Journal of Banking & Finance 35 (2011) 3240–3252

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Term structure modelling with observable state variables

Cristian Huse *

Stockholm School of Economics, Department of Finance, PO Box 6501, SE-113 83 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 29 November 2010
Accepted 7 May 2011
Available online 13 May 2011

JEL classification:
C58
G12
G17

Keywords:
Term structure
Interest rates
Yield curve estimation

ABSTRACT

This paper proposes and implements a parsimonious three-factor model of the term structure whose dynamics is driven uniquely by observable state variables. This approach allows comparing alternative views on the way state variables – macroeconomic variables, in particular – influence the yield curve dynamics, avoids curse of dimensionality problems, and provides more reliable inference by using both the cross-sectional and the time series dimension of the data. I simulate the small-sample properties of the procedure and conduct in- and out-of-sample studies using a comprehensive set of US data. I show that even a parsimonious model where the level, slope and curvature factors of the term structure are driven by, respectively, inflation, monetary policy and economic activity consistently outperforms the (latent-variable) benchmark model in an out-of-sample study.

1. Introduction

The interaction between the term structure of interest rates and macroeconomic variables has been extensively explored in a number of papers in the last decade or so. This momentum in the literature is partly due to the fact that 'yields-only' models based on no-arbitrage were found to do well in fitting the cross-section of yields at a particular point in time (Dai and Singleton, 2000), but poorly in describing the dynamics of the yield curve (Duffee, 2002). Contributions to the literature include Fuhrer and Moore (1995), Rudebusch (1995, 2002), Evans and Marshall (1996), Fuhrer (1996), Ang and Piazzesi (2003), Bikbov and Chernov (2010), Piazzesi (2005), Dewachter and Lyrio (2006), Hördahl et al. (2006), and Rudebusch and Wu (2008).

This paper contributes to this literature by implementing a parsimonious three-factor model of the term structure whose dynamics is driven uniquely by observable state variables – macroeconomic variables in particular-, as opposed to latent variables. It builds upon a three-factor model describing the term structure behaviour first proposed in Nelson and Siegel (1987, NS hereafter) and recently reinterpreted by Diebold and Li (2006, DL hereafter) as a dynamic latent factor model.

The term structure literature can be divided into three main streams. The first one relies on the optimizing behaviour of economic agents, which can thus be cast within the dynamic stochastic general equilibrium (DSGE) framework. Early contributions to the literature of equilibrium pricing include Cox et al. (1985), Campbell (1986) and Dunn and Singleton (1986). The recursive preferences of Epstein and Zin (1989) and Weil (1989) have also been used extensively, as in Campbell (1993, 1996, 1999) and more recently, Piazzesi and Schneider (2006). Currently, this approach still needs unreasonable assumptions about risk aversion and the elasticity of intertemporal substitution (EIS) to deliver satisfactory results. Namely, the risk aversion parameter typically needs to be substantially high, in what Mehra and Prescott (1985) called the equity premium puzzle. Even if Bansal and Yaron (2004) were recently able to match the equity premium using recursive preferences and a plausible value for the risk aversion parameter, this in turn crucially relies on the high persistence of consumption shocks. As for the EIS, even its magnitude is subject to controversy, with Bansal et al. (2007) and Vissing-Jorgensen and Attanasio (2003) estimating it to be above one and Hall (1988) and Campbell (1999) documenting it to be close to zero.

A second stream of the literature adopts only the very basic structure of the DSGE approach, namely the absence of arbitrage opportunities, to study the term structure of interest rates, usually imposing no-arbitrage when estimating a Vector Autoregressive (VAR) model of yields – see, for instance, Ang and Piazzesi (2003), Piazzesi (2005), Bikbov and Chernov (2010). In practice, one needs to, first, specify the instantaneous interest rate and the prices of risk for the factors assumed to affect the yield curve as functions of state variables (such as economic activity and inflation); second, focus on the asset pricing implications of the structure imposed. Although

* Tel.: +46 8 736 9167; fax: +46 8 31 23 27.
E-mail address: cristian.huse@sse.edu

0378-4266/$ - see front matter
doi:10.1016/j.jbankfin.2011.05.004

not usually explored (or even made explicit) in the literature, for each choice of functional form above, in equilibrium, there should be an economy consistent with these choices. The main (empirical) lesson from this stream of literature is that one needs a combination of (observed) state variables plus an unobserved factor to explain the dynamics of the term structure of interest rates.

This paper, however, belongs to a third strand of the literature, going back to NS and DL, in that it decomposes the yield curve into three factors, namely level, slope and curvature, and relates them to observable state variables to forecast the term structure of interest rates without imposing no-arbitrage. The relevance of no-arbitrage for forecasting purposes has been actively studied in recent years: while Duffee (2008) finds that the restrictions on a VAR implied by an arbitrage-free Gaussian dynamic term structure model cannot be rejected against the alternative of an unrestricted VAR, Joslin et al. (2011) show that no-arbitrage is irrelevant for forecasting – in particular, their result implies that forecasts using an arbitrage-free NS model, as in Christensen et al. (2009, 2010), are equivalent to unconstrained VAR forecasts.

The intuition of my modelling strategy is as follows. If the term structure moves as a result of changes in the economic fundamentals, the term structure factors (and, by consequence, the term structure dynamics) should be somehow linked to these state variables. In this paper, I make this link explicit, so that the movements of the term structure are completely exerted by the underlying state variables.

In a number of related studies, observable and latent factors coexist - see for instance Ang and Bekaert (2004), Rudebusch and Wu (2004), Hördahl et al. (2006), Ang et al. (2005). In parallel, however, papers such Ang, Piazzesi, and Wei (2004) and Bekaert, Cho, and Moreno (2004) have focused solely on observable state variables. This approach has a number of advantages. First, the replacement of latent factors with observable state variables as the only drivers of the term structure allows comparing alternative views on the way state variables influence the yield curve dynamics. Besides telling more about the economic fundamentals than latent variables, the use of observable variables might also provide guidance to the construction of theoretical models of the term structure dynamics. Moreover, the explicit link between term structure factors and observable state variables allows performing policy experiments. As a result, one can forecast the term structure by using forecasted variables, or perform stress testing of the term structure using scenarios constructed using state variables. This feature is especially useful to bankers, who are interested in forecasting bond prices and might have a better idea of the expected state of the economy than the expected state of the yield curve. Moreover, it is also of value to financial authorities, as a tool to assess financial stability: for instance, stress testing has become crucial in the risk management toolbox of financial institutions to gauge their potential vulnerability to exceptional – but plausible – events, and has even been highlighted during the Basel II process as a useful tool in assessing banks' internal risk models.

A second advantage of the method here proposed is its robustness to curse of dimensionality problems commonly appearing in traditional models. The curse of dimensionality imposes constraints on the number of yields one can use, which results in especially poor measures of the term structure curvature: while one needs at least three yields for the curvature to be defined, it is not uncommon for researchers to use as few as five yields, making it unlikely to obtain accurate measures of this factor. Here, instead, the dimension of the parameter vector does not increase with the number of yields under study, but with the number of state variables explaining them; this is very much in the spirit of linear regression, where one loses degrees of freedom by including additional covariates, not more observations.

Third, the identification strategy comes out in a natural way. Essentially, the baseline model needs the state variables driving the term structure to be predetermined with respect to yields. When I incorporate a Taylor rule into the model, the identifying assumption made is also standard, requiring the state variables to be predetermined with respect to the monetary policy instrument.

I conduct in- and out-of-sample studies using US data. The in-sample study uses a comprehensive set of macroeconomic variables to compare alternative specifications of the term structure dynamics and suggests two models which I then use in the out-of-sample exercise. The first one is a parsimonious model where the level, slope and curvature factors of the term structure are driven by, respectively, measures of inflation growth, monetary policy, and economic activity (these are, respectively, the Consumer Price Index growth rate, the Fed Funds rate, and the Unemployment Rate). The second specification is a richer one where the level is driven by measures of inflation growth and economic activity, the slope by monetary policy and economic activity, and the curvature by fiscal policy growth. The out-of-sample study shows that both specifications consistently outperform the (latent-variable) benchmark model in the study of the yield curve behaviour during the five NBER-dated recessions which occurred in the last decades. Recessions are of interest not only for being bad states against which economic agents are willing to insure, but also for being periods which tend to be preceded by the inversion of the term structure of interest rates, a feature usually difficult to be quickly captured, if at all, by term structure models.

The paper is organized as follows. Section 2 briefly reviews term structure estimation methods and recent developments of the NS approach. Section 3 presents the model and discusses its identification and implementation. Section 4 presents simulation results, Section 5 performs an empirical exercise using US data and Section 6 concludes.

2. Yield curve estimation

2.1. Static methods

A number of approaches can be used to model the term structure of interest rates. First, one may consider models that make explicit assumptions about the evolution of state variables and use either equilibrium or arbitrage methods, which corresponds to modelling dynamic yield curves. According to this class of models, the evolution of the yield curve is modelled as depending on a small number of (arbitrarily chosen) factors. Since in most of the cases the underlying state variable is the short term interest rate, as in Vasicek (1977) and Cox et al. (1985, CIR), they are frequently labeled as 'short-rate models'. Subsequent extensions to multi-factor models include the two-factor model of Longstaff and Schwartz (1992), and the three-factor one of Balduzzi et al. (1996, BDFS).

When it comes to fitting real data, one-factor models perform poorly: the yield curve corresponding to the Vasicek model does not allow a large range of shapes, whereas the ones corresponding to CIR and extensions allowing time-varying parameters such as Hull and White (1990) tend to evolve unrealistically over time. In what regards multi-factor models, there is an understanding that at least three factors are needed to generate a wide variety of yield curve shapes, although even so the fit close to the long end tends to be poor. On the other hand, Joslin et al. (2011) find evidence that the four- and five-factor models examined in Cochrane and Piazzesi (2008) and Duffee (2011) may be over-parameterized. Moreover, choosing the state variables involves both a certain degree of arbitrariness and a bit of art – direct factors may include the short rate, spot rates of various maturities, forward rates, swap rates, whereas indirect ones may include the short rate volatility, the mean short rate, the latter two rising issues such as the choice of the sample period involved in their calculations. Further, multi-factor models (such as the BDFS) usually lack of explicit formulae and are of difficult calibration to market prices.

Alternatively, one can instead smooth data obtained from asset prices to describe the static yield curve, usually without taking a view on the factors driving it. This corresponds to fitting, the yield curve as a whole. The analysis starts from information on asset prices, from which one extracts the corresponding yields. As there are only a few maturities available for which there are observations on prices (and, thus, yields), it is interesting to somehow 'connect' those points in order to evaluate instruments with maturities different from those of the yields one has already extracted, usually imposing some degree of smoothness. Among the estimation methods most widely used, there are spline techniques (see McCulloch (1971, 1975), Schaefer (1981) and Vasicek and Fong (1982)), kernel methods, but also parametric classes of curves, broadly known as the Nelson–Siegel family of curves.

