

Yield Curve Forecasts and the Predictive Power of Macro Variables in a VAR Framework

Luciano Vereda¹, Hélio Lopes², Jessica Kubrusly³ and Adrian Pizzinga^{3,*}

¹*Economics Department, Fluminense Federal University, Brasil*

²*Department of Informatics, Pontifical Catholic University of Rio de Janeiro, Brasil*

³*Department of Statistics, Institute of Mathematics and Statistics, Fluminense Federal University, Brasil*

Abstract: Recent macro-finance papers have documented the importance of adding information from macro variables in order to improve out-of-sample forecasting performance of bond yields. This paper aims at investigating the reasons for this success. We use Diebold and Li's dynamic version of the Nelson and Siegel exponential approximation of the yield curve to estimate the factors that govern its dynamics. Factors and macro variables are modeled simultaneously in a VAR framework, which is then used to forecast the factors. Our main conclusions are (i) this framework is useful in forecasting slope and curvature factors, but not the level factor; and (ii) to get good results in forecasting the level factor, one needs a macro model which incorporates variables related to long-run trends and expectations.

Keywords: Macroeconomic variables, Nelson and Siegel curve, term structure of interest rates, VAR models, yield curve.

1. INTRODUCTION

Yield curve modeling and forecasting are among the most studied topics in economics and finance due to their remarkable range of applications. For economists, the yield curve is essential because it reflects the prices that consumers pay or receive in order to anticipate or postpone their consumption plans, the costs that firms incur in order to implement their investment projects and the burden faced by the government when issuing public debt instruments. For macroeconomics practitioners, the yield curve is important because it mirrors market expectations of future overnight rates, therefore being capable of signaling future monetary policy decisions. For central bankers, market analysts and other forecasting professionals, the yield curve is a leading indicator of economic activity and inflation, thus being useful for developing short and medium-term macroeconomic scenarios. Financial applications are also many, going from providing the discount factors that are necessary to evaluate the fair value of pure discount and coupon bearing bonds to derivatives pricing and asset liability management.

The issue of developing multivariate models capable of explaining the observed behavior of bond yields has always been in the spotlight of econometric research. Maybe one of the most important papers in

the field is Hall, Anderson and Granger (1992), who show that yields are cointegrated and that, if monetary policy is implemented by means of a short rate (the FED Funds), then the spreads between yields of different maturities define a set of cointegrating vectors. These results lead the authors to propose an error-correction model in which spreads are used as the error-correction terms. Working with a sample of USA data starting right after the Volcker experiment, the model was proved to be stable and useful for forecasting changes in yields. Lots of papers followed the contribution of Hall, Anderson and Granger, generally improving the basic structure in which yields move in tandem by incorporating other characteristics of their dynamics, like the dependence of the model's parameters to the state of the economy (or regime; see Hamilton, 1988 and Ang, Bekaert and Wei, 2008) or the time-varying nature of volatility (see, for example, Engle and Ng, 1993). Despite their increasing sophistication, most of the models had limited success in beating simple random walks in forecasting exercises.

Perhaps this failure comes from the absence of important information embedded in key macroeconomic variables. It is precisely this kind of information (and the ways of incorporating it to a model of the yield curve) that has been studied recently in articles of the so-called macro-finance literature. Understanding the reasons behind this link is straightforward. First, the monetary policy literature asserts that the short rate is the most important policy instrument that central bankers have in their hands to prevent inflationary pressures and stabilize economic

*Address correspondence to this author at the Rua Mario Santos Braga s/n. Instituto de Matemática e Estatística 7º andar – Cep 24020-140 – Centro – Niterói, RJ, Brasil; E-mail: adrianhp@est.uff.br

JEL classification: C53; E43; G12.

activity. The maneuvering of this instrument is usually modeled as a reaction function in which there are short rate increases whenever inflation exceeds a given target and/or real output departs from its long-run potential level (see, for example, Woodford, 2003 and Taylor, 1993). Second, even if the pure expectations hypothesis is not valid (and there are several papers in the economic and financial literature claiming that it is not; see, for example, Campbell, 1995), it is hard to reject the assumption that long rates are risk-adjusted averages of future expected short rates. Therefore, there is a channel between current and expected future states of the economy, which eventually call for short rate changes, and the term structure of interest rates, which ends up reflecting the current stance of monetary policy and its future developments.

Despite the relative freshness of this strand of economic and financial literature, some models pertaining to it achieved very promising forecasting results. The papers of Ang and Piazzesi (2003) and Hordahl, Tristani and Vestin (2006) are good examples of such models. Ang and Piazzesi propose a model in which the joint dynamics of bond yields and macro variables is characterized by means of a VAR that is subject to restrictions that prevent arbitrage opportunities¹. In an application to USA data, the authors conclude that “...*(i) adding term structure restrictions improves forecasts relative to unconstrained VARs, even beating a random walk, and (ii) forecasts can be further improved by including macro variables*”.

The contribution of Hordahl, Tristani and Vestin (2006) stems from replacing the model used by Ang and Piazzesi to describe the dynamic evolution of macro variables (that was a simple VAR involving current and lagged values of two macro factors measuring inflation and real activity) by a more sophisticated one, in which the economy is described by means of a small-scale rational expectations model involving inflation, real activity and the short rate². Bond

yields are affine functions of the state variables of this model, and the relationships between them are also derived assuming that there are no arbitrage opportunities. In an application to German data, the authors conclude that “... *the inclusion of macroeconomic variables within a structural framework contributes to sharpening our ability of forecasting yields accurately out of sample. The improvement is due both to the inclusion of additional information in the model, and to the structural restrictions imposed on its macroeconomic and term structure sections*”.

Therefore, the macro-finance literature reveals that the information content of macro variables is important; without it, future movements of the term structure cannot be forecasted properly. This is precisely the result that we wish to readdress (and confirm) throughout our paper, but here we want to look at this matter from a new perspective that allows answering the following question: after all, what is in the root of this success?

Since Litterman and Scheinkman (1991) it is known that term structure dynamics can be classified into three basic movements (in its level, slope and curvature). In other words, the whole body of bond yields is driven by three fundamental driving forces: the first affects all bond yields in approximately the same way, the second exerts a stronger effect on short-term yields and the third modifies the curvature of the yield curve (that is, it affects medium-term yields by more than it affects its short and long-term counterparts). Therefore, the original question can be converted into the following one: the presence of macro variables gives better forecasting prospects for level, slope or curvature movements? This is the question that our paper seeks to answer.

We will tackle this question by adopting the methodology proposed by Nelson and Siegel (1987), who suggest a three-component exponential approximation for the term structure, and Diebold and Li (2006), who develop a dynamic version of it. Since each component is easily matched to the three driving forces identified above, this framework allows evaluating the strength of the connection between macro variables and the underlying forces of the yield curve in an easy way, especially in forecasting exercises. One drawback of this approach is losing the positive contribution of ruling out arbitrage opportunities, since the Nelson and Siegel framework in its primary form does not incorporate such restrictions. We are aware of this price, and have

¹The restrictions emerge in the context of an affine term structure model in which macro variables, latent factors and their corresponding shocks determine the evolution of the stochastic discount factor that prices all bonds in the economy.

²More specifically, the model comprises three equations: (i) a Phillips curve, which relates current inflation with its own lags, with expectations of future inflation and with the output gap; (ii) an intertemporal IS equation, which relates current output gap with its own lags, with expectations of future output gaps and with the ex-ante real interest rate; and (iii) a monetary policy rule specifying the way central bankers set the short rate in order to keep the economy in the right path. It is worth emphasizing that equations (i) and (ii) come from the first order conditions that characterize the solution of the problems solved by households and firms, who try to maximize the discounted value of utility and profit streams throughout time.

chosen to pay it anyway, because Nelson and Siegel factors and key macroeconomic variables can be easily accommodated into a VAR model, thus avoiding estimation problems and providing a very simple way of generating out-of-sample forecasts. Since simplicity (or parsimony) is an important attribute of any model designed for forecasting purposes, as the famous “KISS principle” of forecasting (Zellner, 1992) attests, and since forecasting is the main goal of the models developed here, the decision of letting them free of restrictions ruling out arbitrage opportunities can be considered reasonable.

In addition, the positive contribution of imposing restrictions avoiding arbitrage opportunities when the main goal is obtaining out-of-sample forecasts of bond yields is still open to debate. The results of Ang and Piazzesi and Hordahl, Tristani and Vestin seem to reinforce their importance (at least when macroeconomic information is also present), but other researchers are more cautious about their positive effects. Diebold and Li, for example, say that “... it is not obvious to us that the use of arbitrage-free models is necessary or desirable for producing good forecasts”. Furthermore, models pertaining to the affine equilibrium term structure literature (which are very concerned with ruling out any arbitrage possibilities) provide very disappointing out-of-sample forecasting results (see Duffee, 2002)

The plan of the remainder of the paper is as follows. In Section 2 we describe and give some theoretical motivation to the VAR models that are tested for forecasting purposes, while in Section 3 we briefly discuss the macro and term structure data with which those models are estimated. In Section 4 we present the forecasting results for four different time horizons (one, six, twelve and eighteen months ahead) and compare them to the ones achieved by other forecasting schemes that are usually taken as benchmarks. Section 5 concludes.

