Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Monetary Economics 50 (2003) 745–787
Journal of MONETARY ECONOMICS
www.elsevier.com/locate/econbase

A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables☆

Andrew Ang[a,b,*], Monika Piazzesi[b,c]

[a] *Columbia Business School, 3022 Broadway, 805 Uris, New York, NY 10027, USA*
[b] *National Bureau of Economic Research, Cambridge, MA 02138, USA*
[c] *Anderson Graduate School of Management, UCLA, 110 Westwood Plaza, Box 951481, Los Angeles, CA 90095-1481, USA*

Received 14 June 2000; received in revised form 29 January 2002; accepted 24 July 2002

Abstract

We describe the joint dynamics of bond yields and macroeconomic variables in a Vector Autoregression, where identifying restrictions are based on the absence of arbitrage. Using a term structure model with inflation and economic growth factors, together with latent variables, we investigate how macro variables affect bond prices and the dynamics of the yield curve. We find that the forecasting performance of a VAR improves when no-arbitrage restrictions are imposed and that models with macro factors forecast better than models with only unobservable factors. Variance decompositions show that macro factors explain up to 85% of the variation in bond yields. Macro factors primarily explain movements at the short

☆We have had helpful discussions with Geert Bekaert, Mike Chernov, John Cochrane, Darrell Duffie, Charles Evans, Mark Ferguson, Joel Hasbrouck, Bob Hodrick, Avi Kamara, Jun Liu, David Marshall, Jun Pan, Thomas Sargent, Stephen Schaefer, Martin Schneider, Ken Singleton, Suresh Sundaresan, Chris Telmer and Jessica Wachter. We thank seminar participants at Columbia University, NYU, the NBER Summer Institute, the American Economics Association, the Western Finance Association, and the Seventh World Congress of the Econometric Society. We thank Marty Eichenbaum (the editor) and we are grateful for extremely helpful comments from an anonymous referee. Andrew Ang acknowledges funding from the National Science Foundation.

*Corresponding author. Professor Andrew Ang, Business School, Columbia University, 3022 Broadway, 805 Uris, New York, NY 10027, USA. Tel.: +1-212-854-9154; fax: +1-212-662-8474.

E-mail address: aa610@columbia.edu (A. Ang).

URL: http://www.columbia.edu/~aa610.

0304-3932/03/$ - see front matter
doi:10.1016/S0304-3932(03)00032-1

end and middle of the yield curve while unobservable factors still account for most of the movement at the long end of the yield curve.

JEL classification: C13; C32; E43; E44; E52

Keywords: Estimation; Time series models; Determination of interest rates; Financial markets and the macroeconomy; Monetary policy

1. Introduction

Describing the joint behavior of the yield curve and macroeconomic variables is important for bond pricing, investment decisions and public policy. Many term structure models have used latent factor models to explain term structure movements, and although there are some interpretations to what these factors mean, the factors are not given direct comparisons with macroeconomic variables. For example, Pearson and Sun (1994)'s factors are labeled "short rate" and "inflation", but their estimation does not use inflation data. The terms "short rate" and "inflation" are just convenient names for the unobserved factors. Another example is Litterman and Scheinkman (1991), who call their factors "level," "slope" and "curvature". Similarly, Dai and Singleton (2000) use the words "level," "slope" and "butterfly" to describe their factors. These labels stand for the effect the factors have on the yield curve rather than describing the economic sources of the shocks.

In the absence of a workhorse general equilibrium model for asset pricing (see Hansen and Jagannathan, 1991), factor models have the advantage that they only impose no-arbitrage conditions and not all other conditions that characterize the equilibrium in the economy. Most existing factor models of term structure are unsatisfactory, however, because they do not model how yields directly respond to macroeconomic variables.[1] In contrast, empirical studies try to directly model the relationships between bond yields and macro variables by using vector autoregressive (VAR) models. Studies like Estrella and Mishkin (1997) and Evans and Marshall (1998) use VARs with yields of various maturities together with macro variables. These studies infer the relationships between yield movements and shocks in macro variables using impulse responses (IRs) and variance decomposition techniques implied from the VAR. For example, Evans and Marshall (2001) associate shocks to economic activity and price levels with level effects across the yield curve. Another type of shock, which can be identified with various schemes, comes from monetary policy.[2]

[1] The exception is Piazzesi (2001), who develops a term structure model with interest-rate targeting by the central bank. In the model, the central bank reacts to macroeconomic variables such as nonfarm payroll employment.

[2] See, for example, Gali (1992), Sims and Zha (1995), Bernanke and Mihov (1998), Christiano, et al. (1996), and Uhlig (2001). For a survey, see Christiano, et al. (1999).

Existing macro VAR studies are characterized by three features. First, only maturities whose yields which have been included in the VAR may have their behavior directly inferred by the dynamics of the VAR. As an unrestricted VAR is generally not a complete theory of the term structure, it says little about how yields of maturities not included in the VAR may move. Second, the implied movements of yields in relation to each other may not rule out arbitrage opportunities when the cross-equation restrictions implied by this assumption are not imposed in the estimation. Finally, unobservable variables cannot be included as all variables in the VAR must be observable. The VAR approach, however, is very flexible, and the implied impulse response functions (IRs) and variance decompositions give insights into the relationships between macro-shocks and movements in the yield curve.

A related asset-pricing literature beginning with Sargent (1979) has tried to estimate VAR systems of yields under the null of the Expectations Hypothesis (see Bekaert and Hodrick, 2001). These systems do not contain macro variables, which is the focus of our paper. Moreover, expected excess returns on US bonds vary over time (see, for example, Campbell and Shiller, 1991). The term structure dynamics in this paper are therefore given by a Gaussian term structure model with time-varying risk premia, consistent with deviations from the Expectations Hypothesis (see Fisher, 1998; Duffee, 2002; Dai and Singleton, 2002).

