Journal of Econometrics 146 (2008) 26–43

Contents lists available at ScienceDirect

Journal of Econometrics

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

Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach☆

Emanuel Moench *

Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045, United States

ARTICLE INFO

Article history:
Received 24 November 2005
Received in revised form
5 May 2007
Accepted 27 June 2008
Available online 6 July 2008

JEL classification:
C13
C32
E43
E44
E52

Keywords:
Yield curve
Factor-augmented VAR
Affine term structure models
Dynamic factor models
Forecasting

ABSTRACT

This paper suggests a term structure model which parsimoniously exploits a broad macroeconomic information set. The model uses the short rate and the common components of a large number of macroeconomic variables as factors. Precisely, the dynamics of the short rate are modeled with a Factor-Augmented Vector Autoregression and the term structure is derived using parameter restrictions implied by no-arbitrage. The model has economic appeal and provides better out-of-sample yield forecasts at intermediate and long horizons than a number of previously suggested approaches. The forecast improvement is highly significant and particularly pronounced for short and medium-term maturities.

Published by Elsevier B.V.

1. Introduction

Traditional models of the term structure decompose yields into a set of latent factors. These models commonly provide a good in-sample fit to the data (e.g. Nelson and Siegel (1987), Knez et al. (1994) and Dai and Singleton (2000)) and can also be used to predict interest rates out-of-sample (e.g. Duffee (2002) and Diebold and Li (2006)). While providing a good statistical fit, however, the economic meaning of such models is limited since they disregard the relationships between macroeconomic variables and interest rates. In this paper, I suggest a model which has both economic appeal and superior predictive ability for yields as compared to traditional approaches.

In a widely recognized paper, Ang and Piazzesi (2003) augment a standard three-factor affine term structure model with two macroeconomic factors that enter the model through a Taylor-rule type of short rate equation. They find that the macro factors account for a large share of the variation in interest rates and also improve yield forecasts. Inspired by this finding, a vivid literature has emerged lately that explores different approaches to jointly model the term structure and the macroeconomy. Examples for such models are Hördahl et al. (2006), Diebold et al. (2006) and Dewachter and Lyrio (2006). While these latter studies consistently find that macroeconomic variables are useful for explaining and/or forecasting government bond yields, they only exploit very small macroeconomic information sets. Yet, by limiting the analysis to only a few variables, other potentially useful macroeconomic information is being neglected.[1]

This is particularly important for term structure modeling as a recent literature argues that the central bank acts in a “data-rich environment” (Bernanke and Boivin, 2003). This means that the monetary policy authority bases its decisions upon a broad set of conditioning information rather than only a few key aggregates. Consistent with this argument, a number of studies have found that factors which by construction summarize the comovement in

☆ The views expressed in this article are those of the author and do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System.

* Tel.: +1 212 720 6625; fax: +1 212 720 1291.
E-mail address: emanuel.moench@ny.frb.org.

[1] Note that the macroeconomic factors in Ang and Piazzesi (2003) are the principal components extracted from a group of four real and three nominal variables, respectively. Accordingly, these authors employ a somewhat larger macroeconomic information set than the other studies referred to above.

0304-4076/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.jeconom.2008.06.002

a large number of macroeconomic time series help to explain and forecast the evolution of short-term interest rates (e.g. Bernanke and Boivin (2003), Giannone et al. (2004) and Favero et al. (2005)). In related work, Bernanke et al. (2005) suggest to combine the advantages of factor modeling and structural VAR analysis by estimating a joint vector-autoregression of the short-term interest rate and factors extracted from a large cross-section of macro time series. They label this approach a "Factor-Augmented VAR" (FAVAR) and use it to analyze the dynamics of the short rate and the effects of monetary policy on a wide range of macroeconomic variables.

In this paper, I take the approach of Bernanke et al. (2005) a step further and employ the FAVAR model to study the dynamics of the entire yield curve within an arbitrage-free model. Precisely, I suggest a model that has the following structure. A Factor-Augmented VAR is used to describe the dynamics of the short-term interest rate conditional on a large macroeconomic information set. Given the dynamics of the short rate, the term structure of interest rates is then derived using parameter-restrictions implied by no-arbitrage. In sum, my model is an affine term structure model that has a Factor-Augmented VAR as the state equation, i.e. the short rate and the common components of a large number of macro time series represent the factors which drive the variation of yields. I label this approach a No-Arbitrage Factor-Augmented VAR.

Estimation of the model is in two steps. First, I extract common factors from a large macroeconomic dataset using the method suggested by Stock and Watson (2002a,b) and estimate the parameters governing their joint dynamics with the monetary policy instrument in a VAR. Second, I estimate a no-arbitrage vector autoregression of yields on the exogenous pricing factors. Specifically, I obtain the price of risk parameters by minimizing the sum of squared fitting errors of the model following the nonlinear least squares approach of Ang et al. (2006). Altogether, estimation of the model is fast and it is thus particularly useful for recursive out-of-sample forecasts.

The results of the paper can be summarized as follows. The No-Arbitrage FAVAR model based on four macro factors and the short rate fits the US yield curve well in-sample. More importantly, the model shows a strikingly good ability to predict yields out-of-sample. In a recursive out-of-sample forecast exercise, the No-Arbitrage FAVAR model is found to provide superior forecasts with respect to a number of benchmark models which have previously been suggested in the literature. Except for extremely short forecast horizons and very long maturities, the model significantly outperforms the random walk, a standard three-factor affine model, the model suggested by Bernanke et al. (2004) which employs individual macroeconomic variables as factors, and the model recently put forth by Diebold and Li (2006) which has been documented to be particularly useful for interest rate predictions. A subsample analysis reveals that the No-Arbitrage Factor-Augmented VAR model performs particularly well in periods when interest rates vary a lot.

The paper is structured as follows. In Section 2, the No-Arbitrage Factor-Augmented VAR model is presented and its parametrization discussed. Section 3 describes the estimation of the model. In Section 4, I document the in-sample fit of the model and then discuss the results of the out-of-sample forecasts in Section 5. Section 6 concludes.

2. The model

Economists typically think of the economy as being affected by monetary policy through the short term interest rate. At the same time, the central bank is often assumed to set the short rate as a function of the overall state of the economy, characterized e.g. by the deviations of inflation and output from their desired levels. Bernanke et al. (2005) point out that theoretical macroeconomic aggregates as output and inflation might not be perfectly observable neither to the policy-maker nor to the econometrician. Instead, they argue that the observed macroeconomic time series should be thought of as noisy measures of economic concepts such as aggregate activity or inflation. Accordingly, these concepts should be treated as unobservable in empirical work so as to avoid confounding measurement error or idiosyncratic dynamics with fundamental economic shocks.

Bernanke et al. (2005) therefore suggest to extract a few common factors from a large number of macroeconomic time series variables and to study the mutual dynamics of monetary policy and the key economic aggregates by estimating a joint VAR of the factors and the policy instrument, an approach which they label "Factor-Augmented VAR" (FAVAR). This approach can be summarized by the following equations:

$$X_t = \Lambda_F F_t + \Lambda_r r_t + e_t \tag{1}$$

$$\begin{pmatrix} F_t \\ r_t \end{pmatrix} = \mu + \Phi(L) \begin{pmatrix} F_{t-1} \\ r_{t-1} \end{pmatrix} + \omega_t. \tag{2}$$

X_t denotes a $M \times 1$ vector of period-t observations of the observed macroeconomic variables, Λ_F and Λ_r are the $M \times k$ and $M \times 1$ matrices of factor loadings, r_t denotes the short-term interest rate, F_t is the $k \times 1$ vector of period-t observations of the common factors, and e_t is an $M \times 1$ vector of idiosyncratic components. Moreover, $\mu = (\mu_f', \mu_r)'$ is a $(k+1) \times 1$ vector of constants, $\Phi(L)$ denotes the $(k+1) \times (k+1)$ matrix of order-p lag polynomials and ω_t is a $(k+1) \times 1$ vector of reduced form shocks with variance covariance matrix Ω. Since affine term structure models are commonly formulated in state-space from, I rewrite the FAVAR in Eq. (2) as

$$Z_t = \boldsymbol{\mu} + \boldsymbol{\Phi} Z_{t-1} + \boldsymbol{\omega}_t, \tag{3}$$

where $Z_t = (F_t', r_t, F_{t-1}', r_{t-1}, \ldots, F_{t-p+1}', r_{t-p+1})'$, and where $\boldsymbol{\mu}$, $\boldsymbol{\Phi}$, $\boldsymbol{\omega}$ and $\boldsymbol{\Omega}$ denote the companion form equivalents of μ, Φ, ω, and Ω, respectively. Accordingly, the short rate r_t can be expressed in terms of Z_t as $r_t = \delta' Z_t$ where $\delta' = (\mathbf{0}_{1\times k}, 1, \mathbf{0}_{1\times(k+1)(p-1)})$.

2.1. Adding the term structure

The term structure model which I suggest is built upon the idea that the Federal Reserve bases its decisions on a large set of conditioning information and that the dynamics of the short-term interest rate are therefore well described by a Factor-Augmented VAR. Accordingly, yields are driven by the policy instrument as well as the main shocks hitting the economy which are proxied by the factors that capture the bulk of common variation in a large number of macroeconomic variables. I thus employ the FAVAR in Eq. (3) as the state equation of my term structure model. To make the model consistent with the assumption of no-arbitrage, I further impose restrictions on the parameters governing the impact of the state variables on the yields of different maturity. More precisely, I model the nominal pricing kernel as

$$\begin{aligned} M_{t+1} &= \exp\left(-r_t - \frac{1}{2}\lambda_t' \Omega \lambda_t - \lambda_t' \omega_{t+1}\right), \\ &= \exp\left(-\delta' Z_t - \frac{1}{2}\lambda_t' \Omega \lambda_t - \lambda_t' \omega_{t+1}\right), \end{aligned} \tag{4}$$

where λ_t are the market prices of risk. Following Duffee (2002), these are commonly assumed to be affine in the underlying state variables Z, i.e.

$$\lambda_t = \lambda_0 + \lambda_1 Z_t. \tag{5}$$

In order to keep the model parsimonious, I restrict the prices of risk to depend only on contemporaneous observations of the model factors. In an arbitrage-free market, the price of a n-months to maturity zero-coupon bond in period t must equal the expected discounted value of the price of an $(n-1)$-months to maturity bond in period $t+1$:

$$P_t^{(n)} = E_t[M_{t+1}\, P_{t+1}^{(n-1)}].$$

Assuming that yields are affine in the state variables, bond prices $P_t^{(n)}$ are exponential linear functions of the state vector:

$$P_t^{(n)} = \exp\left(A_n + B_n' Z_t\right),$$

where the scalar A_n and the coefficient vector B_n depend on the time to maturity n. Following Ang and Piazzesi (2003), I show in Appendix A that no-arbitrage is guaranteed by computing coefficients A_n and B_n according to the following recursive equations:

$$A_n = A_{n-1} + B_{n-1}'(\boldsymbol{\mu} - \boldsymbol{\Omega}\lambda_0) + \frac{1}{2} B_{n-1}' \boldsymbol{\Omega} B_{n-1}, \tag{6}$$

$$B_n' = B_{n-1}'(\boldsymbol{\Phi} - \boldsymbol{\Omega}\lambda_1) - \delta'. \tag{7}$$

Given the price of an n-months to maturity zero-coupon bond, the corresponding yield is thus obtained as

$$\begin{aligned} y_t^{(n)} &= -\frac{\log P_t^{(n)}}{n} \\ &= a_n + b_n' Z_t, \end{aligned} \tag{8}$$

where $a_n = -A_n/n$ and $b_n' = -B_n'/n$.

Altogether, the suggested model is completely characterized by Eqs. (1), (3) and (6)–(8). In a nutshell, it is an essentially affine term structure model that has a FAVAR as the state equation. Accordingly, I will refer to my model as a "No-Arbitrage Factor-Augmented VAR" approach.

3. Estimation of the model

In principle, the Factor-Augmented VAR model can be estimated using the Kalman filter and maximum likelihood. However, this approach becomes computationally infeasible when the number of macro variables stacked in the vector X is large. Bernanke et al. (2005) therefore discuss two alternative estimation methods: a single-step approach using Markov Chain Monte Carlo (MCMC) methods, and a two-step approach in which first principal components techniques are used to estimate the common factors F and then the parameters governing the dynamics of the state equation are obtained via standard classical methods for VARs. Comparing both methods in the context of an analysis of the effects of monetary policy shocks, Bernanke et al. (2005) find that the two-step approach yields more plausible results. Another advantage of this method is its computational simplicity. Since recursive out-of-sample yield forecasts are the main focus of this paper, I therefore employ the principal components based approach in my application of the FAVAR model.

Accordingly, the common factors have to be extracted from the panel of macro data prior to estimating the term structure model. As in Bernanke et al. (2005), this is achieved using standard static principal components following the approach suggested by Stock and Watson (2002a,b). Precisely, let V denote the eigenvectors corresponding to the k largest eigenvalues of the $T \times T$ cross-sectional variance–covariance matrix XX' of the data. Then, subject to the normalization $F'F/T = I_k$, estimates $\hat{F}$ of the factors and $\hat{\Lambda}$ the factor loadings are given by

$$\hat{F} = \sqrt{T}\, V \quad \text{and}$$
$$\hat{\Lambda} = \sqrt{T}\, X'V,$$

i.e. the common factors are estimated as the eigenvectors corresponding to the k largest eigenvalues of the variance-covariance matrix XX'.[2] In practice, the true number of common factors which capture the common variation in the panel X is not known. Bai and Ng (2002) have proposed some panel information criteria which allow to consistently estimate the number of factors. However, in the application of the FAVAR approach suggested here, the number of factors that can feasibly be included in the model is limited due to computational constraints imposed by the market prices of risk. I therefore fix the number of factors instead of applying formal model selection criteria.