2. THE MODELS

All models establish a connection between macroeconomic variables and the term structure of interest rates in a VAR framework. It is assumed that the monetary policy instrument is an overnight rate r_t (the FED funds), which the Central Bank sets by means of the following reaction function:

$$r_t = r(\Omega_t) + \varepsilon_t \quad (1)$$

In equation (1), Ω_t represents the information set available to the Central Bank at date t , r is a linear function that describes the monetary authority's reaction to the state of the economy and ε_t is an exogenous shock to monetary policy (with zero mean and variance σ^2). The policy reaction function r reflects the authority's welfare function in the sense that r is designed to respond to any macroeconomic variable that influences its value. When the welfare function penalizes deviations of inflation and output from some pre-specified targets (for example, zero inflation and the full-employment output level), then it is assumed that Ω_t incorporates current and past values of inflation, output and the monetary policy instrument itself. The residual ε_t reflects random, nonsystematic factors that affect policy decisions, such as political factors and the personalities, views, and composition of the Central Bank's decision board (this framework is well described in Clarida, Galí and Gertler, 1999).

The yield curve is characterized as in Nelson and Siegel (1987), who use a three-component exponential approximation for the spot rate curve, as shown below:

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

where $s_t(T)$ denotes the yield of a zero coupon bond with maturity T . The parameter λ_t controls the decay rate of the function $F(T) = \frac{1 - e^{-\lambda_t T}}{\lambda_t T}$ (which is strictly decreasing for any positive value of λ_t) and the value of T at which the function $G(T) = \frac{1 - e^{-\lambda_t T}}{\lambda_t T} - e^{-\lambda_t T}$ reaches its maximum. The factors $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$ are such that (i) a positive shock on $\beta_{1,t}$ increases all yields by the same amount, (ii) due to the shape of the function $F(T)$, which tends to one when $T \rightarrow 0$ and goes to zero in a monotonic fashion when $T \rightarrow \infty$, a positive shock on $\beta_{2,t}$ increases short-term yields by more than their medium and long-term counterparts, and (iii) due to the shape of the function $G(T)$, which exhibits a global maximum at $T = T^*$ and tends to zero when $T \rightarrow 0$ and $T \rightarrow \infty$, a positive shock on $\beta_{3,t}$ increases medium-term yields (i.e. yields for maturities around T^*) by more than their short and long-term counterparts. These attributes make the Nelson and Siegel representation compatible to the decomposition of the yield curve dynamics into level, slope and

curvature changes, each one being driven by its own factor ($\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$, respectively)³.

The basic model is a VAR where the vector of dependent variables (denoted by y_t) is comprised by some measure of inflation (denoted by π_t), some measure of economic activity (denoted by x_t), the short rate r_t and the three factors $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$ ($y_t = [z_t \ \beta_t]'$, where $z_t = [\pi_t \ x_t \ r_t]'$ and $\beta_t = [\beta_{1,t} \ \beta_{2,t} \ \beta_{3,t}]'$). The monetary policy rule (1) is estimated as one equation within the following VAR⁴:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} z_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} \alpha^z \\ \alpha^\beta \end{bmatrix} + \sum_{k=1}^K \begin{bmatrix} \mathbf{A}_k & \mathbf{B}_k \\ \mathbf{C}_k & \mathbf{D}_k \end{bmatrix} \begin{bmatrix} z_{t-k} \\ \beta_{t-k} \end{bmatrix} + \begin{bmatrix} \epsilon_t^z \\ \epsilon_t^\beta \end{bmatrix} \quad (3)$$

where \mathbf{A} , \mathbf{B} , \mathbf{C} , \mathbf{D} , \mathbf{A}_k , \mathbf{B}_k , \mathbf{C}_k and \mathbf{D}_k ($k = 1, \dots, K$) are 3×3 square matrices; $[\alpha^z \ \alpha^\beta]'$ is a 6×1 column vector of constants (α^z and α^β are 3×1 column vectors) and $[\epsilon_t^z \ \epsilon_t^\beta]'$ is a 6×1 column vector of i.i.d. residuals. The shocks in ϵ_t^z can be interpreted as linear combinations of the structural shocks that affect macro variables; they are serially uncorrelated (i.e. ϵ_t^z and ϵ_{t-j}^z are uncorrelated for every integer $j \geq 1$) and uncorrelated to the shocks in ϵ_t^β (in other words, ϵ_{t-j}^z and ϵ_{t-k}^β are uncorrelated for every integers j and k). We assume that the shocks in ϵ_t^β have the same characteristics.

We propose and test several variants of (3). In the first one, $\mathbf{B} = \mathbf{0}$ and $\mathbf{A} = \mathbf{D} = \mathbf{I}$ (where \mathbf{I} is the 3×3 identity matrix), making each factor a function of its own lags and contemporaneous and lagged values of macro variables. This structure is based on the assumption that shocks pertaining to macro variables are capable of affecting term structure variables instantaneously (because financial markets are swift in incorporating new information about the state of the economy), but shocks pertaining to term structure variables cannot do the same (because nominal and

real rigidities impede output and inflation to respond promptly and monetary authorities are not able to react to financial shocks immediately, perhaps because of decision and recognition lags). The first variant of (3) also explores the predictive power of β_t (specially its second component, $\beta_{2,t}$)⁵ about the future behavior of z_t . Note that lags of $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$ are able to influence the dynamics of π_t , x_t and r_t , as in Diebold, Rudebusch and Aruoba (2006). The first variant of (3) is shown below:

$$z_t = \alpha^z + \sum_{k=1}^K \mathbf{A}_k z_{t-k} + \sum_{k=1}^K \mathbf{B}_k \beta_{t-k} + \epsilon_t^z \quad (4a)$$

$$\beta_t = \alpha^\beta - \mathbf{C}z_t + \sum_{k=1}^K \mathbf{C}_k z_{t-k} + \sum_{k=1}^K \mathbf{D}_k \beta_{t-k} + \epsilon_t^\beta \quad (4b)$$

The second variant of (3) is such that $\mathbf{B} = \mathbf{C} = \mathbf{0}$ and $\mathbf{A} = \mathbf{D} = \mathbf{I}$, making each variable a function of its own lags and the lagged values of other variables in y_t . This structure corresponds to a conventional VAR where contemporaneous effects are ruled out. It is comprised by (4a) and the system below:

$$\beta_t = \alpha^\beta + \sum_{k=1}^{K_2} \mathbf{C}_k z_{t-k} + \sum_{k=1}^{K_2} \mathbf{D}_k \beta_{t-k} + \epsilon_t^\beta \quad (5)$$

The third variant is based on the framework suggested by Evans and Marshall (1998) in their studies about the effects of monetary policy shocks. The authors propose a model in which contemporaneous and lagged values of macro variables are independent of contemporaneous and lagged values of yield curve variables (which implies that the dynamics of macro variables determine the dynamics of bond yields, but not the reverse); more specifically, Evans and Marshall assume that $\mathbf{A} = \mathbf{D} = \mathbf{I}$ and $\mathbf{B} = \mathbf{B}_1 = \dots = \mathbf{B}_K = \mathbf{0}$. Despite its strength, these assumptions are far from being exotic in the literature (see, for example, Bernanke, Gertler, and Watson, 1997 or Ang and Piazzesi, 2003) and can be a satisfactory representation of reality⁶. These restrictions allow splitting the model into two parts: the

³The importance of the Nelson and Siegel framework, its statistical underpinnings, its interpretation in terms of macroeconomic influences and other issues can be seen in Litterman and Scheinkman (1991); Knez, Litterman and Scheinkman (1994); Bliss (1996); Bliss (1997); Evans and Marshall (1998); Fisher (2001); Wu (2001); Evans and Marshall (2002); Wu (2003); Diebold, Rudebusch and Aruoba (2006); Diebold, Li and Yue (2008); and others.

⁴The formulae are written according to the following notation: bold capital letters denote matrices, bold lower case letters denote vectors (column or row ones) and lower case letters denote real numbers.

⁵The predictive power of the slope of the yield curve for future inflation and economic activity was exhaustively discussed in the literature; see, for example, Estrella and Hardouvelis (1991), Campbell (1995), Estrella and Mishkin (1997), Kozicki (1997), Mishkin (1998), Stock and Watson (2000), Ang, Piazzesi and Wei (2006), and Estrella (2005). Furthermore, the yield curve (more specifically, the forward rates implied by it) convey information about future monetary policy decisions; see, for example, Campbell and Shiller (1991) and Svensson (1994).

⁶According to Diebold, Rudebusch and Aruoba (2006), there is "... strong evidence of macroeconomic effects on the future yield curve and somewhat weaker evidence of yield curve effects on future macroeconomic developments".

first one is a mere reduced form of a VAR model and governs the evolution of π_t , x_t and r_t , while the second part (which can also be represented by (4b)) determines the dynamics of β_t as a function of its own lags and the lags of z_t .

$$z_t = \alpha^z + \sum_{k=1}^K A_k z_{t-k} + \varepsilon_t^z \quad (6)$$

Our assumptions guarantee that β_t , β_{t-1} , β_{t-2} , ..., β_{t-K} are unable to influence the dynamics of z_t .⁷ Equation (4b) shows that current and past values of π_t , x_t and r_t are allowed to influence the dynamics of $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$.

The fourth variant is based on some notions about monetary policy implementation and its transmission throughout the economy, in particular the interest rate channel. First, the monetary policy literature asserts that the short-term nominal interest rate is the key policy instrument that the Central Bank sets in order to prevent inflationary pressures and stabilize economic activity. Second, this adjustment is usually modeled as a reaction function like (1), in which the short rate increases whenever inflation exceeds a given target and real output departs from its long run potential (see, for example, Woodford, 2003 and Taylor, 1993). Third, the traditional Keynesian view supports the idea that an increase in the short rate leads to an increase in medium and long-term nominal interest rates, as investors act to arbitrage away differences in risk-adjusted expected returns on bonds of various maturities (which is a consequence of the expectations hypothesis of the term structure). Finally, in the presence of nominal price rigidity, these movements in nominal interest rates translate into movements in real interest rates over all horizons as well, forcing entrepreneurs and consumers to cut back their investment and consumption plans, thus affecting aggregate demand and inflation. The essence of this chain is that the short rate serves as a link between current and expected future states of the economy and the yield curve.