The class of curves first proposed in NS is parsimonious and does well in capturing the overall shape of the yield curve, being popular among practitioners and central banks alike. Their objective is to describe the yield curve, not being consistent with the absence of arbitrage opportunities. For a sample of N bonds measured at a given point in time, the yield curve as a function of time to maturity τ_i is written as

$$y(\tau_i) = \beta_1 + \beta_2\left(\frac{1-e^{-\lambda\tau_i}}{\lambda\tau_i}\right) + \beta_3\left(\frac{1-e^{-\lambda\tau_i}}{\lambda\tau_i} - e^{-\lambda\tau_i}\right) + u(\tau_i), \quad i = 1,\ldots,N$$

providing a parsimonious representation of the term structure which is consistent with a well-behaved discount function i.e. continuous, positive and decreasing in τ, taking value 1 when $\tau = 0$ and approaching β_1 as τgrows large.[1] As we justify below, the parameters $\beta_1, \beta_2, \beta_3$ can be interpreted as, respectively, the level, (the negative of the) slope, and curvature components, whereas the parameter λ_t controls the exponential decay of the yield curve: small values produce slow decay and can better fit the curve at long maturities, while large values generate a fast decay and can better fit the curve at short maturities. Moreover, λ also determines where the loading on β_3 achieves its maximum. The loading on β_1 is a constant, implying that an increase in this factor increases all yields equally, which results in a change in the level of the yield curve. The loading on β_2 is a function that starts at 1 but decays monotonically to zero, implying that an increase in β_2 increases short yields more than long yields, resulting in a change in the slope of the yield curve. As for β_3, this is related to the curvature of the term structure, as an increase in β_3 will have little effect on very short or very long yields, but will increase medium-term yields, thus resulting in an increase of curvature of the yield curve. As first described by DL, this representation can be related to a dynamic three-factor model of, respectively, level, slope, and curvature. It is then no surprise that in parallel with the literature on short-rate models, Diebold et al. (2005) find no evidence that extensions of Nelson–Siegel using four or five factors would do better.

2.2. Nelson–Siegel and beyond

DL reinterprets the NS framework as a dynamic latent-factor model. Essentially, for a given time period t the yield curve follows (1), but since in practice one is often interested in assessing the evolution of the yield curve over time, DL propose a version with time-varying parameters, which evolve following a pre-specified stochastic process. More formally, their model can be written as

$$y_t(\tau_i) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_t\tau_i}}{\lambda_t\tau_i}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_t\tau_i}}{\lambda_t\tau_i} - e^{-\lambda_t\tau_i}\right) + u_t(\tau_i); \quad \times i = 1,\ldots,N, \quad t = 1,\ldots,T \tag{2}$$

with the beta parameters following, for instance, univariate first-order autoregressive processes, as in Diebold and Li (2006).

Estimation could in principle be carried out using Nonlinear Least Squares (NLLS) although the usual practice since NS – and also DL – has been to fix λ_t to a constant value, compute the factor loadings (regressors), and then use OLS to estimate $\{\beta_t\}$. The parameter λ_t determines the maturity τ^* at which the loading on the curvature factor achieves its maximum (usually between 2 and 3 years), and DL simply pick a λ_t such that this maximum is achieved at the midpoint between these maturities – 30 months – and set $\lambda^* = \lambda_t = 0.0609$ due to the small number of yields and the gains from using linear estimatio techniques. After computing the sequence $\{\beta_t\}$ of factors and the pricing errors, they model the factors as a univariate AR (1) models and compare the forecasting power of the model out-of-sample with a number of alternatives, with reasonable performance, especially given the simplicity of the model.

The above framework is intuitive and easy to implement. However, despite the consensus that changes in the yield curve are exerted by changes in macroeconomic conditions (or, more generally, changes in state variables), the factors in the DL framework remain latent, whereas in DRA latent and observable factors (pre-specified by the researchers) coexist. Moreover, the estimation of the stochastic processes driving the level, slope, and curvature factors does not account for the measurement error coming from the fact that $\{\beta_p\}_{p=1}^3$ are estimated rather than observed, so that any asymptotic statements are likely to be misleading.

3. Term structure modelling

This section proposes a term structure modelling approach building upon NS and its reinterpretation by DL. The main contrast with respect to DL is that here the term structure dynamics is solely driven by the dynamics of observable state variables, as opposed to latent factors. The intuition behind this idea is that if the yield curve moves as a result of changes in relevant state variables, the factors should be somehow linked to these state variables. As a result, one can now, for instance, compare alternative hypotheses on the variables driving the term structure factors and state that level, slope and curvature factors are driven by, say, measures of economic activity, inflation, and monetary policy instrument, respectively.

3.1. A model with state variables

I now propose a model according to which movements of the yield curve $y_t(\tau)$ are exerted by *predetermined* observable state variables, which I denote by M_{t-}.[2]

Let $\theta_t := (\beta_t', \lambda_t)'$ be a time-varying parameter vector determining the shape of the yield curve at time t. In what follows I decompose it as a sum of two components, the first, $\bar{\theta} := (\bar{\beta}', \bar{\lambda})'$, being a mean component, and the second being the combination of the state variables M_{t-} and parameters $\underline{\theta} := \left(\sigma_\beta', \sigma_\lambda'\right)'$ measuring their impact on the latent variables. The model then reads

$$y_t(\tau) = X_t(\lambda_t)\beta_t + u_t(\tau) \tag{3}$$

[1] Equivalently, it guarantees positive forward rates at all horizons.

[2] In what follows, given two variables A and B, I write A_{t-} if A is predetermined with respect to B within period t.

$$\begin{bmatrix}\beta_t\\ \lambda_t\end{bmatrix}=\begin{bmatrix}\bar{\beta}\\ \bar{\lambda}\end{bmatrix}+M_{t-}\begin{bmatrix}\sigma_\beta\\ \sigma_\lambda\end{bmatrix}$$

where $y_t(\tau)$ is the column vector of yields observed at date t. The nonlinearity of the model comes from the estimation of λ_t in the $N\times 3$ matrix of factor loadings at period t,

$$X_t(\lambda_t)=\begin{bmatrix}1 & \frac{1-\exp(-\lambda_t\tau_1)}{\lambda_t\tau_1} & \frac{1-\exp(-\lambda_t\tau_1)}{\lambda_t\tau_1}-\exp(-\lambda_t\tau_1)\\ 1 & \frac{1-\exp(-\lambda_t\tau_2)}{\lambda_t\tau_2} & \frac{1-\exp(-\lambda_t\tau_2)}{\lambda_t\tau_2}-\exp(-\lambda_t\tau_2)\\ \cdots & \cdots & \cdots\\ 1 & \frac{1-\exp(-\lambda_t\tau_N)}{\lambda_t\tau_N} & \frac{1-\exp(-\lambda_t\tau_N)}{\lambda_t\tau_N}-\exp(-\lambda_t\tau_N)\end{bmatrix}$$

and the error term is a martingale difference sequence with respect to current and past covariate information, and uncorrelated in the maturity domain, i.e. $E[u_t(\tau)u_t(\tau)']=\sigma^2\mathbf{I}$.[3] This specification can be seen as a random-coefficients one for the term structure factors; the factors vary over time given realizations of the state variables.[4]

Although this model is more costly to be estimated from the numerical point of view than DL in the general case where λ_t is also driven by state variables, this cost is offset by having the dynamics of $\{\beta_t\}$ driven by state variables. Moreover, there are also gains from modelling the dynamics of $\{\lambda_t\}$, apart from a pure generality argument. If the parameter β_{3t} governs the intensity of the curvature of the yield curve, the parameter λ_t governs the locus of its 'tilting point' or, alternatively, where the loading associated to the factor β_{3t} attains its maximum, thus making it unnatural to be disconnected to the analysis of the term structure curvature.

In what regards identification, the argument goes as follows. Data is observed at the monthly frequency, but recorded at different moments within a given month – the state variables M_{t-} are observed at the beginning of each month, whereas the yields are observed at the end of the corresponding month. As a result, the state variables M_{t-} are predetermined with respect to the yields.

Important features of the method are its robustness to errors in variables, its parsimony, and its robustness to the curse of dimensionality. First, as opposed to DL, where (i) the extraction of the $\{\beta_t\}$ sequence of parameters relies solely on the cross-sectional dimension of the data; (ii) the estimation of the AR (1) models for factor dynamics relies solely on the time series dimension of the data; and (iii) the estimation of the factor dynamics uses estimates of $\{\beta_t\}$ as if they were data, incurring in measurement error problems, estimation here relies on both the time series and the cross-sectional dimension of the data and is done in one step. Thus, by working on both T and N, the asymptotic results tend to be much more accurate. Moreover, the fact that the estimation is done simultaneously avoids the measurement error coming from the fact that $\{\beta_p\}_{p=1}^3$ are estimated rather than observed in DL.

Second, parsimony results from the fact that the ultimate parameters of interest are time-invariant. Third, as opposed to traditional VAR models such as in Evans and Marshall (2002), the number of parameters to be estimated does not increase with the number of yields, even after imposing zero restrictions that imply exogeneity of macro variables with respect to yields.

Finally, when it comes to simulate the movements of the term structure - or out-of-sample forecasting, more generally – one just needs to plug-in updated (or forecasted) values of M_{t-} and compute the resulting yields forecasts; alternative models, such as DRA, which contain both latent and observable factors, would need to rely on extra assumptions on the latent part to do so.

3.2. Implementation

Let $y_t(\tau)$ is the vector of yields observed at date t, and $u_t(.)$ is the error term, both of dimension $N\times 1$, $X_t(\cdot)$ is $N\times 3$, β_t and $\bar{\beta}$ are 3×1, λ_t and $\bar{\lambda}$ are scalars, σ_β and σ_λ are, respectively, $k_\beta\times 1$ and $k_\lambda\times 1$, and $M_{t-}=\begin{bmatrix}M_{\beta t-} & 0_{3\times k_\lambda}\\ 0_{1\times k_\beta} & M_{\lambda t-}\end{bmatrix}$ is $4\times k(=k_\beta+k_\lambda)$.[5] The model consists of N yield observations for each one of the T periods, k state variables per period, and $k+4$ parameters to be estimated, regardless of the number of yields or time periods in the sample – the dimension of the parameter vector grows only with the number of state variables in the model (say, at most three per factor, so that most likely $k\leqslant 12$). Since $\lambda_t=\bar{\lambda}+M_{\lambda t-}\sigma_\lambda$, one can write $X_t(\lambda_t)=X_t(\bar{\lambda},\sigma_\lambda)$ but should bear in mind that both $\tau$and $M_{\lambda_{t-}}$ are also arguments of $X_t(\cdot)$ but are omitted for convenience.

