JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **27**: 574–602 (2012)
Published online 22 November 2010 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/jae.1220

TERM STRUCTURE SURPRISES: THE PREDICTIVE CONTENT OF CURVATURE, LEVEL, AND SLOPE

EMANUEL MOENCH*
Federal Reserve Bank of New York, New York, USA

SUMMARY

This paper analyzes the predictive content of the term structure components level, slope, and curvature within a dynamic factor model of macroeconomic and interest rate data. Surprise changes of the three components are identified using sign restrictions, and their macroeconomic underpinnings are studied via impulse response analysis. The curvature factor is found to carry predictive information both about the future evolution of the yield curve and the macroeconomy. In particular, unexpected increases of the curvature factor precede a flattening of the yield curve and announce a significant decline of output more than 1 year ahead. Copyright © 2010 John Wiley & Sons, Ltd.

1. INTRODUCTION

It is widely recognized that the yield curve carries information about the prospective evolution of economic activity, inflation, and monetary policy. Interest rate spreads, for example, are often used to predict recessions. Since the yield curve assumes similar shapes over time, it is common to think of it in terms of the three factors level, slope, and curvature, which together explain almost all of the cross-sectional variation of interest rates. However, despite the factor structure of the yield curve and its informational richness, there is only scattered evidence about the predictive content of each of its components. Therefore, in this paper, I analyze the economic underpinnings of level, slope, and curvature by studying the evolution of key macroeconomic variables subsequent to surprise changes of the three components. While partly confirming conventional wisdom, the most important result of my analysis is that unexpected changes of the curvature factor are more informative about the future evolution of the yield curve and macroeconomic aggregates than has previously been acknowledged.

To carry out this exercise, I employ a Bayesian dynamic factor model of macroeconomic and interest rate data that has the following properties. Yields are decomposed into three latent factors as suggested by Diebold and Li (2006). Their approach is a variant of the Nelson and Siegel (1987) functional form and allows a straightforward interpretation of the latent factors as level, slope, and curvature of the yield curve. Macroeconomic variables are also assumed to have a factor structure. This has two advantages. On the one hand, it allows one to study the dynamic effects of yield curve shocks on various macroeconomic variables. On the other hand, it mitigates the problem related to the use of revised data when studying the informational linkages between macroeconomic and financial variables. Finally, I model both sets of factors−macro and term structure−to share common dynamics within a vector autoregression. Together, these features of the model allow an unrestricted set of interactions between the term structure and the real economy. In this respect, the model goes beyond previous macro-finance models of the term structure such

* Correspondence to: Emanuel Moench, Capital Markets Research, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045, USA. E-mail: emanuel.moench@ny.frb.org

as Ang and Piazzesi (2003) or Hördahl *et al*. (2006), who only offer a unidirectional linkage from macroeconomic variables to the yield curve.

My model is similar in spirit to the one studied in Diebold *et al*. (2006, DRA henceforth). These authors also allow for a bidirectional linkage between macroeconomic variables and use the same factor decomposition of yields. Although related, my model differs in a number of important dimensions from the one studied in DRA. First, DRA include only three individual economic variables in their macro-finance model of the term structure. Employing a factor structure, my model, in contrast, incorporates a broader macroeconomic information set. Second, while the joint dynamics of macro variables and term structure factors in DRA are limited to a vector autoregressive (VAR) of order one, my model includes more lags. Overall, the model in this paper exhibits a richer structure and therefore allows a more comprehensive analysis of macro-term structure dynamics than previous studies.

The additional generality comes at the cost of computational complexity. To estimate the model, I therefore build on recent advantages in Bayesian dynamic factor model analysis. As has been recognized by Kim and Nelson (1999) and others, Gibbs sampling algorithms are well suited to approximate the joint posterior distribution of parameters and unobserved factors in state-space models. In the application of such methods to my model, a complication arises due to the nonstandard distribution of the exponential decay parameter in the Nelson–Siegel spline functions. This is solved by adding a Metropolis step to the Gibbs sampler.

The third crucial difference with respect to the study by DRA regards the identification of innovations or shocks. In order to dissect the informational content of yield curve innovations into level, slope, and curvature, it is important to properly identify the surprise changes of the three components. This is an intricate issue. As has been pointed out by Sarno and Thornton (2004), zero restrictions on impulse responses of financial variables to contemporaneous macroeconomic shocks are inconsistent with the efficient market hypothesis. Hence, an appropriate identification scheme must allow the yield curve factors to contemporaneously react to all macroeconomic shocks. This is not the case in the recursive identification scheme employed by DRA, who order the yield curve factors first. It could be achieved by ordering the yield factors last. Yet this identification would preclude macroeconomic variables from contemporaneously moving when yield curve surprises occur, an assumption that is quite restrictive.

I solve this problem by using identification schemes which make use of sign restriction techniques, as recently suggested by Uhlig (2005) and Mountford and Uhlig (2009). Precisely, to identify a positive surprise change of, say, the curvature factor, the impulse responses of the level and slope are restricted to be zero or small on impact, while the response of the curvature is required to be positive over some periods after the innovation occurs. At the same time, the impulse responses of the macroeconomic variables remain unrestricted. Thus my identification does not treat innovations of the yield curve factors as structural shocks, but rather interprets the surprise changes to these three components as the bond market's reaction to economic shocks. Since different economic shocks which result in different macroeconomic outcomes may affect different sectors of the yield curve in different ways, the objective of my study is to analyze what information surprise changes of the three yield curve components carry about future economic activity.

The results of my analysis can be summarized as follows. Most importantly, I find the curvature factor to be more informative about the future evolution of the yield curve and macroeconomic variables than has previously been acknowledged. In particular, positive surprise changes of the curvature factor announce a strongly significant and persistent hump-shaped movement of the yield curve slope and a persistent decline of the yield curve level. Together, these two features imply a successive flattening of the yield curve which is often associated with an upcoming recession. This is paralleled by a pronounced hump-shaped response of output and a brief decline of inflation.

DOI: 10.1002/jae

In particular, the growth rate of industrial production increases sharply for about 3 months, then slowly declines, and eventually falls below zero about 15 months after the initial surprise. Hence an unexpected surge of the curvature factor–not accompanied by simultaneous moves of the yield curve level or slope–appears to announce economic slowdowns. This result is surprising since the curvature factor has previously been documented to be largely unrelated to macroeconomic variables.

Somewhat more consistent with conventional wisdom, surprise surges of the level factor–not accompanied by simultaneous changes of the slope or curvature–announce strong and persistent movements in inflation. They also anticipate real effects: positive level surprises are followed by a significant hump-shaped response of output growth. In contrast, slope surprises are followed by an immediate but not very pronounced decline of output.[1] These results suggests that, while the level of term spreads may have predictive power for real activity, innovations to the yield curve slope do not carry much information about future output dynamics. Instead, innovations to curvature announce a flattening of the yield curve and a decline in output more than a year ahead.

The paper is organized as follows. I present the empirical model that has been used to study the joint dynamics of the yield curve and the macroeconomy in Section 2. Section 3 briefly discusses identification of the model and its estimation via a Metropolis-within-Gibbs sampling algorithm. In Section 4, I present the different approaches used to identify surprise changes of the yield curve components. I discuss the main empirical results in Section 5. Section 6 concludes the paper. Details on the Metropolis-within-Gibbs sampling algorithm are provided in the Appendix.

2. THE MODEL

I assume that yields of different maturities are driven by three common factors and an idiosyncratic component,

$$Y_t = \Lambda_y F_t^y + e_t^y \tag{1}$$

where Y_t is an $N_y \times 1$ vector of yields of different maturity, Λ_y is an $N_y \times 3$ matrix of factor loadings, F_t^y is a 3×1 vector of factors, and e_t^y is an $N_y \times 1$ vector of idiosyncratic components or pricing errors. I model the loadings on the three yield curve components as

$$\lambda_y^{(n)} = \left[1 \quad \left(\frac{1-e^{-\tau n}}{\tau n}\right) \quad \left(\frac{1-e^{-\tau n}}{\tau n} - e^{-\tau n}\right)\right] \tag{2}$$

where τ denotes a shape parameter and n maturity. Diebold and Li (2006) show that this functional form of the factor loadings which goes back to Nelson and Siegel (1987) implies that the three yield curve factors can readily be interpreted as the level, slope, and curvature of the term structure. Indeed, the loadings on the first factor are constant across the maturity spectrum so that it can clearly be viewed as a level factor. The loadings on the second factor decrease monotonically with maturity. Shocks to the second factor thus affect short-term yields much more strongly than long-term interest rates, implying that it has a straightforward interpretation as the (negative of the) slope of the yield curve. Finally, loadings of yields on the third factor follow a hump-shaped pattern and are close to zero for short and long maturities but larger for intermediate maturities. Accordingly, the third factor can readily be interpreted as a curvature factor. Given these particular functional forms for the loadings on the three yield curve factors, one can thus disentangle movements in the

[1] Note that according to the Diebold–Li formulation of the factor loadings, a positive slope shock is associated with a strong increase in short-term rates and a small increase in long maturities, i.e., a flattening of the yield curve.

DOI: 10.1002/jae

term structure of interest rates into three factors which have a clear-cut interpretation and which have become economic concepts of independent interest.[2]

I further assume that macroeconomic variables in the model are driven by a few common factors and an idiosyncratic component, i.e.

$$X_t = \Lambda_x F_t^x + e_t^x \tag{3}$$

where X_t is an $N_x \times 1$ vector of period-t observations of the variables in the panel, Λ_x is a $N_x \times k_x$ matrix of factor loadings, F_t^x is the $k_x \times 1$ vector of period-t observations of the common factors, and e_t^x is an $N_x \times 1$ vector of idiosyncratic components. By construction, the factors F^x capture the common variation in a large number of economic times series. This allows me to study impulse responses of different macroeconomic time series to surprise changes of the yield curve factors.

Another reason for employing a factor model approach instead of using individual macro variables relates to the problem of data revisions. In fact, a common objection against empirical macro-finance models is that data revisions imply that the information set available to the econometrician is different from the information set available to investors. Hence estimates of the parameters governing the mutual interactions between macroeconomic and financial variables may be biased. Studying the interaction of financial variables with the common components of many macro variables represents one way to address this critique. Indeed, assuming that data revision errors are series-specific (see, for example, Bernanke and Boivin, 2003; Giannone *et al*., 2004), the common factors extracted from different vintages of the same macroeconomic dataset will be the same. Thus factor estimates obtained from revised data likely span the space of information available to investors in real time.

Equations (1) and (3) have the same structure and therefore allow consideration of the unified framework:

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{bmatrix} \Lambda_x & 0 \\ 0 & \Lambda_y \end{bmatrix} \begin{pmatrix} F_t^x \\ F_t^y \end{pmatrix} + \begin{pmatrix} e_x^t \\ e_y^t \end{pmatrix}$$

or

$$Z_t = \Lambda F_t + e_t \tag{4}$$

I assume that the idiosyncratic disturbances are mutually orthogonal and not serially correlated, i.e., $E[ee'] = R$ is a diagonal $N \times N$ matrix where $N = N_x + N_y$. A central feature of my model is the assumption that a few structural innovations cause the common variation in both sets of variables. By construction, the comovement is captured by the two sets of factors, F^x and F^y. Their common dynamics are modeled with a VAR, i.e.,

$$\begin{pmatrix} F_t^x \\ F_t^y \end{pmatrix} = \Phi(L) \begin{pmatrix} F_{t-1}^x \\ F_{t-1}^y \end{pmatrix} + \begin{pmatrix} \omega_t^x \\ \omega_t^y \end{pmatrix}$$

or

$$F_t = \Phi(L) F_{t-1} + \omega_t \tag{5}$$

The reduced-form errors $\omega_t = (\omega_t^{x\prime}, \omega_t^{y})'$ are assumed to have non-diagonal variance–covariance matrix $E[\omega\omega'] = \Omega$. Furthermore, the idiosyncratic disturbances of individual variables, e_t, and the innovations driving the common factors, ω_t, are assumed to be mutually independent.

[2] Consistent with previous evidence, I document in Section 5.3 that these three factors explain almost all of the cross-sectional variation of yields.

DOI: 10.1002/jae

The reduced-form errors can be written as $\omega_t = A\nu_t$, where the structural innovations ν_t are assumed to be mutually orthogonal and normalized to have unit variance. This implies $\Omega = AA'$. In Section 4 I discuss two approaches with which innovations to the three term structure components can be identified that have a specific impact on the three term structure components. This amounts to finding matrices A of which individual columns give rise to impulse responses of the latent yield factors F^y that have a predetermined shape.

3. ESTIMATION OF THE MODEL

Before estimating the model, it needs to be ensured that its parameters and latent factors are uniquely identified. Exact identification is crucial since observationally equivalent sets of factors and parameters may give rise to the same likelihood but lead to different economic conclusions. Hence assumptions need to be made which exclude such indeterminacies. A standard identification approach in factor models of the form (4)–(5) is due to Aguilar and West (2000). These authors show that restricting the upper $k \times k$ block of Λ to be lower-triangular with those on the diagonal uniquely determines the factors and loadings. A variant of this 'hierarchical' identification approach is employed here. A complication arises due to the fact that two separate groups of factors drive the common dynamics of the variables Z. However, as discussed in Appendix A, it is sufficient to restrict the upper $k_x \times k_x$ block of Λ_x to be lower-triangular with fixed diagonal elements in order to ensure unique identification of the model.