Given the factor estimates, estimation of the term structure model is performed using the consistent two-step approach of Ang et al. (2006) which has also been employed in Bernanke et al. (2004). First, estimates of the parameters $(\boldsymbol{\mu}, \boldsymbol{\Phi}, \boldsymbol{\Omega})$ governing the dynamics of the model factors are obtained by running a VAR(p) on the estimated factors and the short term interest rate. Second, given the estimates from the first step, the parameters λ_0 and λ_1 which drive the evolution of the state prices of risk, are estimated by minimizing the sum of squared fitting errors of the model. That is, for a given set of parameter estimates $(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Phi}}, \hat{\boldsymbol{\Omega}})$, the model-implied yields $\hat{y}_t^{(n)} = \hat{a}_n + \hat{b}_n' Z_t$ are computed and the sum S is minimized with respect to λ_0 and λ_1 where S is given by[3]

$$S = \sum_{t=1}^{T} \sum_{n=1}^{N} (\hat{y}_t^{(n)} - y_t^{(n)})^2. \tag{9}$$

Due to the recursive formulation of the bond pricing parameters, S is highly nonlinear in the underlying model parameters. It is thus helpful to find good starting values so as to achieve fast convergence. This is done as follows. I first estimate the parameters λ_0 assuming that risk premia are constant but nonzero, i.e. I set to zero all elements of the matrix λ_1 which governs the time-varying component of the market prices of risk. I then take these estimates of λ_0 as starting values in a second step that allows for time-varying market prices of risk, i.e. I let all elements of λ_0 and λ_1 be estimated freely.

This two-step approach potentially gives rise to an errors-in-variables bias since the estimation of the market price of risk parameters takes as given the estimated evolution of the states. To adjust for this bias, I compute standard errors for λ_0 and λ_1 using a Monte Carlo procedure which is described in Appendix B.

4. Empirical results

4.1. Data

I estimate the model using the following data. The macroeconomic factors are extracted from a dataset which contains about 160 monthly time series of various economic categories for the US. Among others, it includes a large number of time series related

[2] To account for the fact that r is an observed factor which is assumed unconditionally orthogonal to the unobserved factors F in the model (1), its effect on the variables in X has to be isolated from the impact of the latent factors F. This is achieved by regressing all variables in X onto r and extracting principal components from the residuals of these regressions.

[3] The assumption that only contemporaneous factor observations affect the market prices of risk implies a set of zero restrictions on the parameters λ_0 and λ_1. In particular, $\lambda_0 = (\tilde{\lambda}_0', 0_{1\times(k+1)(p-1)})'$ and $\lambda_1 = \begin{pmatrix} \tilde{\lambda}_1 & 0_{(k+1)\times(k+1)(p-1)} \\ 0_{(k+1)(p-1)\times(k+1)} & 0_{(k+1)(p-1)\times(k+1)(p-1)} \end{pmatrix}$ where $\tilde{\lambda}_0$ is of dimension $(k+1)$ and $\tilde{\lambda}_1$ is a $(k+1)\times(k+1)$ matrix. Hence, in practice only $\tilde{\lambda}_0$ and $\tilde{\lambda}_1$ need to be estimated. Bernanke et al. (2004) impose the same set of restrictions in the estimation of their no-arbitrage macro VAR model.

to industrial production, more than 30 employment-related variables, around 30 price indices and various monetary aggregates. It further contains different kinds of survey data, stock indices, exchange rates etc. This dataset has been compiled by Giannone et al. (2004) to forecast US output, inflation, and short term interest rates. Notice that I exclude all interest rate related series from the original panel used by Giannone et al. (2004). The reason is that if the factors of my arbitrage-free model were extracted from a dataset containing yields, restrictions would have to be imposed on the factor loading parameters in (1) so as to make them consistent with the assumption of no-arbitrage. This would imply a non-trivial complication of the estimation process. Accordingly, I exclude the interest rate related series and thus implicitly assume that the central bank does not take into account the information contained in yields when setting the short term rate. Notice also that this assumption implies that long-term interest rates do not affect the evolution of the macroeconomy in my model.[4]

The principal components estimation of the common factors in large panels of time series requires stationarity. I therefore follow Giannone et al. (2004) in applying different preadjustments to the time series in the dataset.[5] Finally, I standardize all series to have mean zero and unit variance.

I use data on zero-coupon bond yields of maturities 1, 3, 6, and 9 months, as well as 1, 2, 3, 4, 5, 7, and 10 years. All interest rates are continuously-compounded unsmoothed Fama–Bliss yields and have been constructed from US treasury bonds using the method outlined in Bliss (1997). I estimate and forecast the model over the post-Volcker disinflation period, i.e. from 1983:01 to the last available observation of the macro dataset, 2003:09.

Table 1
Share of variance explained by factors and factor loadings

Factor 1 (25.1% of total variance)	R^2
Index of IP: Total	0.84
Index of IP: Non-energy, total (NAICS)	0.84
Index of IP: Mfg (SIC)	0.84
Capacity utilization: Total (NAICS)	0.81
Index of IP: Non-energy excl CCS (NAICS)	0.80
Factor 2 (10.9% of total variance)	
CPI: All items less medical care	0.85
CPI: Commodities	0.83
CPI: All items (urban)	0.83
CPI: All items less shelter	0.82
CPI: All items ess food	0.79
Factor 3 (7.8% of total variance)	
CPI: Medical care	0.66
PCE prices: Total excl food and energy	0.48
PCE prices: Services	0.45
M1 (in mil of current $)	0.39
Loans and Securities @ all comm banks: Securities, U.S. govt (in mil of $)	0.37
Factor 4 (5.0% of total variance)	
Employment on nonag payrolls: Financial activities	0.27
Employment on nonag payrolls: Other services	0.23
Employment on nonag payrolls: Service-producing	0.19
Employment on nonag payrolls: Mining	0.18
Employment on nonag payrolls: Retail trade	0.17

This table summarizes R-squares of univariate regressions of the factors extracted from the panel of macro variables on all individual variables. For each factor, I list the five variables that are most highly correlated with it. Notice that the series have been transformed to be stationary prior to extraction of the factors, i.e. for most variables the regressions correspond to regressions on growth rates. The four factors together explain about 50% of the total variation of the time series in the panel.

4.2. Model specification

In the first step of the estimation procedure, I extract common factors from the large panel of macroeconomic time series using the principal components approach of Stock and Watson (2002a,b). Together, the first 10 factors explain about 70% of the total variance of all variables in the dataset. The largest contribution is accounted for by the first four factors, however, which together explain about 50% of the total variation in the panel. Table 1 lists the shares of variance explained by the first four factors as well as the time series in the panel that each of them is most strongly correlated with. Note, however, that the factors estimated by principal components are only identified up to a non-singular rotation and therefore do not have a structural economic interpretation.

As already discussed above, the number of factors that can be included in the No-Arbitrage FAVAR model is limited due to parameterization constraints imposed by the market price of risk specification. Indeed, unless further restrictions are imposed on the market prices of risk, the number of parameters to estimate in the second step of the estimation procedure increases quadratically with the number of factors. For the sake of parsimony, I therefore restrict the number of factors to the first four principal components extracted from the large panel of monthly time series and the short rate. Unreported results with smaller and larger number of factors have shown that this specification seems to provide the best tradeoff between estimability and model fit. A similar choice has to be made regarding the number of lags to include in the Factor-Augmented VAR which represents the state equation of my term structure model. Applying the Hannan–Quinn information criterion with a maximum lag of 12 months indicates an optimal number of four lags for the joint VAR of factors and the short rate. I therefore employ this particular specification for the in-sample estimation of the model. Note that in the recursive out-of-sample forecast exercise documented in Section 5, the lag length of the FAVAR is re-estimated each time a forecast is produced as it would have to be in the context of truly real-time predictions.

4.3. Preliminary evidence

Before estimating the term structure model subject to no-arbitrage restrictions, I run a set of preliminary regressions to check whether the extracted macro factors are potentially useful explanatory variables in a term structure model. First, I use a simple encompassing test to assess whether a factor-based policy reaction function provides a better explanation of monetary policy decisions than a standard Taylor-rule based on individual measures of output and inflation. I then perform unrestricted regressions of yields on the model factors.

4.3.1. Do factors explain the short rate better than output and inflation?

The use of the Factor-Augmented VAR as a model for the dynamic evolution of short-term interest rates has been justified with the argument that central banks base their monetary policy decisions on large sets of macroeconomic conditioning information rather than on individual measures of output and inflation alone. Whether this conjecture holds true empirically can be tested by comparing the fit of a standard Taylor-rule with that of a policy reaction function based on dynamic factors. Bernanke and Boivin (2003) present evidence in favor of this

4 As discussed in Rudebusch et al. (2006a), this assumption is consistent with the predictions of standard New-Keynesian models in which aggregate output is determined by a forward-looking IS curve and therefore only depends on expected future short-term real interest rates.

5 Though with a slight difference as regards the treatment of price series: instead of computing first differences of quarterly growth rates as in Giannone et al. (2004), I follow Ang and Piazzesi (2003) and compute annual inflation rates.

Table 2
Policy rule based on individual variables

c	ρ	ϕ_y	ϕ_π
−0.011	0.955	1.332	2.592
(0.078)	(0.017)	(0.627)	(0.850)

This table reports estimates for a policy rule with partial adjustment based on individual measures of output and inflation, i.e. $r_t = c + \rho r_{t-1} + (1-\rho)(\phi_y y_t + \phi_\pi \pi_t)$, where r denotes the federal funds rate, y the deviation of log GDP from its trend, and π the annual rate of GDP inflation. The sample period is 1983:01 to 2003:09. Standard errors are in parentheses. The R^2 of this regression is 0.967.

Table 3
Policy rule based on factors

c	ρ	ϕ_{F1}	ϕ_{F2}	ϕ_{F3}	ϕ_{F4}
0.198	0.957	0.115	0.076	−0.008	0.006
(0.088)	(0.016)	(0.025)	(0.031)	(0.025)	(0.026)

This table reports estimates for a policy rule with partial adjustment based on the four factors extracted from a large panel of macroeconomic variables, i.e. $r_t = c + \rho r_{t-1} + (1-\rho)(\phi_{F1}F1_t + \phi_{F2}F2_t + \phi_{F3}F3_t + \phi_{F4}F4_t)$, where r again denotes the federal funds rate and $F1$ to $F4$ the four macro factors extracted from a panel of about 160 monthly time series for the US. The sample period is 1983:01 to 2003:09. Standard errors are in parentheses. The R^2 of this regression is 0.97.

claim by showing that the fitted value of the federal funds rate from a factor-based policy reaction function is a significant additional regressor in an otherwise standard Taylor-rule equation. Alternatively, one can separately estimate the two competing policy reaction functions and then perform an encompassing test *à* la Davidson and MacKinnon (1993). This is the strategy adopted by Belviso and Milani (2005). I follow these authors and compare a standard Taylor rule with partial adjustment,[6]

$$r_t = c + \rho r_{t-1} + (1-\rho)(\phi_\pi \pi_t + \phi_y y_t),$$

to a policy reaction function based on the four factors which represent state variables in the No-Arbitrage FAVAR model,

$$r_t = c + \rho r_{t-1} + (1-\rho)\phi_F' F_t.$$

The results from both regressions are summarized in Tables 2 and 3. As indicated by the regression R^2s of 0.967 and 0.970, the factor-based policy rule fits the data slightly better than the standard Taylor rule.

The Davidson and MacKinnon (1993) encompassing test can be used to asses whether this improvement in model fit is statistically significant. Accordingly, I regress the federal funds rate onto the fitted values from both alternative specifications which yields the following result:

$$\begin{aligned} r_t &= 0.207\hat{r}_t^{\text{Taylor}} + 0.793\ \hat{r}_t^{\text{Factors}} \\ &= (0.186) \qquad (0.186). \end{aligned}$$

Hence, the coefficient on the standard Taylor rule is insignificant whereas the coefficient on the factor-based fitted federal funds rate is highly significant.[7] This result can be interpreted as evidence supporting the hypothesis that the Fed reacts to a broad macroeconomic information set.

[6] Inflation π is defined as the annual growth rate of the GDP implicit price deflator (GDPDEF). The output gap is measured as the percentage deviation of log GDP (GDPC96) from its trend (computed using the Hodrick–Prescott filter and a smoothing parameter of 14400). Both quarterly series have been obtained from the St. Louis Fed website and interpolated to the monthly frequency using the method described in Moench and Uhlig (2005). For the interpolation of GDP, I have used industrial production (INDPRO), total civilian employment (CE16OV) and real disposable income (DSPIC96) as related monthly series. CPI and PPI finished goods have been employed as related series for interpolating the GDP deflator.

[7] Unreported results have shown that this finding is robust to alternative specifications of both reaction functions using a larger number of lags of the policy instrument and the macro variables or factors.

Table 4
Correlation of model factors and yields

	$y^{(1)}$	$y^{(6)}$	$y^{(12)}$	$y^{(36)}$	$y^{(60)}$	$y^{(120)}$
Panel A: Contemporaneous correlation of factors and yields						
F1	0.243	0.318	0.351	0.382	0.389	0.379
F2	0.597	0.619	0.617	0.570	0.546	0.537
F3	0.150	0.153	0.161	0.270	0.340	0.407
F4	0.315	0.325	0.331	0.354	0.365	0.380
$y^{(1)}$	1.000	0.987	0.975	0.936	0.899	0.833
Panel B: Correlation of 1 month lagged factors and yields						
F1(−1)	0.296	0.365	0.393	0.409	0.409	0.393
F2(−1)	0.600	0.614	0.610	0.564	0.539	0.531
F3(−1)	0.145	0.152	0.161	0.269	0.342	0.411
F4(−1)	0.296	0.309	0.316	0.346	0.358	0.373
$y^{(1)}(-1)$	0.984	0.974	0.960	0.923	0.888	0.822
Panel C: Correlation of 6 Months lagged factors and yields						
F1(−6)	0.445	0.490	0.502	0.473	0.445	0.412
F2(−6)	0.549	0.535	0.521	0.496	0.479	0.470
F3(−6)	0.128	0.151	0.171	0.286	0.364	0.453
F4(−6)	0.285	0.308	0.318	0.343	0.351	0.342
$y^{(1)}(-6)$	0.899	0.880	0.865	0.850	0.829	0.779
Panel D: Correlation of 12 months Lagged Factors and Yields						
F1(−12)	0.548	0.567	0.564	0.502	0.455	0.390
F2(−12)	0.448	0.405	0.385	0.398	0.400	0.408
F3(−12)	0.145	0.186	0.205	0.303	0.378	0.479
F4(−12)	0.275	0.309	0.329	0.349	0.354	0.348
$y^{(1)}(-12)$	0.742	0.712	0.705	0.738	0.745	0.723

This table summarizes the correlation patterns between the yields and factors used for estimating the term structure model. $F1$, $F2$, $F3$ and $F4$ denote the macro factors extracted form the large panel of monthly economic time series for the US, $y^{(1)}$ to $y^{(120)}$ denote the yields of maturities 1-month to 10-years, respectively.