We incorporate macro variables as factors in a term structure model by using a factor representation for the pricing kernel, which prices all bonds in the economy. This is a direct and tractable way of modelling how macro factors affect bond prices. The pricing kernel is driven by shocks to both observed macro factors and unobserved factors. Since macro factors are correlated with yields, incorporating these factors may lead to models whose forecasts are better than models which omit these factors. We investigate whether the purely unobservable factors of multi-factor term structure models can be explained by macro variables, and we examine how the latent factors change when macro variables are incorporated into such models.

Our methodology gives us several advantages over existing empirical VAR approaches. First, it allows us to characterize the behavior of the entire yield curve in response to macro shocks rather than just the yields included in the VAR. Second, a direct comparison of macro variables with latent yield factors can be made. Third, variance decompositions and other methods can estimate the proportion of term structure movements attributable to observable macro shocks, and other latent variables. Finally, our approach retains the tractability of the VAR approaches because we estimate a VAR subject to nonlinear no-arbitrage restrictions.

Our term structure model is Gaussian, so it is a VAR model, and IRs and variance decompositions can be easily computed. Formally, our model is a special case of discrete-time versions of the affine class introduced by Duffie and Kan (1996), where bond prices are exponential affine functions of underlying state variables. In our model, however, some of the state variables are observed macroeconomic aggregates. With Gaussian processes, the affine model reduces to a VAR with cross-equation restrictions. Our set-up accommodates lags in the driving factors and allows us to compute variance decompositions where we can attribute the proportion of movements in the yield curve to observable and unobservable factors. We can plot

IRs of shocks to various factors on any yield, since the no-arbitrage model gives us bond prices for all maturities.

We obtain our measures of inflation and real activity by extracting principal components of two groups of variables that are selected to represent measures of price changes and economic growth. These factors are then augmented by latent variables. As term structure studies have suggested up to three latent factors as appropriate to capture most salient features of the yield curve, we estimate models with three latent factors in addition to the macro variables. Our main model has three correlated unobservable factors, together with the two macro factors (inflation and real economic activity).

Imposing no-arbitrage restrictions improves out-of-sample forecasts from a VAR. Forecasts can be further improved by incorporating macro factors into models with latent variables. We show that a significant part of the latent factors implied by traditional models with only latent yield variables can be attributed to macro variables. In particular, ''slope'' and ''curvature'' factors can be related to macro factors, while the ''level'' factor survives largely intact when macro variables are incorporated.

We find that macro factors explain a significant amount of the variation in bond yields. Macro factors explain up to 85% of the forecast variance for long forecast horizons at short and medium maturities of the yield curve. The proportion of the forecast variance of yields attributable to macro factors decreases at longer yields. At the long end of the yield curve, 60% of the forecast variance is attributable to macro factors at a 1-month forecast horizon, while at very long forecast horizons, over 60% of the variance is attributable to unobservable factors.

The paper is organized as follows. Section 2 summarizes the data. Section 3 motivates an affine equation for the short rate, which can be interpreted as a Taylor (1993) regression of the short rate on macro factors and an 'unexplained' orthogonal component. Section 4 presents the general model, describes the specific parameterization of the model and discusses the estimation strategy. We present our estimation results in Section 5, and discuss the implied IRs, variance decompositions and forecasting results. Section 6 concludes.

2. Data

2.1. Yield data

We use data on zero coupon bond yields of maturities 1, 3, 12, 36 and 60 months from June 1952 to December 2000. The bond yields (12, 36 and 60 months) are from the Fama CRSP zero coupon files, while the shorter maturity rates (1 and 3 months) are from the Fama CRSP Treasury Bill files. All bond yields are continuously compounded. Fig. 1 plots some of these yields in the upper graph and Table 1 presents some sample statistics. The table shows that our data are characterized by some standard stylized facts. The average postwar yield curve is upward sloping;

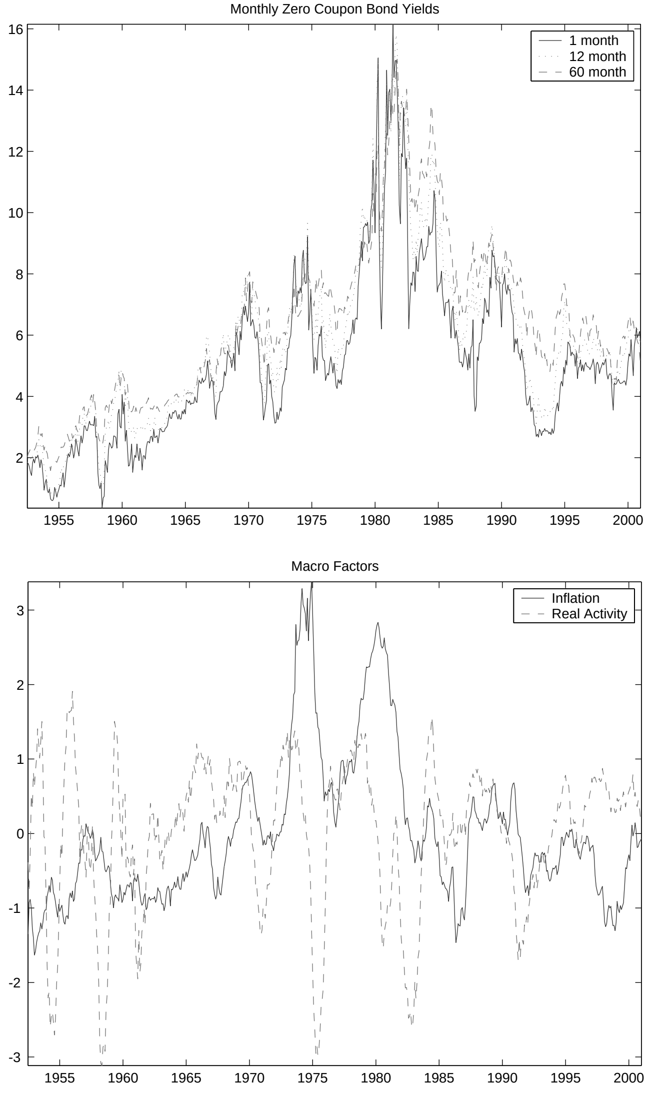


Fig. 1. Bond yields and macro principal components. The top panel shows a plot of (annualized) monthly ZCB yields of maturity 1 month, 12 months and 60 months. The bottom panel plots the two macro factors representing inflation and real activity. The sample period is 1952:06 to 2000:12.

standard deviations of yields generally decrease with maturity; and yields are highly autocorrelated, with increasing autocorrelation at longer maturities.