Estimation of the model via maximum likelihood techniques is infeasible due to the large number of model parameters. An alternative would be to independently extract factors from both sets of variables via, for example, principal components and to study their joint dynamics within a VAR. This two-step estimation approach is inefficient, however, as it does not permit joint estimation of the factors and model parameters. Recently, some authors have advocated estimating large-scale dynamic factor models via Markov chain Monte Carlo (MCMC) methods (see, for example, Bernanke *et al.*, 2005; Kose *et al.*, 2003). Gibbs sampling represents a particularly useful estimation approach for dynamic factor models as it allows approximation of the joint posterior distribution of the model parameters and the latent factors by iteratively sampling from the conditional distributions of the parameters given the factors and vice versa. The latter are straightforward to derive in a linear state space framework as is present here. I thus follow the recent literature on estimating dynamic factor models using MCMC methods to estimate the model presented above. As will be discussed in more detail below, posterior conditional distributions cannot be derived for all parameters of the model and thus a Metropolis-within-Gibbs algorithm is employed.

Estimation of the dynamic factor model (4)–(5) via Gibbs sampling is facilitated by rewriting it in companion form as

$$Z_t = \overline{\Lambda}\ \overline{F}_t + \overline{e}_t \tag{6}$$

$$\overline{F}_t = \overline{\Phi}\ \overline{F}_{t-1} + \overline{\omega}_t \tag{7}$$

where $\overline{F}_t = (F_t, \ldots, F_{t-p+1})$ and where $\overline{\Lambda}$, $\overline{e}_t$, $\overline{\Phi}$, and $\overline{\omega}_t$ denote the companion form equivalents of Λ, e_t, Φ, and ω_t, respectively, and $\overline{R}$ and $\overline{\Omega}$ the corresponding variance covariance matrices.

Let $\theta = (\Lambda_x, \Lambda_y, R, \Phi, \Omega)$ denote the set of model parameters. Moreover, let $\tilde{X}_T = \{X_1, \ldots, X_T\}$ and $\tilde{Y}_T = \{Y_1, \ldots, Y_T\}$ denote the vectors stacking all T observations on yields and macro variables and let $\tilde{Z}_T = \{\tilde{X}_T, \tilde{Y}_T\}$. Analogously, let $\tilde{F}_T = \{\overline{F}_1, \ldots, \overline{F}_t\}$ denote the vector comprising all observations of the factors F. The objective is to generate samples from the joint posterior distribution $p(\theta, \tilde{F}_T|\tilde{Z}_T)$ of model parameters and unobserved factors. If this distribution is not given or is not standard so that drawing from it is infeasible, the Gibbs sampler allows one to

DOI: 10.1002/jae

approximate it by the empirical distributions of simulated values from the conditional posteriors $p(\theta|\tilde{Z}_T, \tilde{F}_T)$ and $p(\tilde{F}_T|\tilde{Z}_T, \theta)$. After finding starting values θ^0, any iteration of the Gibbs sampler involves the following two steps:

- Step 1: Draw $\tilde{F}_T^{(i)}$ from $p(\tilde{F}_T|\tilde{Z}_T, \theta^{(i-1)})$.
- Step 2: Draw $\theta^{(i)}$ from $p(\theta|\tilde{Z}_T, \tilde{F}_T^{(i)})$.

The exact procedures to sample from the conditional distributions of factors and individual parameters are described in Appendix B. The crucial result employed in the Gibbs sampler is that the empirical distribution of draws from the conditional posterior densities converges to the joint marginal posterior distribution as the number of iterations goes to infinity. Accordingly, after discarding an initial number of draws (the 'burn-in'), sampling from the known conditional posterior densities of factors and parameters is equivalent to sampling from their unknown joint posterior distribution.

4. IDENTIFYING SURPRISE CHANGES TO THE YIELD CURVE COMPONENTS VIA SIGN RESTRICTIONS

In this paper I ask the question: What information do innovations to the yield curve components level, slope, and curvature convey about the future evolution of key macroeconomic variables? To answer this question, it is crucial to properly disentangle surprises to the three components. As has been pointed out by Sarno and Thornton (2004), there is a fundamental problem related to the identification of shocks to financial variables in structural VARs. In particular, they argue that zero restrictions on impulse responses of financial variables to macroeconomic shocks are inappropriate under the assumption of efficient markets. Assuming that the US government bond market is efficient and thus reflects all information relevant for the determination of bond prices, the critique by Sarno and Thornton implies that an identification scheme needs to be found which allows the yield curve components to contemporaneously react to all macroeconomic shocks. In principle, this could be achieved by ordering the term structure factors last in a recursive structural VAR. However, the yield curve shocks so identified would be bound to have zero contemporaneous effects on the macro factors. Having in mind a structural interpretation of the yield curve shocks, this could be an appropriate assumption. In this paper, however, I seek to analyze the predictive content of surprise changes of the yield curve, which may very well be driven by macroeconomic shocks and thus coincide with simultaneous changes of macroeconomic variables. Under my identification scheme, I therefore allow the latter to contemporaneously move when a yield curve surprise occurs.

I employ two different approaches to carry out the identification. Both impose restrictions on the sign of the impulse responses of one of the three yield factors for a given number of periods after the surprise occurs, and restrictions on the initial impact of the other two factors. To achieve these identifications, I draw heavily on previous work in Uhlig (2005) and Mountford and Uhlig (2009). Identifying shocks via sign restrictions is based upon prior assumptions about the impact of a certain type of shock on different economic variables. Uhlig (2005), for example, uses prior restrictions on the responses of prices, the federal funds rate and nonborrowed reserves to identify contractionary monetary policy shocks. The economic reasoning behind this approach is that a monetary policy shock should be characterized by a rise of the federal funds rate and a subsequent decline of prices and nonborrowed reserves. In this paper, I apply these methods in order to identify surprise changes of the latent yield curve factors, without, however, attributing a structural interpretation to these surprises. In contrast, I investigate their informational content by

DOI: 10.1002/jae

studying the subsequent dynamic responses of the macro variables stacked in X and of the yield curve components level, slope, and curvature themselves.

4.1. Implementing the Sign Restriction Approach

I start by introducing some useful notation and then turn to explaining the identification strategies in detail. For convenience, I restate the VAR in (5), which represents the state equation of my model:

$$F_t = \Phi_1 F_{t-1} + \Phi_2 F_{t-2} + \ldots + \Phi_p F_{t-p} + \omega_t$$

where $F_t = (F_t^{x\prime}, F_t^{y\prime})'$ is a $k = (k_x + 3)$ vector of common factors and $\Omega = E[\omega_t \omega_t']$ is the constant unconditional variance–covariance matrix of the one-step-ahead prediction errors. The aim is to identify innovations ν that are mutually orthogonal and standardized to have unit variance, i.e., $E[\nu_t \nu_t'] = I_k$. In order to trace impulse responses of the factors F to the innovations ν, one needs to find a matrix A that satisfies $\omega_t = A\nu_t$. Obviously, this implies $\Omega = AA'$. Following Uhlig (2005), I define an impulse vector as the column of some matrix A that has this property. As he shows, any such impulse vector a can be obtained by performing a Cholesky decomposition $\Omega = \tilde{A}\tilde{A}'$ and multiplying $\tilde{A}$ by some $k \times 1$ vector q of unit length, i.e., $a = \tilde{A}q$.

To compute impulse responses of the factors F, stack (5) as

$$\mathbf{F} = \overline{\mathbf{F}}\mathbf{\Phi} + \boldsymbol{w} \tag{8}$$

where $\mathbf{F} = [F_1, \ldots, F_T]'$, $\overline{F}_t = [F_{t-1}', \ldots, F_{t-p}']'$, $\overline{\mathbf{F}} = [\overline{F}_1, \ldots, \overline{F}_T]'$, $\mathbf{\Phi} = [\Phi_1, \ldots, \Phi_p]'$, and $\boldsymbol{w} = [\omega_1, \ldots, \omega_T]'$. For any $k \times 1$ impulse vector a, let $\mathbf{a} = [a', 0_{1,k(p-1)}]'$ and

$$\Gamma = \begin{bmatrix} \mathbf{\Phi}' \\ I_{k(p-1)} \ 0_{k(p-1),k} \end{bmatrix}$$

Then, the response $r_{a,i}(h)$ of factor i to an impulse a at horizon h is given by the ith element of the h-period-ahead prediction of Z after an impulse a, i.e.,

$$r_{a,i}(h) = (\Gamma^h \ \mathbf{a})_i \text{ for } h = 0, \ldots, H \tag{9}$$

In this paper, a joint dynamic factor model of macro and yield data is estimated. This modeling framework allows me to trace responses to the impulses driving the factors F to the individual variables in Z. Hence the economic interpretation of a given innovation can be based on a much broader information set than in usual VAR studies. Denoting by $r_a(h)$ the vector of impulse responses of the factors F to an impulse a at horizon h, the response of the nth variable in Z to that impulse a, denoted $r_a^n(h)$, can be computed as

$$r_a^n(h) = \lambda_n' r_a(h) \quad \text{for} \quad h = 0, \ldots, H \tag{10}$$

where λ_n' is the nth row of the factor loading matrix Λ.

I identify innovations that have a positive impact on the yield curve level, slope, or curvature.[3] This is achieved by imposing a positive response of one of the three factors over a given interval after the shock occurs. In order to separate out the initial effects on level, slope, and curvature,

[3] The positivity assumption is made for normalization reasons and is not restrictive. Since the model is linear, all reported results equivalently apply to negative surprise changes after flipping signs of the impulse responses.

DOI: 10.1002/jae

additional zero restrictions are introduced which ensure that only one of the three components moves on impact. These restrictions are summarized in the following definition:

Definition 1. A *'pure' level* impulse vector is an impulse vector a such that the response of the level factor is positive at horizons $h = 0, \ldots, H$ and such that the responses of the slope and the curvature factor are zero on impact.

Equivalent definitions apply to the slope and curvature factor, respectively.

This definition of impulse vectors involves a combination of zero and sign restrictions on the resulting impulse response functions. To implement these restrictions, I use the penalty function approach suggested by Mountford and Uhlig (2009). This approach amounts to solving

$$a = \arg \min_{a=\tilde{A}q} \Psi(a) \tag{11}$$

where

$$\Psi(a) = \sum_{h=0}^{H} f(-r_j(h)) \tag{12}$$

and where $f(\cdot)$ is some function defined over the real line that penalizes negative impulse responses. I follow Mountford and Uhlig (2009) and define $f(x) = 100x$ if $x \geq 0$ and $f(x) = x$ if $x < 0$. Since this choice of penalty function is to some extent arbitrary, I have also experimented with alternative specifications, which, however, provided qualitatively very similar results.

As my definition of a 'pure' level surprise requires that slope and curvature do not move on impact, the function Ψ needs to be minimized subject to zero restrictions which can be written as linear restrictions on the vector q as follows:

$$R_l q = 0 \tag{13}$$

R_l is a $2 \times k$ matrix with elements

$$R_l = \begin{bmatrix} \tilde{a}_{i,1} & \ldots & \tilde{a}_{i,k} \\ \tilde{a}_{i',1} & \ldots & \tilde{a}_{i',k} \end{bmatrix} \tag{14}$$

where i and i' denote the positions of the slope and curvature factors in the vector F, and where $\tilde{a}_{i,j}$ is the (i, j)-element of the Choleski decomposition $\tilde{A}$ of Ω.

In the analysis below, I identify orthogonal innovations to the level, slope, and curvature factors that are subject to zero restrictions as described above. To impose orthogonality of the three innovations, I first identify the level surprise $a_l = \tilde{A}q_l$ by minimizing $\Psi(a_1)$ subject to $R_l q = 0$. To identify the slope surprise, I then define the matrix R_l accordingly, and augment it such that $\tilde{R}_s = [R_s', q_l]'$. Finally, to identify the curvature surprise, I define the matrix R_c as above and augment it such that $\tilde{R}_c = [R_c', q_l, q_s]'$.

Finding vectors q that minimize the function Ψ subject to the linear restrictions R is a straightforward numerical exercise that can be eased by parameterizing the space of k-dimensional vectors of unit length. In the application studied in this paper, the total number of factors is $k = 7$

DOI: 10.1002/jae

and thus the parameterization

$$q = \begin{bmatrix} \cos(\alpha_1)\cos(\alpha_2)\cos(\alpha_3)\cos(\alpha_4) \\ \cos(\alpha_1)\cos(\alpha_2)\cos(\alpha_3)\sin(\alpha_4) \\ \cos(\alpha_1)\cos(\alpha_2)\sin(\alpha_3) \\ \cos(\alpha_1)\sin(\alpha_2) \\ \sin(\alpha_1)\cos(\alpha_5)\cos(\alpha_6) \\ \sin(\alpha_1)\cos(\alpha_5)\sin(\alpha_6) \\ \sin(\alpha_1)\sin(\alpha_5) \end{bmatrix} \tag{15}$$

can be used. According to this parameterization, any 7×1 vector of unit length is characterized by a set of six angles $\{\alpha_1, \ldots, \alpha_6\}$ defined over the interval $[0, 2\pi]$.