4.3.2. *Unrestricted estimation of the term structure model*

To get a first impression whether the factors extracted from the panel of macro variables also capture predictive information about interest rates of higher maturity, Table 4 summarizes the correlations between the yields and various lags of the factors of the No-Arbitrage FAVAR model. This table shows that the short rate is most strongly correlated with yields of any other maturity. Yet, the four macro factors extracted from the panel of monthly US time series also exhibit considerable correlation with interest rates of higher maturity. While the short rate is contemporaneously most strongly correlated with yields, the correlations between macro factors and yields tend to be higher for longer lags. This indicates that the factors extracted from the panel of macro data might be useful for forecasting interest rates.

To further explore the question whether the model factors have explanatory power for yields, Table 5 provides estimates of an unrestricted regression of yields of different maturities onto a constant, the four macro factors and the 1-month Tbill, i.e. it estimates the pricing equation for yields,

$$Y_t = A + BZ_t + u_t,$$

where no cross-equation restrictions are imposed on the coefficients A and B. The first observation to make is that the R^2 of these regressions are all very high. Together with the short rate, the four factors explain more than 95% of the variation in short yields, and still more than 85% of the variation in longer yields. Not surprisingly, the 1-month Tbill is the most highly significant explanatory variable for short maturity yields. However, in the presence of the macro factors its impact decreases sharply towards the long end of the maturity spectrum. This shows that the factors extracted from the large panel of macro variables exhibit strong explanatory power for longer yields and thus represent potentially useful state variables in a term structure model.

Table 5
Unrestricted regressions of yields on factors

	$y^{(6)}$	$y^{(12)}$	$y^{(36)}$	$y^{(60)}$	$y^{(120)}$
cst	0.65	1.04	2.29	3.18	4.58
	[3.47]	[3.58]	[7.58]	[10.65]	[12.90]
F1	0.23	0.34	0.45	0.50	0.52
	[5.23]	[4.83]	[6.21]	[6.93]	[7.25]
F2	0.19	0.26	0.26	0.30	0.45
	[3.63]	[2.81]	[1.95]	[2.12]	[2.88]
F3	0.04	0.08	0.37	0.55	0.72
	[1.43]	[1.82]	[4.93]	[6.32]	[6.10]
F4	0.10	0.15	0.26	0.33	0.44
	[3.53]	[3.01]	[2.57]	[2.75]	[2.96]
$y^{(1)}$	0.95	0.93	0.82	0.72	0.52
	[28.64]	[17.59]	[11.71]	[9.07]	[5.57]
$\bar{R}^2$	0.98	0.97	0.93	0.91	0.86

This table summarizes the results of an unrestricted VAR of yields of different maturities on the four macro factors extracted from the panel of economic time series, and the short rate. The estimation period is 1983:01 to 2003:09. t-values are in brackets.

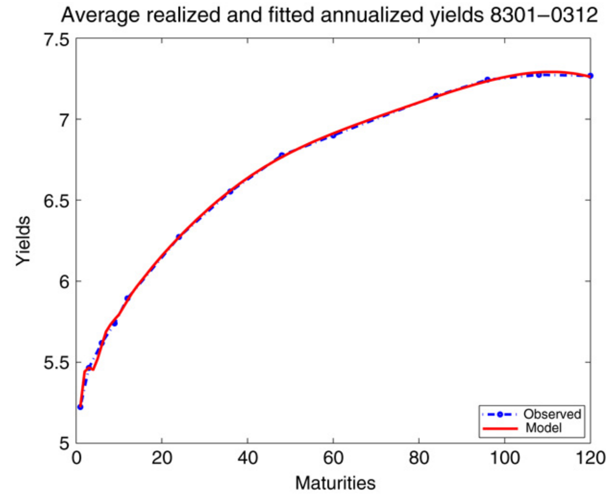

Fig. 1. Observed and model implied average yield curve. This figure plots average observed yields against those implied by the No-Arbitrage FAVAR model.

4.4. Estimating the term structure model

4.4.1. In-sample fit

In this section, I report results obtained from estimating the FAVAR model subject to the cross-equation restrictions (6) and (7) implied by the no-arbitrage assumption. The model fits the data surprisingly well, given that it does not involve any latent yield curve factors as traditional affine models. Table 6 reports the first and second moments of observed and model-implied yields. These numbers show that on average the No-Arbitrage FAVAR model provides a good fit to the yield curve. Fig. 1 provides a visualization of this result by showing the close match between average observed and model-implied yields across the entire maturity spectrum. Notice that the model seems to be missing some of the variation in longer maturities since the standard deviations of fitted interest rates are somewhat lower than the standard deviations of the observed yields, especially at the long end of the curve. This can also be seen in Fig. 2 which plots the time series for a selection of observed and model-implied yields.

Overall, the No-Arbitrage FAVAR model is able to capture the cross-sectional variation of government bond yields quite well, with a somewhat better in-sample fit at the short end of the curve. As will be shown further below, this finding is paralleled by the prediction results obtained from the model. In particular, the forecast improvement of the No-Arbitrage FAVAR model over latent yield factor based term structure models is found to be more pronounced at the short than at the long end of the yield curve.

4.4.2. Parameter estimates

The two-step procedure used to estimate the model implies a potential errors-in-variables bias since it takes as given the state evolution when estimating the market price of risk parameters. To adjust for this bias, I compute standard errors for the market price of risk parameters using a bootstrap procedure which is described in Appendix B.

Table 7 reports the parameter estimates and associated standard errors of the No-Arbitrage FAVAR model. The first panel shows parameter estimates of the Factor-Augmented VAR which represents the state equation of the model. The parameter estimates and corresponding standard errors have been obtained by standard OLS procedures. A noticeable feature of the FAVAR estimates is that most of the off-diagonal elements of the lags of the coefficient matrix Φ are insignificant. Hence, in addition to the unconditional orthogonality of the model factors that follows from the estimation by principal components, there is also little conditional correlation between the factors of the FAVAR model.

The second panel provides the estimates of the state prices of risk which constitute the remaining components of the recursive bond pricing parameters A and B. The estimates show that all elements of the vector $\tilde{\lambda}_0$ governing the unconditional mean of the market prices of risk are large in absolute terms. This suggests that risk premia are characterized by an important constant component. However, the standard errors implied by the bootstrap algorithm are relatively large, so inference on the basis of individual estimates should be exercised with caution. A similar remark applies to the estimates of $\tilde{\lambda}_1$ which govern the dynamic component of the model-implied risk premia. While there are clear signs for time variation in the market prices of risk, only few of the estimated individual coefficients are statistically significant.

As has been noted in previous studies, it is difficult to pin down the market price of risk parameters in affine term structure

Table 6
In-sample fit: Observed and model-implied yields

	$y^{(1)}$	$y^{(3)}$	$y^{(6)}$	$y^{(9)}$	$y^{(12)}$	$y^{(24)}$	$y^{(36)}$	$y^{(48)}$	$y^{(60)}$	$y^{(84)}$	$y^{(120)}$
Panel A: Mean											
$y^{(n)}$	5.22	5.47	5.62	5.74	5.89	6.27	6.55	6.78	6.90	7.14	7.27
$\hat{y}^{(n)}$	5.22	5.47	5.61	5.76	5.88	6.27	6.56	6.77	6.91	7.14	7.26
$\lvert y_t^{(n)} - \hat{y}_t^{(n)}\rvert$	0.00	0.14	0.19	0.24	0.29	0.41	0.46	0.50	0.51	0.56	0.58
Panel B: Standard deviation											
$y^{(n)}$	2.11	2.20	2.25	2.29	2.32	2.33	2.27	2.24	2.21	2.14	2.06
$\hat{y}^{(n)}$	2.12	2.19	2.25	2.28	2.29	2.27	2.21	2.16	2.12	2.04	1.92
$\lvert y_t^{(n)} - \hat{y}_t^{(n)}\rvert$	0.00	0.19	0.25	0.31	0.37	0.50	0.57	0.63	0.65	0.72	0.73

This table summarizes empirical means and standard deviations of observed and fitted yields. Yields are reported in percentage terms. The first and second row in each panel report the respective moment of observed yields and fitted values implied by the No-Arbitrage FAVAR model while in the third row the mean and standard deviation of absolute fitting errors are reported, respectively.

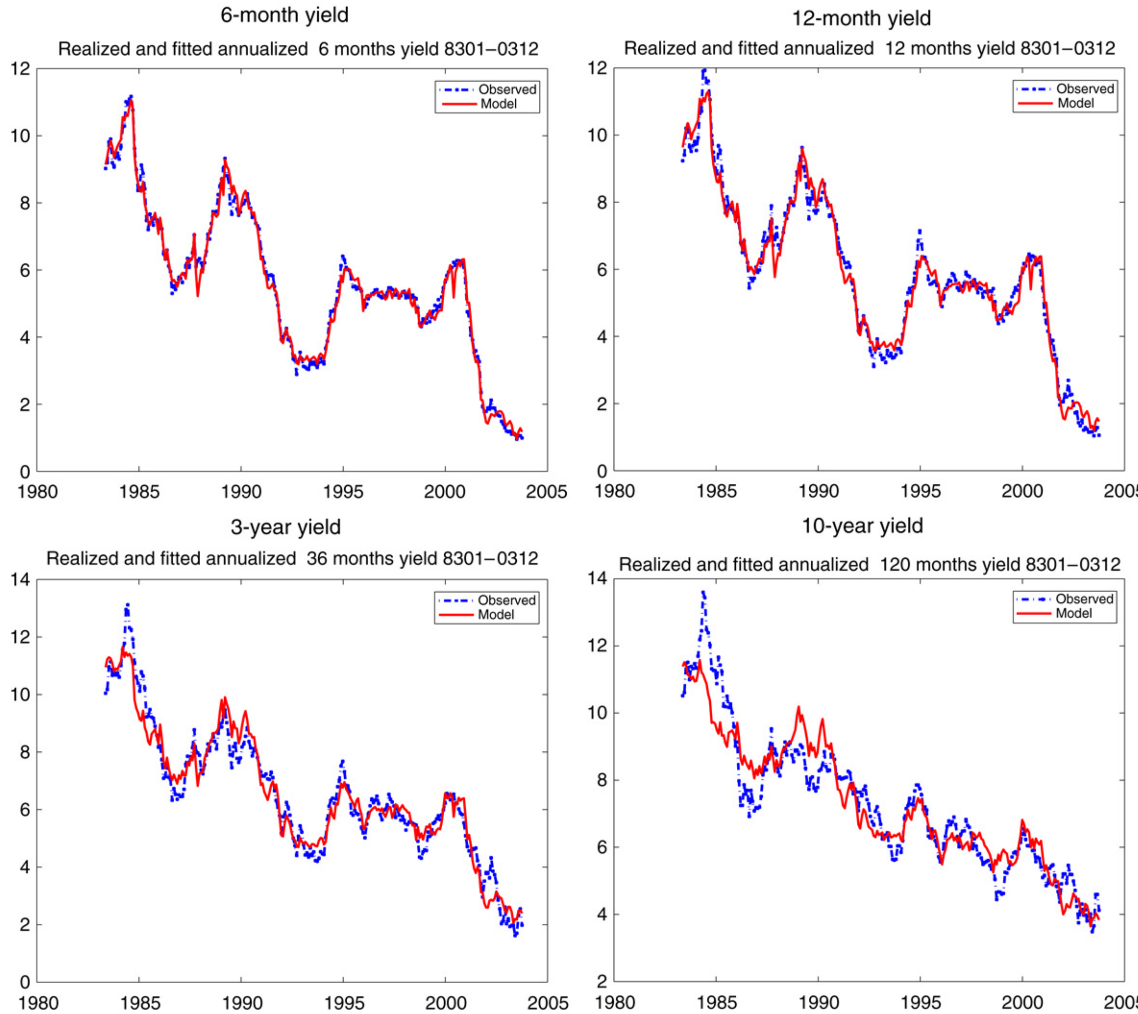

Fig. 2. Observed and model-implied yields. This figure provides plots of observed and model-implied time series for four selected interest rates, the 6-month yield, the 12-month yield and the 3-and 10-year yields.

models.[8] The lack of statistical significance of individual elements of $\tilde{\lambda}_0$ or $\tilde{\lambda}_1$ found here is therefore not necessarily a sign of poor model fit. Yet, economic reasoning based on the significance of individual parameters governing the state prices of risk is unwarranted. Instead, in order to visualize the relation between risk premia and the model factors, Fig. 3 provides a plot of model-implied term premia for the 1-year and the 5-year yield. As indicated by these plots, term premia at the short end of the yield curve are inversely related to the first macro factor which is itself highly correlated with output variables. By contrast, premia for longer yields are more closely related to the second factor which is strongly correlated with inflation indicators.

Fig. 4 provides a plot of the loadings b_n of the yields onto the contemporaneous observations of the model factors. The signs of these loadings are consistent with those obtained from regressing yields onto the model factors without imposing no-arbitrage restrictions, summarized in Table 5. By construction of my arbitrage-free model, the loading of the 1-month yield onto the short rate factor equals unity and those for the macro factors are zero. However, the impact of the short rate on longer yields sharply decreases with maturity. Hence, movements in the short-term interest rate only have a relatively small impact on long-term interest rates. Instead, these are more strongly driven by the macroeconomic factors. Most importantly, the first factor has an equally strong impact on yields of medium and longer maturities.