The assumption that the error term $u_t(\tau)$ is a martingale difference sequence with respect to current and past covariate information W_t implies conditional moment restrictions of the form $E[u_t(\tau)|W_t]=0$. In particular, for every period t, one can use unconditional moments of the form $E[W_t'u_t(\tau)]=0$, with sample counterpart

$$\begin{aligned}0&=(NT)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}u_{it}w_{it}\\ &=(NT)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}(y_t(\tau_i)-X_t(\tau_i,\theta_\lambda)\bar{\beta}-X_t(\tau_i,\theta_\lambda)M_{\beta t-}\sigma_\beta)w_{it}\end{aligned}$$

where $\theta:=\left(\theta_\beta';\theta_\lambda'\right)'=\left(\bar{\beta}',\sigma_\beta';\bar{\lambda},\sigma_\lambda'\right)'$ and where it should be noted that $M_{\lambda t-}$ as an argument of $X_t(\cdot,\cdot)$ and is omitted for convenience.

Before defining the estimation problem, I define a number of matrices. Start by stacking the yields by period to form the $NT\times 1$ vector $y=[y_1(\tau)',y_2(\tau)',\ldots,y_T(\tau)']'$, then stack the $X_t(.)$ matrix to form the $NT\times 3$ matrix $X(\theta_\lambda)=[X_1(\theta_\lambda)',X_2(\theta_\lambda)',\ldots,X_T(\theta_\lambda)']'$ and the $M_{\beta t-}$ matrix to form the $NT\times k$ matrix $XM(\theta_\lambda)=[(X_1(\theta_\lambda)M_{\beta 1-})',(X_2(\theta_\lambda)M_{\beta 2-})',\ldots,(X_T(\theta_\lambda)M_{\beta T-})']'$; then let $Z_t=[X_t(\theta_\lambda),X_t(\theta_\lambda)M_{\beta t-}]$ and its stacked version of dimension $NT\times(3+k)$, $Z=[Z_1(\theta_\lambda)',Z_2(\theta_\lambda)',\ldots,Z_T(\theta_\lambda)']'$, and let $W=[W_1',\ldots,W_T']$ be an instrument matrix of dimension $NT\times r(\geqslant k+4)$. I use a GMM estimator $\hat{\theta}$ of θ, which minimizes the following distance of $G_{NT}(\theta)=\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}u_{it}w_{it}$ from zero:

$$\begin{aligned}\hat{\theta}&=\arg\min_{\theta\in\Theta}[G_{NT}(\theta)]'A_{NT}[G_{NT}(\theta)]\\ &=\arg\min_{\theta\in\Theta}[(NT)^{-1}W'u]'A_{NT}[(NT)^{-1}W'u]\\ &=\arg\min_{\theta\in\Theta}[(NT)^{-1}W'[y-Z(\theta_\lambda)\theta_\beta]]'A_{NT}[(NT)^{-1}W'[y-Z(\theta_\lambda)\theta_\beta]]\end{aligned}$$

where A_{NT} is an $NT\times NT$, possibly random, positive semi-definite weighting matrix with rank at least $k+4$, and the last line shows that nonlinearity comes from the subset of parameters $\theta_\lambda=\left(\bar{\lambda},\sigma_\lambda'\right)'$ governing the locus of the tilting point of the yield curve.

One particular case of the above estimator is when $W_t=\left[Z_t(\theta_\lambda),\frac{\partial Z_t(\theta_\lambda)\theta_\beta}{\partial\theta_\lambda}\right]$ and $A_{NT}=\mathbf{I}_{NT}$, which results in the Nonlinear Least Squares estimator. Here, the associated covariance matrix is given by $\Omega=E(\nabla_\theta'A_0\nabla_\theta)^{-1}E(\nabla_\theta'A_0V_0A_0\nabla_\theta)E(\nabla_\theta'A_0\nabla_\theta)^{-1}$, where the asymptotic variance is $V_0=\lim_{N,T\to\infty}(NT)^{-1}Var$

[3] The assumption of no-correlation is arguably restrictive. Previous versions of the paper have considered more general models allowing for spatial dependence, i.e. dependence across yields of close enough maturities, but without much success, especially out-of-sample.

[4] Decomposing the time-varying parameters as above assumes the relation between state variables and term structure factors is deterministic, which might obviously lead to biased results if this assumption does not hold in practice.

[5] In particular, $M_{t-}=diag\left\{m_t^{\beta_1},m_t^{\beta_2},m_t^{\beta_3},m_t^{\lambda}\right\}$. In the general case, $M_{t-}=[M_{\beta t-},\ \ M_{\lambda t-}]$ is $4\times k(=k_\beta+k_\lambda)$.

$\left((NT)^{-1/2}\sum_{t=1}^{T}\sum_{i=1}^{N}u_{it}w_{it}\right)$, $\nabla_\theta := \frac{\partial G_{NT}(\theta)}{\partial\theta} = \left(\nabla_{\bar{\beta}}, \nabla_{\sigma_\beta}, \nabla_{\bar{\lambda}}, \nabla_{\sigma_\lambda}\right)$, with details given in Appendix A, and A_0 is a positive semi-definite matrix such that $A_{NT} \rightarrow {}^{p}A_0$.

The above estimation strategy highlights a number of desirable properties of the model. First, the model is robust to curse of dimensionality issues, as the dimension of the parameter vector does not increase with the number of yields, but with the number of state variables, which is kept at a manageable size. As a result, one does need to restrain the number of yields used when estimating the yield curve, which takes an especially heavy toll on the estimation of the curvature, thus resulting in a poor estimation of the connection between this factor and any state variables associated to it.

Second, more than just allowing the comparison of alternative specifications, one can test competing theories about variables driving the term structure dynamics using inference tools.

Finally, and in contrast with most of the literature, the estimation makes use of both the cross-sectional and time series dimensions of the data, resulting in much faster convergence of the parameter estimates.[6] This is of special interest given issues commonly raised against VAR models used in the analysis of monetary policy: Rudebusch (1998), for instance, points out that the use of quarterly data, together with the relatively frequent changes in monetary policy in the postwar period results in either short time series or misspecified VAR models, thus making inference unreliable: using quarterly data, the twenty years of the 'Greenspan era' correspond to only 80 observations.

4. Finite-sample performance

This section presents a simulation study investigating the finite-sample performance of the estimation method – we keep the discussion brief, see e.g. Podivinsky (1999) for a survey of the finite-sample behaviour of GMM estimators. I generate state variables M_{t-}, regressors X_t, population parameter values, and errors to generate the variables y_t. For every experiment, I compute the results of 1000 replications, with time-series and cross section dimensions given by, respectively, T (=100, 200, 500) and N (=25, 50, 100). The design of the simulation study is done to reflect dimensions typically seen in empirical studies. In the cross-sectional dimension, while CRSP yields would usually be less than $N = 25$, interpolated yields by McCulloch would typically be in the range $N = 50–100$, depending on the period (the number of yields tends to increase over time). In the time-series dimension, $T = 100$, 200 and 500 reflect 8.33, 16.66 and 41.66 years of data at the monthly frequency, respectively.

The state variables M_{t-} are constructed by taking the exponent of independent standard Gaussian random variables, the regressors X_t are standard Gaussian random variables, whereas the error terms u_t are Gaussian variables with a variance of 0.1.

In what follows, I consider model (3) with each factor driven by one state variable. As in the empirical exercise, I make the curvature-related factors β_{3t} and λ_t, i.e. the curvature intensity and the location where the curve tilts are driven by the same state variable, so that $M_{t-} = diag\{m_{1t-}, m_{2t-}, m_{3t-}, m_{3t-}\}$. In all the experiments, $\bar{\beta} = (0.20, -0.10, -0.20)'$, $\sigma_\beta = (0.95, 0.95, 0.80)'$, $\bar{\lambda} = 0.05$, and $\sigma_\lambda = 0.01$. The parameters reflect an average term structure within the sample period; we have also experimented with alternative parameter values and obtained similar conclusions.[7]

The simulation results reported in Table 1 show the convergence of the θparameter estimates to their population values, with increasing precision in both N and T: fixing one dimension, say $T = 100$, biases (or their absolute values) and especially their standard errors converge towards zero as the cross-sectional dimension increases. Conversely, fixing the cross-sectional dimension (say $N = 25$) biases and their standard errors tend to converge to zero as the time series dimension grows large.

5. Application

5.1. Data

The data I use comprises end-of-month yields from US bonds from January, 1970 to December, 2003 and US macroeconomic variables obtained from the US Federal Reserve macroeconomic database – the FRED – and observed at the monthly frequency.[8] For every given period, the macroeconomic variables used are predetermined with respect to the interest data used. For instance, when using the yield curve of 31 March, 1970, I make sure I only use variables dated prior to that, e.g. 1 March, 1970. In particular, the variables in level used date from 1 March, 1970, and the variables in growth rate are the increment from 1 February 1970 to 1 March, 1970.

5.1.1. Interest rates

The interest rate data used consists of the unsmoothed Fama–Bliss yields described and thoroughly discussed in Bliss (1997).[9] The yields form an unbalanced panel of ranging from 42 to 134 observations per period. The average number of yields for the full sample is 86.944, with a standard error of 26.854, the number of periods in the full sample is $T = 408$ months, and the longest maturity used in the study is 60 months. The main features in the data are the average upward-sloping yield curve, the fact that yield volatility tends to decrease with maturity whereas persistence tends to increase with maturity.[10]

5.1.2. Macroeconomic variables

Based on the existing literature, I consider measures of inflation, economic activity, monetary policy, and fiscal policy. The inflation measures used are the CPI (Consumer Price Index For All Urban Consumers: All Items, seasonally adjusted), PPI1 (Producer Price Index: Finished Goods, seasonally adjusted), PPI2 (Producer Price Index: All Commodities, not seasonally adjusted), PPI3 (Producer Price Index: Industrial Commodities, not seasonally adjusted), and PCE (Personal Consumption Expenditures: Chain-type Price Index, seasonally adjusted) – all measured in growth rates.

The measures of economic activity used are HOUST (Housing Starts: Total: New Privately Owned Housing Units Started, seasonally adjusted), INDPRO (Industrial Production Index, seasonally adjusted), EMP (Civilian Employment, seasonally adjusted) – all measured in growth rates – plus TCU (Capacity Utilization: Total Industry, seasonally adjusted), HELP (Index of Help Wanted Advertising in Newspapers, seasonally adjusted) and UR (Unemployment Rate, seasonally adjusted), measured in levels.

The monetary policy instruments used are FF (Federal funds effective rate), NONBR (Non-Borrowed Reserves of Depository Institutions, seasonally adjusted – the monetary aggregate the Fed targeted during the period from October, 1979 to October,

[6] The following Section illustrates the finite-sample properties of the method.

[7] We checked for robustness in standard ways, such as running pilot experiments with different dimensions of N, T and the number of replications, besides investigating different parameter values, i.e. different shapes of the term structure. In particular, previous versions of the paper reported results of 500 replications with parameter values $\bar{\beta} = (1,1,1)'$, $\sigma_\beta = (1,1,1)'$, $\bar{\lambda} = 0.05$, and $\sigma_\lambda = 0.01$ with similar conclusions.

[8] The dataset is available from http://research.stlouisfed.org/fred2/.

[9] I thank Robert Bliss for making his data available.