4.2. Relaxing the Zero Restrictions

The identification approach discussed above allows me to separate out shocks which on impact move only one of the three yield curve factors. In practice, however, yield curve innovations may rarely be attributed to one of the three components alone. I therefore employ a second identification routine which relaxes the restriction of zero contemporaneous impact. In particular, using sign restrictions as above, I identify positive surprises of the three yield curve factors whose contemporaneous effects on the other two yield curve factors are limited to a fixed number $\overline{a}$. I apply the following definition of a 'strong' level shock:

Definition 2. A *'strong' level* impulse vector is an impulse vector a that maximizes the penalty function Ψ for the level factor, but moves the slope and the curvature factor by a maximum of $\overline{a}$ basis points on impact.

Equivalent definitions apply to the slope and curvature factor, respectively.

This identification is simply achieved by applying the penalty function identification approach described above, but replacing the linear zero restrictions with inequality restrictions; i.e., instead of letting $Rq = 0$, minimize Ψ subject to

$$|Rq| <= \overline{a} \tag{16}$$

In practice, I apply the two identification schemes described above in the following way. I randomly select 1000 draws of the pair (Φ, Ω) from the sampled posterior distribution of the model parameters. For each of these draws, I find impulse vectors a that minimize the penalty function Ψ subject to the respective zero or inequality restrictions. I then compute the corresponding impulse responses according to equation (9) and store them. From the 1000 stored draws I then compute the mean as well as the percentiles of the posterior distribution.

Note that the implementation of my identification approach involves optimization of the penalty function Ψ subject to linear restrictions. While this is a straightforward numerical exercise, alternative empirical approaches are available. For example, Rubio-Ramírez *et al.* (2010) suggest identifying sign restrictions by making use of the QR decomposition. Since this approach does not require numerical optimization to find orthogonal impulse vectors q, it can computationally be more efficient than iteratively searching for orthogonal vectors of unit length. However, for the identification of the yield curve surprises as specified in Definition 1 or as, for example, in Mountford and Uhlig (2009), additional zero linear restrictions are imposed on the candidate impulse vectors. These additional restrictions make the use of numerical optimization routines necessary. Yet, for the identification of 'strong' surprises as defined in Definition 2 which involve

DOI: 10.1002/jae

inequality restrictions, the approach of Rubio-Ramírez *et al*. can be applied in the following way. First, randomly draw independent standard normal matrices $\tilde{X}$ of dimension $k \times k$ and then obtain orthonormal matrices Q via the QR decomposition of $\tilde{X}$. Next check if any three columns of the corresponding impulse matrix $A = \tilde{A}Q$ satisfy the sign restrictions and inequality restrictions of equation (16). If they do, keep the draw, otherwise discard it. Repeat this procedure 1000 times. While this is a feasible alternative to the approach described above, it turns out to be computationally more expansive in practice since many draws need to be discarded. I therefore report results based on the implementation described above which involves numerical optimization.

5. EMPIRICAL RESULTS

This section summarizes the results of the paper. First, I describe the data used and discuss the empirical specification. Then, I document the fit of the joint factor model of macro and yield data before I finally turn to the results of the impulse response analysis.

5.1. Data and Model Specification

I estimate the model using monthly data for the USA from 1983 : 01 until 2003 : 09. This time span covers the post-Volcker disinflation period and can thus be seen as a consistent monetary policy regime. The interest rate data used in this study are unsmoothed Fama–Bliss yields for maturities 1, 3, 6, 9, 12, 15, 18, 21, 24, 30 months and 3, 4, 5, 6, 7, 8, 9, and 10 years. Hence the term structure information is extracted from a wide range of maturities. I have further selected 25 variables of different economic categories in order to exploit the information in a variety of different macroeconomic time series. Table III lists these variables. Note that they are taken from a more comprehensive macroeconomic data panel for the USA that has been compiled by Giannone *et al*. (2004).[4]

As discussed above, the model decomposes yields into three factors which have a straightforward interpretation as level, slope, and curvature. The true number of factors driving the macroeconomic variables in the panel X is not known, however. In addition, the true number of lags in the joint VAR of macro and yield factors which represents the state equation of my model is also unknown. While formal tests for the optimal number of factors in dynamic factor models and the optimal lag length in autoregressions are readily available for classical analyses, at least from a computational perspective, model selection is a more intricate issue in Bayesian dynamic factor analysis. In a Bayesian framework, one commonly selects between two competing models on the basis of their posterior odds ratio. With equal prior probabilities on the two models, this equals the ratio of marginal likelihoods of the two models, which is also called the Bayes factor.

Justiniano (2004) discusses methods for selecting the optimal number of factors in Bayesian factor models that allow for lags in the observation equation. Using a simulation exercise, he documents that the Akaike information criterion (AIC) and the Schwartz information criterion (BIC) select the correct model as often as alternative approaches which involve computation of the marginal data density such as Geweke's Modified Harmonic Mean estimator. This is in line with Kass and Raftery (1995), who show that the BIC approximates the logarithm of the Bayes factor for large sample sizes. Since the BIC is based on the likelihood of a given model and therefore does not involve evaluation of prior densities, it is somewhat easier to compute. Table I provides the BIC criterion for several model specifications that I have considered. This table shows that the BIC is maximized for the combination of $k_x = 4$ macro factors and $p = 2$ lags in the state

[4] For further details on these data, the reader is referred to Giannone *et al*. (2004).

Table I. Bayesian information criterion (BIC) for different model specifications

	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$
$k_x=2$	4108.31	4035.49	3958.79	3880.67	3788.51
$k_x=3$	4638.81	4507.34	4393.65	4268.66	4143.90
$k_x=4$	4964.86	4818.14	4657.71	4503.64	4344.66

Note: This table lists estimates of the BIC for different model specifications of the dynamic factor model. The BIC is computed as $\text{BIC} = L(Z^T|\overline{\theta}_m, m) - \frac{M}{2}\log\ T$, where $L(Z^T|\overline{\theta}_m, m)$ denotes the log-likelihood of model specification m evaluated at the posterior mean of the corresponding parameter vector θ_m. M denotes the number of parameters for model m and T is the sample length. Model specifications with the number of macro factors ranging from $k_x = 2$ to $k_x = 4$ and with numbers of lags in the state equation from $p = 2$ to $p = 6$ have been considered. As the table shows, the BIC criterion is maximized for the model specification ($k_x = 4,\ p = 2$), which I therefore choose as the benchmark specification.

equation. Hence evidence provided by the data is the strongest in favor of this model specification, which I therefore choose to be the benchmark model. Unreported results show that my findings are qualitatively unchanged when I consider alternative model specifications.

5.2. Convergence of the Metropolis-within-Gibbs Sampler

The results reported below are based on 120,000 simulations of the Metropolis-within-Gibbs sampler summarized in Appendix B. In order to ensure convergence to the ergodic distribution, the first 70,000 iterations have been discarded as a burn-in. Of the remaining 50,000 draws, every 5th draw has been saved and from these 10,000 draws the moments and percentiles of the posterior distribution of the latent factors and model parameters have been computed. Figures 1 and 2 provide plots of the mean as well as the 5th and 95th percentiles of the posterior distribution for the two sets of estimated factors. As these figures show, all seven factors are very precisely estimated.

Figure 2 also provides a plot of the mean, the 5th, and the 95th percentile of the posterior distribution of the Nelson–Siegel yield curve loadings Λ_y introduced in equation (2). The loadings on the level factor are equal to one for all maturities and for all draws. The posterior distribution of the loadings on the slope and curvature factors, in turn, are determined by the posterior of the exponential decay parameter τ, which is sampled using the Metropolis step described in Appendix B. The plot in the lower panel of Figure 2 provides a histogram of the posterior distribution of τ. The mean of this posterior distribution is 0.0675, and the 5th and 95th percentiles are 0.0645 and 0.0709, respectively. Note that the posterior distribution of τ estimated here is mainly to the right of the value $\tau = 0.609$ which was suggested by Diebold and Li (2006). The higher posterior mean of τ translates into a peak of the curvature loading at maturity $n = 26$ months, in contrast to a peak of $n = 30$ months implied by Diebold and Li's suggestion. This indicates that their fixed parameter value is largely consistent with the yield data used in my study. In sum, the yield curve factors are very precisely estimated, which is an obvious prerequisite for precise inference in a factor model framework like the one employed here.

I use formal and informal criteria to assess convergence of the sampler. I compute two formal convergence diagnostics: the Raftery–Lewis measure of the number of draws required to achieve a certain precision of the sampler (Raftery and Lewis, 1992), and the relative numerical efficiency (RNE) indicator suggested by Geweke (1992). The parameters for the Raftery–Lewis diagnostic are specified as follows: quantile = 0.025, desired accuracy = 0.0125, and required probability of attaining the desired accuracy = 0.95. Since I compute both diagnostics for each parameter in

DOI: 10.1002/jae

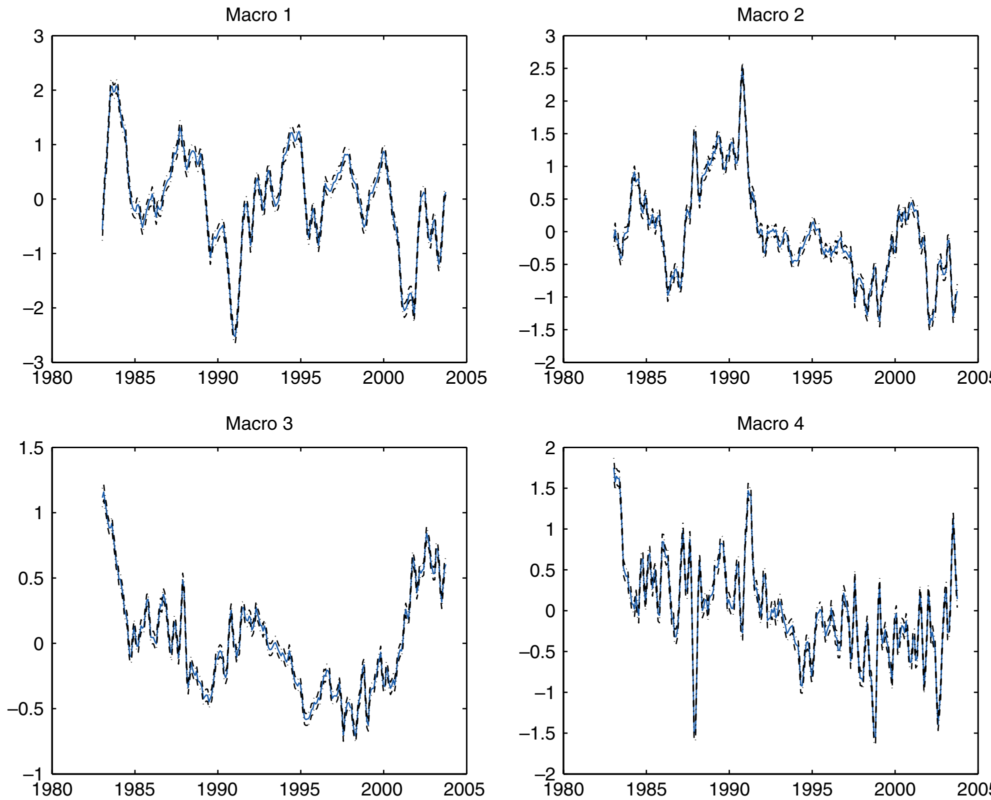

Figure 1. Estimated macro factors. This figure provides plots of the mean as well as the 5th and 95th percentiles of the posterior distribution of the macro factors F^x. This figure is available in color online at wileyonlinelibrary.com/journal/jae

the model, I obtain a cross-section of the Raftery–Lewis diagnostics and a cross-section of the Geweke diagnostics across the parameters of the model.

Table II summarizes the cross-section of these two formal convergence diagnostics across parameters.[5] The maximum of the Raftery–Lewis diagnostic across all parameters is 14,121 draws, which is far less than the 120,000 draws that I actually make. Moreover, for almost all parameters of the model, the Geweke indicator lies well below 20, which is the value considered as small enough to signal good mixing properties of the sampler (see, for example, Justiniano, 2004). The only exception is the exponential decay parameter τ, for which the Geweke measure indicates somewhat mixed convergence results. The likely reason is that this parameter is only weakly identified. Indeed, the middle right plot of Figure 2 shows that relatively large variations in τ only have a very small impact on the shape of the factor loadings Λ_y and hence on the yield curve factors F^y and all other parameters of the model. In that respect, the mixing properties of τ are unlikely to affect any of the conclusions reached in the paper.