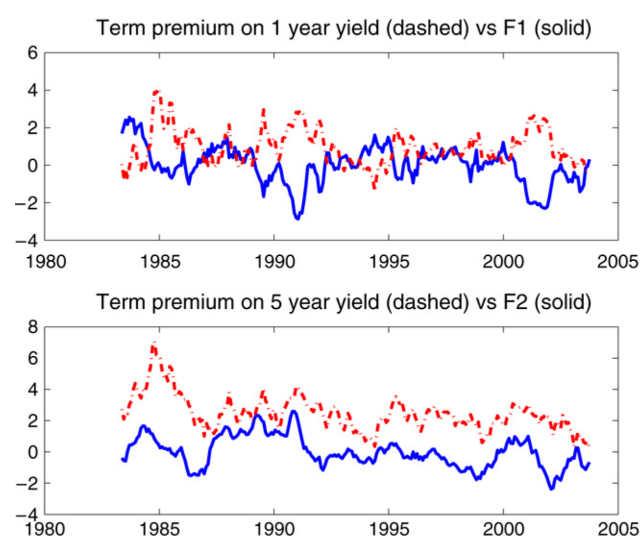

Fig. 3. Risk premia dynamics. This figure provides a plot of the term premia for the 1-year and 5-year yield as implied by the No-Arbitrage FAVAR model. For comparison, they are related to the first and second model factor, respectively.

[8] See e.g. Ang and Piazzesi (2003) or Hördahl et al. (2006).

Table 7
Parameter estimates for no-arbitrage FAVAR model

	Φ_1					Φ_2				
F1	0.977	−0.057	−0.107	−0.103	0.011	0.244	−0.143	0.013	0.143	0.044
	(0.096)	(0.109)	(0.118)	(0.064)	(0.061)	(0.140)	(0.180)	(0.165)	(0.088)	(0.079)
F2	0.196	1.357	0.174	0.038	0.028	−0.055	−0.387	−0.306	0.005	0.086
	(0.064)	(0.073)	(0.079)	(0.043)	(0.041)	(0.094)	(0.121)	(0.111)	(0.059)	(0.053)
F3	−0.160	0.098	0.945	−0.042	−0.043	0.112	−0.340	−0.014	0.072	0.012
	(0.072)	(0.082)	(0.088)	(0.048)	(0.046)	(0.105)	(0.135)	(0.124)	(0.066)	(0.060)
F4	−0.102	−0.172	0.170	1.007	−0.071	−0.068	0.336	0.051	−0.192	−0.044
	(0.123)	(0.140)	(0.151)	(0.082)	(0.079)	(0.179)	(0.231)	(0.212)	(0.112)	(0.102)
$y^{(1)}$	0.140	0.086	−0.045	−0.106	0.860	0.163	−0.057	−0.198	0.147	−0.042
	(0.100)	(0.113)	(0.123)	(0.066)	(0.064)	(0.146)	(0.188)	(0.173)	(0.091)	(0.083)
	Φ_3					Φ_4				
F1	−0.621	0.057	−0.006	−0.089	0.045	0.315	0.079	0.145	0.072	−0.102
	(0.139)	(0.178)	(0.165)	(0.088)	(0.079)	(0.107)	(0.107)	(0.110)	(0.061)	(0.060)
F2	−0.171	−0.049	0.257	−0.028	−0.071	0.084	0.045	−0.117	−0.016	−0.040
	(0.094)	(0.120)	(0.111)	(0.059)	(0.053)	(0.072)	(0.072)	(0.074)	(0.041)	(0.040)
F3	0.013	0.314	−0.350	−0.012	0.039	−0.087	−0.040	0.334	0.041	−0.006
	(0.104)	(0.134)	(0.124)	(0.066)	(0.059)	(0.081)	(0.081)	(0.082)	(0.046)	(0.045)
F4	0.347	−0.358	−0.111	−0.259	0.091	−0.016	0.165	−0.030	0.293	0.040
	(0.178)	(0.228)	(0.212)	(0.113)	(0.101)	(0.138)	(0.138)	(0.141)	(0.078)	(0.077)
$y^{(1)}$	−0.124	0.135	0.293	−0.001	−0.060	−0.022	−0.045	−0.082	−0.005	0.187
	(0.145)	(0.186)	(0.172)	(0.092)	(0.082)	(0.112)	(0.112)	(0.114)	(0.064)	(0.063)
	Ω					μ				
F1	0.100					0.013				
	(0.009)					(0.084)				
F2	−0.020	0.045				−0.003				
	(0.005)	(0.004)				(0.057)				
F3	0.054	−0.013	0.057			−0.036				
	(0.006)	(0.003)	(0.005)			(0.063)				
F4	−0.072	−0.016	−0.044	0.165		−0.091				
	(0.010)	(0.006)	(0.007)	(0.015)		(0.108)				
$y^{(1)}$	0.006	−0.003	−0.005	−0.018	0.109	0.246				
	(0.007)	(0.005)	(0.005)	(0.009)	(0.010)	(0.088)				

The market prices of risk specification is $\lambda_t = \lambda_0 + \lambda_1 Z_t$

$\tilde{\lambda}_0$	$\tilde{\lambda}_1$				
−26.366	0.631	0.215	0.288	3.386	−0.084
(29.514)	(1.806)	(0.920)	(1.325)	(1.416)	(0.544)
−10.220	4.673	0.338	0.870	−1.310	0.025
(139.011)	(5.064)	(3.363)	(3.461)	(3.302)	(1.772)
−88.111	−0.635	−0.982	−0.164	−4.116	0.304
(51.118)	(2.600)	(1.559)	(1.580)	(1.675)	(0.771)
−76.189	−1.612	−1.775	0.256	0.405	0.254
(48.143)	(2.119)	(1.184)	(1.481)	(1.464)	(0.652)
−18.664	−0.208	−0.648	−0.044	−0.800	−0.246
(12.770)	(0.795)	(0.658)	(0.626)	(0.604)	(0.411)

This table provides estimates of the parameters of the FAVAR model obtained using the full sample information, i.e. from estimating the model over the 1983:01–2003:09 period. The state dynamics are given by $Z_t = \mu + \Phi_1 Z_{t-1} + \cdots \Phi_4 Z_{t-4} + \omega_t$, where $E[\omega_t \omega_t'] = \Omega$. The states ($F1 \ldots F4$) of the model have been extracted from the large panel of macro time series using principal components methods.

Interestingly, shocks to the third macro factor appear to have a negative effect on yields of very short maturity and an increasingly strong positive impact on medium-term and long-term rates. This indicates that negative shocks to the third macro factor imply a flattening of the yield curve which is commonly associated with an upcoming recession.

4.5. How are the macro factors related to the components of the yield curve?

In traditional term structure analysis, the yield curve is often decomposed into three factors which together explain almost all of the cross-sectional variation of interest rates. According to their impact on the shape of the term structure, these components are commonly labeled level, slope, and curvature. Since the No-Arbitrage FAVAR model has been shown to explain yields relatively well in-sample, it is interesting to relate the macro factors used in the model to the level, slope, and curvature components of the yield curve. In this section, I thus regress estimates of the latent yield factors onto the macro factors and the short rate. The yield factors are computed as the first three principal components of the interest rates used to estimate the term structure model. Consistent with results reported in previous studies, these explain about 90.8%, 6.4% and 1.6% of the total variance of all yields.

Table 8 summarizes the results of these regressions. The four macro factors and the short-term interest rate explain almost all of the variation in the yield level which captures the most important source of common variation of interest rates. The main contribution comes from the short rate and the first and third macro factor which are correlated with output and inflation-related variables, respectively. Almost 80% of the variation in the slope of the yield curve is explained by the factors of the FAVAR model. Again, the short rate as well as the first and third macro factors are most strongly correlated with the slope. The short rate has a strongly significant negative coefficient in the slope equation which is consistent with the common view that short rate increases lead to a diminishing yield curve slope. Finally note that only about 35% of the variation in the curvature of the yield curve are

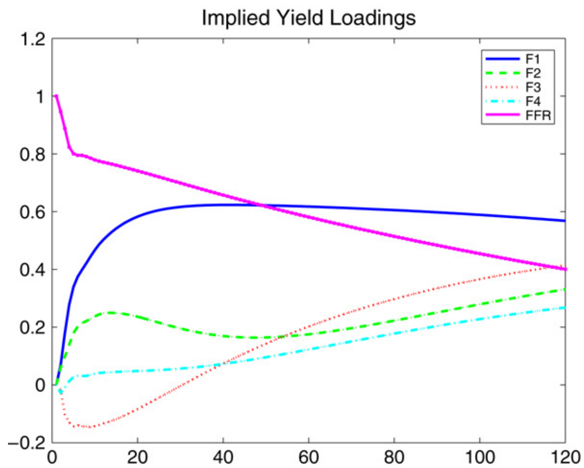

Fig. 4. Implied yield loadings. This figure provides a plot of the yield loadings b_n implied by the No-Arbitrage FAVAR model. The coefficients can be interpreted as the response of the n-month yield to a contemporary shock to the respective factor.

Table 8
Regression of latent yield factors on the model factors

	Level	Slope	Curvature
cst	0.23	1.65	−0.05
	[10.88]	[9.40]	[−0.11]
F1	0.04	0.13	−0.37
	[7.15]	[4.18]	[−3.91]
F2	0.03	0.10	−0.13
	[2.76]	[1.48]	[−0.96]
F3	0.04	0.30	0.02
	[5.89]	[6.23]	[0.30]
F4	0.02	0.14	−0.09
	[2.92]	[2.44]	[−1.26]
$y^{(1)}$	0.07	−0.29	0.02
	[13.41]	[−7.22]	[0.33]
$\bar{R}^2$	0.95	0.77	0.35

This table summarizes the results obtained from a regression of level, slope, and curvature yield factors onto the factors of the FAVAR model. Level, slope, and curvature are computed as the first three principal components extracted from the yields used to estimate the term structure model. They explain 90.8%, 6.4% and 1.6% of the total variance of all yields, respectively. The sample period is 1983:01–2003:9. t-statistics are in brackets.

explained by the macro factors. Hence, variations in the relative size of short, medium and long-term yields seem to be the least related to changes in macroeconomic conditions.

5. Out-of-sample forecasts

The term structure model suggested in this paper is based on the idea that the Federal Reserve uses a large set of conditioning information when setting short-term interest rates and that the FAVAR approach suggested by Bernanke et al. (2005) represents a useful way of capturing this information. While the model can in principle be employed to analyze e.g. the macroeconomic underpinnings of yield curve dynamics, this paper focuses on the usefulness of the No-Arbitrage FAVAR model for predicting the term structure of interest rates.

In the previous section, it has been shown that the model provides a reasonably good in-sample fit to US yield data. In this section, I study the forecast performance of the No-Arbitrage FAVAR model in a recursive out-of-sample prediction exercise. Before documenting the results of the forecasts, I briefly describe how they are computed for the different models studied in this paper. I start with the No-Arbitrage FAVAR model for which model-implied forecasts are obtained according to

$$\hat{y}^{(n)}_{t+h|t} = \hat{a}_n + \hat{b}'_n \hat{Z}_{t+h|t}, \tag{10}$$

where Z contains the contemporaneous and lagged observations of the four factors ($F1$, $F2$, $F3$, $F4$) explaining the bulk of variation in the panel of monthly time series for the US and the 1-month yield, $y^{(1)}$. The four factors are re-estimated via principal components each period t a forecast is produced using data up to t. The coefficients $\hat{a}_n$ and $\hat{b}_n$ are computed according to equations (6) and (7), using as input the estimates $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Phi}}$, and $\hat{\boldsymbol{\Omega}}$ obtained by running a VAR on the states, as well as the estimates $\hat{\lambda}_0$ and $\hat{\lambda}_1$ that result from minimizing the sum of squared fitting errors of the model. Forecasts $\hat{Z}_{t+h|t}$ are obtained by iterating forward the FAVAR Eq. (2), i.e.

$$\hat{Z}_{t+h|t} = \hat{\boldsymbol{\Phi}}^h Z_t + \sum_{i=0}^{h-1} \hat{\boldsymbol{\Phi}}^i \hat{\boldsymbol{\mu}}. \tag{11}$$

Note that the number of lags which enter the FAVAR equation are re-estimated every period a forecast is made on the basis of the Hannan–Quinn criterion with a maximum lag length of 12 months.

5.1. *The competitor models*

I compare the model's forecast performance to that of several competitor models. In particular, these are a No-Arbitrage Macro VAR model, an unrestricted VAR on yield levels, two different specifications of the Nelson–Siegel (1987) three-factor model recently suggested by Diebold and Li (2006), an essentially affine latent yield factor model following Duffee (2002), a simple AR(p) on yield levels, and the random walk. The Nelson–Siegel (1987) model is expected to be the most challenging competitor as Diebold and Li have shown that it outperforms a variety of alternative yield forecasting models. In the following, I briefly sketch the individual competitor forecasting models.

5.1.1. *No-Arbitrage macro VAR model*

In order to analyze whether the forecast performance of the No-Arbitrage FAVAR model can be attributed to the large set of conditioning information incorporated by the model, I compare it to a model that uses as state variables individual macroeconomic indicators. In particular, I compare it to a model that incorporates as states the short rate and four macroeconomic variables which are likely to contain information useful to explain yields.

Such a model has been suggested by Bernanke et al. (2004). In addition to the federal funds rate, these authors use the following four variables as states in their term structure model: a measure of the employment gap (payroll employment detrended by a Hodrick–Prescott filter); inflation over the past year, as measured by the deflator for personal consumption expenditures excluding food and energy; expected inflation over the subsequent year, taken from the Blue Chip survey where inflation is defined in terms of the GDP deflator, and the year-ahead Eurodollar futures rate to capture the expected path of monetary policy over the near-term.

Bernanke et al. (2004) report that their model explains the term structure well over time. This result is confirmed by Rudebusch et al. (2006b) who find that the model even provides a better fit to the cross-section of yields than originally described by Bernanke et al. (2004). However, neither of both studies investigates the forecast performance of the model. Here, I assess the model's ability to predict the yield curve out-of-sample in a recursive pseudo real-time setting. Precisely, I obtain yield forecasts according to

$$\hat{y}^{(n)}_{t+h|t} = \hat{a}_n + \hat{b}'_n \hat{Z}^{\text{VAR}}_{t+h|t}$$

where $Z^{\text{VAR}} = (\text{Emp}, \pi, \pi^e, \text{ED}, y^{(1)})$ contains the cyclical component of payroll employment, PCE inflation, the Blue-Chip survey measure of expected inflation, the year-ahead Eurodollar futures rate, and the 1-month yield. The coefficients $\hat{a}_n$ and $\hat{b}_n$ are obtained from Eqs. (6) and (7) and guarantee the absence of arbitrage opportunities. Estimates of the model parameters based on the entire sample 1983:01-2003:09 are provided in Appendix C. Forecasts $\hat{Z}^{VAR}_{t+h|t}$ are computed as in (11). As for the No-Arbitrage FAVAR model, the lag order of the VAR is re-estimated on the basis of the Hannan–Quinn criterion with a maximum lag length of 12 months every period a forecast is made. The No-Arbitrage Macro VAR model of Bernanke et al. (2004) is denoted "BRS" in the tables below.