The yield levels show mild excess kurtosis at short maturities which decreases with maturity, and positive skewness at all maturities. Excess kurtosis is, however, more pronounced for first-differenced yields (for example, 19.44 for the 1-month yield). Although the distribution of yields in the 1990s seems to exhibit Gaussian tails, the evidence for the long series of monthly postwar yields rejects a normal distribution.

Table 1
Summary statistics of data

	Central moments				Autocorrelations		
	Mean	Stdev	Skew	Kurt	Lag 1	Lag 2	Lag 3
1 mth	5.1316	2.7399	1.0756	4.6425	0.9716	0.9453	0.9323
3 mth	5.4815	2.8550	1.0704	4.5543	0.9815	0.9606	0.9419
12 mth	5.8849	2.8445	0.8523	3.8856	0.9824	0.9626	0.9457
36 mth	6.2241	2.7643	0.7424	3.5090	0.9875	0.9739	0.9620
60 mth	6.4015	2.7264	0.6838	3.2719	0.9892	0.9782	0.9687
CPI	3.8612	2.8733	1.2709	4.3655	0.9931	0.9847	0.9738
PCOM	0.9425	11.2974	1.0352	6.0273	0.9684	0.9162	0.8600
PPI	3.0590	3.6325	1.4436	4.9218	0.9863	0.9705	0.9521
HELP	66.7517	22.0257	−0.1490	1.8665	0.9944	0.9900	0.9830
EMPLOY	1.6594	1.5282	−0.4690	3.2534	0.9378	0.8954	0.8410
IP	3.4717	5.3697	−0.5578	3.6592	0.9599	0.8889	0.7972
UE	5.7344	1.5650	0.4924	3.2413	0.9906	0.9777	0.9595

The 1, 3, 12, 36 and 60 month yields are annual zero coupon bond yields from the Fama–Bliss CRSP bond files. The inflation measures CPI, PCOM and PPI refer to CPI inflation, spot market commodity price inflation, and PPI (Finished Goods) inflation respectively. We calculate the inflation measure at time t using $\log(P_t/P_{t-12})$ where P_t is the inflation index. The real activity measures HELP, EMPLOY, IP and UE refer to the Index of Help Wanted Advertising in Newspapers, the growth rate of employment, the growth rate in industrial production and the unemployment rate respectively. The growth rate in employment and industrial production are calculated using $\log(I_t/I_{t-12})$ where I_t is the employment or industrial production index. For the macro variables, the sample period is 1952:01 to 2000:12. For the bond yields, the sample period is 1952:06 to 2000:12.

For our purposes, the Gaussian assumption made in later sections is a sufficient first approximation to the dynamics of the yield curve, as we are mainly interested in the joint dynamics of yields and macroeconomic variables, rather than modeling yield heteroskedasticity. The Gaussian model we present in Section 4 can be extended to incorporate heteroskedastic dynamics parameterized by discretized square-root processes.

An important stylized fact is that yields of near maturity are extremely correlated—the correlation between the 36-month and 60-month yield is 99%. In our estimations we use all five yields to estimate our models, but we specify that some of the yields are measured with error. We choose the 1, 12 and 60-month yields to be measured without error to represent the short, medium and long ends of the yield curve in our models with 3 unknown factors. (The 3-month yield has a 99% correlation with the 12-month yield, and the 36-month yield has a 99% correlation with the 60-month yield.)

2.2. Macro variables

We use macro variables that can be sorted in two groups. The first group consists of various inflation measures which are based on the CPI, the PPI of finished goods, and spot market commodity prices (PCOM). The second group contains variables that capture real activity: the index of Help Wanted Advertising in Newspapers (HELP), unemployment (UE), the growth rate of employment (EMPLOY) and the growth rate of industrial production (IP). This list of variables includes most variables that have been used in monthly VARs in the macro literature. Among these variables, PCOM and HELP are traditionally thought of as leading indicators of inflation and real activity, respectively. All growth rates (including inflation) are measured as the difference in logs of the index at time t and $t-12$, t in months.

To reduce the dimensionality of the system, we extract the first principal component of each group of variables separately. That is, we extract the first principal component from the inflation measures group, and we extract the first principal component from the real activity measures group. This leaves us with two variables which we call "inflation" and "real activity". More precisely, we first normalize each series separately to have zero mean and unit variance. We then stack the three (four) variables related to inflation (real activity) into a vector Z_t^1 (Z_t^2). For each group i, the vector Z_t^i can be represented as

$$Z_t^i = Cf_t^{o,i} + \varepsilon_t^i, \tag{1}$$

where $Z_t^1 = (\text{CPI}_t \ \text{PPI}_t \ \text{PCOM}_t)$ for the inflation group or $Z_t^2 = (\text{HELP}_t \ \text{UE}_t \ \text{EMPLOY}_t \ \text{IP}_t)$ for the real activity group. The error term ε_t^i satisfies $\text{E}(\varepsilon_t^i) = 0$ and $\text{var}(\varepsilon_t^i) = \Gamma$, where Γ is diagonal. The matrices C and Γ are either 3×1 or 4×1 for the inflation group and the real activity group respectively. The extracted macro factor $f_t^{o,i}$ inherits the zero mean from Z_t^i ($\text{E}(f_t^{o,i}) = 0$) and like any principal component has unit variance ($\text{var}(f_t^{o,i}) = 1$).

