bmatrix}}_{(7\times 1)} = \underbrace{\begin{bmatrix} 0 & 0 & h_{lf} & k_{lf} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & h_{mf} & k_{mf} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & h_{hf} & k_{hf} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}}_{(7\times 12)} \underbrace{\begin{bmatrix} q_{lf,1}(t+1) \\ q_{lf,2}(t+1) \\ q_{lf,3}(t+1) \\ q_{lf,4}(t+1) \\ q_{mf,1}(t+1) \\ q_{mf,2}(t+1) \\ q_{mf,3}(t+1) \\ q_{mf,4}(t+1) \\ q_{hf,1}(t+1) \\ q_{hf,2}(t+1) \\ q_{hf,3}(t+1) \\ q_{hf,4}(t+1) \end{bmatrix}}_{(12\times 1)}$$

$$\underset{(7\times 1)}{\mathbf{y}(t+1)} = \underset{(7\times 12)}{\mathbf{G}} \underset{(12\times 1)}{\mathbf{q}(t+1)} \tag{20}$$

Note that the matrix $\mathbf{H}$ is lower-triangular, so that the eigenvalues governing the dynamics of the input-output state observer are in its principal diagonal. Specifically, the sub-matrices $\mathbf{H}_1$, $\mathbf{H}_2$, $\mathbf{H}_3$ contain three (4×4) blocks of elements in their principal diagonals – corresponding to columns and rows from 1 to 4 of matrix $\mathbf{H}_1$, to columns and rows from 5 to 8 of matrix $\mathbf{H}_2$ and to columns and rows from 9 to 12 of matrix $\mathbf{H}_3$ – which partition the eigenvalues of the input-output state observer and which are in one-to-one correspondence, respectively, with the parameters c_j, d_j, h_j, k_j for $j = lf, mf, hf$. We use 4 eigenvalues to extract and to model the evolution of each of the three frequency components in which we partition the time function $z(t)$. A schematic representation of the dynamic system of the state observer defined by eqs. (19) and (20) is provided by the flow-chart of Figure 4. Recall that the state observer includes the models M_j of eqs. (14) and (15) and that it receives as input the actual time function $z(t)$, which may be affected by measurement errors, and produces as output the vector $\mathbf{y}(t+1)$ of one-step ahead predictions of the three pairs (u_j, z_j) for *j=lf, mf, hf,* and the predicted value $\hat{z}(t+1)$ taken on by the time function z at time $t+1$. To do so, the dynamic filter reacts, through a close loop, to the outputs $z_j(t)$ produced by the systems M_j by means of the inputs $u_j(t)$ in a way that ensures that each modeled frequency component z_j evolves within its frequency bandwidth and that the sum of three frequency components tracks the actual time function $z(t)$ minimizing the residual *we* of the spectral decomposition, as illustrated in section 4. Figure 4 shows these three closed loops: the first is composed by the system M_{lf} which receives a feedback by the $4^{th}-$order low pass-band filter CC_{lf} embedded in the state observer; the second is composed by the system M_{mf} which is

subject to the feedback control stemming from the $4^{th}-$ order low pass-band filter CC_{mf} and, finally, the third is composed by the system M_{hf} which is subject the feedback provided by the $4^{th}-$ order low pass-band filter CC_{lhf}. Through these closed loop reactions, the state observer also corrects for model uncertainty, that is for the deterioration in the model performance caused by model approximations and misspecifications and data measurement errors, thereby ensuring the achievement of a targeted level of accuracy in out-of-sample forecasting the yields.

We go through the loops depicted in Figure 4 more in detail:

1. First, the estimate of the low-frequency component $z_{lf}(t)$ produced by the dynamic system M_{lf} is contrasted with the actual value of $z(t)$. The state observer, by means of the feedback control system CC_{lf} reacts to the difference $z(t) - z_{lf}(t) = e_{lf}(t)$ and computes the input $u_{lf}(t)$ which then steers the dynamics of system M_{lf}. We exogenously impose the value 0.985 to the four eigenvalues of the input-output state observer that are used to extract the low-frequency component. This is equivalent to impose that $u_{lf}(t)$ lies within a bandwidth of angular frequency of 0.015 rad/month. In this way we guarantee that $z_{lf}(t)$ evolves within the low-frequency domain $[0 \div f_{lf}]$.Note that by exogenously placing the eigenvalues, or pole, of the state observer not only do we set its dynamics, but we determine the parameters c_j, d_j, h_j and K_j contained in its matrix $\mathbf{H}$. In such way we specify the closed-loop dynamics of the filter. For details on how to solve the pole-placement problem see, e.g., C.-T. Chen, (1999).