I also monitor convergence of the Gibbs sampler informally. For example, I plot the evolution of draws for all parameters of the model. I also produce trace plots for the rolling mean of each parameter across the draws of the MCMC algorithm. Finally, I initialize the sampler with random

[5] I thank James LeSage for providing Matlab code to compute these statistics on the website www.spatial-econometrics.com.

DOI: 10.1002/jae

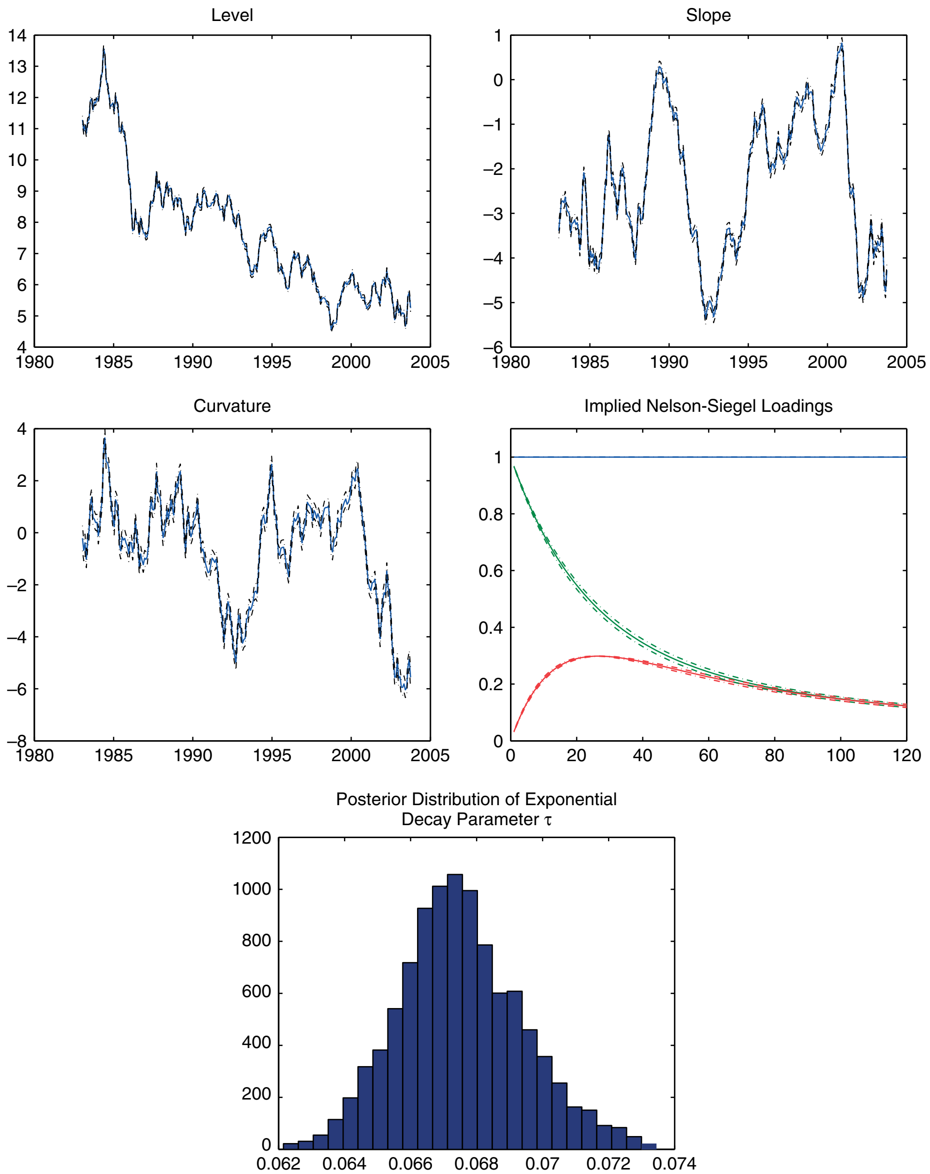

Figure 2. Estimated yield factors, Nelson–Siegel loadings, and decay parameter τ. This figure provides plots of the mean as well as the 5th and 95th percentiles of the posterior distribution of the estimated yield factors F^y and the corresponding yield factor loadings Λ_y. The bottom chart shows the posterior distribution of the exponential decay parameter τ. This figure is available in color online at wileyonlinelibrary.com/journal/jae

starting values and compare the outcomes.[6] In addition to the formal measures of convergence, these informal diagnostics give me confidence that the Markov chain has converged to its ergodic distribution.

[6] In order to conserve space, these informal model diagnostics are not provided here. However, they are available from the author upon request.

DOI: 10.1002/jae

Table II. Raftery–Lewis and Geweke test statistics

	5%	16%	Median	84%	95%	99%	Max.
Raftery–Lewis	584.00	599.00	655.00	834.66	1591.10	1873.52	14121.00
Geweke's RNE	0.76	0.98	1.87	4.10	7.94	14.54	232.68

Note: This table provides the Raftery and Lewis (1992) and Geweke (1992) statistics for assessment of convergence of the MCMC sampler. The parameters for the Raftery and Lewis statistics are specified as follows: quantile = 0.025, desired accuracy = 0.0125, and required probability of attaining the desired accuracy = 0.95. Both statistics are computed for each parameter of the model individually. The table shows the percentiles of the statistics across all parameters.

5.3. Assessing the Model Fit

In this section, I provide a set of results that allow one to assess how well the factor model fits the data. Table III lists for all macro variables in the panel the shares of variance explained by the four estimated factors. According to these figures, most variables exhibit a rather large common component. In particular, variables related to output and inflation are well explained by the four factors. About 70% of the growth rate of industrial production, for example, is captured by the macro factors. Moreover, about 94% of the variation of annual consumer price index (CPI) inflation is explained by the common components. Other business cycle related variables, inflation indicators, and stock price indices are equally well explained by the four factors. However, the variation of some time series such as personal consumption expenditures, the nominal effective exchange rate, and the federal government deficit is largely attributable to idiosyncratic variation. The first two rows of Figure 3 provide plots of some selected macro variables and how well they are explained by the common factors.

As has already been documented in previous studies (e.g., Diebold and Li, 2006; Diebold *et al.*, 2006), the Nelson–Siegel decomposition of yields explains the cross-sectional variation of interest rates of different maturities very well over time. Table IV shows that this result is confirmed by the analysis conducted here. Indeed, except for a few maturities at the very short and the very long end of the curve, all yields are nearly perfectly explained by the three factors. This implies that almost no information about yield curve dynamics is lost by restricting the analysis to surprise changes of its three components. The bottom row of Figure 3 highlights this result visually by providing plots of observed and fitted yields for the 12-month and 10-year maturity.

5.4. Impulse Response Analysis

The main focus of this paper is to shed light on the question of what information innovations of the term structure level, slope, and curvature convey about the future evolution of the economy. As discussed above, this question is approached by studying the impulse responses of macroeconomic variables to surprise changes of the three yield curve components. In this section, I summarize the results obtained from this exercise. I focus on impulse responses to yield curve surprises according to Definition 2, which *mainly* move one of the three components while the other two are restricted to move only little on impact. I also show plots of impulse responses based on Definition 1 which identifies surprise changes that *exclusively* move one of the three yield curve factors on impact. These provide qualitatively very similar results.

I start by studying the evolution of the term structure level, slope, and curvature following surprise changes to each of these three components, respectively. I then analyze the dynamics of macroeconomic variables subsequent to these innovations. For the sake of brevity, I only report impulse responses for industrial production, CPI inflation, and the New York Stock Exchange

DOI: 10.1002/jae

Table III. Macro variables and share of variance explained by estimated factors

Series	Tcode	F_1	F_2	F_3	F_4	All
Purchasing Managers Index	3	0.84	0.00	0.00	0.03	0.88
Outlook: prices paid	0	0.19	0.44	0.06	0.01	0.66
ISM mfg index: new orders	0	0.71	0.03	0.01	0.01	0.80
M2	3	0.04	0.05	0.17	0.03	0.24
Index of IP: total	3	0.68	0.01	0.01	0.04	0.70
Capacity utilization	3	0.09	0.17	0.84	0.05	0.99
Index of help-wanted advertising	3	0.57	0.01	0.00	0.01	0.63
Employment on nonagricultural payrolls	3	0.69	0.02	0.12	0.01	0.75
Avg weekly hours of production workers	3	0.12	0.02	0.00	0.00	0.15
Personal consumption expenditure	0	0.03	0.00	0.00	0.00	0.04
Construction put in place	3	0.28	0.04	0.00	0.00	0.33
Inventories: Mfg & Trade	3	0.27	0.06	0.22	0.05	0.48
NYSE Composite Index	3	0.00	0.08	0.12	0.76	0.98
S&P Composite	3	0.01	0.08	0.14	0.73	0.99
Nominal effective exchange rate	3	0.01	0.00	0.04	0.01	0.06
M1	3	0.00	0.02	0.28	0.03	0.32
M3	3	0.00	0.08	0.01	0.00	0.08
Loans & securities @ all comm. banks	4	0.05	0.00	0.00	0.00	0.05
PPI: finished goods	4	0.06	0.56	0.02	0.04	0.64
CPI: all items	4	0.03	0.85	0.00	0.07	0.94
PCE chain weight price index	4	0.00	0.70	0.03	0.15	0.98
Average hourly earnings	0	0.01	0.07	0.00	0.00	0.08
Philadelphia Fed Business Outlook	0	0.69	0.03	0.00	0.00	0.76
Outlook: prices received	0	0.33	0.26	0.03	0.01	0.63
Federal govt deficit or surplus	3	0.00	0.00	0.00	0.00	0.01

Note: This table lists the 25 macro variables that have been used to estimate the dynamic factor model for macro and yield data. Most series have been subjected to some transformation prior to the estimation, provided in the second column of the table. The transformation codes are: 0 = no transformation, 1 = logarithm, 2 = monthly differences, 3 = quarterly growth rate, 4 = annual growth rate. Columns denoted F_1 to F_4 provide for each of the 25 variables the share of variance explained by the four estimated factors, respectively. The last column shows the share of variance explained by all four factors.

(NYSE) composite stock market index. These three variables represent important economic categories and are all well explained by the common factors, as Table III shows.

5.4.1 *Term Structure Dynamics Following Surprise Changes of Level, Slope, and Curvature*

Figure 4 plots impulse responses of the three yield curve factors subsequent to positive surprise changes of the level, slope, and curvature factors, respectively, when the remaining two components are restricted to move only little on impact. The identification follows Section 4.1 and is based on a value of $H = 5$. Hence responses of the yield curve factors to the impulses defined above are required to be positive for 5 months after the initial surprise. As all three yield curve components are quite persistent, this restriction is actually not binding. Indeed, as Figure 4 shows, the impulse responses of the three yield curve factors remain well above zero for more than the restricted 5 months. Moreover, unreported results show that the conclusions remain qualitatively unchanged when different values for H are imposed. I restrict the responses of the other two yield curve components to be smaller than three basis points on impact. This ensures that the initial movement in yields can mainly be associated with one of the three components, while also allowing some flexibility for the reaction of the other two components. Moreover, it allows easier interpretation of the identified surprise changes in terms of basis points.