5.1.2. VAR(1) on yield levels

In this model, forecasts of yields are obtained according to

$$\hat{y}_{t+h|t} = \hat{c} + \hat{\Gamma} y_t,$$

where $\hat{c}$ and $\hat{\Gamma}$ are estimated by regressing the vector y_t onto a constant and its h-months lag. This model is referred to as "VARylds" in the results below.

5.1.3. Diebold–Li specification of the Nelson–Siegel model

In a recent paper, Diebold and Li (2006) have suggested a dynamic version of the traditional Nelson–Siegel (1987) decomposition of yields and have shown that this model provides superior yield forecasts with respect to a number of benchmark approaches. According to this model, yields are decomposed into three factors with loadings given by exponential functions of the time to maturity n and some shape parameter τ. Precisely, Diebold and Li suggest to obtain yield forecasts according to

$$\hat{y}^{(n)}_{t+h|t} = \hat{\beta}_{1,t+h|t} + \hat{\beta}_{2,t+h|t}\left(\frac{1-e^{-\tau n}}{\tau n}\right) + \hat{\beta}_{3,t+h|t}\left(\frac{1-e^{-\tau n}}{\tau n} - e^{-\tau n}\right)$$

where

$$\hat{\beta}_{t+h|t} = \hat{c} + \hat{\Gamma}\hat{\beta}_t.$$

Diebold and Li (2006) obtain initial estimates of the factors β by regressing yields onto the loadings $\left(1, (\frac{1-e^{-\tau n}}{\tau n}), (\frac{1-e^{-\tau n}}{\tau n} - e^{-\tau n})\right)$ for a fixed value of τ. They set $\tau = 0.0609$ which is the value that maximizes the curvature loading at the maturity of 30 months. Diebold and Li consider two different specifications of their model, one where the factor dynamics are estimated by fitting AR(1) processes and another where the factors follow a VAR(1). In my application of their model, I report results for both specifications. These are denoted as "NS(AR)" and "NS(VAR)", respectively.

5.1.4. Essentially affine latent yield factor model $A_0(3)$

This is a traditional affine model where all the factors are latent and have to be estimated from the yield data. I implement the preferred essentially affine $A_0(3)$ specification of Duffee (2002) who has shown that this model provides superior out-of-sample forecast results with respect to various other affine specifications. The specification of the market prices of risk is therefore similar to the No-Arbitrage FAVAR model. Within the $A_0(3)$ model, yield forecasts are obtained as

$$\hat{y}^{(n)}_{t+h|t} = \hat{a}_n + \hat{b}'_n \hat{Z}^{A_0(3)}_{t+h|t}$$

where $Z^{A_0(3)}$ is composed of three latent yield factors, backed out from the yields using the method by Chen and Scott (1993). In particular, I assume that the 1-month, 1-year and 10-year yield are observed without error. Moreover, the transition matrix Φ in the state equation is assumed to be lower-triangular and the variance-covariance matrix Ω to be an identity matrix so as to ensure exact identification of the model (see Dai and Singleton (2000)). Notice that since the latent factors need to be backed out from the yields, estimation of the model takes considerably longer than estimation of the No-Arbitrage FAVAR and Macro VAR models where the parameters of the state equation are estimated in a first stage of the estimation via OLS.

5.1.5. AR(p) on yield levels

Simple autoregressive processes constitute another natural benchmark for modeling the time variation of bond yields. Assuming that the yield of maturity n follows a p-th order autoregression, its h-step ahead forecast is given by

$$\hat{y}^{(n)}_{t+h|t} = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{y}^{(n)}_{t+h-1|t} + \cdots + \hat{\alpha}_p \hat{y}^{(n)}_{t+h-p|t}$$

where $\hat{y}^{(n)}_{\tau|t} = y^{(n)}_\tau$ for $\tau \le t$.

In the implementation of this model, the lag order p is estimated recursively using the BIC information criterion.

5.1.6. Random walk

Finally, many previous studies have suggested that the evolution of interest rates might be well described by random walk processes. The random walk therefore remains a common benchmark for interest rate prediction models and is also used as a competitor here. Assuming a random walk model for interest rates implies a simple no-change forecast of individual yields. Hence, in this model the h-months ahead prediction of an n-maturity bond yield in period t is simply given by its time t observation:

$$\hat{y}^{(n)}_{t+h|t} = y^{(n)}_t.$$

5.2. Out-of-sample forecast results

The out-of-sample forecasts are carried out over the time interval 1994:01–2003:09. The forecast sample therefore covers a period of almost ten years. The affine models are first estimated over the period 1983:01–1993:12 to obtain starting values for the parameters. All models are then re-estimated recursively using data from 1983:01 to the time that the forecast is made, beginning in 1994:01.

Table 9 summarizes the root mean squared errors obtained from these forecasts. Three main observations can be made. First, the No-Arbitrage FAVAR model clearly outperforms the No-Arbitrage Macro VAR of Bernanke et al. (2004) for most maturities and especially in forecasts 6-months and 12-months ahead. This implies strong support for the use of a broad macroeconomic information set when forecasting the yield curve based on macroeconomic variables. Second, at the 1-month ahead horizon, the VAR(1) in yield levels and the AR(p) model outperform the macro-based FAVAR and VAR models for yields of all maturities, with the AR(p) being slightly superior for intermediate and long yields and the VAR*ylds* model performing best for the short rate. Third and most importantly, however, the No-Arbitrage FAVAR model dominates all considered benchmark models in yield forecasts 6-months and 12-months ahead. Indeed, as panels B and C of Table 9 document, the FAVAR model implies smaller out-of-sample root mean squared forecast errors than the benchmark models except for the 10-year yield that is best predicted by the affine latent yield factor model.

Interestingly, both specifications of the Nelson–Siegel model considered in Diebold and Li (2006) are outperformed by the No-Arbitrage FAVAR model. This result is striking since Diebold and Li have documented their approach to be particularly good at forecasting. This indicates that the combination of a large information set, the rich dynamics of the FAVAR, and the parameter restrictions implied by no-arbitrage together result in a model

Table 9
Out-of-sample RMSEs—Forecast period 1994:01–2003:09

$y^{(n)}$	FAVAR	BRS	VARylds	NS(VAR)	NS(AR)	$A_0(3)$	AR(p)	RW
Panel A: 1-month ahead forecasts								
1	0.534	0.331	**0.249**	0.262	0.275	0.681	0.305	0.305
6	0.502	0.522	0.204	0.218	0.256	0.216	**0.204**	0.222
12	0.516	0.553	**0.250**	0.268	0.293	0.300	0.256	0.259
36	0.630	0.485	0.308	0.313	0.312	0.386	**0.299**	0.309
60	0.685	0.467	0.314	0.316	0.316	0.357	**0.303**	0.307
120	0.717	0.511	0.293	0.289	0.289	0.289	**0.279**	0.282
Panel B: 6-month ahead forecasts								
1	**0.601**	0.854	0.779	0.745	0.838	1.189	0.860	0.856
6	**0.608**	1.094	0.904	0.871	0.931	0.977	0.824	0.853
12	**0.696**	1.170	1.006	0.958	0.981	1.059	0.885	0.876
36	**0.756**	1.073	1.021	0.958	0.922	0.962	0.897	0.873
60	**0.794**	0.930	0.969	0.915	0.870	0.848	0.847	0.830
120	0.825	0.773	0.872	0.764	0.720	**0.671**	0.702	0.696
Panel C: 12-month ahead forecasts								
1	**0.919**	1.539	1.366	1.448	1.357	1.741	1.397	1.395
6	**0.981**	1.723	1.613	1.569	1.458	1.487	1.400	1.417
12	**1.055**	1.729	1.728	1.633	1.495	1.506	1.407	1.391
36	**1.066**	1.473	1.599	1.504	1.349	1.264	1.253	1.236
60	**1.063**	1.210	1.464	1.359	1.233	1.076	1.132	1.138
120	1.071	0.932	1.313	1.108	1.022	**0.853**	0.917	0.942

This table summarizes the root mean squared errors obtained from out-of-sample yield forecasts. The models have been estimated using data from 1983:01 until the period when the forecast is made. The forecasting period is 1994:01–2003:09. "FAVAR" refers to the No-Arbitrage Factor-Augmented VAR model; "BRS" denotes the arbitrage-free Macro VAR model of Bernanke et al. (2004); "VARylds" refers to an unrestricted VAR(1) on yield levels; "NS(VAR)" and "NS(AR)" denote the Diebold–Li (2006) version of the three-factor Nelson–Siegel model with VAR and AR factor dynamics, respectively; "$A_0(3)$" refers to the essentially affine latent yield factor model of Duffee (2002); "AR(p)" denotes an AR model where the lag order p is recursively estimated; "RW" refers to the random walk forecast.

Table 10
RMSEs relative to random walk—Forecast period 1994:01–2003:09

$y^{(n)}$	FAVAR	BRS	VARylds	NS(VAR)	NS(AR)	$A_0(3)$	AR(p)
Panel A: 1-month ahead forecasts							
1	1.751	1.085	**0.816**	0.859	0.900	2.232	1.000
6	2.266	2.355	0.921	0.984	1.154	0.972	**0.918**
12	1.993	2.135	**0.964**	1.034	1.131	1.160	0.987
36	2.039	1.571	0.996	1.013	1.011	1.250	**0.969**
60	2.232	1.522	1.022	1.029	1.031	1.165	**0.988**
120	2.547	1.815	1.039	1.028	1.025	1.027	**0.989**
Panel B: 6-month ahead forecasts							
1	**0.702**	0.997	0.910	0.870	0.979	1.389	1.004
6	**0.712**	1.283	1.059	1.022	1.092	1.145	0.966
12	**0.795**	1.336	1.148	1.094	1.119	1.209	1.010
36	**0.866**	1.229	1.171	1.098	1.056	1.103	1.028
60	**0.956**	1.121	1.167	1.102	1.048	1.022	1.021
120	1.186	1.112	1.254	1.099	1.035	**0.964**	1.010
Panel C: 12-month ahead forecasts							
1	**0.659**	1.103	0.979	1.038	0.973	1.249	1.002
6	**0.692**	1.216	1.139	1.107	1.029	1.049	0.988
12	**0.759**	1.243	1.242	1.174	1.075	1.083	1.012
36	**0.863**	1.192	1.293	1.217	1.091	1.023	1.014
60	**0.934**	1.063	1.287	1.194	1.084	0.946	0.995
120	1.136	0.989	1.393	1.175	1.085	**0.905**	0.973

This table summarizes the root mean squared errors relative to those implied by the random walk. The models have been estimated using data from 1983:01 until the period when the forecast is made. The forecasting period is 1994:01–2003:09. "FAVAR" refers to the No-Arbitrage Factor-Augmented VAR model; "BRS" denotes the arbitrage-free Macro VAR model of Bernanke et al. (2004); "VARylds" refers to an unrestricted VAR(1) on yield levels; "NS(VAR)" and "NS(AR)" denote the Diebold–Li (2006) version of the three-factor Nelson–Siegel model with VAR and AR factor dynamics, respectively; "$A_0(3)$" refers to the essentially affine latent yield factor model of Duffee (2002); "AR(p)" denotes an AR model where the lag order p is recursively estimated; "RW" refers to the random walk forecast.

which is particularly useful for out-of-sample predictions. In the subsample analysis carried out in the next section, I will analyze this result in more detail.

Table 10 reports RMSEs of all considered models relative to the random walk forecast. These results show that the improvement in terms of root mean squared forecast errors implied by the FAVAR model is particularly pronounced for short and medium term maturities. At the one-month forecast horizon, all yield-based models outperform the affine models based on macro variables. However, at forecast horizons beyond one month, the No-Arbitrage FAVAR model outperforms all other models for maturities from one month to five years. Relative to the random walk, the suggested model reduces root mean squared forecast errors up to 35% at the short end of the yield curve and improves forecasts of medium-term yields up to 15%. While all considered competitor models outperform the random walk in 6-months and 12-months ahead forecasts only for some maturities, the No-Arbitrage FAVAR model consistently outperforms the Random Walk except for the 10-year yield.

Table 11
White's reality check test—Forecast period 1994:01–2003:09

$y^{(n)}$	BRS	VARylds	NS(VAR)	NS(AR)	$A_0(3)$	AR(p)	RW
Panel A: 1-month ahead forecasts							
1	1.881	2.390	2.315	2.241	−1.915	2.055	2.057
6	**−0.201**	2.274	2.212	2.019	2.222	2.277	2.195
12	**−0.388**	2.214	2.115	1.963	1.916	2.180	2.164
36	1.798	3.298	3.266	3.266	2.719	3.349	3.293
60	2.774	4.055	4.041	4.032	3.745	4.116	4.101
120	2.797	4.693	4.716	4.720	4.716	4.766	4.761
Panel B: 6-month ahead forecasts							
1	**−3.966**	**−2.727**	**−2.062**	**−3.581**	**−11.050**	**−4.024**	**−3.946**
6	**−8.916**	**−4.883**	**−4.268**	**−5.311**	**−6.266**	**−3.365**	**−3.943**
12	**−9.525**	**−5.788**	**−4.823**	**−5.148**	**−6.892**	**−3.337**	**−3.223**
36	**−6.324**	**−5.235**	**−3.936**	**−3.073**	**−3.974**	**−2.774**	**−2.314**
60	**−2.608**	**−3.524**	**−2.448**	**−1.489**	−1.186	**−1.260**	**−0.878**
120	0.816	**−1.051**	0.756	1.546	2.246	1.693	1.862
Panel C: 12-month ahead forecasts							
1	**−15.919**	**−10.630**	**−12.891**	**−10.170**	**−22.505**	**−11.589**	**−11.504**
6	**−21.081**	**−17.180**	**−15.790**	**−12.063**	**−13.331**	**−10.608**	**−11.404**
12	**−19.634**	**−19.526**	**−16.494**	**−11.750**	**−12.519**	**−9.291**	**−9.248**
36	**−11.096**	**−14.876**	**−12.092**	**−7.259**	**−5.390**	**−5.132**	**−4.821**
60	**−3.814**	**−10.630**	**−7.755**	**−4.200**	−0.770	**−2.314**	**−2.214**
120	2.625	**−6.008**	**−1.121**	0.915	4.040	2.502	2.348

This table summarizes "White's Reality Check" test statistics based on a squared forecast error loss function. I choose the No-Arbitrage FAVAR model as the benchmark model and compare it bilaterally with the competitor models. Negative test statistics indicate that the average squared forecast loss of the FAVAR model is smaller than that of the respective competitor model. Bold figures indicate significance which is checked by comparing the average forecast loss differential with the 5% percentile of the empirical distribution of the loss differential series approximated by applying a block bootstrap with 1000 resamples and a smoothing parameter of 1/12.