2. Next, the residual $e_{lf}(t)$ is contrasted with the medium-frequency component $z_{mf}(t)$ produced by the dynamic system M_{mf}. Through the input $u_{mf}(t)$, the feedback control system CC_{mf} guarantees that $z_{mf}(t)$ evolving within the selected medium-frequency domain tracks $e_{lf}(t)$. This is achieved by setting the four corresponding eigenvalues of the input-output state observer equal to 0.6, which is equivalent to impose that $u_{mf}(t)$ lies within a bandwidth of angular frequency $[0.015 \div 0.51]$ rad/month.

3. The residual $e_{lf}(t) - z_{mf}(t) = e_{mf}(t)$ belongs to a higher frequency domain, which we contrast with the high-frequency component $z_{hf}(t)$. To compute the input $u_{hf}(t)$ that guarantees that the system output $z_{hf}(t)$ tracks $e_{mf}(t)$ within the pre-specified high-frequency bandwidth, we use the feedback control system CC_{mhf} and we assign the value 0.2 to the four related eigenvalues of the input-output state observer. As a result, the power spectrum $u_{hf}(f)$ belongs to the angular frequency range $[0.5 \div 1.6]$ rad/month.

4. The residual $e_{mf}(t) - z_{hf}(t) = we(t)$, which we do not investigate, receives the power of the time function $z(t)$ that lies within the angular frequency domain $[1.6 \div 3.14]$ rad/month.

Table 1: Summary Statistics: the Yields

Maturity	Central moments				Autocorrelations		
(*months*)	Mean	Std. Dev.	Skewness	Kurtosis	Lag (1)	Lag (12)	Lag (30)
6	5.29	2.29	0.19	3.00	0.98	0.66	0.19
12	5.45	2.30	0.23	3.05	0.98	0.66	0.20
18	5.60	2.30	0.31	3.12	0.98	0.66	0.23
24	5.74	2.28	0.39	3.19	0.98	0.66	0.25
30	5.86	2.26	0.47	3.25	0.97	0.66	0.26
36	5.97	2.24	0.53	3.31	0.97	0.66	0.28
42	6.08	2.23	0.60	3.36	0.97	0.66	0.29
48	6.17	2.21	0.65	3.41	0.97	0.65	0.30
54	6.25	2.20	0.69	3.45	0.97	0.65	0.31
60	6.31	2.18	0.72	3.47	0.97	0.65	0.32
66	6.38	2.17	0.76	3.48	0.97	0.65	0.33
72	6.44	2.17	0.80	3.50	0.97	0.65	0.33
78	6.49	2.16	0.82	3.52	0.97	0.65	0.34
84	6.54	2.14	0.83	3.52	0.97	0.65	0.34
90	6.58	2.11	0.83	3.50	0.97	0.65	0.35
96	6.62	2.10	0.83	3.48	0.97	0.65	0.36
102	6.65	2.09	0.83	3.47	0.97	0.65	0.36
108	6.69	2.08	0.84	3.47	0.97	0.65	0.37
114	6.72	2.07	0.85	3.46	0.97	0.65	0.37
120	6.76	2.05	0.85	3.44	0.97	0.65	0.37

This table reports the mean, standard deviation, skewness and kurtosis of the month-end U.S. annual zero coupon bond yields of maturities from 6 to 120 months, for the period 31 January 1984 to 31 December 2007. The last three columns show the autocorrelation of the yields at displacements of 1, 12 and 30 months.