Several observations can be made based on Figure 4. First, positive surprise changes of the yield curve level are followed by hump-shaped responses of the slope and curvature factors, respectively. This implies that while initially all maturities move up in lockstep, in the months following a level

DOI: 10.1002/jae

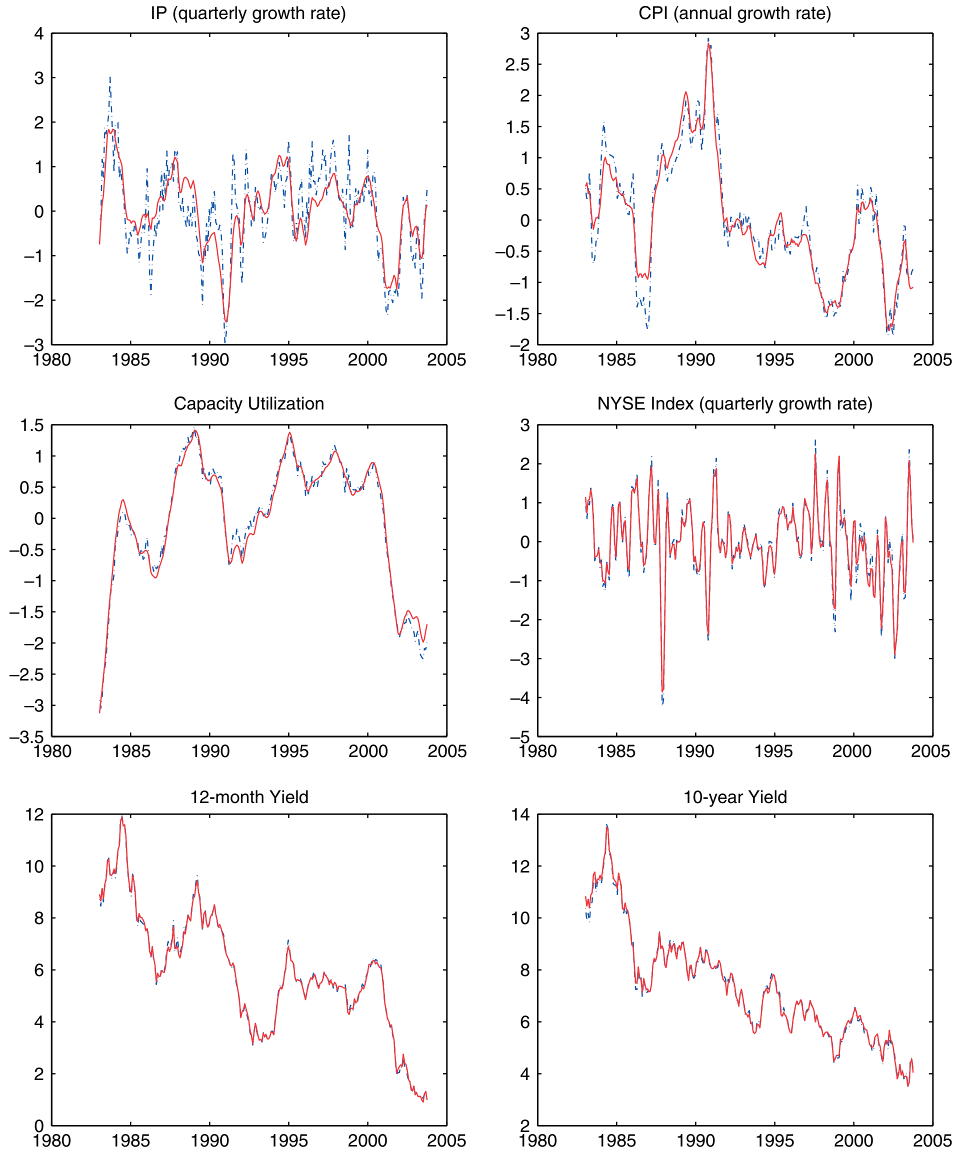

Figure 3. Model fit: macro variables and yields. This figure provides plots of observed (dashed) and fitted (solid) values for a selection of macro variables and zero coupon Treasury yields. This figure is available in color online at wileyonlinelibrary.com/journal/jae

surprise short- and medium-term maturities rise somewhat more strongly than longer maturities. Second, positive surprise changes to the yield curve slope are followed by an initial brief decline of the level of interest rates, which reverses after about 5 months and a hump-shaped positive response of the curvature factor. These dynamics imply that the short end of the yield curve moves up sharply, whereas medium- and longer-dated maturities only rise slightly in the months following the innovation. Third, positive surprise changes to the curvature factor are followed by a persistent decline of the level factor and by a strongly significant hump-shaped response of the slope factor. More precisely, the slope factor increases for about 15 months subsequent to a

DOI: 10.1002/jae

Table IV. Yields and share of variance explained by estimated factors

Maturity	L	L, S	L, S, C
1 month	0.544	0.987	0.987
3 months	0.566	0.996	0.998
6 months	0.595	0.988	0.999
9-months	0.619	0.980	0.999
12 months	0.631	0.973	0.999
15 months	0.655	0.969	0.999
18 months	0.681	0.967	1.000
21 months	0.705	0.964	1.000
24 months	0.725	0.961	1.000
30 months	0.762	0.962	1.000
3 years	0.793	0.965	0.999
4 years	0.847	0.969	0.999
5 years	0.879	0.974	0.999
6 years	0.910	0.980	0.999
7 years	0.928	0.983	0.999
8 years	0.944	0.987	0.998
9 years	0.955	0.989	0.998
10 years	0.963	0.989	0.997

Note: This table lists the 18 maturities of unsmoothed Fama–Bliss government bond yields that have been used for the estimation. The second to fourth columns provide, for each maturity, the cumulative shares of variance explained by the level (L), slope (S), and curvature (C) factor, respectively.

curvature surprise, and then gradually falls towards its initial level. The level factor, after rising briefly, falls below its initial level about 6 months after the curvature surprise and remains there for a long period of time. Hence, following a surprise surge of the curvature factor, long rates remain largely unchanged while short rates increase strongly. This implies a flattening of the yield curve in subsequent periods. This effect is visualized in Figure 5, which shows the evolution of the Treasury yield curve on impact, as well as 6 and 12 months after the initial 'strong' surprises to level, slope, and curvature.

As the third column of Figure 5 documents, a 'strong' curvature surprise initially shifts intermediate maturities up by about 15 basis points, whereas short- and long-term maturities are impacted much less. This is followed by a much stronger increase of short-term rates than long-term rates 6 and 12 months after the initial innovation, implying that curvature surprises precede a flattening of the Treasury yield curve. A strong flattening or inversion of the yield curve is commonly associated with an upcoming recession.

The upper left chart of the first column of the figure documents that a 'strong' positive surprise leads to a parallel upward shift of all yields by about 15 basis points. This is followed by a further upward shift of all maturities 6 and 12 months out. Finally, the middle column of Figure 5 shows that identified 'strong' slope surprises move short-term rates up by about 15 basis points on impact, but by construction only have a moderate initial impact on medium-term maturities and little effect on longer yields. To summarize, 'strong' level slope and curvature surprises move different sectors of the yield curve by about 15 basis points on impact. For comparison, the unconditional standard deviations of the monthly changes of yields in my dataset range from 29 to about 36 basis points and are thus fairly similar across the curve. The identified surprise changes thus roughly correspond to a one-half standard deviation of interest rates in a given sector of the curve on impact.

5.4.2 Macro Dynamics Following Surprise Changes of Level, Slope, and Curvature

I now discuss the macroeconomic dynamics following level, slope, and curvature surprises, in that order. Recall that the surprise changes to the three yield curve factors have been identified

DOI: 10.1002/jae

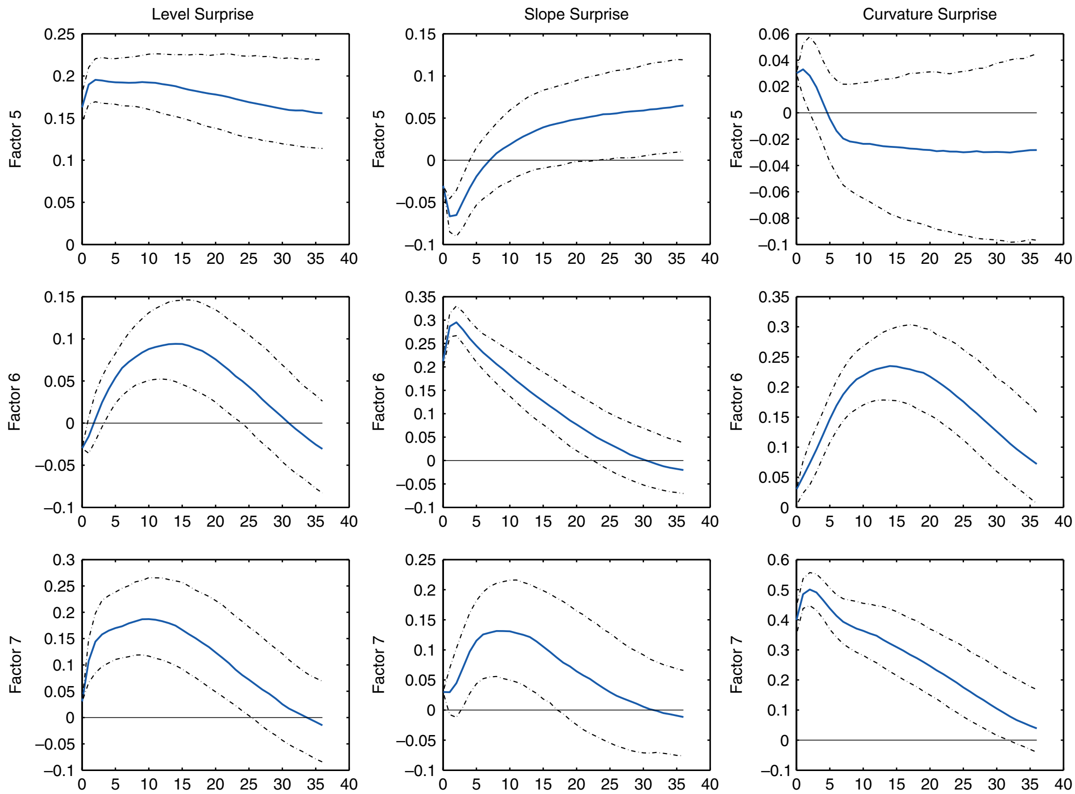

Figure 4. Level, slope, and curvature surprises and yield dynamics. This figure plots impulse responses of the yield factors level (upper row), slope (middle row), and curvature (lower row) to surprises as specified by Definition 2. The initial impact of the remaining two yield curve factors is restricted to be less than 3 basis points in absolute terms, respectively. Dash-dotted lines represent the 16% and the 84% quantiles of the posterior distribution of impulse responses. This figure is available in color online at wileyonlinelibrary.com/journal/jae

without imposing any restrictions on the responses of the macro factors. The reasoning behind this approach is that yield curve surprise changes may be associated with macroeconomic shocks and thus coincide with unexpected changes in macroeconomic variables. Figure 6 shows the dynamics of the quarterly growth rate of industrial production, annual CPI inflation, and the quarterly growth rate of the NYSE stock market index following the yield curve surprises.

Consistent with conventional wisdom, positive surprise changes of the yield curve level indicate a significant and persistent subsequent rise of inflation. As the middle left panel of Figure 6 reveals, inflation rises for about 5 months after a level surprise and then slowly reverts towards its initial level. Unexpected surges of the yield curve level also announce significant changes in output growth. According to the upper left panel of Figure 6, the growth rate of industrial production increases for about 3 months and then quickly reverts towards zero, eventually turning negative about 1 year after the initial level surprise. Moreover, level surprises appear to be accompanied by a brief simultaneous rise in equity prices.

Before turning to the impulse responses following a surprise change of the yield curve slope, it should be kept in mind that term spreads are traditionally considered to be strong predictors of recessions. Indeed, as has been documented by Estrella and Hardouvelis (1991), Estrella and Mishkin (1998), and many others, spreads between long-term and short-term interest rates forecast

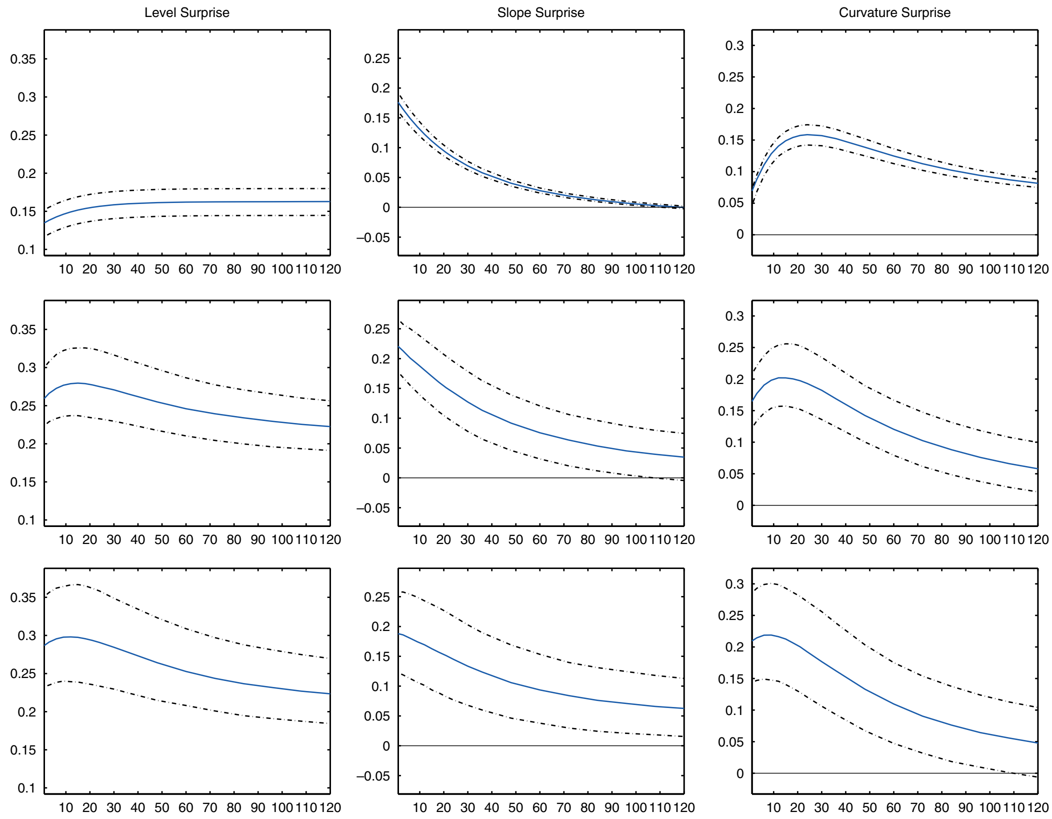

Figure 5. Level, slope, and curvature surprises: future yield curve shapes. This figure plots the differential effects on the shape of the Treasury yield curve on impact (upper row), 6 months (middle row), and 12 months (lower row) after the initial surprises to level (left column), slope (middle column), and curvature (right column), respectively. Dash-dotted lines represent the 16% and the 84% quantiles of the posterior distribution of impulse responses. This figure is available in color online at wileyonlinelibrary.com/journal/jae

recessions better than most other leading indicator variables. By studying the responses of different economic variables to *innovations* of the yield curve slope, the exercise conducted in this paper complements the mostly regression-based evidence on the link between the term spread and real activity. The middle column of Figure 6 summarizes impulse responses to an unexpected increase of the slope factor. The upper left plot shows the impulse response of industrial production (IP) growth. The conventional view that a rising slope factor (a decreasing spread between long-term and short-term bonds) announces an economic downturn seems to be confirmed here since the growth rate of industrial production–after a brief initial rise–falls below zero. Yet the link is rather weak compared to the initial rise of output growth following a surprise of the level or curvature factor. This result suggests that, controlling for the yield curve level and curvature, changes of the slope are not very informative about future economic growth.[7] When discussing the effects of curvature surprises further below, I provide evidence that reconciles these findings with previous results. The reaction of inflation to an unexpected increase of the slope factor is also hump-shaped and looks qualitatively similar to the response to a level surprise.