The ideas discussed above lead to a fourth variant of (3) in which $A = D = I$, $B = 0$ and:

$$C = \begin{bmatrix} 0 & 0 & c_{13} \\ 0 & 0 & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix} \quad C_k = \begin{bmatrix} 0 & 0 & c_{13}^k \\ 0 & 0 & c_{23}^k \\ 0 & 0 & c_{33}^k \end{bmatrix} \quad (7)$$

Equation (4b) turns out to be:

$$\beta_t = \alpha^\beta - v r_t + \sum_{k=1}^K v_k r_{t-k} + \sum_{k=1}^K D_k \beta_{t-k} + \varepsilon_t^\beta \quad (8)$$

where $v^{rr} = [c_{13} \quad c_{23} \quad c_{33}]$ and $v_k^{rr} = [c_{13}^k \quad c_{23}^k \quad c_{33}^k]$. We can see that the assumptions in (7) serves to make β_t a function of its own lags and current and lagged values of r_t . Finally, we can also build a fifth variant in which the VAR model (6) and the system (8) are put together.

Why imposing so many restrictions? The main reason is simple: since parsimony is an important feature of any model which is designed to generate good out-of-sample forecasts, we decided to sequentially impose additional restrictions to the model formed by (4a) and (4b) with the hope of improving its forecasting performance. The importance of preserving economic content made us recur to economic theory to design the extra restrictions whenever possible. Of course restrictions can be false, and Appendix A brings a brief discussion about their validity suggesting that this is precisely the case. Anyway, they will be maintained for the sake of saving parameters and providing good forecasting results⁸.

In this paper we use models (4a), (4b); (4a), (5); (6), (4b); (4a), (8) and (6), (8) to forecast the future paths of $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$. The first step is estimating the models described in (4a) and (6) with information available until date t ; then we use the outcomes to calculate forecasts of π_t , x_t and r_t ($E_t[\pi_{t+\tilde{T}}]$, $E_t[x_{t+\tilde{T}}]$ and $E_t[r_{t+\tilde{T}}]$, \tilde{T} being the forecasting horizon). Finally, these forecasts are inserted into (4b), (5) and (8),

⁷As a (desirable) collateral effect, the separation between (6) and (4b) makes the second system capable of absorbing any exogenous paths of π_t , x_t and r_t that we wish (in particular, not only those generated by (6)). System (4b) can even absorb the true values of π_t , x_t and r_t , thus generating a kind of "first best results" that tell us what would be the improvement in forecasting performance if forecasting errors in macro variables were set to zero. We will have more to say about this later.

⁸The search for parsimony when the basic interest is in out-of-sample forecasts is recommended by Diebold and Li (2006), who say that "... *unrestricted VARs tend to produce poor forecasts of economic variables ... due to the large number of included parameters and the resulting potential for in-sample overfitting*". Furthermore, Diebold and Li invoke the so-called "shrinkage principle", which relies on the idea that the degradation of in-sample fit can be traded against the benefits in out-of-sample forecasting performance by imposing additional restrictions to the model. In fact, Diebold and Li say that the shrinkage principle can be interpreted "... *as the insight that imposition of restrictions, which will of course degrade in-sample fit, may nevertheless be helpful for out-of-sample forecasting, even if the restrictions are false*".

which are then used to generate forecasts of $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$ ($E_t[\beta_{1,t+\bar{t}}]$, $E_t[\beta_{2,t+\bar{t}}]$ and $E_t[\beta_{3,t+\bar{t}}]$).

3. THE DATA

In this section we introduce the data. The yield curve is taken from FRED (Federal Reserve Economic Data), a database provided by the Federal Reserve Bank of Saint Louis. We take 11 monthly series of constant maturity interest rates (identified as GS1M, GS3M, GS6M, GS1, GS2, GS3, GS5, GS7, GS10, GS20 and GS30) to proxy yields of zero-coupon bonds of several maturities (1, 3 and 6 months; 1, 2, 3, 5, 7, 10, 20 and 30 years). Its public character was the main reason behind our decision of using this database instead of the CRSP government bonds files, which are used in most of yield curve studies. Nevertheless, the same general picture emerges when one engages on studying their main characteristics⁹.

Several approaches are available to estimate the factors and parameters of the Nelson and Siegel representation. The most straightforward approach, which is used in Diebold and Li (2006), consists of assuming that the parameter λ_t does not vary over time and involves calibrating its value in order to make the function $G(T)$ attain its maximum at $T = T^*$, where T^* is fixed at some reasonable medium-term maturity. This strategy allows applying ordinary least squares to estimate the values of level, slope and curvature at period t . Repeating this procedure for every t allows recovering the three time series $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$.

Another approach, which can also be applied to the cross-section of yields at a given period t , consists of estimating λ_t and the factors altogether. This approach incurs in augmented complexity due to the nonlinear nature of the problem that must be undertaken¹⁰. There is yet another approach (suggested by Diebold, Rudebusch and Aruoba (2006)), which consists of imposing a dynamic model for $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$ (for example, a VAR(1)), writing a model for observable yields and latent factors in state-space form (according to the structure imposed by (2)), and estimating its

parameters (together with a fixed λ) by means of the Kalman filter. Here we pursue the first (and simplest) option for two reasons: (i) in our view, the “KISS principle” of forecasting (Zellner, 1992) may be extended to rule out complicated estimation strategies; and (ii) De Pooter (2007) shows that the three time series $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$ that are uncovered by means of different estimation procedures are quite similar¹¹.

Therefore, we follow Diebold and Li (2006) and set $\lambda_t = \bar{\lambda} = 0.0609$, implying that the maximum of $G(T)$ occurs at $T^* = 30$ months. After calibrating λ_t , factor loadings are calculated according to (2) and their values are used as independent variables in the ordinary least squares regressions that provide level, slope and curvature estimates. Repeating this procedure for every month allows obtaining the three time series $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$.

Tables 1 and 2 present some descriptive statistics for the period 1985:01 to 2008:03¹². We can see that $\beta_{1,t}$, $\beta_{2,t}$, $\beta_{3,t}$ and their empirical proxies¹³ are quite close (correlations between factors and empirical proxies are 0.998 for level, 0.989 for slope and 0.978 for curvature), which indicates that our choice of calibrating λ_t did not come up with unreasonable results. Table 1 shows that: (i) $\beta_{1,t}$ is positive on average, its evolution throughout time is very persistent and its value varies only moderately when compared to its mean (the ratio between its mean and standard deviation is 4.34); (ii) $\beta_{2,t}$ is negative on average (suggesting that yields are an increasing function of maturity), its evolution is not very persistent (indeed, its persistence is smaller than that of any other individual yield), and its value varies significantly when compared

¹¹See the discussion in Section 6.2 and Figures 7, 8 and 9 in De Pooter's paper for more details.

¹²We use a sample which starts in January 1985 and ends in March 2008. We chose January 1985 as the starting point for two reasons: (i) comparability: the same starting point was chosen by Diebold and Li; and (ii) economic uniformity: since we wanted to avoid changes in the data generating process, observations coming from periods before, during and right after the Volcker experiment (which was a somewhat noisy transition period) were eliminated. We chose March 2008 as the ending point to avoid data disturbed by the subprime crisis, which reached its peak in the end of the same year. In fact, the Federal National Mortgage Association (FNMA), known as “Fannie Mae”, and the Federal Home Loan Mortgage Corporation (FHLMC), nicknamed “Freddie Mac”, were both nationalized by the US government in August and September 2008. Shortly thereafter, financial markets were even more disrupted by the bankruptcy of the traditional investment bank Lehman Brothers.

¹³In the spirit of the definitions commonly found in the literature, the empirical proxies of level, slope and curvature are defined as the yield of a thirty-year zero-coupon bond ($s_{(360)}$), the negative of the difference between the yields of a thirty-year and an one-month zero-coupon bonds ($-(s_{(360)} - s_{(1)})$), and the sum of the yields of a thirty-year and an one-month zero-coupon bonds subtracted by twice the value of the yield of a three-year zero-coupon bond ($(s_{(360)} + s_{(1)}) - 2 \times s_{(36)}$), respectively.

⁹For example: (i) the average yield curve is upward sloping, (ii) long rates are less volatile and more persistent than short rates, (iii) all yields are highly correlated to each other, (iv) this correlation varies inversely with the distance between them, and (v) all yields show excess kurtosis and have a long right tail (as indicated by positive measures of skewness). It is worth emphasizing that any details of preliminary statistical exercises are available upon request.

¹⁰The main difficulties come from the somewhat challenging optimization problem that must be solved and the possibility of arriving at very extreme factor estimates, generating outliers that complicate the task of estimating dynamic models for $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$ (see De Pooter, 2007 for more details).