One can formally assess whether the improvement of the FAVAR model over the benchmark models in terms of forecast error is significant by applying White's (2000) "reality check" test. This test uses bootstrap resamples of the forecast error series to derive the empirical distribution of the forecast loss differential of a model with respect to some benchmark model. It can thus be employed to evaluate the superior predictive ability of a model as compared to one or more competitor models. Here, I test whether the No-Arbitrage FAVAR model has superior predictive accuracy with respect to the seven considered competitors. The test statistics are reported in Table 11. Negative numbers indicate that the average squared forecast loss of the No-Arbitrage FAVAR model is smaller than that of the respective competitor model while positive test statistics indicate the opposite. I perform 1000 block-bootstrap resamples from the prediction error series to compute the significance of the forecast improvement at the 5% level which are indicated by bold figures. As the results in panels B and C of Table 11 show, the documented improvement in terms of root mean squared forecast errors is significant at the 5% level for all but very long maturities at forecast horizons of 6-months and 12-months ahead. This underscores the observation made above that the No-Arbitrage FAVAR model predicts interest rates considerably better than all studied competitor models, including the Nelson–Siegel model and the $A_0(3)$ model.

5.3. Subsample analysis of forecast performance

The results documented in the previous section show that the No-Arbitrage FAVAR model exhibits strong relative advantages over a variety of benchmark models which have been documented powerful tools in forecasting the yield curve. This result somewhat challenges the recent findings of Diebold and Li (2006) and therefore a closer look at the predictive ability of the different models is warranted. In this section, I thus perform a subsample analysis of the out-of-sample prediction results. In particular, I analyze the relative performance of the No-Arbitrage FAVAR model with respect to the Nelson–Siegel model over exactly the sample period that has been studied by Diebold and Li (2006).

Table 12 provides the root mean squared forecast errors of the different models for the out-of-sample prediction period 1994:01–2000:12. At the 1-month ahead horizon, both specifications of the Nelson–Siegel model outperform the other models except for the AR(p) model and the random walk which predict maturities from 6 months to 5-years better. The absolute size of the RMSEs is very similar to those documented by Diebold and Li (2006). For example, based on the NS(AR) model Diebold and Li report RMSEs of 0.236, 0.292, and 0.260 for the 1-year, 5-year and 10-year yields at the 1-month ahead horizon whereas I find values of 0.249, 0.280, and 0.249, respectively, for the same maturities. The small deviations are likely due to differences in the choice of data and the set of maturities used to estimate the models. Turning to the results for 6-months ahead predictions, the picture becomes less favorable for the Nelson–Siegel model. Only for the 1-month yield, the VAR specification of the Nelson–Siegel model performs best. In contrast, the No-Arbitrage FAVAR model outperforms all other models for the range of maturities between 6-months and 5-years. Again, the absolute size of the RMSEs found here is very similar to those reported by Diebold and Li. For example, while these authors document RMSEs of 0.669, 0.777, and 0.721 for the 1-year, 5-year and 10-year yields, I find values of 0.711, 0.764, and 0.694, respectively. The results again change somewhat if one considers 12-months ahead predictions for the sample period studied in Diebold and Li (2006). In this case, there appears to be a clearer advantage of their preferred NS(AR) specification which outperforms all other models except for the 10-year maturity.

To visualize these results, Figs. 5 to 7 show the actual yields and those predicted by the No-Arbitrage FAVAR, the BRS, the NS(AR), and the $A_0(3)$ model for some selected maturities. Fig. 5 plots the outcomes for the 1-month ahead forecast horizon. According to this plot, the NS(AR) and the $A_0(3)$ model forecast the persistent movements of yields quite well while the FAVAR model predicts more variation than actual yields exhibit. In particular, the model's predictions appear to be particularly poor around turning points of yield dynamics. Interestingly, the same observation applies to the BRS model. Hence, both models that are based on macroeconomic information tend to overstate the volatility of interest rates. This confirms the relatively poor predictive ability of the two models

Table 12
Out-of-sample RMSEs—Forecast Period 1994:01–2000:12

$y^{(n)}$	FAVAR	BRS	VARylds	NS(VAR)	NS(AR)	$A_0(3)$	AR(p)	RW
Panel A: 1-month ahead forecasts								
1	0.380	0.313	0.255	0.265	**0.249**	0.722	0.299	0.297
6	0.357	0.479	0.194	0.186	0.215	0.209	**0.180**	0.192
12	0.395	0.559	0.242	0.238	0.249	0.280	**0.236**	0.239
36	0.566	0.488	0.281	0.286	0.272	0.368	**0.267**	0.277
60	0.683	0.468	0.290	0.289	0.280	0.343	0.276	**0.275**
120	0.804	0.555	0.270	0.256	**0.249**	0.254	0.254	0.253
Panel B: 6-month ahead forecasts								
1	0.625	0.893	0.696	**0.509**	0.532	1.140	0.683	0.635
6	**0.594**	1.108	0.799	0.660	0.648	0.936	0.719	0.655
12	**0.656**	1.199	0.898	0.778	0.711	0.999	0.815	0.742
36	**0.650**	1.142	0.947	0.877	0.747	0.938	0.895	0.834
60	**0.742**	0.994	0.949	0.885	0.764	0.834	0.868	0.821
120	0.879	0.826	0.911	0.793	0.694	**0.637**	0.760	0.730
Panel C: 12-month ahead forecasts								
1	0.900	1.537	1.025	0.899	**0.812**	1.654	1.104	0.945
6	0.910	1.687	1.179	1.002	**0.908**	1.430	1.177	0.977
12	0.952	1.702	1.268	1.078	**0.932**	1.414	1.205	1.017
36	0.983	1.532	1.333	1.168	**0.937**	1.188	1.228	1.078
60	1.055	1.284	1.331	1.179	**0.979**	1.007	1.176	1.072
120	1.160	1.005	1.333	1.089	0.941	**0.775**	1.028	0.985

This table summarizes the root mean squared errors obtained from out-of-sample yield forecasts. The models have been estimated using data from 1983:01 until the period when the forecast is made. The forecasting period is 1994:01–2003:09. "FAVAR" refers to the No-Arbitrage Factor-Augmented VAR model; "BRS" denotes the arbitrage-free Macro VAR model of Bernanke et al. (2004); "VARylds" refers to an unrestricted VAR(1) on yield levels; "NS(VAR)" and "NS(AR)" denote the Diebold–Li (2006) version of the three-factor Nelson–Siegel model with VAR and AR factor dynamics, respectively; "$A_0(3)$" refers to the essentially affine latent yield factor model of Duffee (2002); "AR(p)" denotes an AR model where the lag order p is recursively estimated; "RW" refers to the random walk forecast.

Table 13
Out-of-sample RMSEs—Forecast period 2000:01–2003:09

$y^{(n)}$	FAVAR	BRS	VARylds	NS(VAR)	NS(AR)	$A_0(3)$	AR(p)	RW
Panel A: 1-month ahead forecasts								
1	0.762	0.385	0.300	**0.296**	0.349	0.648	0.371	0.366
6	0.699	0.566	**0.214**	0.256	0.316	0.228	0.226	0.257
12	0.674	0.493	**0.250**	0.297	0.348	0.326	0.271	0.281
36	0.726	0.420	0.336	0.341	0.359	0.396	**0.329**	0.342
60	0.660	0.406	0.339	0.344	0.360	0.366	**0.328**	0.342
120	0.488	0.362	0.310	0.330	0.336	0.325	**0.306**	0.312
Panel B: 6-month ahead forecasts								
1	**0.581**	0.761	0.896	1.027	1.231	1.391	1.120	1.165
6	**0.617**	1.058	1.038	1.148	1.298	1.085	0.978	1.118
12	**0.735**	1.090	1.153	1.213	1.327	1.169	0.988	1.078
36	**0.879**	0.885	1.150	1.095	1.158	0.972	0.887	0.956
60	0.830	**0.745**	1.019	0.969	1.020	0.841	0.797	0.868
120	0.656	0.586	0.798	0.716	0.759	0.709	**0.562**	0.634
Panel C: 12-month ahead forecasts								
1	**0.939**	1.571	1.896	2.191	2.079	1.927	1.911	2.052
6	**1.116**	1.848	2.297	2.382	2.221	1.654	1.823	2.108
12	**1.257**	1.829	2.459	2.461	2.288	1.749	1.805	2.030
36	**1.224**	1.385	2.085	2.084	1.974	1.471	1.372	1.601
60	1.069	**1.066**	1.738	1.711	1.660	1.252	1.111	1.329
120	0.842	0.760	1.275	1.175	1.185	1.022	**0.716**	0.891

This table summarizes the root mean squared errors obtained from out-of-sample yield forecasts. The models have been estimated using data from 1983:01 until the period when the forecast is made. The forecasting period is 1994:01–2003:09. "FAVAR" refers to the No-Arbitrage Factor-Augmented VAR model; "BRS" denotes the arbitrage-free Macro VAR model of Bernanke et al. (2004); "VARylds" refers to an unrestricted VAR(1) on yield levels; "NS(VAR)" and "NS(AR)" denote the Diebold–Li (2006) version of the three-factor Nelson–Siegel model with VAR and AR factor dynamics, respectively; "$A_0(3)$" refers to the essentially affine latent yield factor model of Duffee (2002); "AR(p)" denotes an AR model where the lag order p is recursively estimated; "RW" refers to the random walk forecast.

at very short forecast horizons documented above. Yet, at the 6-months ahead forecast horizon the picture looks strikingly different. In particular, as Fig. 6 shows, the No-Arbitrage FAVAR model predicts the surge of yields in 1999 and 2000 quite well. More impressively, it forecasts the strong decline of yields starting in late 2000 very precisely. By contrast, both the NS(AR) and the $A_0(3)$ models miss the particular dynamics in this episode by a few months. The affine macro VAR model of Bernanke et al. (2004) forecasts the strong decline of interest rates somewhat earlier than these two models, but overstates short and medium term maturities at the end of the sample. Although less pronounced, a similar pattern can be seen for the 12-months ahead forecasts, provided in Fig. 7.

Altogether, these results show that the No-Arbitrage FAVAR model performs particularly well compared to yield-based prediction models in periods when interest rates exhibit strong variation. To provide a more quantitative assessment of this finding, Table 13 displays the root mean squared forecast errors of the different

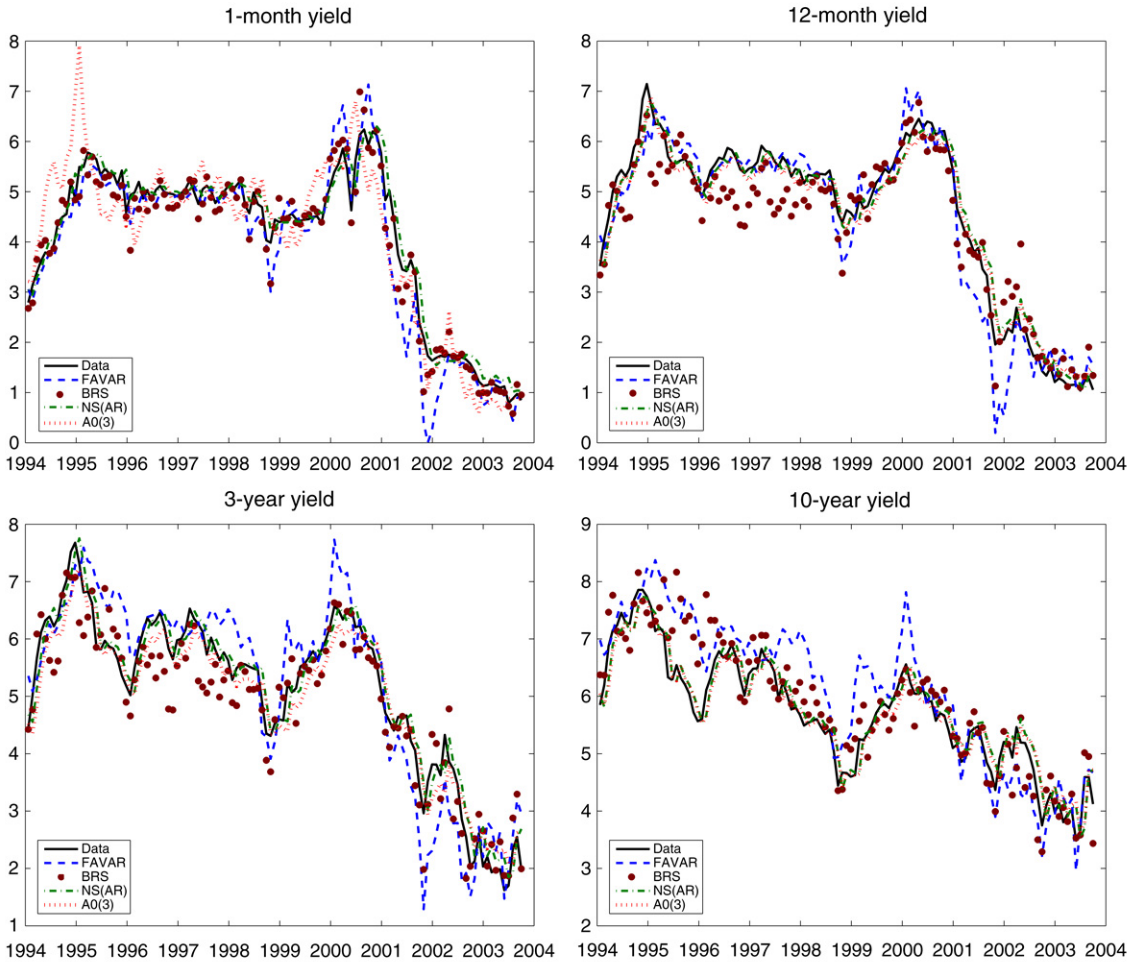

Fig. 5. Observed and predicted yields—1 month ahead. This figure provides plots of the observed and 1-month ahead predicted time series for the 1-month, the 12-month, the 3- and 10-year maturities. The observed yields are plotted by solid lines, whereas dashed, dash-dotted, and dotted lines correspond to predictions of the No-Arbitrage FAVAR model, the NS(AR) model, and the $A_0(3)$ model, respectively.

models for the subperiod 2000:01–2003:09. As can be seen from the plots above, this period was characterized by an initial surge of yields which was then followed by a sharp and persistent decline of interest rates of all maturities. The results of Table 13 show that over this particular sample period, the No-Arbitrage FAVAR model strongly outperforms all competitor models at forecast horizons 6-months and 12-months ahead for maturities up to 3 years. More precisely, the reduction in RMSEs relative to the random walk amounts to a striking 50% for very short maturities. Over the same subperiod, the 5-year and 10-year yields are best predicted by the BRS and AR(p)-models, respectively.