[7] Somewhat consistent with this result, Ang *et al*. (2006) find that the short rate forecasts GDP growth better than term spreads.

DOI: 10.1002/jae

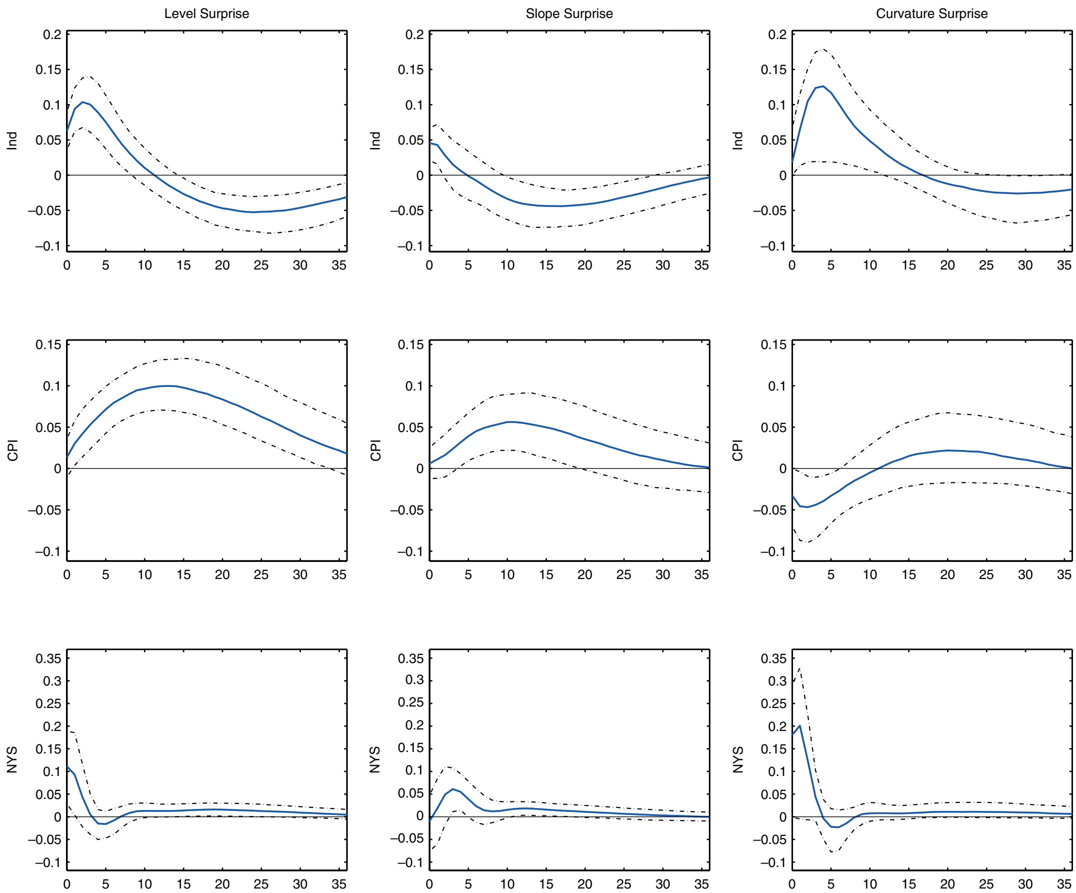

Figure 6. Level, slope, and curvature surprises and macro dynamics. This figure plots impulse responses of three selected macroeconomic variables to surprise changes of the yield curve level, slope, and curvature as specified by Definition 2. The initial impact of the remaining two yield curve factors is restricted to be less than 3 basis points in absolute terms, respectively. Dash-dotted lines represent the 16% and the 84% quantiles of the posterior distribution of impulse responses. This figure is available in color online at wileyonlinelibrary.com/journal/jae

Before discussing the impulse responses following a surprise change of the curvature factor, it should be noted that the previous literature has not analyzed in great detail the economic content of curvature. Only few studies have looked at the interaction of curvature with macroeconomic variables. DRA, for example, report negligible responses of macro variables to shocks in the curvature factor, an observation restated in Diebold *et al.* (2005). Yet their results are based on a recursive identification that does not allow yield curve factors to contemporaneously respond to macroeconomic shocks. Dewachter and Lyrio (2006) estimate latent yield factors within an affine term structure model and find some evidence suggesting that the curvature factor is related to real interest rate movements that are uncorrelated with macroeconomic variables. However, they base their results on regression analysis and do not investigate the information carried by yield curve innovations. Evans and Marshall (2004) study the responses of yield curve factors to specific macroeconomic shocks which they obtain based on estimations of theoretical models. They conclude that the curvature is largely unaffected by macroeconomic shocks. However, their

DOI: 10.1002/jae

approach does not allow one to study the evolution of macroeconomic variables subsequent to surprise changes of the yield curve factors, which is the focus of this paper.

The dynamics of the macroeconomic variables following a curvature surprise change are shown in the last column of Figure 6. These document that the growth rate of industrial production strongly rises for about 5 months and then slowly reverts, significantly falling below zero about 15 months after the surprise. Hence curvature surprises appear to announce a significant decline of output more than a year ahead. This is consistent with the strong positive response of the slope factor following a curvature surprise discussed above. This response is equivalent to a fall of the term spread, which in turn is commonly associated with lower future economic growth. Note further that, in contrast to level and slope surprises, inflation significantly falls for a few periods following a curvature surprise. It then bounces back after about 9 months. Finally, the NYSE equity market return shows a very strong and significant positive initial response when the curvature surprise hits, only falling below zero about 5 months after the initial surprise. Together, these findings imply that there is a rather strong connection between innovations of the curvature factor and subsequent macroeconomic dynamics. This is surprising in light of the fact that the curvature factor has previously been documented to be largely unrelated to economic variables.

Note that these results are not qualitatively altered by imposing the zero restriction that surprise changes of level, slope, and curvature do not contemporaneously move the other yield curve factors, respectively. Indeed, as Figures 7 and 8 show, the impulse responses follow identical patterns as for the case of the surprise changes identified under inequality restrictions, but exhibit slightly larger posterior uncertainty.[8]

5.4.3 Interpretation

Since my finding of a strong connection between curvature innovations and subsequent macroeconomic dynamics appears to be new to the literature, some further discussion is warranted. As mentioned earlier, my analysis does not attribute a structural interpretation to the term structure surprises identified with the sign restriction approach. Instead, my objective is to provide a descriptive account of the dynamic interactions between shifts in the components of the yield curve and key macroeconomic aggregates. Rather than representing exogenous economic shocks themselves, these shifts may be thought of as the market reaction to a certain type of economic shock or a combination of different economic shocks occurring at the same time.

As discussed above, I find that positive curvature innovations are associated with an immediate decline of inflation, a strong positive reaction of the stock market, and are followed by a persistent hump-shaped evolution of output growth. These dynamics are consistent with the following interpretation of curvature surprises. When market participants observe signals that indicate positive future output growth, they likely expect the monetary authority to eventually start raising the short-term interest rate. When, however, investors see inflation fall at the same time, they may expect this tightening to occur later in the future. Abstracting from term premia, this anticipation of a policy tightening in the medium term will contemporaneously lift interest rates of intermediate maturities relative to short- and long-term interest rates. If these expectations are realized ex post, we see output growth rise and inflation fall for a few periods after the positive surprise to curvature occurs. Moreover, the combination of higher expected output growth, lower inflation,

[8] As has been noted by Hamilton *et al*. (2007), normalizations in econometric models can affect the posterior distribution of the estimated parameters and potentially result in ill-behaved impulse response functions. Here, I use two normalizations. First, as discussed above, I impose that the diagonal elements of the upper $k_x \times k_x$ block of Λ_x be unity. Second, in the identification of 'pure' level, slope, and curvature surprises I impose zero restrictions on impulse response functions. To check whether these normalizations lead to ill-behaved posterior distributions of the reported impulse response functions, I generate histogram plots for the initial responses $r_a^n(0)$ shown in Figures 4–8. These histograms confirm that all reported posterior distributions are unimodal and well behaved.

DOI: 10.1002/jae

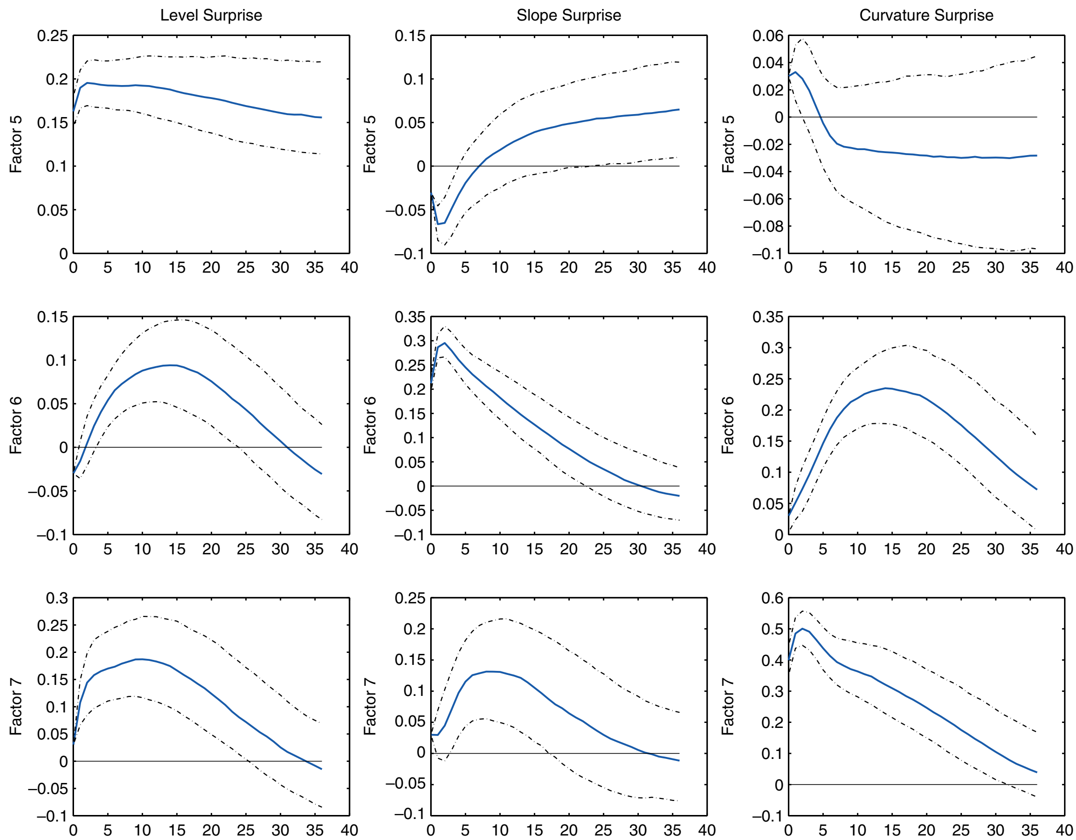

Figure 7. 'Pure' Level, slope, and curvature surprises and yield dynamics. This figure plots impulse responses of the yield factors level (upper row), slope (middle row), and curvature (lower row) to surprise changes identified with the 'pure' sign restriction approach as specified in Definition 1, respectively. Dash-dotted lines represent the 16% and the 84% quantiles of the posterior distribution of impulse responses. This figure is available in color online at wileyonlinelibrary.com/journal/jae

and a delayed tightening of monetary policy is also consistent with the positive stock market response that follows the curvature surprise.

6. CONCLUSION

The yield curve is known to convey information about the future course of the economy. Moreover, most of the yield curve variation is captured by its three components level, slope, and curvature. Although interest rate spreads are commonly used as predictors of recessions, the informational content of innovations of the three yield curve components is little understood to date. Therefore, in this paper, I provide a systematic account of the predictive information carried by innovations of the yield curve level, slope, and curvature.