[10] While the Fama–Bliss dataset has a large number of yields, one may be less comfortable in using its longer maturities due to interpolation issues. Results available from the author upon request show the robustness of the results when considering the CRSP dataset, which has fewer yields and longer maturities.

Table 1
Simulation results for single-variable factor specification. The table reports results from the simulation experiments to assess finite-sample performance of the method. For each cell, the bias and the standard error (in square brackets) of 1000 replications are reported.

	$T = 100$			$T = 200$			$T = 500$		
	$N = 25$	$N = 50$	$N = 100$	$N = 25$	$N = 50$	$N = 100$	$N = 25$	$N = 50$	$N = 100$
$\overline{\beta}_1$	−0.0220	−0.0003	−0.0005	−0.0177	−0.0025	−0.0001	−0.0048	−0.0020	−0.0007
	[0.1099]	[0.0474]	[0.0201]	[0.0943]	[0.0418]	[0.0156]	[0.0788]	[0.0341]	[0.0114]
$\overline{\beta}_2$	0.0207	0.0019	0.0005	0.0173	0.0017	−0.0005	0.0047	0.0023	0.0006
	[0.1119]	[0.0518]	[0.0295]	[0.0942]	[0.0444]	[0.0242]	[0.0777]	[0.0366]	[0.0196]
$\overline{\beta}_3$	0.0275	−0.0050	−0.0035	0.0212	0.0011	0.0040	0.0099	0.0019	−0.0012
	[0.1973]	[0.1169]	[0.0826]	[0.1743]	[0.0095]	[0.0665]	[0.1364]	[0.0763]	[0.0575]
$\bar{\lambda}$	−0.0004	−0.0001	−0.0001	−0.0003	−0.0001	0.0001	0.0001	0.0000	−0.0001
	[0.0061]	[0.0031]	[0.0021]	[0.0043]	[0.0024]	[0.0016]	[0.0030]	[0.0017]	[0.0010]
σ_{β_1}	0.0006	−0.0002	0.0000	0.0016	0.0004	0.0000	0.0010	0.0003	0.0001
	[0.0161]	[0.0097]	[0.0047]	[0.0124]	[0.0080]	[0.0031]	[0.0100]	[0.0057]	[0.0019]
σ_{β_2}	−0.0003	−0.0002	0.0000	−0.0015	−0.0002	0.0001	−0.0010	−0.0004	−0.0001
	[0.0168]	[0.0114]	[0.0073]	[0.0132]	[0.0088]	[0.0050]	[0.0101]	[0.0062]	[0.0035]
σ_{β_3}	−0.0029	0.0016	0.0006	−0.0028	−0.0002	−0.0005	−0.0017	0.0003	0.0001
	[0.0407]	[0.0285]	[0.0204]	[0.0305]	[0.0208]	[0.0143]	[0.0216]	[0.0136]	[0.0103]
σ_{λ}	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	[0.0012]	[0.0006]	[0.0005]	[0.0007]	[0.0005]	[0.0003]	[0.0004]	[0.0003]	[0.0002]

1982), and M1 (Money Stock, in Billions of Dollars, seasonally adjusted).

All the above variables are recorded at the monthly frequency, and were obtained from the FRED database. Finally, following Dai and Philippon's (2005) finding that fiscal policy affects the term structure, I introduce the variable DEBT, which is their quarterly fiscal policy variable interpolated to the monthly frequency and divided by INDPRO, a *proxy* variable for GDP at the monthly frequency. Panel A in Table 2 summarizes the macroeconomic variables used.

5.2. On the economic determinants of the yield curve

This section starts by selectively reviewing the literature addressing the relation between macroeconomic variables and the yield curve factors, thus paving the way for the empirical strategy I implement next. It goes without saying that with a set of macroeconomic variables as big as the one available from the FRED, there are countless alternative specifications to be compared ($15^3 = 3375$ using only the contemporaneous variables described above), so that a pragmatic starting point would be to consider specifications based on the existing literature and summarized on Panel B in Table 2. The evidence documented in the literature is used to construct alternative configurations of M_{t-} which are then compared. For the sake of parsimony, I devote a section to single-variable (SV) specifications – the ones where each factor is driven by one state variable only – before addressing the general multi-variable (MV) case. I then use the 'best' SV and MV specifications in the out-of-sample comparison with the benchmark DL model.

Much of the work in macro-finance gained momentum in the late 1990s (see, for instance, Diebold et al. (2005), and references therein). One of the early papers is Evans and Marshall (1996) – to which Evans and Marshall (1998) also relates – where, using a VAR framework, the authors study the impact of shocks of measures of monetary policy, employment and inflation on the nominal term structure of interest rates. Their results suggest that the main effect of both employment and inflation measures is to induce a parallel shift of the yield curve, whereas (short-run) fluctuations in the slope and curvature of the yield curve are primarily attributed to the monetary policy shocks.

Table 2
Macroeconomic variables and term structure factors. The table summarizes findings from the literature in what concerns the economic factors driving the level, slope and curvature factors of the term structure of interest rates.

Economic activity	Inflation	Monetary policy	Fiscal policy
Panel A: Macroeconomic variables by group			
$\text{UR}^{\mathcal{L}}$	CPI	$\text{FF}^{\mathcal{L}}$	DEBT
$\text{TCU}^{\mathcal{L}}$	PCE	NONBR	
$\text{HELP}^{\mathcal{L}}$	PPI1	M1	
IP	PPI2		
EMP	$\text{PPI3}^{\mathcal{N}}$		
HOUST			
Reference	Level factor	Slope factor	Curvature factor
Panel B: Macroeconomic variables driving term structure factors			
Evans and Marshall (1996)	Employment Inflation	Monetary policy	Monetary policy
Ang and Piazzesi (2003)	–	Inflation	Output
Piazzesi (2005)	–	Monetary policy	–
DRA (2005)	Inflation	Output	–

Also within the VAR framework, but imposing no-arbitrage restrictions, Ang and Piazzesi (2003) construct inflation and economic growth indices which they address as macro factors. By a factor representation of the pricing kernel they obtain a tractable way to examine how those macro factors affect the yield curve dynamics. However, in their study macro factors are able to explain only the short end and the middle of the yield curve. Due to difficulty to deal with the long end they introduce latent factors, now allowing the pricing kernel to be driven by both macro and latent factors. By relying on a Gaussian assumption and on the affine specification, they find that the slope and curvature factors can be explained by the macro factors, whereas the level factor can be only dealt with by using latent factors. In a related paper, but within a different framework, Piazzesi (2005) finds that monetary policy shocks change the slope of the yield curve, since they affect short rates more than long ones.

More recently, DRA examine the correlations between NS factors and macroeconomic variables under a VAR framework and find that the level factor is highly correlated with inflation and the slope factor is highly correlated with real activity, whereas the curvature factor does not appear to be related to any of the macroeconomic variables used.

5.3. In-sample analysis

5.3.1. SV specification

I start estimating SV specifications, where each factor is driven by one state variable only. These can be seen either as a parsimonious way of approaching the problem or as a first step before considering more complex (and potentially difficult to compute) specifications for M_{t-}, besides providing additional out-of-sample benchmarks for those more complex specifications. A simplifying assumption made throughout the exercise is that the curvature intensity β_{3t} and the parameter governing the location of the tilting point of the yield curve are the same.

Given two competing specifications with the same number of variables, I compare them using the Mean Absolute Error criterion (both the average and the median of the MAE's across time are reported). The MAE is of special interest here for providing a model selection criterion, an idea of goodness-of-fit, and of mispricing of the specifications. Given the large number of specifications one can estimate, Panel A in Table 3 reports results of *selected* specifications from an exercise designed to select the best forecasting variables from the different categories.[11] The results summarized in Table 3 provide some insights on the forecasting ability of the state variables.

First, specifications 1–5 suggest that CPI and PCE have a slightly better forecasting ability for the level factor than PPI-related variables. Second, specifications 6–8 suggest that the best monetary policy variable explaining the slope is FF. Third, specifications 9–14 suggest that UR is the best economic activity variable explaining the curvature factor. UR is also the economic activity variable doing the best job at explaining the level factor (see specifications 15–20) and the slope (see specifications 21–26). Further, the best monetary policy variable explaining curvature is FF (see specifications 27–29), and the best economic activity variable is UR (see specifications 9–14). The better performance of specifications 31 and 33 when compared to 32 and 34, respectively, suggest that CPI does a better job at explaining the level than PCE, while specifications 35–44 fit the data rather poorly.

The investigation of the variable DEBT in a number of specifications comes from recent evidence that fiscal policy does play a role at explaining the curvature factor of the term structure (Dai and Philippon, 2005). However, the fact that specifications with UR instead of DEBT as drivers of the curvature factor seems to be robust (in particular, compare specifications 30 and 33). All in all, there seems to be a dominance of specifications for which inflation (either CPI or PCE) explains the level, monetary policy (FF) explains the slope, and economic activity (UR) explains the curvature of the term structure i.e. specification 33 is the best-performing one. This specification shares with DRA the fact that the level factor is driven by an inflation measure and with Evans and Marshall (1996) and Piazzesi (2005) the fact that the slope is driven by a monetary policy variable (since monetary policy instruments typically have more effect over the short- than the long-term of the term structure). Finally, it relates to Ang and Piazzesi (2003) in that the curvature factor is driven by an output-related variable.

In what follows, I refer to the best SV specification (specification 33 in Table 3, Panel A) as SV – see Panel B in Table 3 for the corresponding parameter estimates. The parameter estimates for the SV

Table 3
Results for single-variable specifications. The table reports, respectively, average and median MAEs in the last two columns. The preferred specification is marked with *.