Table 2: Summary Statistics: the Yield Fitting Errors

Maturity	Mean	Std. Dev.	Max	Min	MAE	RMSE	Autocorrelations		
(*months*)	(*basis points*)	(*basis points*)	(*basis points*)	(*basis points*)	(*basis points*)	(*basis points*)	Lag (1)	Lag (12)	Lag (30)
6	-0.184	4.258	0.124	-0.158	3.363	4.262	0.784	0.272	-0.168
12	-0.091	2.503	0.116	-0.107	1.755	2.505	0.721	0.127	-0.108
18	0.262	3.684	0.100	-0.122	2.899	3.693	0.742	0.162	-0.102
24	0.107	3.379	0.082	-0.097	2.709	3.380	0.754	0.226	-0.122
30	-0.090	2.893	0.081	-0.090	2.287	2.894	0.798	0.342	-0.221
36	0.092	2.435	0.082	-0.100	1.796	2.437	0.755	0.204	-0.178
42	0.593	1.991	0.079	-0.059	1.503	2.077	0.685	0.043	-0.103
48	0.684	2.371	0.092	-0.057	1.796	2.468	0.766	0.186	-0.161
54	0.072	2.896	0.117	-0.110	2.093	2.897	0.765	-0.041	-0.041
60	-0.675	3.445	0.107	-0.141	2.631	3.510	0.793	-0.081	0.006
66	-0.653	3.667	0.157	-0.100	2.886	3.725	0.802	0.339	-0.170
72	-0.250	5.036	0.310	-0.089	3.243	5.042	0.893	0.207	-0.128
78	-0.035	5.086	0.267	-0.086	3.076	5.087	0.889	0.048	-0.062
84	-0.182	3.531	0.226	-0.092	2.467	3.535	0.779	-0.022	-0.112
90	-0.470	2.933	0.061	-0.158	2.158	2.970	0.757	0.061	-0.247
96	-0.591	2.689	0.070	-0.130	1.926	2.753	0.798	0.231	-0.140
102	-0.528	2.324	0.051	-0.100	1.800	2.383	0.766	0.320	0.039
108	-0.248	2.439	0.096	-0.082	1.861	2.451	0.729	0.249	-0.073
114	0.474	3.129	0.103	-0.094	2.454	3.164	0.733	0.204	-0.163
120	1.714	5.079	0.145	-0.166	4.212	5.361	0.772	0.127	-0.044
Average	0.000	3.288	0.123	-0.107	2.446	3.330	0.774	0.160	-0.115

This table reports the mean, the standard deviation, the maximum and the minimum of the fitting errors $\varepsilon(m,t)=\tilde{y}(m,t)-y(m,t)$ where $y(m,t)$ are the model-based yields obtained from eq. (11) and $\tilde{y}(m,t)$ are the actual zero-coupon bond yields for the period 31 January 1984 to 31 December 2007. The table reports also the mean average errors, $MEA=\frac{1}{288}\sum_{t=1}^{288}|\varepsilon(m,t)|$, and the root mean squared fitting errors $RMSE=\sqrt{\frac{1}{288}\sum_{t=1}^{288}[\varepsilon(m,t)]^2}$. The last three columns show the autocorrelation of the fitting errors at displacements of 1, 12 and 30 months. The row at the bottom of the table reports the arithmetic average of the columns.

Table 3: Summary Statistics: the Latent Factors

Initial	Mean	Std. Dev.	Max	Min	Unit Root	Prob	Autocorrelations		
State	(*percent*)	(*percent*)	(*percent*)	(*percent*)	Test		Lag (1)	Lag (12)	Lag (30)
X2 (0,t)	5.122	2.268	11.272	0.777	-2.330	0.416	0.985	0.642	0.160
X1 (0,t)	5.948	2.802	14.677	-0.444	-3.554	0.036	0.948	0.566	0.228
X0 (0,t)	7.285	2.011	13.981	4.127	-3.226	0.081	0.973	0.637	0.388

This table reports the mean, standard deviation, maximum and minimum of the three latent factors for calendar-time t going from 31 January 1984 to 31 December 2007. The Phillips-Perron test (using the Bartlett kernel) and the associated one-sided p-values indicate that the null hypothesis of unit root cannot be rejected at the 1 percent level for all three latent factors, and at the 5 percent level for $x_2(0,t)$ and $x_0(0,t)$. This is consistent with the autocorrelation values shown in the last three columns of the table.

Table 4: The Frequency Bandwidths

Frequency Component	Low Passband Filter Eigenvalues	Frequency Bandwidth (*rad/month*)	Time Resolution Interval (*months*)
Low-frequency	0.985	[0 - 0.015]	T > 415.5
Medium-frequency	0.6	[0.015 - 0.51]	T >12.29
High-frequency	0.2	[0.51 - 1.6]	T > 3.90

This table reports the eigenvalues assigned to the gain matrix of the dynamic filters to perform the frequency decomposition, the associated finite frequency resolution intervals partitioning the frequency domain, and the related four finite time resolution intervals. Note that when measured in rad/month the time series complete frequency range is $[0 \div f_{max}] = 3.14$, whereas it equals $[0 \div f_{max}] = 0.5$ when measured in cycles/month.