The main result of my analysis is that unexpected changes of the curvature are quite informative about the future evolution of the term structure and the prospective dynamics of output, inflation, and stock prices. In particular, positive surprise changes of the curvature factor precede a pronounced flattening of the yield curve. Output growth follows a hump-shaped pattern, significantly falling below zero about a year and 3 months after the curvature surprise. At the same

DOI: 10.1002/jae

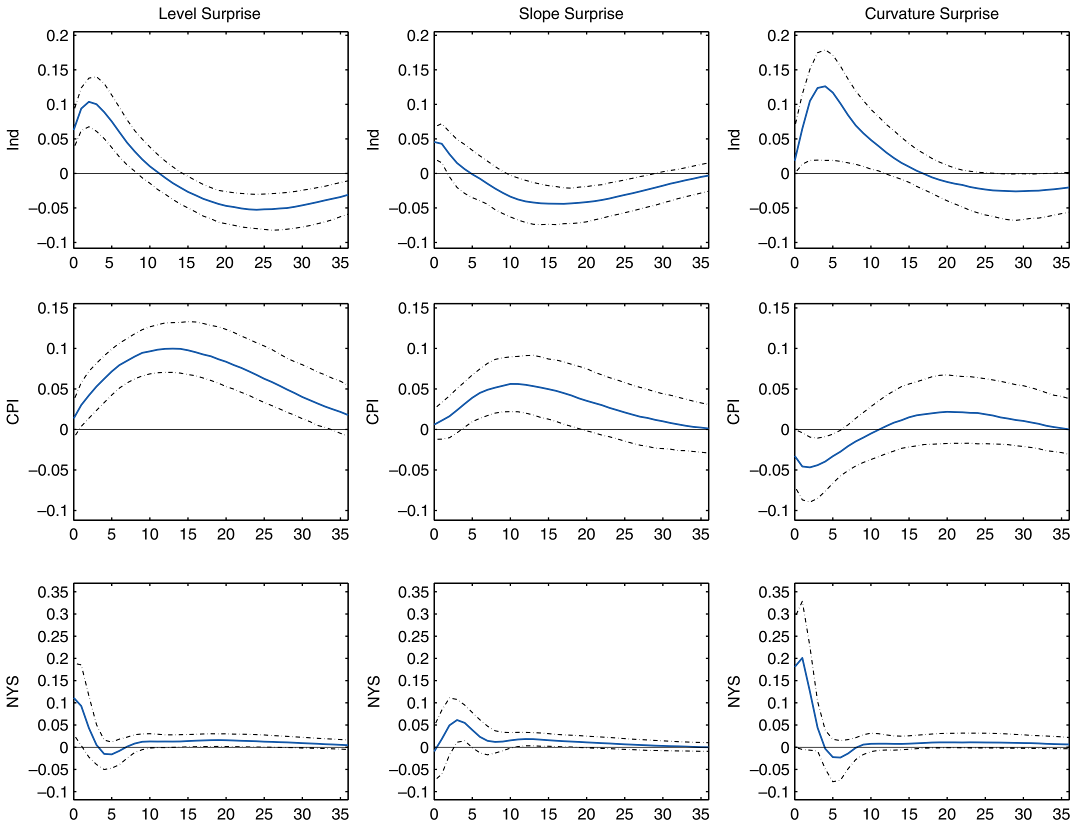

Figure 8. 'Pure' level, slope, and curvature surprises and macro dynamics. This figure plots impulse responses of three selected macroeconomic variables to surprise changes of the yield curve level, slope, and curvature identified with the 'pure' sign restriction approach as specified in Definition 1, respectively. Dash-dotted lines represent the 16% and the 84% quantiles of the posterior distribution of impulse responses. This figure is available in color online at wileyonlinelibrary.com/journal/jae

time, inflation falls and the stock market surges. These dynamic interactions are consistent with the interpretation that when the curvature of the Treasury term structure strongly rises while the yield curve level and slope are little changed, market participants may anticipate output to rise and a monetary policy tightening to occur in the medium term. In contrast, unexpected increases of the slope factor–tantamount to diminishing yield spreads–are followed by an immediate but not very pronounced decline of output. According to these findings, a rising slope factor is associated with a future decline of output, but is itself announced by innovations of the curvature factor. The results obtained for the level factor are more consistent with conventional wisdom. In particular, surprise surges of the yield curve level are followed by a strong and persistent increase of inflation rates. All results are qualitatively robust to variations of the identification scheme. In sum, these results suggest that innovations to the term structure components level, slope, and curvature carry information about future macroeconomic dynamics. Since we observe yield curves at very high frequencies but macroeconomic variables at much lower frequencies, my analysis may prove useful in interpreting the economic content of yield curve dynamics in real time.

My findings also suggest a few promising lines for future research. First, the model can be employed to forecast yields and macroeconomic variables making use of the common dynamics of term structure and macroeconomic factors. In particular, it would be interesting to compare its forecast performance with that of models which exclusively employ term structure (Diebold and Li, 2006) or macroeconomic dynamics (Moench, 2008). Second, to gain a deeper structural understanding of the dynamic correlations described here, one could try to identify specific economic shocks and assess the response of the yield curve components to these shocks. Ahmadi and Uhlig (2009) use sign restrictions to identify monetary policy shocks in a factor-augmented VAR model. It would be interesting to augment their model with a term structure in order to study the reaction of the yield curve components to this specific economic shock. A related recent contribution has been made by Kurmann and Otrok (2010), who find that a shock which maximizes the forecast error variance of the yield curve slope shares strongly similar characteristics with productivity news shocks identified in the macroeconomic literature. They argue that this provides an explanation of the predictive power of the term spread for future economic activity. Kurmann and Otrok do not, however, analyze the dynamic interaction of curvature with macroeconomic variables.

ACKNOWLEDGEMENTS

I am grateful to Samad Sarferaz, Almuth Scholl, Christian Stoltenberg, Harald Uhlig, two anonymous referees as well as seminar participants at the Econometric Society North American and European Summer Meetings for fruitful discussions, comments and advice. I also thank Lucrezia Reichlin and Robert Bliss for sharing their data with me. Earlier versions of this paper were written with financial support by the German National Academic Foundation and the Deutsche Forschungsgemeinschaft through the National Research Center 649 'Economic Risk'. The views expressed in this paper are those of the author and do not necessarily reflect the views of the Federal Reserve Bank of New York or the Federal Reserve System.

REFERENCES

Aguilar O, West M. 2000. Bayesian dynamic factor models and portfolio allocation. *Journal of Business and Economic Statistics* **18**(3): 338–357.

Ahmadi PA, Uhlig H. 2009. Measuring the dynamic effects of monetary policy shocks: a Bayesian FAVAR approach with sign restrictions. Mimeo, University of Chicago.

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

Ang A, Piazzesi M, Wei M. 2006. What does the yield curve tell us about GDP growth? *Journal of Econometrics* **127**(1–2): 359–403.

Bernanke BS, Boivin J. 2003. Monetary policy in a data-rich environment. *Journal of Monetary Economics* **50**(3): 525–546.

Bernanke B, Boivin J, Eliasz PS. 2005. Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics* **120**(1): 387–422.

Carter CK, Kohn R. 1994. On Gibbs sampling for state space models. *Biometrika* **81**: 541–553.

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

Diebold FX, Li C. 2006. Forecasting the term structure of government bond yields. *Journal of Econometrics* **127**(1–2): 337–364.

Diebold FX, Piazzesi M, Rudebusch G. 2005. Modeling bond yields in finance and macroeconomics. *American Economic Review Papers and Proceedings* **95**(2): 415–420.

Diebold FX, Rudebusch GD, Aruoba BS. 2006. The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics* **127**(1–2): 309–338.

Estrella A, Hardouvelis GA. 1991. The term structure as a predictor of real economic activity. *Journal of Finance* **46**(2): 555–576.

Estrella A, Mishkin FS. 1998. Financial variables as leading indicators predicting U.S. recessions. *Review of Economics and Statistics* **80**(1): 45–61.

Evans CL, Marshall D. 2004. Economic determinants of the nominal Treasury yield curve. Mimeo, Federal Reserve Bank of Chicago.

Geweke J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian Statistics*, Vol. 4, Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds). Oxford University Press: Oxford; 169–194.

Giannone D, Reichlin L, Sala L. 2004. Monetary policy in real time. *NBER Macroeconomics Annual*. MIT Press: Cambridge, MA; 161–225.

Hamilton J, Waggoner DF, Zha T. 2007. Normalization in econometrics. *Econometric Reviews* **26**(2–4): 221–252.

Hördahl P, Tristani O, Vestin D. 2006. A joint econometric model of macroeconomic and term-structure dynamics. *Journal of Econometrics* **127**(1–2): 405–444.

Justiniano A. 2004. Estimation and model selection in dynamic factor analysis. Mimeo, Princeton University.

Kadiyala KR, Karlsson S. 1997. Numerical methods for estimation and inference in Bayesian VAR-models. *Journal of Applied Econometrics* **12**(2): 99–132.

Kass RE, Raftery AE. 1995. Bayes factors. *Journal of the American Statistical Association* **90**(430): 773–795.

Kim C-J, Nelson CR. 1999. *State Space Models with Regime Switching*. MIT Press: Cambridge, MA.

Kose MA, Otrok C, Whiteman CH. 2003. International business cycles: world, region, and country-specific factors. *American Economic Review* **93**(4): 1216–1239.

Kurmann A, Otrok C. 2010. News shocks and the slope of the term structure of interest rates. Mimeo, University of Virginia.

Moench E. 2008. Forecasting the yield curve in a data-rich environment: a no-arbitrage factor-augmented VAR approach. *Journal of Econometrics* **146**: 26–43.

Mountford A, Uhlig H. 2009. What are the effects of fiscal policy shocks? *Journal of Applied Econometrics* **24**(6): 960–992.

Nelson CR, Siegel AF. 1987. Parsimonious modeling of yield curves. *Journal of Business* **60**(4): 473–489.

Raftery AE, Lewis S. 1992. How many iterations in the Gibbs sampler? In *Bayesian Statistics*, Vol. 4, Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds). Oxford University Press: Oxford; pp. 763–773.

Rubio-Ramírez JF, Waggoner DF, Zha T. 2010. Structural vector autoregressions: theory of identification and algorithms for inference. *Review of Economic Studies* **77**(2): 665–696.

Sarno L, Thornton DL. 2004. The efficient market hypothesis and identification in structural VARs. *Federal Reserve Bank of St Louis Review* **86**(1): 49–60.

Uhlig H. 2005. What are the effects of monetary policy on output? Results from an agnostic identification procedure. *Journal of Monetary Economics* **52**(2): 381–419.

DOI: 10.1002/jae

APPENDIX A: IDENTIFICATION AND NORMALIZATION OF THE FACTOR MODEL

Before discussing how the model (4)–(5) can be estimated, the restrictions needed to ensure unique identification of the model parameters and the unobserved factors have to be stated. Factor models suffer from the well-known problem of rotational indeterminacy, meaning that different rotations of the factors and model parameters may be observationally equivalent. Any feasible rotation of the factors must preserve the block-diagonal structure of the factor loading matrix Λ. This implies that the rotation matrix must also be block-diagonal. Accordingly, identification of the joint macro-finance factor model reduces to separate identification problems for each subset of factors.

Consider first the lower N_y equations of (4) corresponding to the observation equations for the yields. We know from equation (2) that the yield loadings are given by $\lambda_n^y = \left[1 \;\; \left(\frac{1-e^{-\tau n}}{\tau n}\right) \;\; \left(\frac{1-e^{-\tau n}}{\tau n} - e^{-\tau n}\right)\right]$. Any rotation of the yield factors must preserve this particular structure. It is straightforward to show that this only holds true for an identity matrix. Hence the tight parametric structure imposed by the Nelson–Siegel decomposition of yields ensures unique identification of the level, slope, and curvature factors. It thus remains to fix Λ_x such that unique identification of the macro factors F_x is guaranteed. Here, I can build on existing results by Aguilar and West (2000), who show that this is the case when the the upper left $k_x \times k_x$ block of Λ_x is lower-triangular with those on the diagonal, leaving the factor variances unrestricted.

APPENDIX B: ESTIMATION BY ESTIMATION BY MCMC METHODS

Initializing the Metropolis-within-Gibbs Sampler

If the model is exactly identified, the algorithm should converge to the ergodic distribution of the model parameters independently of the choice of initial parameter values. However, to achieve fast convergence of the sampler, it is advisable to choose a meaningful set of initial parameter values. A sensible choice for the macroeconomic factors would be the estimates obtained from principal components analysis. This approach, also pursued by Bernanke *et al*. (2005), has been employed here. Moreover, recall that the yield factor loadings are completely pinned down by the exponential decay parameter τ. I use the initialization $\tau = 0.0609$, which is the value suggested by Diebold and Li (2006). Given the latter, I obtain starting values of the yield factors F_y by regressing the yields Y onto the starting values of Λ_y implied by $\tau = 0.0609$. Given the two sets of initial factor estimates, starting values of the parameters (Φ, Ω) governing the dynamics of the state equation are obtained by estimating the VAR in (5) via ordinary least squares (OLS). Moreover, I get initial values of the parameters (Λ, R) in the observation equation from running univariate regressions of the variables in Z onto the initial factor estimates.