Table 1: The table provides summary statistics for level, slope and curvature calculated according to the Nelson and Siegel framework. The last three columns contain sample autocorrelations at displacements of 1, 12, and 24 months. Raw data come from monthly observations of constant maturity interest rates taken from January 1985 until March 2008

	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis	(A)	(B)	(C)
$\beta_{1,t}$	6.95	6.61	12.02	4.49	1.60	0.69	3.14	0.97	0.63	0.59
$\beta_{2,t}$	-2.26	-2.00	0.65	-5.28	1.62	-0.11	1.83	0.98	0.44	-0.05
$\beta_{3,t}$	-0.99	-0.65	3.97	-7.19	2.24	-0.86	3.52	0.95	0.46	0.15

to its mean (the absolute value of the ratio between its mean and standard deviation is 1.40); (iii) $\beta_{3,t}$ is the least persistent of all factors and the most highly variable when compared to its mean (the absolute value of the ratio between its mean and standard deviation is equal to 0.44). Table 2 shows that factors $\beta_{1,t}$ and $\beta_{3,t}$ are significantly correlated (the correlation coefficient between them is around 0.4); this conclusion is in contrast with the one taken from Diebold and Li (2006), who find that level, slope and curvature are not highly correlated with each other, but it is compatible with other results found in the literature. In Christiansen and Lund (2005), for example, the authors document a significant correlation between their measures of level and curvature, this one being defined as the return of a butterfly position with zero duration (see their Table 2 for more details).

Table 2: The table provides correlation coefficients between level, slope and curvature calculated according to the Nelson and Siegel framework. Raw data come from monthly observations of constant maturity interest rates taken from January 1985 until March 2008

	$\beta_{1,t}$	$\beta_{2,t}$	$\beta_{3,t}$
$\beta_{1,t}$	1.00	-0.28	0.41
$\beta_{2,t}$	-0.28	1.00	0.47
$\beta_{3,t}$	0.41	0.47	1.00

We study results coming from two different groups of macro variables. The first one embraces the twelve-month inflation rate calculated by means of the consumer price index for all urban consumers (seasonally adjusted, ex-food and energy, identified as CPILFESL at the FRED database), the twelve-month growth rate of the industrial production index (seasonally adjusted, identified as INDPRO at the FRED database) and the FED Funds rate. The second group is built according to Ang and Piazzesi (2003), who extract "good" measures of inflation and economic activity from two sets of individual measures¹⁴.

¹⁴Ang and Piazzesi assert that individual measures of inflation, economic activity and so on reflect not only their "true" values (which are the theoretical objects that they aim to capture), but also noise. Since only the "true" values are of interest, the analyst or researcher should apply some method to extract the "true" values from the individual measures. Ang and Piazzesi adopt principal component analysis to disentangle the "true" values of inflation and economic activity from a bunch of individual measures, and we repeat this strategy here.

Therefore, the alternative measure of π_t that we use throughout the paper is the first principal component of four individual measures of annual inflation, while the alternative measure of x_t is the first principal component of six individual measures of economic activity¹⁵. The measure of the overnight rate r_t is still given by the FED Funds.

Figures 1, 2 and 3 compare the pairs $\beta_{1,t}, \pi_t$; $\beta_{2,t}, x_t$ and $\beta_{3,t}, r_t$. All macro variables are measured as described in the second group. Figure 1 shows that $\beta_{1,t}$ and π_t follow the same downward trend, confirming the validity of the Fisher equation for all yields. Figure 2 shows the pair $\beta_{2,t}$ and x_t , suggesting that this factor reflects the cyclical dynamics of the economy. These results are in line with those discussed in Diebold, Rudebusch and Aruoba (2006) (Section 2.2). Finally, Figure 3 indicates that there is a close relationship between $\beta_{3,t}$ and r_t . This result is in line with Litterman and Scheinkman (1991), who document the existence of a link between the curvature of the yield curve and interest rate volatility, and Chan *et al.* (1992), who show that interest rate volatility increases with the value assumed by the short rate.

Before moving forward it is worth discussing the nature of the six time series, that is, we want to determine if they are stationary or not. Table 3 provides the results. The SIC criterion was applied to choose the lags in the augmented Dickey-Fuller unit-root test, in

¹⁵The four measures of inflation are the twelve-month growth rate of personal consumption expenditures, chain-type price index (identified as PCEPI at the FRED database); the twelve-month growth rate of personal consumption expenditures, chain-type price index less food and energy (identified as PCEPILFE); the twelve-month inflation rate obtained from the consumer price index calculated for all urban consumers (identified as CPIAUCSL); and the twelve-month inflation rate obtained from the consumer price index calculated for all urban consumers and including all items less food and energy (identified as CPILFESL). The six measures of economic activity are housing starts (defined as new privately owned housing units started; identified as HOUST), the industrial production index (identified as INDPRO), the total capacity utilization (identified as TCU), the capacity utilization of the manufacturing sector (identified as MCUMFN), the NAPM index of economic activity (calculated by the Institute for Supply Management; identified as NAPM) and the civilian unemployment rate (calculated for persons of 16 years of age and older; identified as UNRATE). All measures are seasonally adjusted. Results show that over 90% (50%) of the variance of the inflation (economic activity) measures is explained by the first principal component of the group.

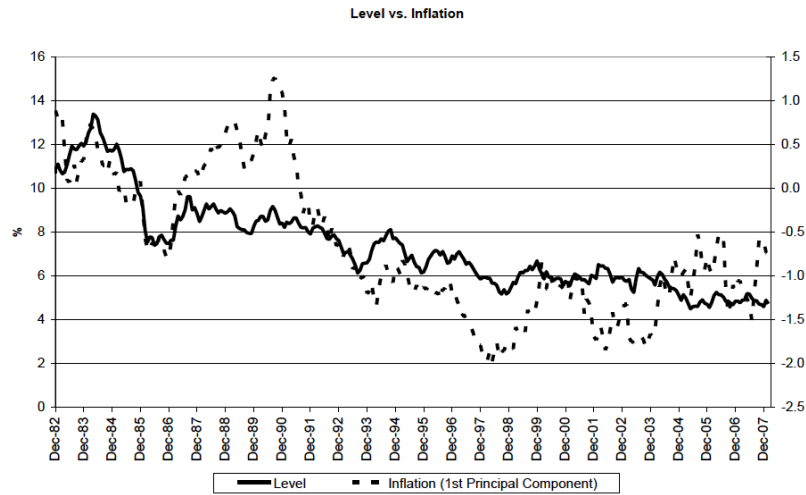


Figure 1: Comparisons between term structure factors and macro variables. Figure 1 shows $\beta_{1,t}$ and the first principal component of four individual measures of twelve-month inflation rate.

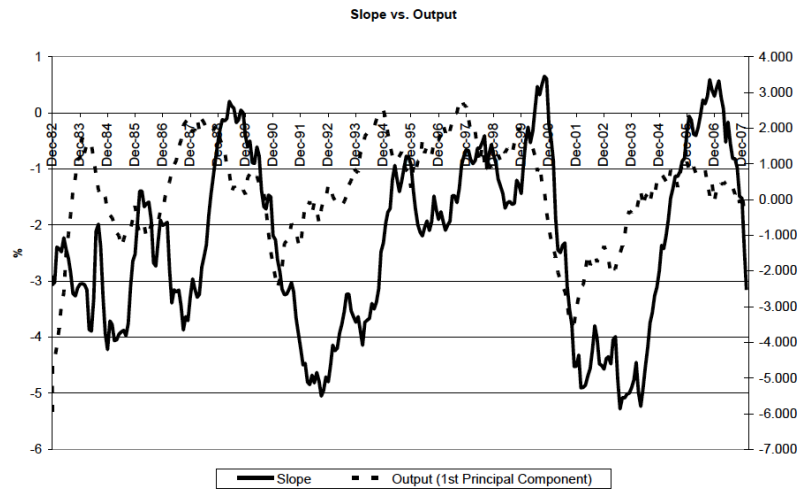


Figure 2: Comparisons between term structure factors and macro variables. Figure 2 compares $\beta_{2,t}$ and the first principal component of six individual measures of economic activity.

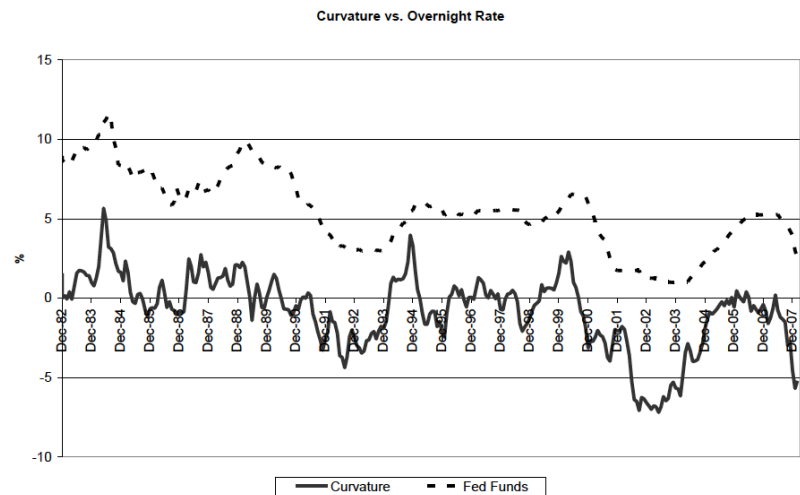


Figure 3: Comparisons between term structure factors and macro variables. Figure 3 makes a comparison between $\beta_{3,t}$ and the FED Funds rate. Raw data come from monthly observations taken from December 1982 until March 2008.