Table 5: The Parameters s_j and r_j of the Dynamic Models M_j

Initial State	System M_j for j=	Parameter	
		s	*r*
$X_0(0,t)$	*lf*	7.55E-03	0.00E+00
	mf	1.46E-01	2.51E-02
	hf	2.00E-01	4.33E-02
$X_1(0,t)$	*lf*	8.38E-03	0.00E+00
	mf	1.46E-01	3.29E-02
	hf	5.52E-01	1.97E-01
$X_2(0,t)$	*lf*	5.00E-04	0.00E+00
	mf	0.00E+00	4.29E-03
	hf	3.10E-01	7.01E-02

This table reports the estimated values of the time-invariant parameters s_j and r_j of the systems M_j for $j = lf, mf, hf$.

Table 6: Summary Statistics: the Yield Forecast Errors

1-month Forecast Horizon

Maturity	Mean (*basis points*)	Std. Dev. (*basis points*)	Autocorrelations			RMSE (*basis points*)	RW RMSE (*basis points*)	RMSE ratio
			Lag (1)	Lag (6)	Lag (12)			
6 months	-7.52	17.33	0.20	0.00	-0.02	18.89	21.92	0.86
1 year	-3.97	20.52	0.11	0.02	-0.04	20.90	24.48	0.85
3 years	-4.40	25.54	-0.04	-0.02	-0.09	25.92	26.79	0.97
5 years	-4.42	23.97	-0.03	-0.03	-0.13	24.37	28.13	0.87
10 years	7.87	19.96	0.08	-0.02	-0.18	21.46	28.87	0.74

6-month Forecast Horizon

Maturity	Mean (*basis points*)	Std. Dev. (*basis points*)	Autocorrelations			RMSE (*basis points*)	RW RMSE (*basis points*)	RMSE ratio
			Lag (1)	Lag (6)	Lag (12)			
6 months	-18.07	35.31	0.49	0.07	0.02	39.66	84.95	0.47
1 year	-19.84	43.11	0.50	0.07	-0.02	47.46	85.12	0.56
3 years	-22.04	49.44	0.57	0.08	-0.07	54.13	85.00	0.64
5 years	-15.02	42.06	0.60	0.01	-0.11	44.66	84.43	0.53
10 years	11.09	28.90	0.59	-0.32	-0.20	30.96	83.24	0.37