Specification	Level	Slope	Curvature	Avg (MAE)	Med (MAE)
Panel A: SV Specifications					
1	CPI	FF	M1	1.010	0.830
2	PCE	FF	M1	1.031	0.825
3	PPI1	FF	M1	1.012	0.836
4	PPI2	FF	M1	1.025	0.830
5	PPI3	FF	M1	1.023	0.833
6	PCE	FF	DEBT	1.069	0.869
7	PCE	NONBR	DEBT	1.822	1.395
8	PCE	M1	DEBT	1.823	1.415
9	CPI	FF	UR	0.833	0.696
10	CPI	FF	TCU	1.022	0.851
11	CPI	FF	HELP	1.058	0.893
12	CPI	FF	IP	1.036	0.873
13	CPI	FF	EMP	1.042	0.898
14	CPI	FF	HOUST	1.050	0.889
15	UR	FF	DEBT	0.877	0.736
16	TCU	FF	DEBT	1.041	0.854
17	HELP	FF	DEBT	1.069	0.873
18	IP	FF	DEBT	1.053	0.881
19	EMP	FF	DEBT	1.063	0.851
20	HOUST	FF	DEBT	1.062	0.865
21	PCE	UR	M1	1.629	1.284
22	PCE	TCU	M1	1.817	1.449
23	PCE	HELP	M1	1.668	1.279
24	PCE	IP	M1	1.815	1.448
25	PCE	EMP	M1	1.817	1.437
26	PCE	HOUST	M1	1.811	1.401
27	UR	FF	FF	0.854	0.738
28	UR	FF	NONBR	0.884	0.746
29	UR	FF	M10	0.872	0.739
30	CPI	FF	DEBT	1.048	0.886
31	CPI	FF	M1	1.010	0.830
32	PCE	FF	M1	1.031	0.825
33*	CPI	FF	UR	<u>0.833</u>	<u>0.696</u>
34	PCE	FF	UR	0.836	0.701
35	UR	FF	UR	0.873	0.727
36	CPI	UR	DEBT	1.048	0.886
37	PCE	UR	DEBT	1.468	1.203
38	UR	UR	DEBT	1.521	1.272
39	CPI	UR	M1	1.462	1.199
40	PCE	UR	M1	1.467	1.170
41	UR	UR	M1	1.526	1.287
42	CPI	UR	UR	1.462	1.211
43	PCE	UR	UR	1.467	1.179
44	UR	UR	UR	1.523	1.262

Specification		$\left(\hat{\bar{\beta}}', \bar{\lambda}\right)'$		$(\hat{\sigma}_{\beta}, \hat{\sigma}_{\lambda})'$		Avg-Med (MAE)
Panel B: Parameter estimates for best SV specification						
SV:	β_{1t}:CPI	10.655	[0.091]	55.820	[1.317]	0.833 – 0.696
	β_{2t}:FF	−10.148	[0.089]	0.892	[0.003]	
	β_{3t}:UR	−13.661	[0.169]	2.177	[0.019]	
	λ_t:UR	0.007	[0.001]	0.004	[0.001]	

model show the positive impact of CPI on the level of the term structure, the impact of the monetary policy instrument FF on the slope (actually defined as $-\beta_{2t}$), and the impact of UR on both the intensity and the locus of the curvature, all of them found to be significant using Newey and West (1987) standard errors to account for the time dependence in the data. Interestingly, neither the CPI nor the UR are revised, which makes them attractive as predetermined variables with respect to yields. When coupled with the real time Taylor rule proposed in Evans (1998), the findings are consistent with what one would intuitively expect, in the sense that the yield curve tends to invert for values of FF above the Taylor rule, but remaining upward-sloping for values below the threshold.

5.3.2. MV specifications

Based on the findings in the literature, SV specifications are probably too simple to account of the term structure dynamics in

[11] The results are robust with respect to choice of selection criterion used – using the minimum value of the criterion function, the AIC or the BIC criteria gives the same results in terms of the preferred specifications.

Table 4
Results for multi-variable factor specifications. The table reports estimates for alternative MV specifications and their respective MAEs and Schwarz's BIC criteria, in the last two columns.

Specification	Level	Slope	Curvature	Avg (MAE)	BIC
Panel A: MV specifications					
1	CPI, PCE$^{\varnothing}$, UR	FF, UR	DEBT, FF$^{\varnothing}$, UR$^{\varnothing}$	0.759	49.191
2	CPI, UR	FF, UR	DEBT, FF$^{\varnothing}$, UR	0.759	45.097
3	CPI, PCE$^{\varnothing}$, UR	BT, UR	DEBT, UR$^{\varnothing}$	0.844	45.100
4	CPI, PCE, URFF,UR	FF, UR	DEBT$^{\varnothing}$, FF$^{\varnothing}$	0.800	45.099
5	CPI, UR	FF, UR	DEBT, FF	0.972	41.025
6	CPI, PCE$^{\varnothing}$, UR	FF, UR	DEBT, FF$^{\varnothing}$, UR$^{\varnothing}$	0.819	41.004
7[*]	CPI, URFF, UR	FF, UR	DEBT	0.820	36.910

Specification		$\left(\hat{\beta}', \bar{\lambda}\right)'$		$(\hat{\sigma}_\beta, \hat{\sigma}_\lambda)'$		Avg-Med (MAE)
Panel B: Parameter estimates for best MV specification						
MV:	β_{1t}:*CPI*	−2.017	[0.304]	47.564	[1.330]	
	β_{1t}:*UR*	2.094	[0.284]	1.791	[0.014]	0.820
	β_{2t}:*FF*			0.886	[0.004]	
	β_{2t}:*UR*			−1.727	[0.017]	
	β_{3t}:*DEBT*	2.011	[0.551]	−6.029	[0.153]	
	λ_t:*DEBT*	0.031	[0.001]	−0.035	[0.002]	

[*] The best specification according to the BIC.
$^{\varnothing}$ Non-significance (at the 5 percent significance level) of the corresponding parameter.

a satisfactory way. The natural next step is then to study the more general MV specifications. Based on the results reported in Table 3, I employ a general-to-specific approach starting with a specification where CPI, PCE and UR drive the level, FF and UR drive the slope, and DEBT, FF and UR drive the curvature. The alternative specifications compared in Panel A of Table 4 show that several coefficients in the larger models are statistically insignificant. The model with the smaller BIC and with all of the parameters statistically significant is specification 7 – which I from now on refer to as MV, which has the level driven by CPI and UR, the slope by FF and UR, and the curvature by DEBT. Albeit more parsimonious than the full model the average MAE is only slightly larger.

The parameter estimates for the MV model are reported on Panel B of Table 4. The findings are in line with previous results in that economic activity and inflation drive the level factor (as in Evans and Marshall, 1996), economic activity (as in DRA) and monetary policy (as in Evans and Marshall, 1996; Piazzesi, 2005) drive the slope, and fiscal policy drives the curvature factor, consistent with findings of Dai and Philippon (2005). Note, however, the performance of the model in terms of MAE is very similar to the SV model.

The parameter estimates – all of which significant – show the upward impact of inflation on the term structure, as expected. The parameters related to the slope also have the expected sign, with FF affecting shorter rates more strongly, but UR having the opposite effect, as one would expect given the typical relation between these two variables.

5.4. Incorporating economic relations

So far, the model presented considers only state variables which are predetermined with respect to the yield curve, not exploring (i) any interdependence among them; (ii) any forecasts of their future values, both of which are expected to play a role at explaining future realizations of the yield curve. In this Section 1 discuss how to incorporate into the model information on the joint behaviour of the state variables. Intuitively, by informing the model that certain variables are related one should expect to get more accurate results, provided the relation imposed holds.

I now inform the model about the joint behaviour of the state variables using a feedback interest rule, or *Taylor rule*. Taylor (1993) suggested a simple formula describing how the US Federal Open Market committee has set the Federal funds rate since 1987 as a response to measures of inflation and output gaps – this relationship has been dubbed the Taylor rule and has been extensively studied and developed since then. Despite its simplicity, the Taylor rule has a number of appealing properties. Woodford (2001) shows how it incorporates several features of an optimal monetary policy in a class of optimizing models, and provides conditions under which the Taylor rule has a stabilizing effect on the economy. More recent developments such as Clarida et al. (2000, CGG hereafter) propose and estimate a Taylor rule incorporating both forward- and backward-looking elements. The former account for the fact that the monetary authority is considering future paths of the output and inflation gaps when setting the current value of the monetary policy instrument, whereas the latter arises as a consequence of interest rate smoothing conducted by the monetary policy authority – see also Rudebusch (1995). In what follows, I estimate both forward- and backward-looking versions of the Taylor rule. Instead of using quarterly data, as in CGG, I use monthly observations and find that, by and large, their results follow through to the monthly frequency.

The results of this section provide the ground for alternative ways of computing out-of-sample forecasts, in the sense that one can plug into the model estimated quantities generated by a model inspired by (or consistent with) economic theory to obtain estimates of the future behaviour of the term structure.

5.4.1. Taylor rules

In what follows I consider the following specification proposed and estimated in CGG, which nests both forward- and backward-looking versions of the Taylor rule and posits that the target rate each period is a linear function of the gaps between expected inflation and output and their respective target levels,

$$r_t^* = rr^* + \pi^* + \gamma^\pi[E(\pi_{t,l_\pi}|\psi_t) - \pi^*] + \gamma^g E(g_{t,l_g}|\psi_t)$$

where ψ_t is the information set available at time t, π_{t,l_π} denotes the percent change in the price level between periods t and $t + l_\pi$ (expressed in annual rates), π^* is the target for inflation, $rr^*(=r^* - \pi^*)$ is the long-run equilibrium real rate, with r^* being, by definition, the desired nominal rate when both output and inflation are at their target values. g_{t,l_g} is a measure of the average output gap between periods t and $t + l_g$, with the output gap being defined as the percent deviation between actual GDP and the corresponding target.[12]

[12] Typically, the information set at time t contains past values of the Fed Funds rate and other economic variables, and usually no information on current inflation and output measures.

Table 5
Taylor rule results. The table reports estimates alternative Taylor rules. The symbols l_π, l_g, l_r denote, respectively, the lags (or forward shifts, for negative values) of the inflation, economic activity and interest rate variables. Specification BWTR1 uses a constant, current values of CPI and UR, and lagged values of CPI, UR, and FF as instruments. Specification BWTR2–4 use the same instruments as BWTR1 plus 2–4 lagged versions of CPI, UR, and UR. Newey–West standard errors with 12 lags are reported inside square brackets. Specification FWTR1 uses a constant, current values of CPI and UR, and lagged values of CPI, UR, and FF as instruments. Specification FWTR2–4 use the same instruments as FWTR1 plus 2–4 lagged versions of CPI, UR, and UR. Newey-West standard errors with 12 lags are reported inside square brackets.

	Leads/lags of macroeconomic variables		
	l_π	l_g	l_r
Panel A: Taylor rule specifications			
(Backward) Taylor rule	<0	<0	–
(Backward) Taylor rule with interest rate smoothing	<0	<0	<0
Clarida-Galí-Gertler	>0	>0	<0

	BWTR1	BWTR2	BWTR3	BWTR4	FWTR1	FWTR2	FWTR3	FWTR4
Panel B: Results for backward- and forward-looking Taylor rules								
l_{CPI}	0	−2	−6	−10	1	1	1	1
l_{UR}	0	−2	−6	−10	1	1	1	1
l_{FF}	−1	−1	−1	−1	−1	−1	−1	−1
Instrument lags	1	1	1	1	1	2	3	4
γ^{CPI}	2.264	2.273	2.355*	2.866*	2.427*	1.929**	1.702**	1.605**
	[2.127]	[1.408]	[1.388]	[1.691]	[1.358]	[0.870]	[0.770]	[0.746]
γ^{UR}	1.158	0.410	0.598	0.352	0.454	0.473	0.618	0.615
	[1.380]	[0.504]	[0.431]	[0.416]	[0.498]	[0.381]	[0.394]	[0.404]
CPI^*	1.257	2.270	1.715	2.089	2.020	1.640	1.085	0.911
	[6.491]	[3.305]	[2.177]	[1.882]	[2.898]	[3.054]	[3.461]	[4.084]
ρ	0.986***	0.964***	0.951***	0.949***	0.965***	0.952***	0.948***	0.950***
	[0.001]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
R^2	0.993	0.993	0.993	0.993	0.993	0.993	0.993	0.993
J-statistic	26.153***	6.326*	1.128	0.860	0.377	7.787	10.654	12.894
df	2	2	2	2	2	5	8	11

* Significance at the 10 percent level.
** Significance at the 5 percent level.
*** Significance at the 1 percent level.