Table 3: This table provides the results achieved when evaluating the null hypothesis of term structure factors and macro variables being I (1). The SIC criterion was applied to choose the lags in the augmented Dickey-Fuller unit-root test (ADF), in which the presence of an intercept was assumed. In the test proposed by Phillips-Perron (PP), the spectral estimation method was accomplished by means of the Bartlett kernel, while the Newey-West bandwidth was chosen automatically. The MacKinnon critical values for rejection of the null are -3.4537 at the 1% level, -2.8717 at the five percent level and -2.5723 at the 10% level. Observations belong to the period January 1985 to March 2008

	$\beta_{1,t}$	$\beta_{2,t}$	$\beta_{3,t}$	r_t	π_t	x_t
ADF	-2.6614	-2.8229	-3.1052	-1.3125	-1.4155	-2.7656
PP	-2.6327	-2.6147	-2.4967	-1.8803	-1.7292	-3.0854

which the presence of an intercept was assumed. In the test proposed by Phillips-Perron, the spectral estimation method was accomplished by means of the Bartlett kernel, while the Newey-West bandwidth was chosen automatically. The MacKinnon critical values for rejection of the null (a unit root exists) are -3.4537 at the 1% level, -2.8717 at the five percent level and -2.5723 at the 10% level. The hypothesis that term structure factors are not stationary is frequently rejected, but the same does not occur for macro variables (where the twelve-month inflation rate calculated by means of the consumer price index for all urban consumers and the twelve-month growth rate of the industrial production index stand for π_t and x_t , respectively). Although results do not recommend modeling $\beta_{1,t}$, $\beta_{2,t}$, $\beta_{3,t}$, π_t , x_t and r_t all in levels, this is precisely what we are going to do for three reasons: (i) the majority of theoretical macroeconomic models pertaining to the New-Keynesian tradition incur in equilibrium equations in which macro variables appear in levels¹⁶, (ii) the same decision was taken in other articles (for example, Diebold and Li, 2006 and Diebold, Rudebusch and Aruoba, 2006)¹⁷, and (iii) some authors (Sims and Uhlig, 1991 and Sims, Stock and Watson, 1990) argue that classical unit roots asymptotics is of little practical value and that the common practice of attempting to transform models into a stationary form by first differencing or applying cointegration operators whenever it appears likely that the data are integrated is, in many cases, unnecessary.

4. ECONOMETRIC RESULTS

In this section we use models (4a),(4b); (4a),(5); (6),(4b); (4a),(8) and (6),(8) to generate out-of-sample forecasts of level, slope and curvature for four different forecasting horizons, and compare the results to those achieved by other forecasting schemes proposed in the

literature. Our benchmarks are: (a) a VAR model in which the vector of explanatory variables contains $\beta_{1,t}$,

$\beta_{2,t}$ and $\beta_{3,t}$ (that is, $\beta_t = \alpha^\beta + \sum_{k=1}^K \mathbf{D}_k \beta_{t-k} + \varepsilon_t^\beta$), and (b)

three univariate autoregressive models (one for each factor, that is, $\beta_{i,t} = \alpha^{\beta_i} + \sum_{k=1}^K d_{11}^k \beta_{i,t-k} + \varepsilon_t^{\beta_i}$, where $i = 1, 2, 3$)¹⁸. Our procedure for examining out-of-sample forecasts is very conventional and involves the following steps:

- 1) Estimate models (4a),(4b); (4a),(5); (6),(4b); (4a),(8) and (6),(8), together with benchmarks (a) and (b), using observations taken from period t until $t+T$. We fix $K=1$ and $K=2$ ¹⁹.
- 2) Calculate \tilde{T} -month-ahead forecasts ($\tilde{T}=1, 6, 12$ and 18 months) of level, slope and curvature (in other words, calculate $E_{t+T}[\beta_{i,t+T+\tilde{T}}]$, $i=1, 2, 3$) and the errors ($\beta_{i,t+T+\tilde{T}} - E_{t+T}[\beta_{i,t+T+\tilde{T}}]$, $i=1, 2, 3$).
- 3) Add one more observation to the previous sample (that is, now the sample starts at t and ends at $t+T+1$), go back to step 1 and obtain another set of \tilde{T} -month-ahead forecasts ($E_{t+T+1}[\beta_{i,t+T+\tilde{T}+1}]$, $i=1, 2, 3$) and their respective errors ($\beta_{i,t+T+\tilde{T}+1} - E_{t+T+1}[\beta_{i,t+T+\tilde{T}+1}]$, $i=1, 2, 3$).

The process is undertaken using a sample coming from January 1985 until March 2008. The first estimation round is executed using a sub-sample

¹⁶In this literature, first differences are exceptions or special cases. The accelerationist Phillips curve, which emerges in the context of perfect indexation, is the most well-known example.

¹⁷Diebold and Li, who deal with data coming from January 1985 until December 2000, find that the null of level and slope being I(1) cannot be rejected at the 10% level. In spite of this result, both variables appear in levels (and not first differences) in their proposed univariate models.

¹⁸These restrictions produce other versions of the original model in which relationships between term structure factors and macro variables (and links between the term structure factors themselves) are completely ignored.

¹⁹These choices come from an exercise in which systems (4a),(4b); (4a),(5); (6),(4b); (4a),(8) and (6),(8) were estimated recursively (i.e. month by month). The procedure was initiated with a sample coming from January 1985 until June 1995, and it was interrupted when the sample comprised observations taken from the same point until September 2006. Results revealed that these choices almost always minimized the value assumed by the BIC statistics. In addition, there is no point in trying greater values of K because this would "violate" the "KISS principle".

starting at January 1985 and ending at June 1995. The last 18 observations of the sample are not used for estimation purposes, since they are necessary to compute the out-of-sample forecasting errors pertaining to the last execution. Estimations were made by means of the Eviews software and applying the OLS technique. Forecasts and their respective errors were also obtained by means of Eviews. The criterion used to judge forecasting performance is the mean squared difference between actual and forecasted values of term structure factors²⁰.

Tables 5, 6 and 7 show two kinds of information:

- The mean squared difference between actual and forecasted values of term structure factors for each forecasting model²¹.
- The statistical significance of the difference between the mean squared errors achieved by the model that serves as benchmark and the others.

The statistical significance is assessed by means of the test proposed by Diebold and Mariano (1995), in which the benchmark is compared to its competing models one at a time (that is, we execute a sequence of bilateral comparisons). The null hypothesis is that the forecasts calculated by means of the benchmark and the competitors have the same mean squared error. Positive (negative) values of the test statistics indicate superiority (inferiority) of the competing model, and results above 1.96 (below -1.96) denote that the difference is significant at the 5% level.

Tables 5, 6 and 7 display the results achieved by models (4a),(4b); (4a),(5); (6),(4b); (4a),(8); (6),(8); (a) and (b) when forecasting the future path of the level, slope and curvature factors, respectively. Due to space limitations, only the results achieved when $K=1$ are shown²². Each table is divided into 2 blocks, one for each group of inflation and economic activity measures. The first line of each block informs the MSEs achieved by model (b) for each forecasting

horizon. Since model (b) benchmarks the calculations needed to obtain the Diebold and Mariano test statistics, the second line is empty. The next pair of lines informs the MSEs and the values of the Diebold and Mariano test statistics pertaining to model (a) for each forecasting horizon. The other lines should be read conformably. The cells containing the smallest MSEs for each forecasting horizon and Diebold and Mariano test statistics that indicate significant differences in forecasting performance have a coloured background.

We also test systems (4b) and (8) using the true values of macro variables instead of their forecasted values. This is possible because, from the point of view of the term structure block, the macro variables π_t , x_t and r_t are exogenous in two of our modeling strategies. Thus, the symbol (*) refers to systems (4b) and (8) having the true values of π_t , x_t and r_t as their inputs. The exercise allows evaluating the potential advantages of substituting the VAR model described in (6), which is a very simple representation of the economy, by another (perhaps more sophisticated) macroeconomic model, that is probably capable of generating better forecasts of the future paths of macro variables.

We note the following points regarding the forecasting performance of the models. First, it is hard to beat models (a) and (b) if one is interested in forecasting the future path of the level factor, since they appear 6 times (out of 8) among the group of best predictors. A full understanding of this result asks for some words about the information embedded in model (6). We know that the main role of monetary policy is counterbalancing the negative effects of structural shocks coming from different sources, that different shocks can exert different effects on macro variables and that VAR models like (6) are frequently used to study the consequences of transitory supply, demand and monetary shocks, taking into account the systematic response of monetary authorities in order to reestablish the desired path of the economy. By contrast, VAR models like (6) have little to say about persistent influences, such as long-run perceptions of inflation and the steady-state real interest rate, this one reflecting the pace of long-run economic growth. Long-run expectations of economic growth depend basically on technology developments and the accumulation of capital and labor, factors that are not contemplated in conventional VAR models. Furthermore, long-run expectations of inflation are implicitly or explicitly determined by credible central bankers (this is quite

²⁰We could also use the mean squared error between actual and forecasted yields, but we will not do so because our main interest is in forecasting the future paths of the factors. Furthermore, there is always the risk of large offsetting errors, that is, we want to avoid situations in which forecasting errors pertaining to level, slope and curvature are such that, when using (2) to "move back" to forecasted yields, they compensate each other in such a way that "net" forecasting errors are small.

²¹Results are multiplied by 100000 to facilitate comparisons.

²²Nevertheless, we want to emphasize that the insights provided by the analysis and interpretation of Tables 5, 6 and 7 are not sensitive to the value of K . Furthermore, the results achieved when $K=2$ are available upon request.