1-year Forecast Horizon

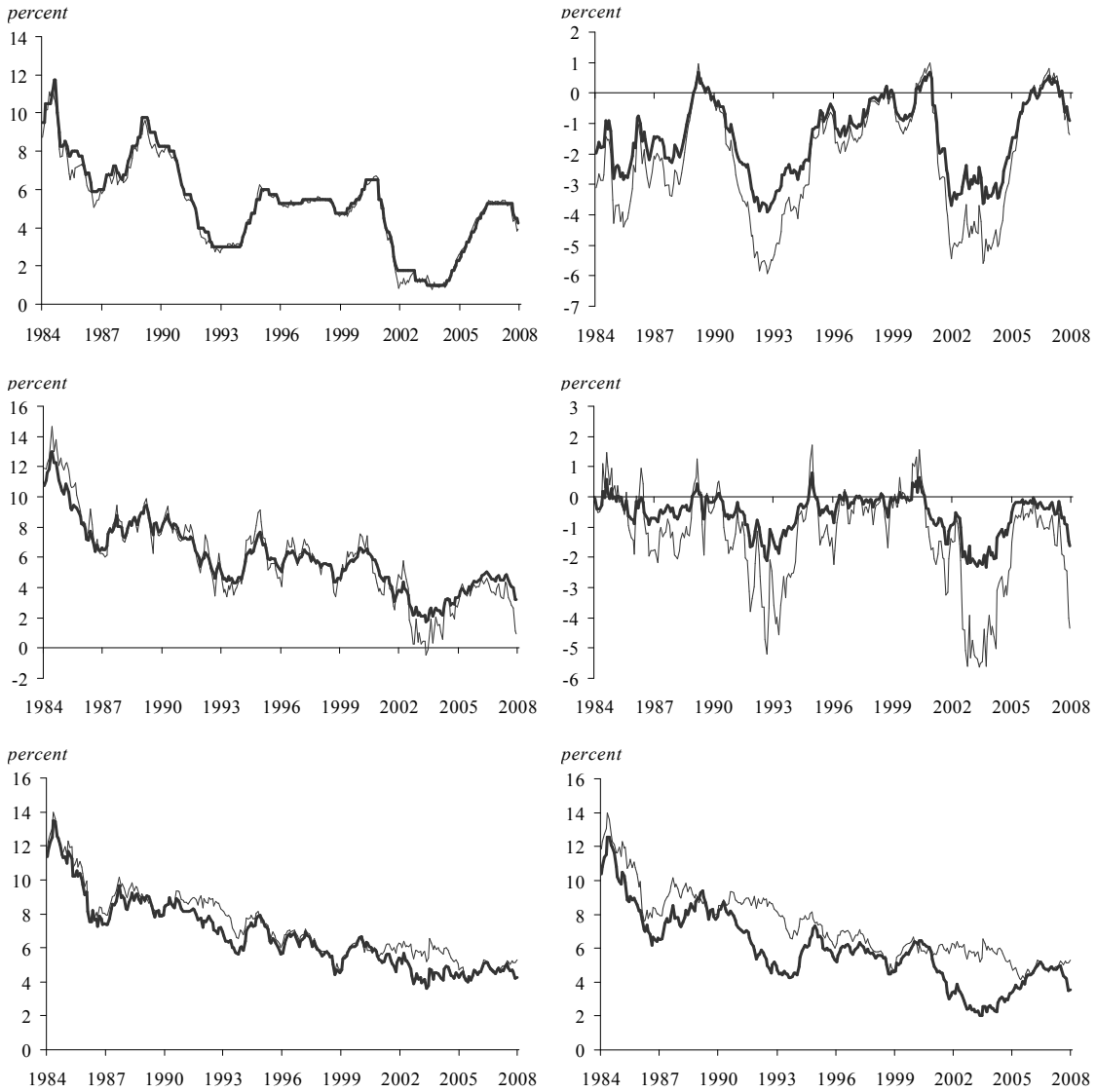
Maturity	Mean (*basis points*)	Std. Dev. (*basis points*)	Autocorrelations			RMSE (*basis points*)	RW RMSE (*basis points*)	RMSE ratio
			Lag (1)	Lag (6)	Lag (12)			
6 months	-19.67	41.67	0.53	0.14	0.03	46.08	147.64	0.31
1 year	-23.81	50.07	0.47	0.12	-0.03	55.44	143.26	0.39
3 years	-26.26	54.12	0.51	0.13	-0.15	60.16	137.85	0.44
5 years	-15.43	42.61	0.57	0.09	-0.20	45.32	132.33	0.34
10 years	17.88	23.99	0.63	-0.18	-0.19	29.92	126.93	0.24

2-year Forecast Horizon

Maturity	Mean (*basis points*)	Std. Dev. (*basis points*)	Autocorrelations			RMSE (*basis points*)	RW RMSE (*basis points*)	RMSE ratio
			Lag (1)	Lag (6)	Lag (12)			
6 months	-3.17	44.15	0.57	0.12	0.00	44.27	233.67	0.19
1 year	-1.05	49.26	0.53	0.11	0.01	49.27	223.22	0.22
3 years	-3.48	49.47	0.47	0.05	0.00	49.59	209.54	0.24
5 years	-3.41	39.85	0.38	-0.17	0.00	39.99	195.26	0.20
10 years	9.86	27.25	0.61	-0.30	0.01	28.98	181.76	0.16