Table 2 shows the loadings of the first three (four) principal components, and the factor loadings for using only one principal component to explain the variation in each group. Over 70% (50%) of the variance of nominal variables (real variables) is explained by just the first principal component of the group. The first principal component of the inflation measures loads negatively on CPI, PPI, and PCOM. Since negative shocks to this variable represent positive shocks to inflation, we multiply it by -1 so that we can interpret it as an "inflation" factor. The first principal component of real activity measures loads negatively on HELP, EMPLOY, and IP and positively on UE. Again, we multiply this variable by -1 to interpret positive shocks to this factor as positive shocks to economic growth. We call this factor "real activity". We plot these macro factors in the bottom plot in Fig. 1.

To obtain some intuition about these constructed measures of inflation and real activity, Table 3 lists the correlation between the principal components and the original macro series in each group. These correlations show that the inflation factor is most closely correlated with PPI and CPI (97% and 93% respectively) and less correlated with commodity prices (59%). The real activity factor is most closely correlated with employment growth (91%) and industrial production (87%).

The unconditional correlation between the two macro factors is tiny, one tenth of 1%, as reported in Table 3. Although the unconditional correlation is weak, the lower plot in Fig. 1 of the macro factors indicates that some conditional correlations might be important. In fact, when we estimate a VAR for the macro factors, the

Table 2
Principal component analysis

	Principal components: inflation		
	1st	2nd	3rd
CPI	−0.6343	−0.3674	0.6802
PCOM	−0.4031	0.9080	0.1145
PPI	−0.6597	−0.2015	−0.7240
% variance explained	0.7143	0.9775	1.0000

	Principal components: real activity			
	1st	2nd	3rd	4th
HELP	−0.3204	−0.7365	−0.5300	0.2719
UE	0.3597	−0.6283	0.6871	0.0612
EMPLOY	−0.6330	−0.1648	0.2444	−0.7158
IP	−0.6060	0.1886	0.4327	0.6403
% variance explained	0.5202	0.7946	0.9518	1.0000

We take the three (four) macro variables representing inflation (real activity) and normalize them to zero mean and unit variance. For each group i, the normalized data Z_t^i follows the 1 factor model:

$$Z_t^i = Cf_t^{o,i} + \varepsilon_t^i,$$

where C is the factor loading vector, $\mathrm{E}(f_t^{o,i}) = 0$, $\mathrm{cov}(f_t^{o,i}) = I$, $\mathrm{E}(\varepsilon_t^i) = 0$, and $\mathrm{cov}(\varepsilon_t^i) = \Gamma$, where Γ is a diagonal matrix. The columns titled "principal components" list the principal components corresponding to the first to smallest eigenvalue. The % variance explained for the nth principal component gives the cumulative proportion of the variance explained by the first up to the nth eigenvalue. IP refers to the growth in industrial production, CPI to CPI inflation, PCOM to commodity price inflation and PPI to PPI inflation, HELP refers to the Index of Help Wanted Advertising in Newspapers, UE to the unemployment rate, EMPLOY to the growth in employment. The sample period is 1952:01 to 2000:12

conditional correlation is significant. Specifically, we estimate a bivariate process with 12 lags for the macro factors $f_t^o = (f_t^{o,1}\ f_t^{o,2})'$:

$$f_t^o = \rho_1 f_{t-1}^o + \cdots + \rho_{12} f_{t-12}^o + \Omega u_t^o, \tag{2}$$

where ρ_1 to ρ_{12} and Ω are 2×2 matrices with $u_t^o \sim$ IID N(0,I). The estimation results (not reported) show that the coefficient on the seventh lag of real activity in the inflation equation is significant and the coefficient on the first two lags of inflation in the equation for real activity are significant. This can also be seen from the IRs from a VAR(12) fitted to the macro factors, plotted in Fig. 2.[3] The response of inflation to shocks in real activity is positive and hump-shaped, while the response of real activity to inflation shocks is initially weakly positive, and then turns slightly negative before dying out.

[3] The IRs are computed using a Cholesky orthogonalization. It makes little difference reversing the ordering of the variables.

Table 3
Selected correlations

	CPI	PCOM	PPI	
Inflation	0.9286	0.5901	0.9657	
	HELP	UE	EMPLOY	IP
Real activity	0.4622	−0.5188	0.9131	0.8742
	Inflation	Real activity	1 mth	12 mth
Real activity	0.0017			
1 mth	0.6666	0.0627		
12 mth	0.6484	0.0510	0.9771	
60 mth	0.5614	−0.0270	0.9191	0.9639

The table reports selected correlations for the inflation factor extracted from the first principal component of PCI, PCOM and PPI, the real activity factor extracted from the first principal component of HELP, UE, EMPLOY and IP, and the 1, 12 and 60 month bond yields, which are used in the estimation. IP refers to the growth in industrial production, CPI to CPI inflation, PCOM to commodity price inflation, PPI to PPI inflation, HELP refers to the Index of Help Wanted Advertising in Newspapers, UE to the unemployment rate, EMPLOY to the growth in employment. The sample period is 1952:06 to 2000:12.