Tables 4 and 5: The tables provide the results of Wald tests aiming at examining the validity of the restrictions imposed to systems (4a) and (4b). Each block shows the results achieved when comparing systems (4b) (unrestricted) and (5) (restricted), (4a) (unrestricted) and (6) (restricted) and (4b) (unrestricted) and (8) (restricted). The sample comprises observations taken from January 1983 until March 2008. Tables 4a, 4b and 4c (5a, 5b and 5c) display the values achieved by the test statistics when applying the correction proposed by White (Newey-West). The number of lags K equals 2

Equation for...		Equation for...	
$\beta_{1,t}$	$\beta_{2,t}$	$\beta_{1,t}$	$\beta_{2,t}$
4a	16.573	27.515	23.504
	0.001	0.000	0.000
	(4b) x (5)		
4b	25.585	14.552	22.790
	0.000	0.024	0.001
	(4b) x (8)		
4c	π_t	x_t	r_t
	8.040	23.330	41.823
	0.235	0.001	0.000
	(4a) x (6)		
Equation for...		Equation for...	
$\beta_{1,t}$	$\beta_{2,t}$	$\beta_{1,t}$	$\beta_{2,t}$
5a	15.645	31.864	19.917
	0.001	0.000	0.000
	(4b) x (5)		
5b	32.871	17.706	17.915
	0.000	0.007	0.006
	(4b) x (8)		
5c	π_t	x_t	r_t
	13.930	22.137	41.714
	0.030	0.001	0.000
	(4a) x (6)		

WHITE

NEWHEY-WEST

Table 5: Level

		Errors evaluated from 1995:07 to 2008:03			
		1 step	6 steps	12 steps	18 steps
1st Group CPILFESL INDPRO FF	(b)	0.407	2.561	4.922	7.126
	(a)	0.420	2.852	5.747	8.525
		-2.904	-5.980	-7.309	-8.619
	(4a), (4b)	0.427	2.848	4.910	6.205
		-1.567	-1.700	0.039	2.322
	(4a), (5)	0.427	2.863	4.961	6.193
		-1.551	-1.768	-0.129	2.318
	(6), (4b)	0.420	2.650	4.411	5.829
		-1.427	-0.719	1.904	2.821
	(6), (4b) (*)	0.472	3.177	6.474	8.158
		-2.027	-2.075	-2.610	-1.167
	(6), (8)	0.451	3.668	8.087	12.847
		-2.735	-6.209	-7.484	-7.906
	(6), (8) (*)	0.460	3.194	7.471	11.707
		-2.196	-3.512	-6.484	-6.115
(4a), (8)	0.444	3.496	7.761	12.284	
	-3.505	-7.083	-9.306	-9.390	
	0.407	2.561	4.411	5.829	
2nd Group INFL_PC1 PROD_PC1 FF	(b)	0.407	2.561	4.922	7.126
	(a)	0.420	2.852	5.747	8.525
		-2.904	-5.980	-7.309	-8.619
	(4a), (4b)	0.435	3.232	6.611	9.603
		-1.410	-3.227	-3.564	-4.669
	(4a), (5)	0.435	3.252	6.657	9.669
		-1.390	-3.279	-3.611	-4.745
	(6), (4b)	0.439	3.311	6.913	10.387
		-1.342	-3.180	-3.787	-5.061
	(6), (4b) (*)	0.497	3.422	7.527	11.240
		-2.941	-2.498	-3.035	-2.914
	(6), (8)	0.444	3.683	8.695	14.261
		-3.027	-6.756	-10.102	-9.984
	(6), (8) (*)	0.460	3.194	7.471	11.707
		-2.196	-3.512	-6.484	-6.115
(4a), (8)	0.438	3.483	8.015	12.771	
	-3.109	-7.058	-10.397	-9.230	
	0.407	2.561	4.922	7.126	

Tables 6 and 7: These tables provide the mean squared differences between actual and forecasted values of the level, slope and curvature factors, together with the values of the Diebold and Mariano test statistics, calculated by taking model (b) as the benchmark. Each table is divided into 2 blocks, one for each group of inflation and economic activity measures. Out-of-sample errors are calculated using a sub-sample starting at July 1995 and ending at March 2003. The last line of each table informs the smallest mean squared difference that was obtained for each forecasting horizon. Results achieved by the best model for each forecasting horizon are highlighted. Diebold and Mariano test statistics indicating significant differences in forecasting performance are also highlighted.

		Errors evaluated from 1995:07 to 2008:03				
		1 step	6 steps	12 steps	18 steps	
	(b)	0.881	10.345	26.434	40.864	
	(a)	0.793	8.150	25.379	48.973	
		1.250	2.668	0.509	-2.570	
	(4a), (4b)	0.793	7.427	19.963	33.634	
		1.327	3.214	3.093	2.761	
	1st Group CPILFESL INDPRO FF	(4a), (5)	0.789	7.411	20.065	33.868
			1.323	3.096	2.929	2.621
		(6), (4b)	0.803	8.472	25.245	41.030
		1.229	1.972	0.546	-0.056	
		(6), (4b) (*)	0.709	5.348	10.705	13.449
		2.835	5.478	5.719	6.881	
(6), (8)		0.786	7.986	25.477	44.641	
		1.527	2.589	0.433	-1.232	
(6), (8) (*)		0.677	4.685	10.385	15.679	
		3.173	5.831	5.736	6.485	
(4a), (8)		0.780	7.261	21.510	38.377	
		1.450	3.449	2.261	0.891	
		0.677	4.685	10.385	13.449	
		(b)	0.881	10.345	26.434	40.864
	(a)	0.793	8.150	25.379	48.973	
		1.250	2.668	0.509	-2.570	
	(4a), (4b)	0.799	7.252	18.445	29.655	
		1.054	2.894	2.932	3.223	
	2nd Group INFL_PC1 PROD_PC1 FF	(4a), (5)	0.805	7.383	18.647	29.873
			0.955	2.702	2.796	3.129
		(6), (4b)	0.803	7.760	21.126	34.157
		1.051	2.416	1.993	1.768	
		(6), (4b) (*)	0.734	5.329	11.108	15.791
		1.842	4.456	4.913	5.861	
(6), (8)		0.763	6.680	19.360	34.175	
		1.849	4.198	3.141	2.102	
(6), (8) (*)		0.677	4.685	10.385	15.679	
		3.173	5.831	5.736	6.485	
(4a), (8)		0.771	6.698	18.775	33.195	
		1.564	4.007	3.307	2.512	
		0.677	4.685	10.385	15.679	

Table 6: Slope

		Errors evaluated from 1995:07 to 2008:03			
		1 step	6 steps	12 steps	18 steps
	(b)	3.425	27.386	51.665	76.831
	(a)	3.474	29.602	57.780	87.008
		-1.685	-3.180	-4.368	-5.674
	(4a), (4b)	3.494	29.410	56.945	89.638
		-0.749	-2.611	-3.545	-5.984
1st Group CPILFESL INDPRO FF	(4a), (5)	3.497	29.458	56.921	89.650
		-0.767	-2.644	-3.455	-5.799
	(6), (4b)	3.739	30.094	51.256	75.866
	-2.105	-1.629	0.120	0.263	

	(6), (4b) (*)	3.661	23.058	35.406	44.063
		-1.065	3.145	6.174	10.017
	(6), (8)	3.676	28.980	48.446	70.184
		-2.016	-0.980	0.954	1.827
	(6), (8) (*)	3.566	23.504	36.751	46.242
		-0.793	3.572	6.568	10.126
	(4a), (8)	3.471	28.265	52.441	78.917
		-0.492	-1.135	-0.546	-1.171
		3.425	23.058	35.406	44.063
	(b)	3.425	27.386	51.665	76.831
	(a)	3.474	29.602	57.780	87.008
		-1.685	-3.180	-4.368	-5.674
	(4a), (4b)	3.344	25.222	44.369	68.572
		0.661	1.836	3.279	3.375
2nd Group	(4a), (5)	3.346	25.231	44.348	68.285
		0.645	1.818	3.262	3.441
INFL_PC1	(6), (4b)	3.476	24.586	39.836	61.839
PROD_PC1		-0.342	1.383	3.059	3.329
FF	(6), (4b) (*)	3.589	20.473	28.816	34.388
		-0.698	3.925	5.985	8.898
	(6), (8)	3.445	24.676	41.264	63.092
		-0.184	1.634	3.203	3.726
	(6), (8) (*)	3.566	23.504	36.751	46.242
		-0.793	3.572	6.568	10.126
	(4a), (8)	3.335	25.623	46.728	70.144
		0.898	1.838	2.654	3.055
		3.335	20.473	28.816	34.388

Table 7: Curvature

clear in inflation targeting regimes), and VAR models like (6) are usually mute about the way monetary authorities “adjust” these expectations.

Furthermore, the recognition that only persistent shocks are capable of affecting the level factor is implicit in the macro-finance literature. This point is made clear in Dewachter and Lyrio (2006), who build a model in which inflation π_t , output gap y_t and the instantaneous real interest rate ρ_t wander around long-run macroeconomic attractors (or long-run expectations) π_t^* , y_t^* and ρ_t^* . Assuming that π_t^* and y_t^* are random walks and that ρ_t^* is an affine function of π_t^* and y_t^* , the authors show that inflation expectations are crucial for long-term maturities, while actual macro variables are important for short-term maturities. As they write, “... *this macroeconomic decomposition allows us to interpret the standard level, slope and curvature factors typically found in the standard finance literature. We find the level factor to be closely linked to the long-run inflation expectation, the slope factor to be an aggregate series for the business cycle condition, and the curvature factor to be related to the monetary stance of the central bank*”.

These arguments suggest that there is no advantage in incorporating information about π_t , x_t

and r_t that comes from a VAR like (6), or any other model designed to describe mainly the short-run (or business cycle) evolution of the economy. If the macro block does not incorporate information about some other key macroeconomic variables (like the steady-state real interest rate and inflation expectations) and is unable to describe the response of the economy to changes in its fundamentals (which are important in determining long-term yields and, consequently, the level factor), then poor forecasting results for the level factor will inevitably follow. In fact, the task of improving the quality of level forecasts demands a model capable of characterizing the evolution of variables that are important to determine these key variables.