I also ran the sampler various times with randomly generated starting values. These were obtained as follows. First, I randomly generate time series of pseudo macro factors F_x. I then obtain starting values for the factor loadings Λ_x and the variances of the idiosyncratic components by regressing the macro variables X onto the randomly generated factors. Moreover, I randomly generate an exponential decay parameter τ from a normal distribution with mean 0.0609 and variance 0.01. Given the latter, I obtain starting values of the yield factors F_y by regressing the yields Y onto the randomly generated loadings Λ_y. I then obtain starting values for the variances of the idiosyncratic yield components from the residuals of these regressions. Finally, given starting values for the macro and yield factors, I obtain starting values for the state equation parameters

DOI: 10.1002/jae

Φ and Ω by estimating a VAR(p) via OLS. Unreported results confirm that the posterior means of the factors are indistinguishable for different sets of such randomly chosen starting values.

Sampling from $p(\tilde{F}_T|\tilde{Z}_T, \theta^{(i-1)})$

Based on a result of Carter and Kohn (1994), Kim and Nelson (1999) show that draws from the conditional distribution of the latent factors F can be obtained by performing the following steps. First run the Kalman filter forward to obtain estimates $\overline{F}_{T|T}$ of the factors in period T and their variance–covariance matrix $P_{T|T}$ based on all available sample information. Then, for $t = T-1, \ldots, 1$ proceed backwards to generate draws $\overline{F}_{t|T}$ from

$$\overline{F}_{t|T}|\overline{F}_{t+1|T}, \tilde{Z}_T, \theta \sim N(\overline{F}_{t|t,\overline{F}_{t+1|T}}, P_{t|t,\overline{F}_{t+1|T}}) \tag{17}$$

where

$$\overline{F}_{t|t,\ \overline{F}_{t+1|T}} = \overline{F}_{t|t} + P_{t|t}\overline{\Phi}'\ (\overline{\Phi}\ \overline{P}_{t|t}\ \overline{\Phi}' + \overline{\Omega})^{-1}\ (\overline{F}_{t+1} - \overline{\Phi}\ \overline{F}_{t|t})$$

and

$$P_{t|t,\ P_{t+1|T}} = P_{t|t} - P_{t|t}\ \overline{\Phi}'\ (\overline{\Phi}\ \overline{P}_{t|t}\ \overline{\Phi}' + \overline{\Omega})^{-1}\ \overline{\Phi}\ P_{t|t}$$

Kim and Nelson show that this algorithm needs to be slightly modified if $\overline{\Omega}$ is singular. This is the case here since the state equation includes more than one lag and the model is written in companion form. Denote by $\overline{\Omega}^*$ the upper left $k \times k$ block of $\overline{\Omega}$ which is positive-definite and let $\overline{F}_t^*$ and $\overline{\Phi}^*$ be the first k rows of $\overline{F}_t$ and $\overline{\Phi}_t$. As before, run the Kalman filter forward to obtain $\overline{F}_{T|T}$ and $P_{T|T}$. Then, for $t = T-1, \ldots, 1$ proceed backwards to generate draws $\overline{F}_{t|T}$ from

$$\overline{F}_{t|T}|\overline{F}^*_{t+1|T}, \tilde{Z}_T, \theta \sim N(\overline{F}^*_{t|t,\overline{F}_{t+1|T}}, P^*_{t|t,\overline{F}_{t+1|T}})$$

where

$$\overline{F}^*_{t|t,\ \overline{F}_{t+1|T}} = \overline{F}_{t|t} + P_{t|t}\overline{\Phi}'^*\ (\overline{\Phi}^*\ \overline{P}_{t|t}\ \overline{\Phi}'^* + \overline{\Omega}^*)^{-1}\ (\overline{F}^*_{t+1} - \overline{\Phi}^*\ \overline{F}_{t|t})$$

and

$$P^*_{t|t,\ P_{t+1|T}} = P_{t|t} - P_{t|t}\ \overline{\Phi}^{*\prime}\ (\overline{\Phi}^*\ \overline{P}_{t|t}\ \overline{\Phi}^{*\prime} + \overline{\Omega}^*)^{-1}\ \overline{\Phi}^*\ P_{t|t}$$

Sampling from $p(\theta^{(i)}|\tilde{Z}_T, \tilde{F}_T^{(i)})$

Conditional on the data and draws of the unobserved factors, the observation equation (4) of the state-space model amounts to a set of $N_x + N_y$ regressions. Since the errors are assumed to be mutually orthogonal, one can sample each equation's parameters independently. Moreover, conditional on the factor draws, equation (5) is just a VAR(p) in the factors. Hence standard results from Bayesian VARs can be applied to estimate the parameters of the state equation. I start with describing the algorithm to sample from the parameters of the observation equation and then move to the procedure used to draw the parameters of the state equation.

Sampling Λ and R

Consider first the set of equations characterizing the decomposition of the macroeconomic variables X into common components F^x and idiosyncratic components e^x. Except for the lower-triangularity imposed on the upper left $k_x \times k_x$ block in order to ensure exact identification of the factors, the loadings Λ_x are unrestricted. Since the idiosyncratic components are assumed to be independent

(R is diagonal), the parameters of the macro observation equations can be obtained by estimating N_x independent univariate linear regression models. Using natural conjugate priors

$$p\ (\lambda_i^x|R_{ii}) = N(\lambda_{i0}^x, R_{ii}V_{i0}^{-1})$$

and

$$p\ (R_{ii}) = iG(\nu_{i0}, \nu_{i0}\ \sigma_{i0}^2)$$

standard Bayesian results show that the conditional posterior distributions of λ_i^x and R_{ii} are given by

$$p\ (\lambda_i^x|R_{ii}, \tilde{X}_T, \tilde{F}_T, \theta_{-\lambda_i^x}) = N(\overline{\lambda}_i^x, R_{ii}V_i^{-1}) \tag{18}$$

and

$$p\ (R_{ii}|\tilde{X}_T, \tilde{F}_T, \theta_{-R_{ii}}) = iG(\nu_i, \nu_i\sigma_i^2) \tag{19}$$

where

$$\begin{aligned} \overline{\lambda}_i^x &= V_i^{-1}(V_{i0}^{-1}\ \lambda_{i0}^x + F^{x\prime}F^x\ \hat{\lambda}_i^x), \\ V_i &= V_{i0} + F^{x\prime}F^x, \\ \nu_i &= \nu_{i0} + T, \text{ and} \\ \nu_i\ \sigma_i^2 &= \nu_{i0}\sigma_i^2 + (T - k_x)\ S_i^2 + \\ &\quad + (\hat{\lambda}_i^x - \lambda_{i0})'[V_{i0}^{-1} + (F^{x\prime}F^x)^{-1}]^{-1}(\hat{\lambda}_i^x - \lambda_{i0}) \end{aligned}$$

where $S_i^2 = \dfrac{1}{T-k_x}(x_i - F^x\hat{\lambda}_i^{x\prime})'(x_i - F^x\hat{\lambda}_i^{x\prime})$ is the sum of squared fitting errors of the ith equation estimated via OLS. Following Bernanke *et al.* (2005), I use proper but diffuse priors ($\lambda_{i0}^x = 0_{k_x \times 1}$, $\nu_{i0} = 0.001$, $\nu_{i0}\ \sigma_{i0}^2 = 3$, $V_{i0}^{-1} = I$) for the estimation.

Matters are more complicated when it comes to estimating Λ^y. Recall that the Diebold–Li loadings of the yield of maturity n on the three factors F^y are given by 1, $\left(\dfrac{1-e^{-\tau n}}{\tau n}\right)$, and $\left(\dfrac{1-e^{-\tau n}}{\tau n} - e^{-\tau n}\right)$. Hence the exponential decay parameter τ is the only unknown parameter in Λ_y. The observation equations related to the N_y yields therefore amount to a nonlinear least squares problem that has to be solved for τ. As a consequence, the distribution of τ is non-standard and cannot be sampled from directly. I therefore set up the following random walk Metropolis algorithm to sample τ:

1. Conditional on τ^{j-1}, generate a proposal from $\tau^* = \tau^{(j-1)} + c\varepsilon$, where $\varepsilon \sim N(0, 1)$; i.e., draw τ^* from $N(\tau^{(j-1)}, c^2)$.
2. Accept τ^* with probability

$$\begin{aligned} \alpha &= \min\left\{1, \frac{p\ (\tau^*|\tilde{Z}_T, \tilde{F}_T, \theta_{-\tau})\ q(\tau^*|\tau^{(i-1)})}{p\ (\tau^{(i-1)}|\tilde{Z}_T, \tilde{F}_T, \theta_{-\tau})\ q(\tau^{(i-1)}|\tau^*)}\right\} \\ &= \min\left\{1, \frac{p\ (\tau^*|\tilde{Z}_T, \tilde{F}_T, \theta_{-\tau})}{p\ (\tau^{(i-1)}|\tilde{Z}_T, \tilde{F}_T, \theta_{-\tau})}\right\} \\ &= \min\left\{1, \frac{p\ (\tilde{Z}_T|\tilde{F}_T, \tau^*, \theta_{-\tau})\ p(\tau^*)}{p\ (\tilde{Z}_T|\tilde{F}_T, \tau^{(i-1)}, \theta_{-\tau})\ p(\tau^{(i-1)})}\right\} \end{aligned} \tag{20}$$

DOI: 10.1002/jae

Note that the q-terms drop since the proposal density is symmetric around $\tau^{(j-1)}$. Hence, with a flat prior on τ, drawing τ amounts to generating a proposal value from a normal distribution centered around the last iteration's value, and to accept that proposal with probability given by the ratio of likelihoods implied by the two candidates, $\frac{p\ (\tilde{Z}_T|\tilde{F}_T, \tau^*, \theta_{-\tau})}{p\ (\tilde{Z}_T|\tilde{F}_T, \tau^{(i-1)}, \theta_{-\tau})}$. This ratio can be computed easily using factor loadings $\Lambda_y(\tau^*)$ and $\Lambda_y(\tau^{(i-1)})$. The scaling parameter c is calibrated so that acceptance ratios between 0.2 and 0.5 are obtained. Further, as mentioned above, I initialize the sampler using the value of $\tau = 0.0609$.

Conditional on a draw of τ and the latent yield factors F^y, the variances R^y_{ii} of the pricing errors represent mutually independent regression residuals that can be drawn individually. Specifying a natural conjugate inverse Gamma prior distribution, the posterior is given by a conjugate inverse Gamma distribution as in (19).

Sampling Φ *and* Ω

The state equation of the model (4)–(5) is a VAR(p) in the factors F. To estimate the parameters of this VAR, I follow Bernanke *et al*. (2005) by imposing a diffuse conjugate Normal–Wishart prior:

$$p\ (\text{vec}(\Phi)|\Omega) = N(0, \Omega \otimes Q_0)$$
$$p\ (\Omega) = iW(\Omega_0, \nu_0)$$

where diagonal elements of Ω_0 are set so as to equal the residual variances of the corresponding p-lag univariate autoregressions, σ_i^2. Further, in the spirit of the Minnesota prior, diagonal elements of Q_0 are set so that the prior variance of the l-lagged jth variable in equation i equals $\sigma_i^2/(l\sigma_j^2)$. Using a result from Kadiyala and Karlsson (1997), the conditional posterior distributions are then given by

$$p\ (\Omega|\ \tilde{Z}_T, \tilde{F}_T, \theta_{-\Omega}) = iW(\overline{\Omega}, T + \nu_0) \tag{21}$$
$$p\ (\text{vec}(\Phi)|\ \Omega, \tilde{Z}_T, \tilde{F}_T, \theta_{-\text{vec}(\Phi)}) = N(\text{vec}(\overline{\Phi}), \Omega \otimes \overline{Q}) \tag{22}$$

where

$$\overline{\Omega} = \Omega_0 + \hat{\omega}'\hat{\omega} + \hat{\Phi}'(\tilde{F}_T'\tilde{F}_T)\hat{\Phi} - \overline{\Phi}'[Q_0^{-1} + (\tilde{F}_T'\tilde{F}_T)]\overline{\Phi},$$
$$\overline{\Phi} = \overline{Q}\ (\tilde{F}_T'\tilde{F}_T)\ \hat{\Phi},$$

and

$$\overline{Q} = [Q_0^{-1} + (\tilde{F}_T'\tilde{F}_T)]^{-1}$$

and where $\hat{\omega}$ denotes the matrix of OLS residuals. Stationarity of the VAR parameters is enforced by discarding draws of Φ that have eigenvalues greater than 1.00001 in absolute value.

DOI: 10.1002/jae