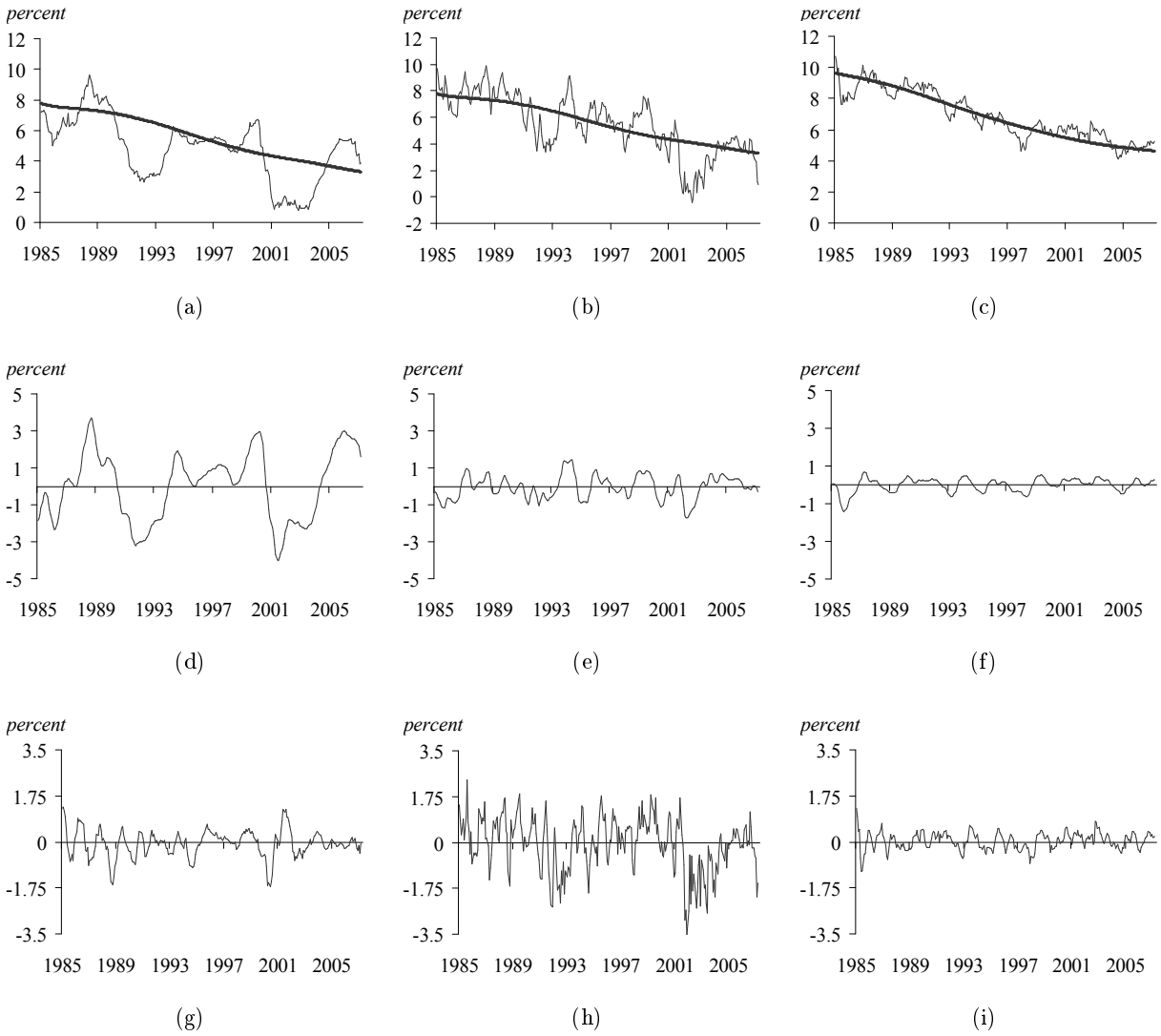
The table reports the mean, the standard deviation, the autocorrelation at displacements of 1, 6 and 12 months and the root mean squared error (RMSE) of the errors defined by the difference between the actual yields at time t and their levels as forecasted out of sample 1, 6, 12 and 24 months before t using yield curve model of eq. (11). The RW RMSE column reports the root mean squared error for the forecasts obtained when using the random walk model to predict the yields, while the last column shows the ratio between the values in the RMSE column and the values in the RW RMSE column.