The same conclusion does not emerge if one is interested in forecasting the future evolution of the slope factor, since models (a) and (b) never appear among the group of best predictors. Note that this result does not depend on the proxies chosen for π_t , x_t and r_t and the forecasting horizon. This result cannot come as a surprise because the macro-finance literature has always revealed a strong connection between monetary impulses and the slope factor (see Evans and Marshall, 1998 and Wu, 2001).

If one is interested in forecasting the future path of the curvature factor, then the advantage of using the

information content of macro variables is also clear. Note that model (b) appears only 1 time (out of 8) among the group of best predictors. This result is only slightly worsened if one excludes the versions of models (6),(4b) and (6),(8) that have the true values of macro variables as their inputs because, in this case, model (b) appears 2 times among the group of best predictors. Now the conclusion seems to depend on the forecasting horizon, because model (b) beats all the others in one-month-ahead forecasts, and on the proxies chosen for π_t , x_t and r_t , because model (b) is excluded from the group of best predictors when we use the first principal components of four individual measures of annual inflation and six individual measures of economic activity to proxy for π_t and x_t .

It is also worth discussing the statistical significance of the difference between the mean squared errors achieved by the benchmark model and the others in forecasting exercises of the level factor. First, the differences between the mean squared errors computed for model (b) and its competing models are frequently statistically significant at the 5% level. Unfortunately, the differences are always negative and indicate a strong superiority of model (b), which is a simple univariate model²³. This conclusion seems to be independent of the proxies chosen for the macro variables (first or second groups) and the forecasting horizon ($\hat{T} = 1, 6, 12$ and 18).

The same conclusion does not prevail in forecasting exercises of the slope factor. Note that the differences between the mean squared errors computed for the benchmark and its competitors are frequently positive and statistically significant at the 5% level, indicating a strong inferiority of model (b). Again, this conclusion seems to be independent of the proxies chosen for macro variables, but now the forecasting horizon appears to be important. More specifically, the differences seem to be unfavorable in one-month-ahead forecasting exercises.

A different picture emerges when forecasting the future path of the curvature factor. This time results seem to depend on the proxies chosen for macro variables and the forecasting horizon. The differences between the mean squared errors computed for the benchmark and its competitors are frequently slightly positive, or even negative, in one and six-month-ahead forecasts. This situation changes for twelve and

eighteen-month-ahead forecasts, when differences are frequently positive and statistically significant at the 5% level, especially when macro variables are represented by the first principal components of the two groups of inflation and economic activity measures. This result suggests that improving the quality of short rate forecasts is critical for improving the quality of curvature forecasts, and that this improvement is easier if one does not rely on specific measures of macro variables.

Results also show that substituting forecasts of π_t , x_t and r_t by their true values dramatically improves forecasts of slope and curvature movements, although there is no positive effect in forecasting the level factor. This can be seen by comparing the forecasting performance of models (6), (4b) and (6), (8) with their counterparts marked with a (*). This indicates that there is plenty of room for improvements if one substitutes the VAR model described in (6) by other (perhaps more sophisticated) macroeconomic model, which is potentially capable of generating better forecasts of macro variables.

Finally, Tables 5, 6 and 7 also show a remarkable advantage in using the information content of macro variables when forecasting slope and curvature movements, although no specific model following this strategy seems to beat the others. Furthermore, there is no indication that imposing extra restrictions to diminish the number of coefficients is always capable of improving forecasting performance. For example, model (4a),(4b) (which is the least restricted version of model (3)) beats all others if one is interested in calculating twelve and eighteen-month-ahead forecasts of the slope factor. It is worth stressing that the role of macro variables as important sources of information is also discussed in Ang and Piazzesi (2003) and Hördahl, Tristani and Vestin (2006), but these authors focus on forecasting yields directly. The great advantage of our methodology is revealing what is in the heart of this predictive power, which is the ability of improving forecasts of the slope and curvature factors.

The results above suggest that the task of forecasting the future path of the level factor is not simple at all. In fact, it seems that calculating better forecasts of the level factor by means of models in which the links between macro variables and factors are made explicit is not straightforward, especially when models are unable to capture the relationship between the state of the economy and long-run trends and expectations. This result is somewhat unpleasant,

²³In fact, this univariate model behaves almost as a random walk because its autoregressive coefficient is near one.

since the task of predicting the future values of the level factor is crucial for generating good forecasts of the entire range of yields (recall that the level factor is by far the most important one) and we are not able to beat simple autoregressive models, or VAR models in which factors are modeled alone, in most of the cases. On the other hand, the hopes of refining the forecasts of the slope and curvature factors by incorporating macroeconomic information coming from simple VAR models (or any other model capable of characterizing the dynamics of macro variables throughout the business cycle) seem to be well justified; it is a pity that this is far from being enough.

5. CONCLUSION

In this paper we develop a number of models that have the common characteristic of establishing a connection between macroeconomic variables and the term structure of interest rates in a VAR framework. The models are designed to describe the dynamics of three macro variables (inflation, economic activity and the overnight rate) and three term structure factors (level, slope and curvature). These factors are obtained by means of the dynamic version of the Nelson and Siegel representation, which was firstly proposed by Diebold and Li.

The models are designed to respect theoretical relationships between macro and term structure variables. The obedience to economic theory allows imposing additional restrictions to otherwise complex versions in order to save parameters and reinforce the ability of generating good forecasts. The models are then used to generate out-of-sample forecasts of level, slope and curvature factors for four different forecasting horizons (one, six, twelve and eighteen months). Errors are computed in order to calculate the mean squared difference between actual and forecasted factors, which is the criterion used to judge forecasting performance. A model is considered good if its forecasts are better (on average) than those computed by means of other simple forecasting schemes that have little economic content. More specifically, the benchmarks are a VAR model comprising current and past values of term structure factors and univariate autoregressive models in which factors are modeled alone. We also discuss the statistical significance of the differences between mean squared errors, that is, we compute the value assumed by the Diebold-Mariano test statistics by taking simple univariate autoregressive models as benchmarks.

We also consider two models in which the true values of macro variables serve as inputs of the system describing the dynamics of term structure factors. This is possible because, in these models, the macro block does not depend on term structure variables. The exercise allows evaluating the potential advantages of substituting a simple VAR model by other representations of the economy. Although they are unable to reach “perfect foresight” results, the alternatives are possibly better in generating good macroeconomic forecasts, which are essential to improve the quality of term structure forecasts (especially slope and curvature, which are sensitive to short rate forecasts).

In a nutshell, the main results are as follows:

- (1) It is hard to beat simple univariate models if one is interested in forecasting the future path of the level factor. This result does not depend on the specific measures of macro variables and the forecasting horizon. This conclusion stems in the absence of other relevant macro variables, especially long-run expectations of inflation and the steady-state real interest rate. This finding is somewhat unfortunate, since the task of predicting future values of the level factor is crucial for generating good forecasts of the entire range of yields.
- (2) Things change if one is interested in forecasting the future evolution of the slope and curvature factors. This conclusion seems to be independent of the proxies effectively chosen to represent macro variables, but now the forecasting horizon is an issue if one is interested in forecasting curvature movements. It is worth mentioning that better forecasts of the overnight rate are critical for improving the quality of curvature forecasts. Improvements seem to be easier if one does not rely on specific measures of macro variables, but this assertive demands further research to be confirmed. Finally, even simple macroeconomic models (like the VARs used here) can add macroeconomic information capable of improving the quality of slope and curvature forecasts. Further improvements can definitely be achieved by substituting these simple macro models by more sophisticated ones, which are probably capable of generating better forecasts of macro variables.
- (3) Last but not least, the role of macro variables as important sources of information for forecasting

purposes is confirmed. Our methodology allows revealing what is the heart of this success, which is the ability of achieving better forecasts of the slope and curvature factors.

ACKNOWLEDGEMENTS

We are grateful to the Referee for the invaluable comments that helped us in achieving an improved final version of the paper.

APPENDIX A

This Appendix discusses the validity of the restrictions imposed to the model comprised by (4a) and (4b). Taking this model as the benchmark for comparisons (because it is the least restricted version of the general econometric model described in (3)), we proceed by executing a number of Wald tests to check if the restrictions imposed to the reference model are confirmed or not by the data. More specifically, we make comparisons between:

- 1) Systems (4b) and (5), to check if the hypothesis that there are no contemporaneous links between macro variables and term structure factors is true or false.
- 2) Systems (4a) and (6), to check if the hypothesis that term structure factors are unable to affect macro variables is true or false (i.e. we repeat the exercise found in Diebold, Rudebusch and Aruoba, 2006).
- 3) Systems (4b) and (8), to check if the connection between macro variables and term structure factors is established by means of the short rate alone.

The procedure starts by examining the residuals obtained after estimating each equation belonging to (4a) and (4b) by OLS. If residuals are serially uncorrelated, then Wald statistics that are robust to heteroskedasticity of unknown form are calculated according to the White methodology. If residuals display serial correlation, then Wald statistics that are robust to both heteroskedasticity and autocorrelation of unknown forms are calculated according to the Newey-West methodology.