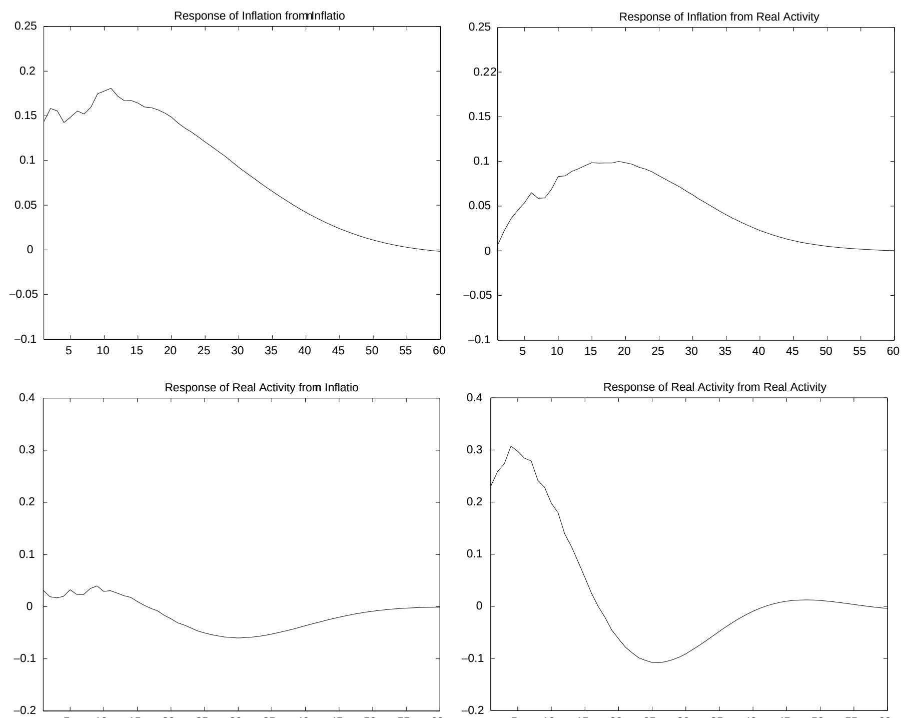


Fig. 2. Impulse responses from the VAR(12) on macro factors. We fit a VAR(12) to the inflation and real activity macro factors, where inflation is ordered first. The plot shows the impulse responses to a Cholesky one standard deviation innovation to each variable. Time is measured in months on the x-axis.

Some preliminary information about the relationship between the macro factors and the yield curve can be gained from the correlation matrix in Table 3. The inflation factor is highly correlated with yields. This correlation is highest for short yields (67% correlation between inflation and the 1-month yield), and somewhat smaller for long yields (56% correlation between inflation and the 60-month yield). Real activity is only weakly correlated with yields. This correlation does not exceed 6% for any maturity. This weak relationship is not representative for all measures of real activity. For example, the correlation of HELP and 1-month yield is 63%, but our real activity factor loads mostly on EMPLOY and IP. Hence, at least for measures of economic activity, it may matter whether the particular variable in question is a leading indicator of business cycles. This implies that in our analysis we may potentially understate the impact of real activity on the yield curve by the construction of our real activity factor.

3. A first look at short rate dynamics

3.1. Policy rules and short rate dynamics in affine models

According to the policy rule recommended by Taylor (1993), movements in the short rate r_t are traced to movements in contemporaneous macro variables f_t^o and a component which is not explained by macro variables, an orthogonal shock v_t:

$$r_t = a_0 + a_1' f_t^o + v_t. \tag{3}$$

The shock v_t may be interpreted as a monetary policy shock following identifying assumptions made in Christiano et al. (1996). Taylor's original specification uses two macro variables as factors in f_t^o. The first variable is an annual inflation rate, similar to our inflation factor, and the second variable is the output gap. GDP data are only available at a quarterly frequency, while our real activity factor is constructed using various monthly series such as EMPLOY and IP.

Another type of policy rule that has been proposed by Clarida et al. (2000) is a forward-looking version of the Taylor rule. According to this rule, the central bank reacts to *expected* inflation and the *expected* output gap. This implies that any variable that forecasts inflation or output will enter the right-hand side of (3). In the hope of capturing the information underlying macro forecasts, we add lagged macro variables as arguments in Eq. (3).[4] This is done by writing $X_t^o = (f_t^{o\prime} f_{t-1}^{o\prime}, \ldots, f_{t-p-1}^{o\prime})'$ for some lag length p and including the lags as arguments in the policy rule:

$$r_t = b_0 + b_1' X_t^o + v_t. \tag{4}$$

[4] Clarida et al. (2000) implement their forward-looking rule by redefining the shock term v_t to include forecast errors $f_{t+1}^o - \mathrm{E}_t(f_{t+1}^o)$. This allows them to use future values of macro variables f_{t+1}^o as arguments on the right-hand side of (3). We could in principle adopt the same approach by including these forecast errors into some latent variables, but this would mean that we would have to drop the assumption that latent and macro variables are orthogonal. Our focus is assigning as much explanatory power to macro factors as possible, so we specify the latent variables as orthogonal.

Affine term structure models (Duffie and Kan, 1996) are based on a short rate equation just like Eq. (3) together with an assumption on risk premia. The difference between the short rate dynamics in affine term structure models and the Taylor rule is that in affine term structure models the short rate is specified to be an affine (constant plus linear term) function of underlying latent factors X_t^u:

$$r_t = c_0 + c_1' X_t^u. \tag{5}$$

The unobserved factors themselves follow affine processes, of which a VAR is a special Gaussian case. The prices of bonds of longer maturities are explicit exponential affine functions (dependent upon parameters) of X_t^u if pricing is risk neutral. In the more general case that we consider, the risk adjustment needs to be specified carefully to obtain similar closed-form solutions for bond yields (this is explained in the next section).

Eqs. (3)–(5) are very similar: they all specify the short rate as affine functions of factors. We can combine them by writing:

$$r_t = \delta_0 + \delta_{11}' X_t^o + \delta_{12}' X_t^u. \tag{6}$$

The approach we take in this paper is to specify the latent factors X_t^u as orthogonal to the macro factors X_t^o. In this case, the short rate dynamics of the term structure model can be interpreted as a version of the Taylor rule with the errors $v_t = \delta_{12}' X_t^u$ being unobserved factors. We use the restrictions from no-arbitrage to separately identify the individual latent factors.

3.2. Estimating the short rate dynamics

The coefficients on inflation and real activity in the short rate equation (6) can be estimated by ordinary least squares because of the independence assumption on X_t^o and X_t^u. Table 4 reports the estimation results from two regressions: the original Taylor rule (3) and the forward-looking version of the Taylor rule (4), which incorporates lags of the macro variables. These regression results give a preliminary view as to how much of the yield movements macro factors may explain with respect to the unobservable variables. The R^2 of the estimated Taylor rule is 45%, while the estimated forward-looking version of the Taylor raises the R^2 to 53%. These numbers suggest that macro factors should have explanatory power for yield curve movements.