In sum, the results of the subsample analysis show that some of the strong forecast performance of the Nelson–Siegel model documented by Diebold and Li may be due to their choice of forecast period. In addition, the superior predictive ability of the model partly vanishes when confronted with the No-Arbitrage FAVAR model which strongly outperforms all benchmark models in periods when interest rates move a lot.

6. Conclusion

This paper presents a model of the term structure based on the idea that the central bank uses a large set of conditioning information when setting the short term interest rate and that this information can be summarized by a few factors extracted from a large panel of macroeconomic time series. Precisely, the Factor-Augmented VAR (FAVAR) approach suggested by Bernanke et al. (2005) is used to model the dynamics of the short-term interest rate. Given this dynamic characterization of the short rate, the term structure is then built up using restrictions implied by no-arbitrage. This setup is labeled a "No-Arbitrage Factor-Augmented VAR" approach. In contrast to most previously proposed macro-finance models of the term structure, the model suggested in this paper does not contain latent yield factors, but is entirely built upon macroeconomic information.

Fitting the model to US data, I document that it explains the dynamics of yields quite well. This underlines that most of the variation of interest rates is captured by macroeconomic variables. Most importantly, I find that the No-Arbitrage FAVAR model exhibits a strikingly good ability to predict the yield curve out-of-sample. In particular at intermediate and long forecast horizons, the model outperforms various benchmarks including the essentially affine three factor model of Duffee (2002) and the dynamic variant of the Nelson–Siegel model that Diebold and Li (2006) have recently suggested as a prediction model. A subsample analysis of the forecast results documents that the No-Arbitrage FAVAR model performs particularly well in periods when interest rates exhibit pronounced dynamics.

Based on the findings of the paper, there are a number of interesting directions for future research. First, while this paper has focused on the predictive ability of the No-Arbitrage FAVAR

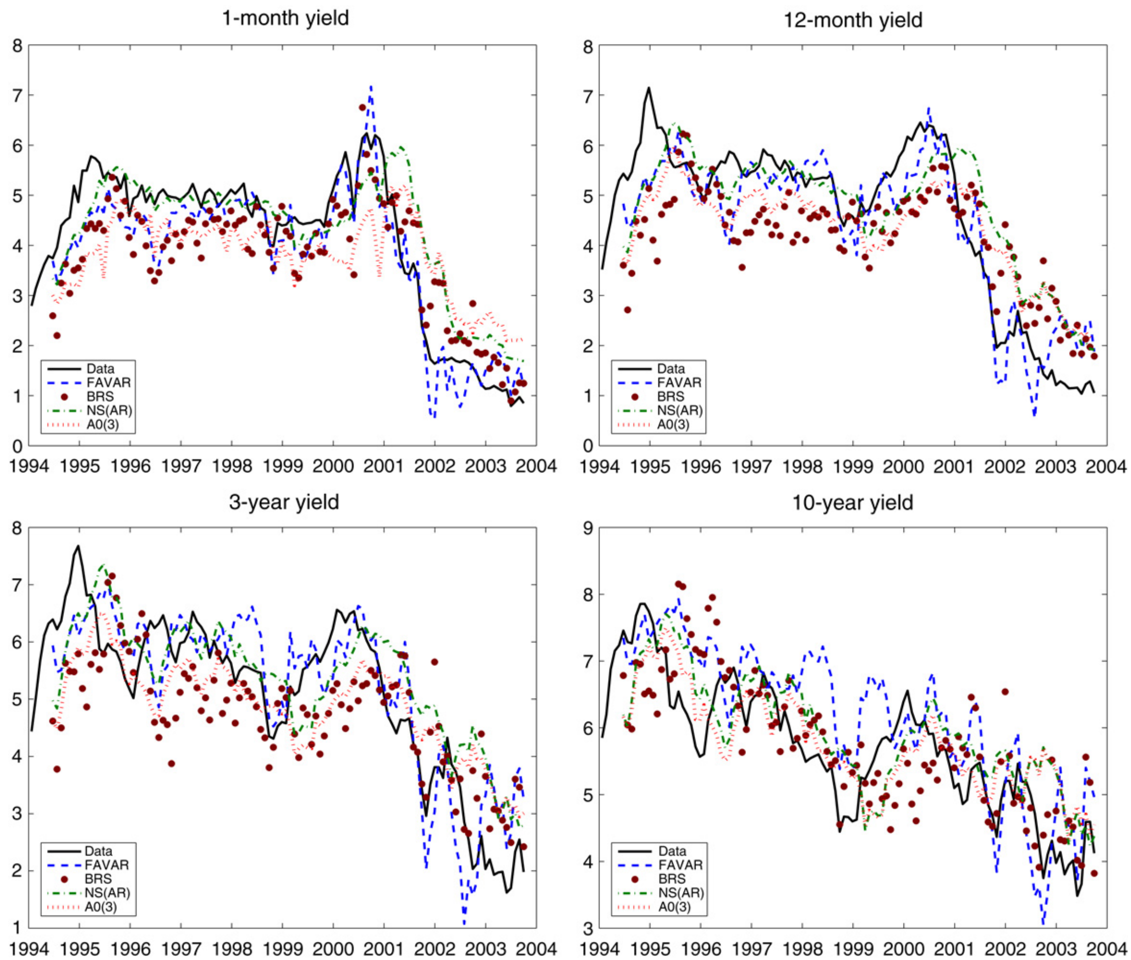

Fig. 6. Observed and predicted yields—6 months ahead. This figure provides plots of the observed and 6-months ahead predicted time series for the 1-month, the 12-month, the 3- and 10-year maturities. The observed yields are plotted by solid lines, whereas dashed, dash-dotted, and dotted lines correspond to predictions of the No-Arbitrage FAVAR model, the NS(AR) model, and the $A_0(3)$ model, respectively.

approach, the model can also be used for structural economic analysis. For example, it would be interesting to identify monetary policy shocks as in Bernanke et al. (2005) and study their impact on the yield curve. Second, based on estimates of term premia, one could use the model to analyze the risk-adjusted expectations of future monetary policy conditional on all macro information available. Finally, estimating the model using a one-step likelihood based Bayesian approach, one could easily add latent yield factors and assess to what extent these enhance the explanatory and predictive power of the model.

Acknowledgements

I would like to thank two anonymous referees, the editor Arnold Zellner, Theofanis Archontakis, Jean Boivin, Claus Brand, Albert Lee Chun, Sandra Eickmeier, Lars Hansen, Philipp Hartmann, Peter Hördahl, Manfred Kremer, Monika Piazzesi, Diego Rodriguez Palenzuela, Glenn Rudebusch, Oreste Tristani, Harald Uhlig, David Vestin, Thomas Werner, Jonathan Wright and various seminar participants for helpful comments. I am further grateful to Robert Bliss, Lucrezia Reichlin, and Brian Sack for sharing their data with me. Parts of this work were done when the author was visiting the ECB and the University of Pennsylvania. Financial support by the German National Academic Foundation, the Deutsche Forschungsgemeinschaft through the "SFB 649", and the Fritz-Thyssen Foundation is gratefully acknowledged. All remaining errors are the author's responsibility.

Appendix A. Derivation of the bond pricing parameters

The absence of arbitrage between bonds of different maturity implies the existence of the stochastic discount factor M such that

$$P_t^{(n)} = E_t[M_{t+1}\, P_{t+1}^{(n-1)}],$$

i.e. the price of a n-months to maturity bond in month t must equal the expected discounted price of an $(n-1)$-months to maturity bond in month $(t+1)$. Following Ang and Piazzesi (2003), the derivation of the recursive bond pricing parameters starts by assuming that the nominal pricing kernel M takes the form

$$M_{t+1} = \exp\left(-r_t - \frac{1}{2}\lambda_t' \boldsymbol{\Omega} \lambda_t - \lambda_t' \boldsymbol{\omega}_{t+1}\right)$$

and by guessing that bond prices P are exponentially affine in the state variables Z, i.e.

$$P_t^{(n)} = \exp(A_n + B_n' Z_t).$$

Plugging the above expressions for P and M into the first relation, one obtains

$$\begin{aligned} P_t^{(n)} &= E_t[M_{t+1}\, P_{t+1}^{(n-1)}] \\ &= E_t\left[\exp\left(-r_t - \frac{1}{2}\lambda_t' \boldsymbol{\Omega} \lambda_t - \lambda_t' \boldsymbol{\omega}_{t+1}\right) \exp(A_{n-1} + B_{n-1}' Z_{t+1})\right] \\ &= \exp\left(-r_t - \frac{1}{2}\lambda_t' \boldsymbol{\Omega} \lambda_t + A_{n-1}\right) \end{aligned}$$

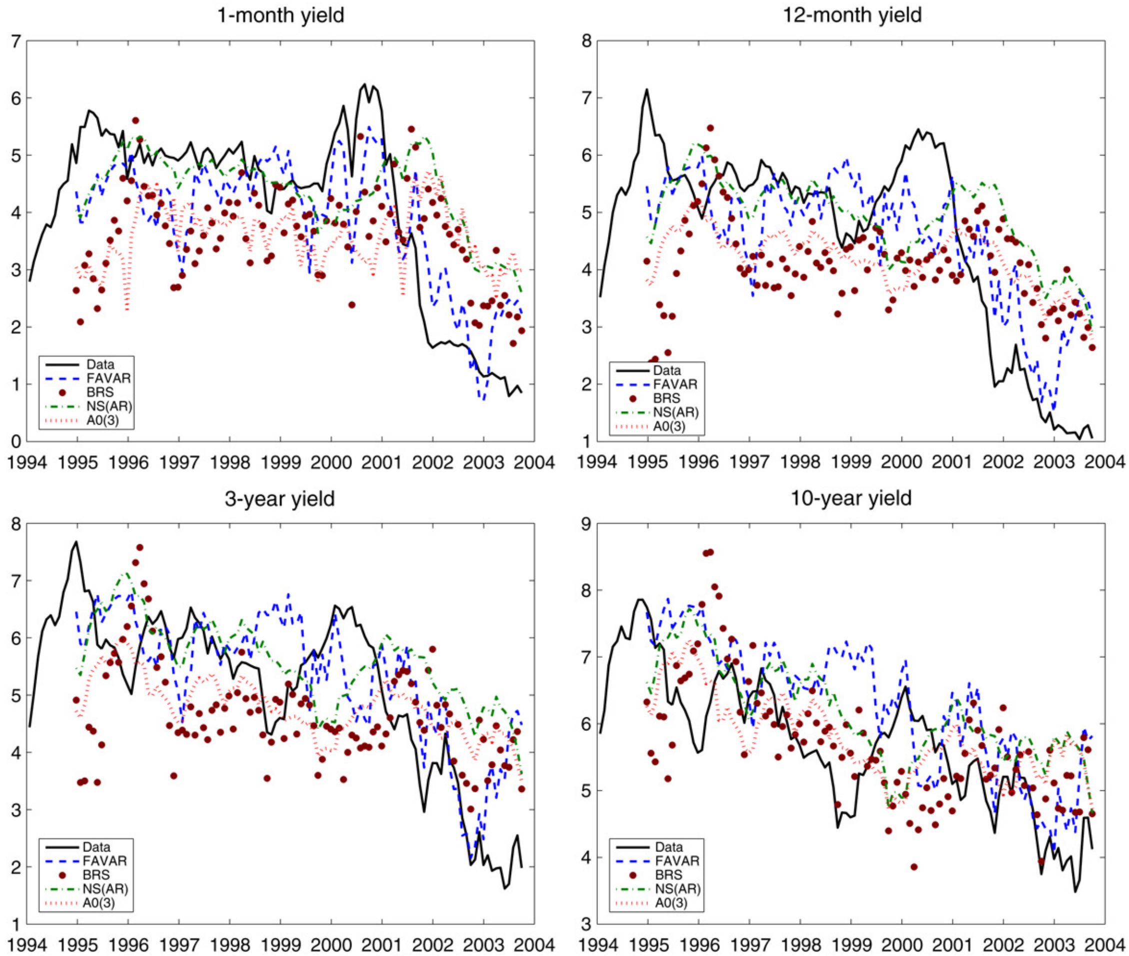

Fig. 7. Observed and predicted yields—12 months ahead. This figure provides plots of the observed and 12-months ahead predicted time series for the 1-month, the 12-month, the 3- and 10-year maturities. The observed yields are plotted by solid lines, whereas dashed, dash-dotted, and dotted lines correspond to predictions of the No-Arbitrage FAVAR model, the NS(AR) model, and the $A_0(3)$ model, respectively.