Tables 4a, 4b, 4c, 5a, 5b and 5c show the results. Tables 4a and 5a inform the values of the test statistics (calculated according to White and Newey-West methodologies, respectively) when evaluating the

hypothesis that there is no contemporaneous effect of macro variables on term structure factors (that is, we compare systems (4b) and (5) equation by equation). Tables 4b and 5b display the same information regarding the hypothesis that term structure factors are unable to affect macro variables (that is, we compare systems (4a) and (6) equation by equation). Finally, Tables 4c and 5c show the same information regarding the hypothesis that connections between term structure factors and macro variables are made by the short rate alone (that is, we compare systems (4b) and (8)). Each table shows the value of the test statistics (that follows a chi-square distribution with N degrees of freedom, where N equals the number of restrictions imposed to the unrestricted versions), and the p-value. Unrestricted equations are estimated using a sample starting in January 1985 and ending in March 2008. The measures of inflation and economic activity are taken from the second group. Tests are executed by setting $K = 2$. One can see that restricted versions are strongly rejected by the data in almost all the cases, the exception being the impossibility of rejecting the null that terms related to term structure factors are not important in the inflation specification. We consider this result spurious because there is evidence of serially correlated residuals in the inflation equation. If one applies the Newey-West methodology instead of the method proposed by White, then the null is also rejected. It is worth noting that we confirm the results achieved by Diebold, Rudebusch and Aruoba, who find that term structure factors do influence the dynamics of inflation, economic activity and the short rate.

REFERENCES

- Ang, A. and Bekaert, G., 2002, Regime switches in interest rates, *Journal of Business & Economic Statistics*, vol. 20(2), 163-182.
<http://dx.doi.org/10.1198/073500102317351930>
- Ang, A. and Piazzesi, M., 2003, A no-Arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, *Journal of Monetary Economics*, vol. 50, no. 4, 745-787.
[http://dx.doi.org/10.1016/S0304-3932\(03\)00032-1](http://dx.doi.org/10.1016/S0304-3932(03)00032-1)
- Ang, A., Piazzesi, M. and Wei, M., 2006, What does the yield curve tell us about GDP growth?, *Journal of Econometrics*, 131, 359-403.
<http://dx.doi.org/10.1016/j.jeconom.2005.01.032>
- Ang, A., Bekaert, G. and Wei, M., 2008, The term structure of real rates and expected inflation, *The Journal of Finance*, vol. 63(2), 797-849.
<http://dx.doi.org/10.1111/j.1540-6261.2008.01332.x>
- Bernanke, B., Gertler, M. and Watson, M., 1997, Systematic monetary policy and the effects of oil price shocks, Working Paper 97-25, C.V. Starr Center for Applied Economics, New York University.
- Campbell, J., 1995, Some lessons from the yield curve, *Journal of Economic Perspectives*, vol. 9(3), 129-152.
<http://dx.doi.org/10.1257/jep.9.3.129>

- Campbell, J. and Shiller, R., 1991, Yield spreads and interest rate movements: a bird's eye view, *Review of Economic Studies*, vol. 58(3), 495-514.
<http://dx.doi.org/10.2307/2298008>
- Chan, K., Karolyi, A., Longstaff, F. and Sanders, A., 1992, An empirical comparison of alternative models of the term structure of interest rates, *The Journal of Finance*, vol. 47(3), 1209-1227.
<http://dx.doi.org/10.1111/j.1540-6261.1992.tb04011.x>
- Christiansen, C. and Lund, J., 2005, Revisiting the shape of the yield curve: the effect of interest rate volatility, EFA 2002 Berlin Meetings Presented Paper.
- Clarida, R., Galí, J. and Gertler, M., 1999, The science of monetary policy: a new Keynesian perspective, *Journal of Economic Literature*, vol. 37(4), 1661-1707.
<http://dx.doi.org/10.1257/jel.37.4.1661>
- Dewachter, H. and Lyrio, M., 2006, Macro factors and the term structure of interest rates, *Journal of Money, Credit, and Banking*, vol. 38 (1), 119 -140.
<http://dx.doi.org/10.1353/mcb.2006.0014>
- De Pooter, M., 2007, Examining the Nelson-Siegel class of term structure models, Tinbergen Institute Discussion Papers 07-043/4.
- Diebold, F. and Li, C., 2006, Forecasting the term structure of government bond yields, *Journal of Econometrics*, vol. 130, 337-364.
<http://dx.doi.org/10.1016/j.jeconom.2005.03.005>
- Francis Diebold, F., Li, C. and Yue, V., 2008, Global yield curve dynamics and interactions: a generalized Nelson and Siegel approach, *Journal of Econometrics*, 146, 351-363.
<http://dx.doi.org/10.1016/j.jeconom.2008.08.017>
- Diebold, F. and Mariano, R., 1995, Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13, 253-263.
- Diebold, F., Rudebusch, G. and Aruoba, S., 2006, The macroeconomy and the yield curve: a dynamic latent factor approach, *Journal of Econometrics*, vol. 127(1-2), 309-338.
<http://dx.doi.org/10.1016/j.jeconom.2005.01.011>
- Duffee, G., 2002, Term premia and interest rate forecasts in affine models, *Journal of Finance*, 57, 405-443.
<http://dx.doi.org/10.1111/1540-6261.00426>
- Engle, R. and Ng, V., 1993, Time-varying volatility and the dynamic behavior of the term structure, *Journal of Money, Credit and Banking*, vol. 25(3), 336-349.
<http://dx.doi.org/10.2307/2077766>
- Estrella, A., 2005, Why does the yield curve predict output and inflation?, *Economic Journal*, vol. 115, 722-744.
<http://dx.doi.org/10.1111/j.1468-0297.2005.01017.x>
- Estrella, A. and Hardouvelis, G., 1991, The term structure as a predictor of real economic activity, *Journal of Finance*, 46, 555-576.
<http://dx.doi.org/10.1111/j.1540-6261.1991.tb02674.x>
- Estrella, A. and Mishkin, F., 1997, The predictive power of the term structure of interest rates in Europe and the United States: implications for the European Central Bank, *European Economic Review*, 41, 1375-1401.
[http://dx.doi.org/10.1016/S0014-2921\(96\)00050-5](http://dx.doi.org/10.1016/S0014-2921(96)00050-5)
- Evans, C. and Marshall, D., 1998, Monetary policy and the term structure of nominal interest rates: evidence and theory, *Carnegie-Rochester Conference Series on Public Policy*, vol. 49, 53-111.
[http://dx.doi.org/10.1016/S0167-2231\(99\)00004-4](http://dx.doi.org/10.1016/S0167-2231(99)00004-4)
- Evans, C. and Marshall, D., 2002, Economic determinants of the nominal treasury yield curve, Working Paper 2001-16, Federal Reserve Bank of Chicago.
- Hall, A., Anderson, H. and Granger, C., 1992, A cointegration analysis of treasury bill yields, *Review of Economics and Statistics*, vol. 74, no. 1, 116-126.
<http://dx.doi.org/10.2307/2109549>
- Hamilton, J., 1988, Rational-expectations econometric analysis of changes in regime: an investigation of the term structure of interest rates, *Journal of Economic Dynamics and Control* 12, 385-423.
[http://dx.doi.org/10.1016/0165-1889\(88\)90047-4](http://dx.doi.org/10.1016/0165-1889(88)90047-4)
- Hördahl, P., Tristani, O. and Vestin, D., 2006, A joint econometric model of macroeconomic and term structure dynamics, *Journal of Econometrics*, vol. 127, 405-444.
<http://dx.doi.org/10.1016/j.jeconom.2005.01.012>
- Knez, P., Litterman, R. and Scheinkman, J. 1994, Explorations into factors explaining money market returns, *The Journal of Finance*, vol. 49(5), 1861-1882.
<http://dx.doi.org/10.1111/j.1540-6261.1994.tb04784.x>
- Kozicki, S., 1997, Predicting real growth and inflation with the yield spread, *Federal Reserve Bank of Kansas City Economic Review*, 82, 39-57.
- Litterman, R. and Scheinkman, J., 1991, Common factors affecting bond returns, *Journal of Fixed Income*, vol. 1, 44-51.
<http://dx.doi.org/10.3905/jfi.1991.692347>
- Mishkin, F., 1998, The information in the term structure: some further results, *Journal of Applied Econometrics*, Volume 3, Issue 4, 307-314.
- Nelson, C. R. and Siegel, A. F. (1987). Parsimonious Modeling of Yield Curves. *The Journal of Business*, vol. 60(4), 473-489.
- Sims, C., Stock, J. and Watson, M., 1990, Inference in linear time series models with some unit roots, *Econometrica*, volume 58, issue 1, 113-144.
<http://dx.doi.org/10.2307/2938337>
- Sims, C. and Uhlig, H., 1991, Understanding unit rooters: a helicopter tour, *Econometrica*, volume 59, issue 6, 1591-1599.
<http://dx.doi.org/10.2307/2938280>
- Stock, J. and Watson, M., 2000, Forecasting output and inflation: the role of asset prices, manuscript, Kennedy School of Government (revised January 2003).
- Svensson, L., 1994, Estimating and interpreting forward interest rates: Sweden 1992-1994, IMF Working Paper 94/114, available at SSRN: <http://ssrn.com/abstract=883856>
- Taylor, J., 1993, Discretion versus policy rules in practice, *Carnegie-Rochester Conference Series on Public Policy*, 39, 195-214.
[http://dx.doi.org/10.1016/0167-2231\(93\)90009-L](http://dx.doi.org/10.1016/0167-2231(93)90009-L)
- Woodford, M., 2003, *Interest and prices*, Princeton University Press.
- Wu, T., 2001, Stylized facts on nominal term structure and business cycles: an empirical VAR approach, Working Paper 2002-08, Federal Reserve Bank of San Francisco.
- Zellner, A. (1992). *Statistics Science and Public Policy* (ASA Presidential Address). *Journal of the American Statistical Association*, 87, 1-6.

Received on 12-05-2014

Accepted on 18-09-2014

Published on 24-10-2014

DOI: <http://dx.doi.org/10.6000/1929-7092.2014.03.29>© 2014 Vereda *et al.*; Licensee Lifescience Global.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.