Figure 1: The Latent Factors and Empirical Counterparts



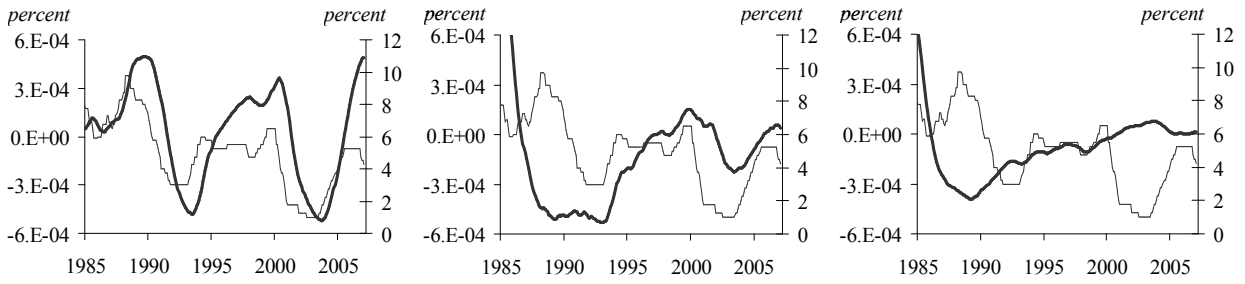
In this figure, the left-hand side column, from top to bottom, shows: 1) the *starting value* $x_2(0,t)$ with the U.S. federal funds rate target (thick line); 2) the *shape factor* $x_1(0,t)$ with the middle-range maturity transformation of the yield curve $[y(24,t)+y(36,t)+y(60,t)]/3$, (thick line); the *final value* $x_0(0,t)$ with the 10-year maturity yield (thick line). The column on right-hand side, from top to bottom, shows: 4) the *slope factor* $\beta_2(t)$ with the slope-transformation of the yields $[y(6,t)-y(120,t)]$, (thick line); 5) the *curvature factor* $\beta_3(t)$ with the curvature-transformation of the yields $[2*y(24,t)-y(120,t)-y(6,t)]$, (thick line); and 6) the *level factor* $\beta_1(t)$ with the level-transformation of the yields $[y(6,t)+y(24,t)+y(120,t)]/3$, (thick line). The calendar-time considered t goes from 31 January 1984 to 31 December 2007.

Figure 2: The Frequency Decompositions of the Latent Factors $x_0(0,t)$, $x_1(0,t)$ and $x_2(0,t)$



This figure shows the frequency decompositions of the latent factors over 31 October 1985 to 31 December 2007. The first column, shows the starting value $x_2(0,t)$ together with its low-frequency component (LF) (thick line) in panel (a), its medium-frequency component (MF) in panel (d), and its high-frequency component (HF) in panel (g). The second column presents the shape factor $x_1(0,t)$ and its LF (thick line) in panel (b), its MF in panel (e), and its HF in panel (h). The last column plots the final value $x_0(0,t)$ and its LF (thick line) in panel (c), its MF in panel (f), and its HF in panel (i).

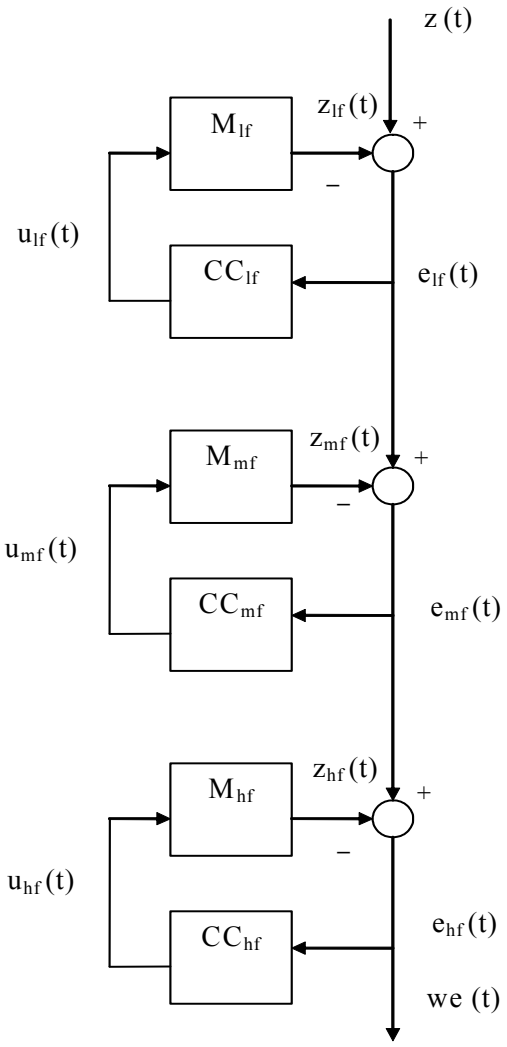
Figure 3: The Long-Run Shocks Driving the Dynamics of the Latent Factors



(a) (b) (c)

This figure contrasts the target for the U.S. federal funds rate (thin line, scale on the right-hand side) with the long-run forces (thick lines) driving the calendar-time evolution of the starting value $x_2(0,t)$ in panel (a), of the shape factor $x_1(0,t)$ in panel (b), and of the final value $x_0(0,t)$ in panel (c) over the period from 31 October 1985 to 31 December 2007.