$$\times E_t\left[\exp(-\lambda_t'\omega_{t+1} + B_{n-1}'(\mu + \Phi Z_t + \omega_{t+1}))\right]$$
$$= \exp\left(-r_t - \frac{1}{2}\lambda_t'\Omega\lambda_t + A_{n-1} + B_{n-1}'\mu + B_{n-1}'\Phi Z_t\right)$$
$$\times E_t\left[\exp((-\lambda_t' + B_{n-1}')\omega_{t+1})\right].$$

Since the innovations ω of the state variable process are assumed Gaussian with variance-covariance matrix Ω, it is obvious that

$$\ln E_t\left[\exp((-\lambda_t' + B_{n-1}')\omega_{t+1})\right] = E_t\left[\ln(\exp((-\lambda_t' + B_{n-1}')\omega_{t+1}))\right]$$
$$+ \frac{1}{2}\mathrm{Var}_t\left(\ln(\exp((-\lambda_t' + B_{n-1}')\omega_{t+1}))\right)$$
$$= \frac{1}{2}\left[\lambda_t'\Omega\lambda_t - 2B_{n-1}'\Omega\lambda_t + B_{n-1}'\Omega B_{n-1}\right]$$
$$= \frac{1}{2}\lambda_t'\Omega\lambda_t - B_{n-1}'\Omega\lambda_t + \frac{1}{2}B_{n-1}'\Omega B_{n-1}.$$

Hence, $E_t\left[\exp((-\lambda_t' + B_{n-1}')\omega_{t+1})\right] = \exp(\frac{1}{2}\lambda_t'\Omega\lambda_t - B_{n-1}'\Omega\lambda_t + \frac{1}{2}B_{n-1}'\Omega B_{n-1})$ and thus

$$P_t^{(n)} = \exp\left(-r_t - \frac{1}{2}\lambda_t'\Omega\lambda_t + A_{n-1} + B_{n-1}'\mu + B_{n-1}'\Phi Z_t + \cdots + \frac{1}{2}\lambda_t'\Omega\lambda_t - B_{n-1}'\Omega\lambda_t + \frac{1}{2}B_{n-1}'\Omega B_{n-1}\right).$$

Using the relations $r_t = \delta' Z_t$ and $\lambda_t = \lambda_0 + \lambda_1 Z_t$, and matching coefficients finally yields

$$P_t^{(n)} = \exp(A_n + B_n' Z_t),$$

where

$$A_n = A_{n-1} + B_{n-1}'(\mu - \Omega\lambda_0) + \frac{1}{2}B_{n-1}'\Omega B_{n-1},$$

and $B_n' = B_{n-1}'(\Phi - \Omega\lambda_1) - \delta'$.

These are the recursive equations of the pricing parameters stated in (6) and (7).

Appendix B. Computation of standard errors by Monte Carlo

The two-step approach used to estimate the No-Arbitrage FAVAR model implies a potential errors-in-variables bias since the estimation of the market price of risk parameters takes as given the estimated evolution of the states. To adjust for this bias, I compute standard errors for λ_0 and λ_1 using the following Monte Carlo procedure.

From the estimation of the model, I save the model-implied pricing errors and state innovations, respectively. Using the stationary block bootstrap of Politis and Romano (1994a,b),[9] I then

[9] This algorithm delivers blocks of time indexes that are of random length and distributed according to the Geometric distribution with mean block length equal to $1/q$ where q is a smoothing parameter to be chosen. In the implementation of this algorithm, I set $q = 1/12$ which implies a mean block length of 12 months.

Table 14
Parameter estimates for the BRS model: State dynamics

	Φ_1					Φ_2				
Emp	0.842	−0.146	0.370	−0.188	−0.362	0.324	0.369	−1.703	−0.078	−0.024
	(0.151)	(0.417)	(0.820)	(0.161)	(0.232)	(0.116)	(0.536)	(1.090)	(0.193)	(0.318)
π	0.009	0.878	0.064	0.026	−0.044	−0.030	−0.090	−0.067	−0.012	0.025
	(0.010)	(0.132)	(0.144)	(0.021)	(0.036)	(0.011)	(0.086)	(0.182)	(0.038)	(0.033)
π^e	−0.003	0.054	1.030	0.057	0.042	−0.006	0.015	−0.115	−0.068	−0.048
	(0.004)	(0.033)	(0.078)	(0.014)	(0.023)	(0.006)	(0.046)	(0.110)	(0.025)	(0.032)
ED	−0.045	0.326	−0.296	1.106	0.104	−0.000	−0.026	0.463	−0.299	−0.063
	(0.019)	(0.176)	(0.290)	(0.059)	(0.107)	(0.027)	(0.178)	(0.396)	(0.099)	(0.142)
$y^{(1)}$	−0.024	0.074	−0.073	0.243	0.814	−0.025	−0.120	−0.244	−0.144	−0.005
	(0.020)	(0.111)	(0.369)	(0.052)	(0.093)	(0.024)	(0.132)	(0.500)	(0.087)	(0.104)
	Φ_3					Φ_4				
Emp	0.031	0.332	1.899	−0.239	0.063	−0.261	−0.209	−0.760	0.424	0.224
	(0.073)	(0.596)	(1.160)	(0.388)	(0.288)	(0.057)	(0.549)	(0.807)	(0.273)	(0.206)
π	0.010	0.152	0.063	−0.029	−0.037	0.006	0.050	−0.056	0.025	0.036
	(0.011)	(0.079)	(0.179)	(0.043)	(0.040)	(0.009)	(0.075)	(0.105)	(0.027)	(0.036)
π^e	0.002	−0.007	−0.024	0.030	0.033	0.007	0.014	0.000	−0.010	−0.027
	(0.007)	(0.042)	(0.103)	(0.021)	(0.023)	(0.004)	(0.032)	(0.062)	(0.012)	(0.016)
ED	0.038	−0.237	−0.028	0.134	−0.025	0.010	−0.063	−0.034	−0.077	0.076
	(0.032)	(0.151)	(0.486)	(0.079)	(0.084)	(0.020)	(0.163)	(0.331)	(0.067)	(0.073)
$y^{(1)}$	0.021	0.139	0.553	0.020	−0.087	0.024	−0.003	−0.350	−0.027	0.160
	(0.023)	(0.129)	(0.362)	(0.081)	(0.078)	(0.017)	(0.129)	(0.339)	(0.052)	(0.060)
	Ω					μ				
Emp	1.485					0.552				
	(0.505)					(0.445)				
π	0.003	0.028				0.034				
	(0.013)	(0.006)				(0.069)				
π^e	−0.005	0.002	0.006			0.045				
	(0.007)	(0.001)	(0.001)			(0.025)				
ED	−0.037	0.006	0.003	0.159		0.078				
	(0.033)	(0.004)	(0.002)	(0.017)		(0.145)				
$y^{(1)}$	−0.016	−0.000	0.004	0.038	0.108	0.054				
	(0.027)	(0.003)	(0.002)	(0.010)	(0.022)	(0.139)				

This table provides estimates of the parameters of the BRS model obtained using the full sample information, i.e. from estimating the model over the 1983:01–2003:09 period. The state dynamics are given by $Z_t = \mu + \Phi_1 Z_{t-1} + \cdots \Phi_4 Z_{t-4} + \omega_t$, where $E[\omega_t \omega_t'] = \Omega$. The states of the BRS model are given by the cyclical component of payroll employment (Emp), PCE inflation (π), the Blue-Chip survey measure of expected inflation (π^e), the year-ahead Eurodollar futures rate (ED), and the 1-month yield ($y^{(1)}$).

Table 15
Parameter estimates for the BRS model: Market prices of risk

$\tilde{\lambda}_0$	$\tilde{\lambda}_1$				
−18.781	−0.028	0.479	−0.670	−0.083	−0.004
(30.717)	(0.220)	(0.921)	(1.063)	(0.443)	(0.368)
−94.238	−0.310	0.456	1.059	−1.258	0.761
(120.554)	(1.341)	(4.178)	(4.491)	(2.651)	(2.010)
−64.335	0.171	2.701	1.381	0.408	0.532
(288.949)	(2.351)	(7.274)	(10.629)	(4.186)	(2.960)
−42.076	0.054	−0.162	0.101	−0.610	0.459
(38.405)	(0.266)	(1.235)	(1.386)	(0.620)	(0.443)
15.458	−0.034	0.783	−1.501	0.181	−0.274
(21.464)	(0.247)	(1.361)	(2.055)	(0.810)	(0.716)

The market price of risk specification is $\lambda_t = \lambda_0 + \lambda_1 Z_t$ where $\lambda_0 = (\tilde{\lambda}_0', 0_{1\times(k+1)(p-1)})'$ and $\lambda_1 = \begin{pmatrix} \tilde{\lambda}_1 & 0_{(k+1)\times(k+1)(p-1)} \\ 0_{(k+1)(p-1)\times(k+1)} & 0_{(k+1)(p-1)\times(k+1)(p-1)} \end{pmatrix}$.

generate 1000 artificial samples of the state vector and the vector of yields. In particular, in each bootstrap iteration I first simulate the state equation (3) and then generate a set of artificial yields by adding bootstrapped pricing errors to the model-implied yields. The latter are obtained by simulating the term structure model using the respective vector of sampled states and the estimated model parameters.

For each of the 1000 generated samples of yields and factor observations, I re-estimate the term structure model in two steps. That is, I first estimate the parameters governing the state dynamics using simple OLS. In a second step, I then take these estimates as given and infer the market price of risk parameters by minimizing the sum of squared fitting errors. Finally, I report as parameter estimates and standard errors the means and standard deviations from the so generated sample of model parameters.

Appendix C. In-sample estimation of the BRS model

See Tables 14 and 15.

References

Ang, Andrew, Piazzesi, Monika, 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. Journal of Monetary Economics 50 (4), 745–787.

Ang, Andrew, Piazzesi, Monika, Wei, Min, 2006. What does the yield curve tell us about gdp growth? Journal of Econometrics 127 (1–2), 359–403.

Bai, Jushan, Ng, Serena, 2002. Determining the number of factors in approximate factor models. Econometrica 70 (1), 191–221.

Belviso, Francesco, Milani, Fabio, 2005. Structural factor-augmented var (sfavar) and the effects of monetary policy. March, Princeton University.

Bernanke, Ben, Boivin, Jean, Eliasz, Piotr S., 2005. Measuring the effects of monetary policy: A factor-augmented vector autoregressive (favar) approach. The Quarterly Journal of Economics 120 (1), 387–422.

Bernanke, Ben S., Boivin, Jean, 2003. Monetary policy in a data-rich environment. Journal of Monetary Economics 50 (3), 525–546.

Bernanke, Ben S., Reinhart, Vincent R., Sack, Brian P., 2004. Monetary policy alternatives at the zero bound: An empirical assessment. Brookings Papers on Economic Activity 2, 1–78.

Bliss, Robert R., 1997. Testing term structure estimation methods. Advances in Futures and Options Research 9, 197–231.

Chen, Ren-Raw, Scott, Louis, 1993. Maximum likelihood estimation for a multi-factor equilibrium model of the term structure of interest rates. Journal of Fixed Income 3, 14–31.

Dai, Qiang, Singleton, Kenneth J., 2000. Specification analysis of affine term structure models. Journal of Finance 55 (5), 1943–1978.

Davidson, Russell, MacKinnon, James G., 1993. Estimation and Inference in Econometrics. Oxford University Press, New York.

Dewachter, Hans, Lyrio, Marco, 2006. Macro factors and the term structure of interest rates. Journal of Money, Credit, and Banking 38 (1), 119–140.

Diebold, Francis X., Li, Canlin, 2006. Forecasting the term structure of government bond yields. Journal of Econometrics 127 (1–2), 337–364.
Diebold, Francis X., Rudebusch, Glenn D., Aruoba, Boragan S., 2006. The macroeconomy and the yield curve: A dynamic latent factor approach. Journal of Econometrics 127 (1–2), 309–338.
Duffee, Gregory R., 2002. Term premia and interest rate forecasts in affine models. Journal of Finance 57 (1), 405–443.
Favero, Carlo A., Marcellino, Massimiliano, Neglia, Francesca, 2005. Principal components at work: The empirical analysis of monetary policy with large data sets. Journal of Applied Econometrics 20 (5), 603–620.
Giannone, Domenico, Reichlin, Lucrezia, Sala, Luca, 2004. Monetary policy in real time. NBER Macroeconomics Annual 161–225.
Hördahl, Peter, Tristani, Oreste, Vestin, David, 2006. A joint econometric model of macroeconomic and term-structure dynamics. Journal of Econometrics 127 (1–2), 405–444.
Knez, Peter J., Litterman, Robert, Scheinkman, José A., 1994. Explorations into factors explaining money market returns. Journal of Finance 49 (5), 1861–1882.
Moench, Emanuel, Uhlig, Harald, 2005. Towards a monthly business cycle chronology for the euro area. Journal of Business Cycle Measurement and Analysis 2 (1), 43–69.
Nelson, Charles R., Siegel, Andrew F., 1987. Parsimonious modeling of yield curves. Journal of Business 60 (4), 473–489.
Politis, Dimitris N., Romano, Joseph P., 1994a. Limit theorems for weakly dependent hilbert space valued random variables with applications to the stationary bootstrap. Statistica Sinica 4 (2), 461–476.
Politis, Dimitris N., Romano, Joseph P., 1994b. The stationary bootstrap. Journal of the American Statistical Association 89 (428), 1303–1313.
Rudebusch, Glenn D., Sack, Brian P., Swanson, Eric T., 2006a. Macroeconomic implications of changes in the term premium. Federal Reserve Bank of San Francisco.
Rudebusch, Glenn D., Swanson, Eric T., Wu, Tao, 2006b. The bond yield "conundrum" from a macro-finance perspective. Federal Reserve Bank of San Francisco.
Stock, James H., Watson, Mark W., 2002a. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97, 1167–1179.
Stock, James H., Watson, Mark W., 2002b. Macroeconomic forecasting using diffusion indexes. Journal of Business & Economic Statistics 20 (2), 147–162.
White, Halbert, 2000. A reality check for data snooping. Econometrica 68 (5), 1097–1126.