Following CGG, the *actual* Fed funds rate follows

$$r_t = \rho(L)r_{t-1} + (1-\rho)r_t^*$$

where $\rho(L) = \rho_1 + \rho_2 L + \cdots + \rho_{l_r} L^{l_r-1}$ and $\rho = \rho(1) = \sum_{j=1}^{l_r}\rho_j$, which postulates a partial adjustment of the Fed funds rate to the target r_t^*, with ρbeing an indicator of the degree of smoothing of interest changes by the monetary policy authority.

Combining the target rate and Fed funds equations above results in the Taylor rule

$$r_t = (1-\rho)[rr^* + (1-\gamma^\pi)\pi^* + \gamma^\pi \pi_{t,l_\pi} + \gamma^g g_{t,l_g}] + \rho(L)r_{t-1} + \varepsilon_t$$

where $\varepsilon_t = (1-\rho)(\gamma^\pi[E(\pi_{t,l_\pi}|\psi_t) - \pi_{t,l_\pi}] + \gamma^g[E(g_{t,l_g}|\psi_t) - g_{t,l_g}])$ is a linear combination of forecast errors, thus being orthogonal to any variable in the information set ψ_t. As one can only identify the term rr^* + $(1-\gamma^\pi)\pi^*$, but not rr^* or γ^π separately, and the inflation target is of interest, CGG assume that the equilibrium real rate rr^* equals its sample average. This specification allows a number of choices regarding the lead/lag periods of inflation and output, l_π and l_g, respectively, and lags for the Fed funds, l_r. The parameters of interest are π^*, γ^π, γ^g, $\{\rho_j\}_{j=1}^{l_r}$, so that the dimension of the parameter vector is 3 + l_r. It also nests a number of specifications, as shown in Panel A in Table 5.[13]

The regression equation above implies the set of moment conditions

$$E([r_t - (1-\rho)[rr^* + (1-\gamma^\pi)\pi^* + \gamma^\pi \pi_{t,l_\pi} + \gamma^g g_{t,l_g}] - \rho(L)r_{t-1}]z_t) = 0$$

where z_t is a vector of instruments known when r_t is set ($z_t \in \psi_t$) and π_{t,l_π}, g_{t,l_g}, and r_{t-1} also belong in ψ_t.

Table 6
NBER-dated recessions considered. The table reports the five NBER-dated recessions considered, their starting and ending months, and their duration in months.

Recession code	Start date	End date	Duration (months)
R1	November, 1973	March, 1975	16
R2	January, 1980	July, 1980	6
R3	July, 1981	November, 1982	16
R4	July, 1990	March, 1991	8
R5	March, 2001	November, 2001	8

The above moment conditions are used to obtain parameter estimates using the Generalized Method of Moments. As in CGG, I set the equilibrium rate rr^* to its sample average, so I can identify the inflation target. To make the feedback rule consistent with the SV specification, I replace r with FF, π with CPI, and I also follow Evans' (1998) implementation of the Taylor rule, replacing the output gap with the unemployment gap using Okun's law, besides setting the natural rate of unemployment to $UR_t^* = UR^* = 6$.[14] Moreover, I assume that current inflation and unemployment are not observed when setting the Fed Funds rate, i.e. neither of them belongs in ψ_t. The moment conditions thus become $E(\varepsilon_t^* z_t) = 0$, where$\varepsilon_t^* = [FF_t - (1-\rho)[rr^* + (1-\gamma^{CPI})CPI^* + \gamma^{CPI}CPI_{t-1,l_{CPI}} + +\gamma^{UR}3(6 - UR_{t-1,l_{UR}})] - \rho(L)FF_{t-1}]$.

Interestingly, given that in the SV specification the slope is driven by the Fed funds rate, the above specification can be linked to the interest-rate rule proposed in McCallum (2005), according to

[13] Note that the Taylor rule is usually applied to quarterly data, whereas I consider monthly data.

[14] Arthur Okun observed that a one percent fall in the unemployment rate from its full employment level tended to produce a three percent increase in real GDP relative to trend. See Evans (1998) for a discussion and robustness checks.

Table 7
Overall accuracy of alternative specifications. The table reports the overall accuracy of alternative specifications. The underlined quantities are the smaller values for a given time period and episode.

Month	Recession R1					Recession R2					Recession R3				
	DL	SV	SV-TR	MV	MV-TR	DL	SV	SV-TR	MV	MV-TR	DL	SV	SV-TR	MV	MV-TR
Panel A: Recessions 1–3															
1st	1.31	1.56	*0.65*	1.71	1.00	4.48	*0.88*	1.38	1.03	1.26	1.81	*0.88*	5.36	1.01	4.98
2nd	*0.81*	1.16	0.82	1.22	1.13	4.57	*1.69*	1.77	1.82	*1.69*	3.50	2.47	2.58	2.47	*1.89*
3rd	*0.62*	1.17	0.79	1.16	1.08	4.28	1.48	1.35	1.45	*1.28*	5.18	2.61	2.65	2.58	*2.25*
4th	0.60	0.56	*0.39*	0.68	0.59	*1.99*	3.22	2.26	3.25	2.55	4.98	2.32	2.31	2.40	*2.05*
5th	1.09	0.45	*0.33*	0.52	0.41	1.81	*0.81*	1.08	0.90	1.18	3.72	1.71	1.68	1.65	*1.52*
6th	1.48	0.95	0.85	0.87	*0.67*	1.60	0.91	0.74	0.83	*0.66*	4.84	2.43	2.36	2.22	*2.14*
7th	1.37	1.21	1.13	*0.94*	1.00						4.94	2.32	2.28	*1.93*	2.00
8th	1.57	1.67	1.60	1.39	*1.15*						4.85	1.21	1.26	1.25	*0.90*
9th	1.74	1.69	1.67	1.32	*1.15*						4.58	2.49	2.52	2.39	*2.30*
10th	1.88	1.12	1.13	0.83	*0.67*						4.74	1.03	1.08	*0.91*	*0.91*
11th	1.14	1.05	1.06	0.92	*0.63*						4.79	0.69	0.72	0.74	*0.59*
12th	1.38	0.71	0.79	0.61	*0.24*						5.72	1.54	1.59	1.53	*1.40*
13th	1.05	0.34	0.38	0.40	*0.28*						5.02	1.55	1.58	1.62	*1.51*
14th	0.77	*0.37*	0.40	0.57	0.43						4.47	2.67	2.66	2.53	*2.49*
15th	0.99	0.71	*0.65*	1.79	1.07						2.98	1.76	1.75	*1.59*	1.63
16th	*0.73*	1.09	1.03	1.68	1.65						2.81	1.87	1.86	1.71	*1.69*

Month	Recession R4					Recession R5				
	DL	SV	SV-TR	MV	MV-TR	DL	SV	SV-TR	MV	MV-TR
Panel B: Recessions 4–5										
1st	1.30	0.46	*0.35*	0.40	0.38	1.79	0.84	3.45	*0.69*	3.75
2nd	0.97	0.90	0.93	0.91	*0.82*	2.52	0.64	0.65	0.60	*0.56*
3rd	1.02	0.71	0.72	0.73	*0.64*	2.98	0.71	0.72	0.78	*0.55*
4th	1.11	0.64	0.64	0.67	*0.59*	3.24	0.93	0.94	0.91	*0.74*
5th	1.52	0.58	0.58	0.59	*0.53*	3.71	0.78	0.79	0.75	*0.71*
6th	1.77	0.57	0.57	0.56	*0.51*	4.03	0.96	0.97	0.90	*0.86*
7th	1.86	0.86	0.86	0.87	*0.82*	4.72	1.29	1.30	1.32	*1.21*
8th	1.88	1.00	0.99	0.98	*0.95*	4.79	1.23	1.23	*1.17*	1.34

which the monetary authority reacts to term premia – the slope in particular – when setting the monetary policy instrument.[15]

The parameter estimates of the forward-looking Taylor rule are reported on Panel B in Table 5. Although only the former is statistically significant, the responses to CPI inflation and unemployment rate are consistent with the results in CGG, which uses 1960:1–1996:4 data at the quarterly – as opposed to monthly – frequency. The closest inflation target level to their estimates is given by FWTR1, although not significant, and the interest rate smoothing parameter is more persistent than theirs. The goodness-of-fit of the specifications is very similar and none of them is rejected when testing for overidentifying restrictions.

Forward-looking Taylor rules might give accurate descriptions in-sample, but if the aim is to do out-of sample forecasting, one needs backward-looking ones. Panel B in Table 5 also reports estimates for alternative specifications of backward-looking Taylor rules regarding the choice of l_{CPI} and l_{UR}, the horizons at which the monetary policy authority looks when setting the monetary policy instrument.

The results for the backward-looking Taylor rules are robust to alternative horizons, and suggest that – at least at the monthly frequency – the monetary authority looks mostly at past inflation and past values of the monetary policy instrument when setting its current value. The persistence in the Fed funds rate is shown to be high, and even the non-significant parameters γ^{UR} and CPI^{*} tend to gravitate across a relatively narrow interval, at least for non-zero values of l_{CPI} and l_{UR}. The *J*-statistics suggest that the horizon at which the Fed looks is at least six months back. When compared to the forward-looking estimates, the responses to inflation seem to be tougher, and both the response to unemployment and the inflation target level are found not to be statistically significant.

5.5. Out-of-sample analysis

This Section reports the results of an out-of-sample study considering five episodes of economic interest: the five NBER-dated US recessions which have entirely occurred during the period 1970–2003. Recessions are of economic interest *per se* being bad states of nature, characterized by reduced economic activity and increased lay-off of workers, thus being events against which economic agents are willing to insure. Moreover, within the term structure literature, recessions are of interest for being periods which tend to be preceded by the inversion of the yield curve, a feature often difficult to be quickly captured – if at all – by term structure models, making the exercise both more interesting and challenging. The recessions considered are described in Table 6.[16]

For every month in each of the five recessions, I compare the forecasts of the alternative specifications using two measures of accuracy. I also report results for specifications SV-TR and MV-TR, which incorporate the Taylor rule in an attempt to improve forecasting ability.

[15] The McCallum interest-rate rule also allows rationalizing the empirical failure of the expectations hypothesis – see also Gallmeyer et al. (2005).

[16] The NBER-dated recession going from December 1969 to November 1970 is not considered here since the dataset starts on January, 1970.

Table 8
Maturity-disaggregated accuracy. The table reports the accuracy of alternative specifications disaggregated at the maturity level. The underlined quantities are the smaller values for a given time period and episode.