The behavior of the residuals, however, provides some intuition about what to expect from a model with unobservable factors. First, the residuals from both versions of the Taylor rule are highly autocorrelated. The autocorrelation of residuals from the short rate equation with only contemporaneous macro factors is 0.945, while the autocorrelation from the equation that includes lagged macro factors is slightly lower, 0.937. The short rate itself has an autocorrelation of 0.972, indicating that macro variables do explain some of the persistent shocks to the short rate. Second, unless a variable which mimics the short rate itself is placed on the RHS of Eq. (3), the residuals will follow the same broad pattern as the short rate. This can be seen from Fig. 3, which plots the residuals together with the de-meaned

Table 4
The dependence of the short rate on macro variables

Coeff	Constant	Inflation	Real activity	Adj R^2
Panel A: y_t^1 on constant, inflation and real activity				
t	0.4250	0.1535	0.0143	0.4523
	(0.0070)**	(0.0070)**	(0.0070)*	
Panel B: y_t^1 on constant, 12 lags of inflation and real activity				
t	0.4296	0.0037	0.0398	0.5337
	(0.0065)**	(0.0534)	(0.0306)	
$t-1$		0.0659	0.0150	
		(0.0828)	(0.0452)	
$t-2$		−0.0435	0.0105	
		(0.0830)	(0.0450)	
$t-3$		0.0062	−0.0054	
		(0.0833)	(0.0444)	
$t-4$		0.0233	−0.0172	
		(0.0828)	(0.0441)	
$t-5$		−0.0088	0.0145	
		(0.0825)	(0.0442)	
$t-6$		−0.0245	−0.0213	
		(0.0825)	(0.0438)	
$t-7$		0.0175	0.0062	
		(0.0821)	(0.0435)	
$t-8$		0.0080	0.0196	
		(0.0825)	(0.0438)	
$t-9$		−0.0049	0.0121	
		(0.0821)	(0.0441)	
$t-10$		−0.0079	0.0005	
		(0.0820)	(0.0439)	
$t-11$		0.1427	−0.0069	
		(0.0522)**	(0.0299)	

In Panel A we regress the 1 month yield y_t^1 on a constant, the inflation factor and the real activity factor. In Panel B we regress y_t^1 on a constant, inflation, real activity and 11 lags of inflation and real activity. We report OLS standard errors in parenthesis. Standard errors significant at the 5% (1%) level are denoted * (**). Sample period is 1952:06 to 2000:12.

short rate. This suggests that the ''level'' factor found by earliest term structure studies (see Vasicek, 1977), may still reappear when macro variables are added in a linear form to the short rate in a term structure model.

The coefficients on inflation and real activity in the simple Taylor rule are both significant and positive. This is consistent with previous estimates of the Taylor rule in the literature, and also the parameter values proposed by Taylor (1993)'s original specification. However, these coefficients are highly sensitive to the sample period selected, as structural changes (or regime shifts) cause the coefficients in (6) to be time-varying (see Ang and Bekaert, 2002).

The sign of the Taylor-rule coefficient on real activity crucially depends on the inclusion of the two NBER recessions in 1954 and 1958. This is evident from the

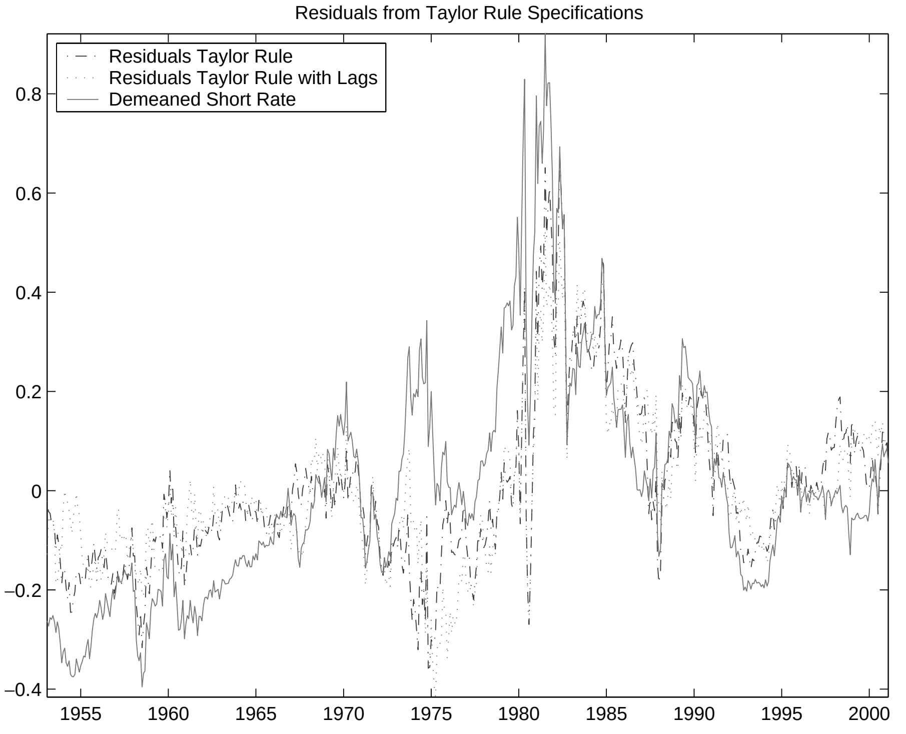


Fig. 3. Residuals from the Taylor rule regressions. We show the residuals from the Taylor rule regressions, together with the de-meaned short rate (1 month yield). We show the residuals from the Taylor rule with no lags, which have 0.9458 autocorrelation, and the residuals from the Taylor rule with 11 lags, which have 0.9370 autocorrelation. For comparison, the autocorrelation of the short rate is 0.9716.

plots of real activity and the 1-month yield in Fig. 1. Both these recessions go hand in hand with decreases in the 1-month rate and make the Taylor rule coefficient on output positive. If we start the estimation of the Taylor rule later, say in 1960 or 1970, the coefficient on real activity is negative. Only if we start the estimation after the monetary experiment of 1982 is the coefficient positive. Interestingly enough, the coefficient on output is not significant for the whole post-1982 period, but it is significant for the Greenspan years (post-1987). The large and significantly positive coefficient on inflation is much more robust across different sample periods. However, we assume that during our sample period, the Taylor rule relationships are stable, just as in Gali (1992), Christiano et al. (1996), and Cochrane (1998).