Figure 4: Frequency Decomposition of the Time Function $z(t)$



This figure provides a schematic representation of the input-output state observer recursions to perform the frequency decomposition of the generic time function $z(t)$.

European Central Bank Working Paper Series

For a complete list of Working Papers published by the ECB, please visit the ECB's website (http://www.ecb.europa.eu).

888 "House prices, money, credit and the macroeconomy" by C. Goodhart and B. Hofmann, April 2008.

889 "Credit and the natural rate of interest" by F. De Fiore and O. Tristani, April 2008.

890 "Globalisation, domestic inflation and global output gaps: evidence from the euro area" by A. Calza, April 2008.

891 "House prices and the stance of monetary policy" by M. Jarociński and F. Smets, April 2008.

892 "Identification of New Keynesian Phillips Curves from a global perspective" by S. Dées, M. H. Pesaran, L. V. Smith and R. P. Smith, April 2008.

893 "Sticky wages: evidence from quarterly microeconomic data" by T. Heckel, H. Le Bihan and M. Montornès, May 2008.

894 "The role of country-specific trade and survey data in forecasting euro area manufacturing production: perspective from large panel factor models" by M. Darracq Pariès and L. Maurin, May 2008.

895 "On the empirical evidence of the intertemporal current account model for the euro area countries" by M. Ca'Zorzi and M. Rubaszek, May 2008.

896 "The Maastricht convergence criteria and optimal monetary policy for the EMU accession countries" by A. Lipińska, May 2008.

897 "DSGE-modelling when agents are imperfectly informed" by P. De Grauwe, May 2008.

898 "Central bank communication and monetary policy: a survey of theory and evidence" by A. S. Blinder, M. Ehrmann, M. Fratzscher, J. De Haan and D.-J. Jansen, May 2008.

899 "Robust monetary rules under unstructured and structured model uncertainty" by P. Levine and J. Pearlman, May 2008.

900 "Forecasting inflation and tracking monetary policy in the euro area: does national information help?" by R. Cristadoro, F. Venditti and G. Saporito, May 2008.

901 "The usefulness of infra-annual government cash budgetary data for fiscal forecasting in the euro area" by L. Onorante, D. J. Pedregal, J. J. Pérez and S. Signorini, May 2008.

902 "Fiscal consolidation in the euro area: long-run benefits and short-run costs" by G. Coenen, M. Mohr and R. Straub, May 2008.

903 "A robust criterion for determining the number of static factors in approximate factor models" by L. Alessi, M. Barigozzi and M. Capasso, May 2008.

904 "Does money matter in the IS curve? The case of the UK" by B. E. Jones and L. Stracca, June 2008.

905 "A persistence-weighted measure of core inflation in the euro area" by L. Bilke and L. Stracca, June 2008.

906 "The impact of the euro on equity markets: a country and sector decomposition" by L. Cappiello, A. Kadareja and S. Manganelli, June 2008.

907 "Globalisation and the euro area: simulation based analysis using the New Area Wide Model" by P. Jacquinot and R. Straub, June 2008.

908 "3-step analysis of public finances sustainability: the case of the European Union" by A. Afonso and C. Rault, June 2008.

909 "Repo markets, counterparty risk and the 2007/2008 liquidity crisis" by C. Ewerhart and J. Tapking, June 2008.

910 "How has CDO market pricing changed during the turmoil? Evidence from CDS index tranches" by M. Scheicher, June 2008.

911 "Global liquidity glut or global savings glut? A structural VAR approach" by T. Bracke and M. Fidora, June 2008.

912 "Labour cost and employment across euro area countries and sectors" by B. Pierluigi and M. Roma, June 2008.

913 "Country and industry equity risk premia in the euro area: an intertemporal approach" by L. Cappiello, M. Lo Duca and A. Maddaloni, June 2008.

914 "Evolution and sources of manufacturing productivity growth: evidence from a panel of European countries" by S. Giannangeli and R. Gómez-Salvador, June 2008.

915 "Medium run redux: technical change, factor shares and frictions in the euro area" by P. McAdam and A. Willman, June 2008.

916 "Optimal reserve composition in the presence of sudden stops: the euro and the dollar as safe haven currencies" by R. Beck and E. Rahbari, July 2008.

917 "Modelling and forecasting the yield curve under model uncertainty" by P. Donati and F. Donati, July 2008.

ISSN 1561-0810