Maturity	Recession R1					Recession R2					Recession R3				
	DL	SV	SV-TR	MV	MV-TR	DL	SV	SV-TR	MV	MV-TR	DL	SV	SV-TR	MV	MV-TR
Panel A: Recessions 1–3															
1mo	2.45	0.90	0.96	0.90	*0.73*	2.89	1.61	2.32	*1.57*	2.17	3.73	*1.44*	1.89	1.49	1.73
2mo	1.36	0.86	0.90	0.86	*0.70*	2.57	1.60	2.30	*1.55*	2.14	1.64	*1.43*	1.88	1.48	1.71
3mo	*0.63*	0.83	0.85	0.82	0.67	2.81	1.57	2.26	*1.51*	2.09	*0.79*	1.44	1.90	1.48	1.71
4mo	0.95	0.80	0.81	0.80	*0.65*	3.09	1.52	2.20	*1.46*	2.02	1.65	*1.46*	1.91	1.51	1.71
5mo	1.49	0.57	0.57	*0.54*	0.58	3.42	1.45	2.12	*1.39*	1.95	2.72	*1.40*	1.85	*1.40*	1.60
6mo	1.93	0.60	*0.57*	0.58	0.58	3.67	1.33	1.98	*1.31*	1.81	3.57	1.48	1.91	*1.45*	1.66
7mo	2.23	0.62	*0.56*	0.63	0.59	3.94	1.25	1.86	*1.24*	1.69	4.17	1.55	1.97	*1.50*	1.71
8mo	2.36	0.66	0.56	0.66	*0.55*	4.12	*1.19*	1.74	*1.19*	1.58	4.58	1.51	1.93	*1.43*	1.66
9mo	2.37	0.71	0.59	0.71	*0.58*	4.19	1.15	1.63	*1.14*	1.49	4.84	1.57	1.97	*1.51*	1.72
10mo	2.27	0.75	0.63	0.76	*0.61*	4.21	1.13	1.54	*1.09*	1.40	5.01	1.62	2.01	*1.58*	1.77
11mo	2.12	0.78	0.65	0.80	*0.63*	4.22	*1.03*	1.49	1.10	1.37	5.11	1.68	2.05	*1.65*	1.82
12mo	1.94	0.81	*0.66*	0.90	0.72	4.22	*1.01*	1.47	1.13	1.33	5.18	1.78	2.14	*1.71*	1.83
24mo	1.27	0.91	*0.76*	1.00	*0.76*	3.06	1.02	0.97	1.04	*0.85*	4.57	1.37	1.74	*1.24*	1.40
36mo	0.84	0.91	0.80	1.01	*0.79*	2.80	*0.92*	0.96	0.93	1.03	5.01	2.16	2.32	*2.02*	2.06
48mo	*0.70*	1.02	0.86	1.03	0.77	1.89	1.85	*1.56*	1.89	1.68	3.63	1.66	1.90	*1.61*	1.71
60mo	*0.93*	2.13	1.95	2.21	2.01	4.00	3.38	2.95	3.45	*2.87*	*4.28*	4.61	4.81	4.57	4.69

Maturity	Recession R4					Recession R5				
	DL	SV	SV-TR	MV	MV-TR	DL	SV	SV-TR	MV	MV-TR
Panel B: Recessions 4–5										
1mo	6.90	0.68	0.79	*0.65*	0.75	7.01	1.36	1.73	*1.21*	2.02
2mo	5.77	0.71	0.82	*0.69*	0.78	6.45	1.42	1.79	*1.28*	2.07
3mo	4.83	0.71	0.82	*0.70*	0.78	5.98	1.45	1.82	*1.31*	2.09
4mo	4.07	0.70	0.80	*0.69*	0.76	5.59	1.47	1.84	*1.33*	2.10
5mo	3.47	*0.67*	0.78	*0.67*	0.72	5.28	1.47	1.87	*1.33*	2.09
6mo	3.00	*0.67*	0.77	*0.67*	0.71	5.03	1.51	1.87	*1.36*	2.10
7mo	2.64	*0.66*	0.70	*0.66*	*0.66*	4.84	1.50	1.47	*1.35*	2.08
8mo	2.35	0.39	0.45	*0.33*	0.39	4.69	1.11	1.47	*0.99*	1.71
9mo	2.13	0.39	0.45	*0.35*	0.39	4.57	1.10	1.47	*0.97*	1.69
10mo	1.93	0.40	0.44	*0.37*	0.38	4.45	1.08	1.47	*0.94*	1.65
11mo	1.77	0.41	0.44	0.41	*0.39*	4.35	1.08	1.44	*0.93*	1.63
12mo	1.63	0.43	0.45	0.44	*0.40*	4.24	1.06	1.42	*0.91*	1.59
24mo	1.35	0.57	0.56	0.53	*0.48*	3.33	*0.54*	0.87	0.61	0.87
36mo	0.98	*0.47*	0.50	0.49	0.50	2.79	*0.78*	0.95	0.80	*0.78*
48mo	*0.42*	1.25	1.17	1.26	1.11	2.73	0.87	1.24	*0.85*	1.02
60mo	0.84	0.69	0.61	0.70	*0.60*	2.35	*1.51*	1.85	1.59	1.64

The ways I compute the out-of-sample forecasts are as follows. For the macro-based specifications, assume the estimation sample has observations from periods $t = 1, \ldots, T$, where $t = 1$ is January, 1970 and $t = T$ is the month preceding the recession of interest. After obtaining parameter estimates using the estimation sample, the yield curve forecast for period $t^* > T$, denoted as $\widehat{y}_{t^*|T}$ are obtained either from observed or estimated values M_{t^*} of the state variables using the previously estimated parameters – in the case of the SV and MV specifications, I keep the parameter estimates fixed and keep updating the matrix M_{t^*} of state variables every period.

When it comes to the SV-TR and MV-TR specifications, I estimate the Taylor rules as above, and just update the information on *CPI*, *UR*, and *FF* every period, thus obtaining a Taylor rule-based estimate of the value of *FF* the following period.

Finally, for the DL model, I estimate the model for every period $t = 1, \ldots, T$, compute the AR (1) processes describing the dynamics of each factor, and re-estimate the model at every period $t > t^*$.[17]

[17] See Diebold and Li (2006) for a thorough out-of-sample comparison of their model and previously existing ones.

As a measure of 'overall accuracy', I compute average MAEs for the entire duration of each recession, i.e. for every month t^* of a given recession, I compute

$$OA_{t^*} = \frac{1}{N_{t^*}} \sum_{j=1}^{N_{t^*}} \left|\hat{y}_{t^*}(\tau_j) - y_{t^*}(\tau_j)\right|, \qquad t^* \in \text{Recession}$$

where N_{t^*} refers to the number of yield at period t^* in the recession. The results reported in Table 7.

The results in Table 7 show that the macro-based specifications consistently outperform the latent variable model. In fact, DL cannot beat its competitors for any month in Recessions R3–5. When comparing SV and MV specifications, the former tends to perform better in the first two or three months of the recessions, being then outperformed by the latter. This suggests that it might take time for all the state variables to work in favour of the MV specification in such periods.

The results for R1 show the dominance of the MV-TR model, especially during the second half of recession R1. Its performance is followed by the SV-TR model, which suggests that Taylor rules convey information about the future state of the term structure. It also shows the potential effect of a change in policy regime on the forecasting ability of the TR specifications – R2 was the first recession following the monetary policy experiment, right after its introduction.[18] As a result, the SV and MV-TR models perform closely, and the Taylor rule does not seem to provide a substantial gain to the models incorporating it.

The results for Recessions R3–5 show a clear dominance of the MV-TR specification, which might suggest two things. First, that SV specifications are way too simple to describe the term structure dynamics. Second, that incorporating Taylor rules does indeed play a role, improving the accuracy of the forecasts.

As a measure of 'maturity-disaggregated accuracy', I report time-averaged MAEs for fixed maturities, i.e. for a given maturity τ_j I calculate

$$MDA_{\tau_j} = \frac{1}{\#t^*} \sum_{t^* \in \text{Recession}} \left| \hat{y}_{t^*}(\tau_j) - y_t(\tau_j) \right|, \qquad j = 1, \ldots, N$$

where $\#t^*$ denotes the number of periods in the recession. The results are reported in Table 8.

Table 8 reports results that confirm the view that macro-based specifications outperform the benchmark DL model. For R1, the results suggest a superior performance of the MV-TR specification up to the 36-month maturity, after which the DL specification tends to do better.

The results for Recessions R2–3 show the superior performance of the MV and, to a lesser extent, SV specifications, most likely due to the change in the policy regime resulting from the monetary policy experiment. For R4–5, there seems to be a dominance of the MV specification, at least for maturities up to 10–12 months. For Recession R4, the better performing specification from the 11-month maturity towards the long end of the curve is MV-TR, whereas for Recession R5 it is specification SV which performs better between the 24- and 60-month maturities.

Although it is not obvious which macro-based specification performs best throughout the exercises, all of them consistently outperform the latent-variable benchmark.

6. Conclusion

This paper proposes a term structure model with factors driven uniquely by observable – as opposed to latent – state variables. The explicit link between the term structure factors and the state variables allows comparing alternative views on the drivers of its dynamics and competing economic hypotheses.

The method is robust to curse of dimensionality issues commonly appearing in the literature. This happens because instead of increasing with the number of observations (yields) used, the dimension of the parameter vector increases with the number of state variables. As a result, the method is in a position to deliver more accurate measures of the curvature factor, thus better explaining intermediate maturities i.e. the 'belly' of the curve.

The estimation method uses both the cross-sectional and time series dimensions of the data, which results in faster convergence of the parameter estimates and more reliable inference. This is in stark contrast with VAR models, which are subject to the criticism that they make researchers choose between either short time series or misspecified models, thus making inference unreliable – a direct consequence of the frequent changes in policy regimes in the postwar period (Rudebusch, 1998).

The empirical exercise uses a comprehensible set of US macroeconomic data to compare alternative specifications of the term structure. In the in-sample study, the baseline (SV) specification is such that the level, slope and curvature factors are driven by, respectively, measures of inflation (CPI growth), monetary policy (the Fed Funds rate), and economic activity (the unemployment rate). The out-of-sample study compares macro-based models to a latent-variable benchmark model for five NBER-dated recessions in the last three decades, showing that the former consistently outperforms the latter.

This paper raises a number of questions for future research. First, how does the method perform using alternatives such as expectations variables obtained in consensus forecasts as state variables. Second, how it performs as a risk management tool, which should be relevant to financial institutions and regulators. Third, how it performs when coupled with VAR models feeding it with macroeconomic variables, or measures such as the Bernanke and Mihov (1998) monetary policy indicator.

Despite its reliance on the Nelson–Siegel family of curves, it is by no means restricted to it. Nelson–Siegel curves are used here due to their intuitive appeal, well-known properties, and the common understanding that they are a reasonable first-order approximation to the yield curve. Alternative methods can be used and are left for future research.

Acknowledgements

I thank Caio Almeida, Marcelo Fernandes, Antonio Mele, Alex Michaelides, Richard Payne, Alberto Salvo, and Sarquis J. B. Sarquis for insightful comments and helpful discussions as well as Oliver Linton for advice and audiences at the London School of Economics, the 2006 North American Summer Meetings of the Econometric Society, Bank of England, Bocconi, Tilburg, Stockholm School of Economics and Cambridge for their feedback. Any remaining errors are my own.