In contrast to the simple Taylor rule estimation, Table 4 reports that most parameter estimates for the forward-looking version of the Taylor rule are not significant, except for the 11th lag on inflation. This suggests that using many lags in the Taylor rule may lead to an over-parameterized and potentially poorly behaved system. However, a likelihood ratio tests rejects the null of the simple Taylor rule (with only contemporaneous inflation and real activity) in favor of the alternative of the Taylor rule with lags with a p-value less than 1%. On the other hand, the optimal

Schwartz (BIC) choice is the simple Taylor rule. We present models with both specifications.

4. A term structure model with macro factors

Based on the macro dynamics (2) and the short rate equation (6), we now develop a discrete-time term structure model. The model combines observable macroeconomic variables with unobservable or latent factors. Risk premia in our set-up are time-varying, because they are taken to be affine in potentially all of the underlying factors. Section 4.1 presents the general model and Section 4.2 parameterizes the latent variables and risk premia. We outline our estimation procedure in Section 4.3. Section 4.4 summarizes our parameterization.

4.1. General setup

4.1.1. State dynamics

Suppose there are K_1 observable macro variables f_t^o and K_2 latent variables f_t^u. The vector $F_t = (f_t^{o'}, f_t^{u'})'$ follows a Gaussian VAR(p) process:

$$F_t = \Phi_0 + \Phi_1 F_{t-1} + \cdots + \Phi_p F_{t-p} + \theta u_t \tag{7}$$

with $u_t \sim$ IID N$(0, I)$. The latent factors f_t^u are AR(1) processes, so that we set the coefficients $\Phi_2 \dots \Phi_p$ in Eq. (7) corresponding to $X_t^u = f_t^u$ equal to zero. The state of the economy is then described by a K-dimensional vector of state variables X_t, where $K = K_1 \cdot p + K_2$. We partition the state vector X_t into $K_1 \cdot p$ observable variables X_t^o and K_2 unobservable variables X_t^u. The observable vector contains current and past levels of macroeconomic variables $X_t^o = (f_t^{o'}, f_{t-1}^{o'}, \dots, f_{t-p-1}^{o'})'$, while $X_t^u = f_t^u$ only contains contemporaneous latent yield factors. We take the bivariate VAR(12) in Eq. (2) as the process for inflation and real activity so set $p = 12$.

We write the dynamics of $X_t = (X_t^{o'} \; X_t^{u'})'$ in compact form as a first order Gaussian VAR:

$$X_t = \mu + \Phi X_{t-1} + \Sigma \varepsilon_t \tag{8}$$

with $\varepsilon_t = (u_t^{o'}, 0, \dots, 0, u_t^{u'})'$, where u_t^o (u_t^u) are the shocks to the observable (unobservable) factors. In the first order companion form, there are blocks of zeros in the $K \times K$ matrix Σ to accommodate higher order lags in F_t.

4.1.2. Short rate equation

The one-period short rate r_t is assumed to be an affine function of all state variables:

$$r_t = \delta_0 + \delta_1' X_t. \tag{9}$$

We work with monthly data, so we use the one-month yield y_t^1 as an observable short rate r_t. By constraining the coefficient δ_1 to depend only on contemporaneous factor values, we obtain the Taylor rule (3). We call this the "Macro Model." We also consider the case where δ_1 is unconstrained, which correspond to the forward-looking Taylor rule incorporating lags. We refer to this formulation as the "Macro Lag Model," because it uses lags of macro variables in the short rate equation.

4.1.3. Pricing kernel

To develop the term structure model, we use the assumption of no-arbitrage (Harrison and Kreps, 1979) to guarantee the existence of an equivalent martingale measure (or risk-neutral measure) Q such that the price of any asset V_t that does not pay any dividends at time $t+1$ satisfies $V_t = \mathrm{E}_t^Q(\exp(-r_t)V_{t+1})$, where the expectation is taken under the measure Q. The Radon–Nikodym derivative (which converts the risk-neutral measure to the data-generating measure) is denoted by ξ_{t+1}. Thus, for any $t+1$ random variable Z_{t+1} we have that $\mathrm{E}_t^Q(Z_{t+1}) = \mathrm{E}_t(\xi_{t+1}Z_{t+1})/\xi_t$. The assumption of no-arbitrage, or equivalently the assumption of the existence of ξ_{t+1}, allows us to price any asset in the economy, in particular all nominal bond prices.

Assume that ξ_{t+1} follows the log-normal process:

$$\xi_{t+1} = \xi_t \exp\left(-\frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\varepsilon_{t+1}\right), \tag{10}$$

where λ_t are the time-varying the market prices of risk associated with the sources of uncertainty ε_t. We parameterize λ_t as an affine process:[5]

$$\lambda_t = \lambda_0 + \lambda_1 X_t \tag{11}$$

for a K-dimensional vector λ_0 and a $K \times K$ matrix λ_1. Eqs. (10) and (11) relate shocks in the underlying state variables (macro and latent factors) to ξ_{t+1} and therefore determine how factor shocks affect all yields. Parameters in λ_0 and λ_1 that correspond to lagged macro variables are set to zero. We do this for parsimony while ensuring that both the macro and unobservable factors are priced. This means that the K-vector λ_0 contains a total of $K_1 + K_2$ free parameters: the upper $K_1 \times 1$ row and the bottom $K_2 \times 1$ row. The matrix λ_1 contains $(K_1 + K_2)^2$ free parameters: the upper-left $K_1 \times K_1$ corner together with the upper-right $K_1 \times K_2$ corner, and the lower-left $K_2 \times K_1$ corner together with the lower-right $K_2 \times K_2$ corner.