Appendix A. Covariance matrix derivation

As in the text, $\nabla_\theta = (\nabla_{\bar{\beta}}, \nabla_{\sigma_\beta}, \nabla_{\bar{\lambda}}, \nabla_{\sigma_\lambda})$, where

$$\nabla_{\bar{\beta}} = X(\theta_\lambda)$$
$$\nabla_{\sigma_\beta} = XM(\theta_\lambda)$$
$$\nabla_{\bar{\lambda}} = \frac{\partial X(\theta_\lambda)}{\partial \bar{\lambda}} \bar{\beta} = \begin{bmatrix} \frac{\partial X_1(\theta_\lambda)}{\partial \bar{\lambda}} \\ \frac{\partial X_t(\theta_\lambda)}{\partial \bar{\lambda}} \\ \frac{\partial X_T(\theta_\lambda)}{\partial \bar{\lambda}} \end{bmatrix} \bar{\beta}$$

with general element

$$\frac{\partial X_t(\theta_\lambda)}{\partial \bar{\lambda}} = \begin{bmatrix} 0 & \phi_{1t-} & \phi_{1t-} + \exp(-(\bar{\lambda} + M_{\lambda t-}\sigma_\lambda)\tau_1) \\ 0 & \phi_{2t-} & \phi_{2t-} + \exp(-(\bar{\lambda} + M_{\lambda t-}\sigma_\lambda)\tau_2) \\ \cdots & \cdots & \cdots \\ 0 & \phi_{Nt-} & \phi_{Nt-} + \exp(-(\bar{\lambda} + M_{\lambda t-}\sigma_\lambda)\tau_N) \end{bmatrix}$$

and

$$\phi_{it-} = \frac{\exp(-(\bar{\lambda} + M_{\lambda t-}\sigma_\lambda)\tau_i)}{(\bar{\lambda} + M_{\lambda t-}\sigma_\lambda)\tau_i} \bar{\lambda} - \frac{1 - \exp(-(\bar{\lambda} + M_{\lambda t-}\sigma_\lambda)\tau_i)}{[(\bar{\lambda} + M_{\lambda t-}\sigma_\lambda)\tau_i]^2} \tau_i,$$
$$i = 1, \ldots, N$$

Finally,

$$\nabla_{\sigma_\lambda} = \frac{\partial X(\theta_\lambda)}{\partial \bar{\lambda}} \begin{bmatrix} M_{\beta_1}\sigma_\beta \\ M_{\beta_2}\sigma_\beta \\ \cdots \\ M_{\beta_N}\sigma_\beta \end{bmatrix} \begin{bmatrix} M_{\lambda_1} \\ M_{\lambda_2} \\ \cdots \\ M_{\lambda N} \end{bmatrix}$$

[18] See Clarida et al. (2000) for a study of how the Taylor rule changed with this regime change.

References

Ang, A., Piazzesi, M., 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. Journal of Monetary Economics 50, 745–787.

Balduzzi, P., Das, S.R., Foresi, S., Sundaram, R., 1996. A simple approach to three-factor affine models of the term structure. Journal of Fixed Income 6, 43–53.

Bansal, R., Yaron, A., 2004. Risks for the long run: a potential resolution of asset pricing puzzles. Journal of Finance 59, 1481–1509.

Bansal R., Kiku, D., Yaron, A., 2007. Risks for the Long Run: Estimation and Inference. Working Paper. Duke University.

Bernanke, B.S., Mihov, I., 1998. Measuring monetary policy. Quarterly Journal of Economics 113, 869–902.

Bikbov, R., Chernov, M., 2010. No-arbitrage macroeconomic determinants of the yield curve. Journal of Econometrics 159, 166–182.

Bliss, R.R., 1997. Movements in the Term Structure of Interest Rates. Federal Reserve Bank of Atlanta Economic Review 82, 16–33.

Campbell, J.Y., 1986. Bond and stock returns in a simple exchange model. Quarterly Journal of Economics 101, 785–804.

Campbell, J.Y., 1993. Intertemporal asset pricing without consumption data. American Economic Review 83, 487–512.

Campbell, J.Y., 1996. Understanding risk and return. Journal of Political Economy 104, 298–345.

Campbell, J.Y., 1999. Asset prices, consumption and the business cycle. In: Taylor, J.B., Woodford, M. (Eds.), The Handbook of Macroeconomics. Elsevier Science, North-Holland, Amsterdam.

Christensen, J.H.E., Diebold, F.X., Rudebusch, G.D., 2009. An arbitrage-free generalized Nelson–Siegel term structure model. The Econometrics Journal 1, 1–31.

Christensen, J.H.E., Diebold, F.X., Rudebusch, G.D., 2010. The affine arbitrage-free class of Nelson–Siegel term structure models. Journal of Econometrics, forthcoming.

Clarida, R., Galı, J., Gertler, M., 2000. Monetary policy rules and macroeconomic stability: evidence and some theory. Quarterly Journal of Economics 115, 147–180.

Cochrane, J., Piazzesi, M., 2008. Decomposing the Yield Curve. Working Paper. University of Chicago.

Cox, J.C., Ingersoll, J.E., Ross, S.A., 1985. A theory of the term structure of interest rates. Econometrica 53, 385–407.

Dai, Q., Philippon, T., 2005. Fiscal Policy and the Term Structure of Interest Rates. Working Paper. Stern School of Business.

Dai, Q., Singleton, K.J., 2000. Specification analysis of affine term structure models. Journal of Finance 55, 1943–1978.

Dewachter, H., Lyrio, M., 2006. Macro factors and the term structure of interest rates. Journal of Money, Credit and Banking 38, 119–140.

Diebold, F.X., Li, C., 2006. Forecasting the term structure of government bond yields. Journal of Econometrics 130, 337–364.

Diebold, F.X., Piazzesi, M., Rudebusch, G.D., 2005. Modeling bond yields in finance and macroeconomics. American Economic Review 95, 415–420.

Duffee, G.R., 2002. Term premia and interest rate forecasts in affine models. Journal of Finance 57, 405–443.

Duffee, G.R., 2008. Forecasting with the Term Structure: The Role of No-arbitrage. Working Paper. UC Berkeley.

Duffee, G.R., 2011. Information in (and not in) the term structure. Review of Financial Studies, forthcoming.

Dunn, K., Singleton, K.J., 1986. Modeling the term structure of interest rates under non-separable utility and durability of goods. Journal of Financial Economics 17, 27–55.

Epstein, L.G., Zin, S.E., 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. Econometrica 57, 937–968.

Evans, C.L., 1998. Real-time Taylor rules and the federal funds futures market. Federal Reserve Bank of Chicago Economic Perspectives 22, 44–55.

Evans, C.L., Marshall, D.A., 1996. Economic Determinants of the Term Structure of Nominal Interest Rates. Working Paper. Federal Reserve Bank of Chicago.

Evans, C.L., Marshall, D.A., 1998. Monetary policy and the term structure of nominal interest rates: evidence and theory. Carnegie-Rochester Conference Series on Public Policy 49, 53–111.

Evans, C.L., Marshall, D., 2002. Economic Determinants of the Nominal Treasury Yield Curve. Federal Reserve Bank of Chicago working paper.

Fuhrer, J., 1996. Monetary policy shifts and long-term interest rates. Quarterly Journal of Economics 111, 1183–1209.

Fuhrer, J., Moore, G., 1995. Monetary policy trade-offs and the correlation between nominal interest rates and real output. American Economic Review 85, 219–239.

Gallmeyer, M.F., Hollifield, B., Zin, S.E., 2005. Taylor rules, McCallum rules, and the term structure of interest rates. Journal of Monetary Economics 52, 921–950.

Hall, R.E., 1988. Intertemporal substitution in consumption. Journal of Political Economy 96, 339–357.

Hördahl, P., Tristani, O., Vestin, D., 2006. A joint econometric model of macroeconomic and term structure dynamics. Journal of Econometrics 131, 405–444.

Hull, J., White, A., 1990. Pricing interest-rate derivative securities. Review of Financial Studies 3, 573–592.

Joslin, S., Singleton, K.J., Zhu, H., 2011. A new perspective on Gaussian DTSMs. Review of Financial Studies 24, 926–970.

Longstaff, F., Schwartz, E., 1992. Interest rate volatility and the term structure: a two-factor general equilibrium model. Journal of Finance 47, 1259–1282.

McCallum, B.T., 2005. Monetary policy and the term structure of interest rates. Federal Reserve Bank of Richmond Economic Quarterly 91, 1–21.

McCulloch, J., 1971. Measuring the term structure of interest rates. Journal of Business 44, 19–31.

McCulloch, J., 1975. The tax-adjusted yield curve. Journal of Finance 30, 811–830.

Mehra, R., Prescott, E.C., 1985. The equity premium: a puzzle. Journal of Monetary Economics 15, 145–161.

Nelson, C.R., Siegel, A.F., 1987. Parsimonious modelling of yield curves. Journal of Business 60, 473–489.

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Piazzesi, M., 2005. Bond yields and the Federal Reserve. Journal of Political Economy 113, 311–344.

Piazzesi, M., Schneider, M., 2006. Equilibrium yield curves. In: Woodford, M., Acemoglu, D., Rogoff, K. (Eds.), NBER Macroeconomics Annual. MIT Press, Cambridge.

Podivinsky, J.M., 1999. Finite sample properties of GMM estimators and tests. In: Matyas, L. (Ed.), Generalized Method of Moments Estimation. Cambridge University Press, Cambridge.

Rudebusch, G.D., 1995. Federal Reserve interest rate targeting, rational expectations, and the term structure. Journal of Monetary Economics 35, 245–274.

Rudebusch, G.D., 1998. Do measures of monetary policy in a VAR make sense? International Economic Review 39, 907–931.

Rudebusch, G.D., 2002. Term structure evidence on interest rate smoothing and monetary policy inertia. Journal of Monetary Economics 49, 1161–1187.

Rudebusch, G.D., Wu, T., 2008. Macro-finance model of the term structure, monetary policy, and the economy. Economic Journal 118, 906–926.

Schaefer, S., 1981. Measuring a tax-specific term structure of interest rates in the market for British government securities. Economic Journal 91, 415–431.

Taylor, J.B., 1993. Discretion versus policy rules in practice. Carnegie-Rochester Conference Series on Public Policy 39, 195–214.

Vasicek, O., 1977. An equilibrium characterization of the term structure. Journal of Financial Economics 5, 177–188.

Vasicek, O., Fong, H., 1982. Term structure modelling using exponential splines. Journal of Finance 37, 339–341.

Vissing-Jorgensen, A., Attanasio, O.P., 2003. Stock market participation, intertemporal substitution and risk aversion. American Economic Review 93, 383–391.

Weil, Ph., 1989. The equity premium puzzle and the risk-free rate puzzle. Journal of Monetary Economics 24, 401–421.

Woodford, M., 2001. The Taylor rule and optimal monetary policy. American Economic Review 91, 232–237.