We define the pricing kernel m_{t+1} as

$$m_{t+1} = \exp(-r_t)\xi_{t+1}/\xi_t. \tag{12}$$

Substituting $r_t = \delta_0 + \delta_1' X_t$ we have

$$m_{t+1} = \exp\left(-\frac{1}{2}\lambda_t'\lambda_t - \delta_0 - \delta_1' X_t - \lambda_t'\varepsilon_{t+1}\right). \tag{13}$$

4.1.4. Bond prices

We take Eq. (13) to be a nominal pricing kernel which prices all nominal assets in the economy. This means that the total gross return process R_{t+1} of any nominal

[5] The specification (11) has been used in continuous time by Constantinides (1992), El Karoui et al. (1992), Fisher (1998), Liu (1999), Duffee (2002), and Dai and Singleton (2002), among many others.

asset satisfies

$$\mathrm{E}_t(m_{t+1}R_{t+1}) = 1. \tag{14}$$

If p_t^n represents the price of an n-period zero coupon bond, then Eq. (14) allows bond prices to be computed recursively by

$$p_t^{n+1} = \mathrm{E}_t(m_{t+1}p_{t+1}^n). \tag{15}$$

The state dynamics of X_t (Eq. (8)) together with the dynamics of the short rate r_t (Eq. (9)) and the Radon–Nikodym derivative (Eq. (10)) form a discrete-time Gaussian K-factor model with $K_1 \cdot p$ observable factors and K_2 unobservable factors, where p is the number of lags in the autoregressive representation of the observable factors. It falls within the affine class of term structure models because bond prices are exponential affine functions of the state variables. More precisely, bond prices are given by

$$p_t^n = \exp(\bar{A}_n + \bar{B}_n' X_t), \tag{16}$$

where the coefficients $\bar{A}_n$ and $\bar{B}_n$ follow the difference equations:

$$\begin{aligned}\bar{A}_{n+1} &= \bar{A}_n + \bar{B}_n'(\mu - \Sigma\lambda_0) + \frac{1}{2}\bar{B}_n'\Sigma\Sigma'\bar{B}_n - \delta_0,\\ \bar{B}_{n+1}' &= \bar{B}_n'(\phi - \Sigma\lambda_1) - \delta_1'\end{aligned} \tag{17}$$

with $\bar{A}_1 = -\delta_0$ and $\bar{B}_1 = -\delta_1$. These difference equations can be derived by induction using Eq. (15), and details are provided in the appendices.[6]

The continuously compounded yield y_t^n on an n-period zero coupon bond is given by

$$\begin{aligned}y_t^n &= -\frac{\log p_t^n}{n}\\ &= A_n + B_n' X_t,\end{aligned} \tag{18}$$

where $A_n = -\bar{A}_n/n$ and $B_n = -\bar{B}_n/n$. Note that yields are affine functions of the state X_t, so that Eq. (18) can be interpreted as being the observation equation of a state space system. Additional observation equations come from the observable variables X_t^o. Most examples of discrete-time affine models have not incorporated lagged state variables. However, by treating the lagged variables as state variables in X_t, the affine form is still maintained. Despite time-varying risk premia, our system is still Gaussian, and IRs, variance decompositions and other techniques can be handled as easily as an unrestricted VAR.

4.2. Choice of parameterization

4.2.1. Latent variables

Empirical studies have concluded that three unobserved factors explain much of yield dynamics (see Knez et al., 1994). To compare models with only latent variables with models incorporating both latent and macro factors we use three unobservable

[6] See the techniques in Campbell et al. (1997).

factors. Hence our most comprehensive model consists of two macro ($K_1 = 2$) and three latent factors ($K_2 = 3$).

Since there are unobservable variables present, normalizations can be made that give observationally equivalent systems. The idea behind these normalizations in a VAR setting is that affine transformations and rotations of the unobservable factors lead to observationally equivalent yields. These normalizations are discussed in Dai and Singleton (2000). We estimate the most general parameterization for the unobserved variables in this paper.

We estimate the following system for the unobservable factors:

$$f_t^u = \rho f_{t-1}^u + u_t^u, \tag{19}$$

with 3-dimensional shock vector $u_t^u \sim \text{IID N}(0, I)$ and a lower-triangular 3×3 companion matrix ρ. This is the most general identified representation for a Gaussian specification. The unit variance of u_t^u implies that the lower-right corner 3×3 matrix in θ of Eq. (7) and in Σ of Eq. (8) is just equal to $\text{var}[u_t^u] = I$. A multifactor Vasicek (1977) model with correlated unobservable factors consists of (19), an affine short rate equation (5), and the assumption that $\lambda_1 = 0$. In a Vasicek model, specifying the companion form and holding fixed the covariances is equivalent to holding the companion form fixed and specifying the covariances. Numerous papers in the term structure literature have used independent factors as a first-cut modeling approach, including Longstaff and Schwartz (1992) and Chen and Scott (1993). At the estimated parameters, however, the latent factors usually turn out to violate the independence assumption. We therefore estimate a correlated latent factor model to give the latent variables a fair chance to explain the yield curve by themselves, without the inclusion of macro variables.

We impose independence between latent and macro factors, so that the upper-right 24×3 corner and the lower-left 3×24 corner of Φ and Σ in the compact form in (8) contain only zeros. This approach to including observed macro factors in a pricing kernel specifies all uncertainties arising in the latent factors as orthogonal to the macro variables. This independence assumption has two main drawbacks. First, it contradicts empirical evidence that the term structure predicts movements in macro economic activity (see Harvey, 1988; Estrella and Hardouvelis, 1991). Second, monetary policy has no impact on future inflation or real activity. In other words, the Fed is conducting monetary policy using the Taylor rule in an environment where policy has no effects on the variables to which the Fed is responding. Extensions of this model can be done by freeing up the companion matrix to allow feedback (so Φ does not contain zero corner blocks), and looking at contemporaneous correlations of macro and latent factors (θ_0 does not contain zero corner blocks). We leave these extensions to future research.

4.2.2. *Risk premia*

The data-generating and the risk neutral measures coincide if $\lambda_t = 0$ for all t. This case is usually called the "Local Expectations Hypothesis," which differs from the traditional Expectations Hypothesis by Jensen inequality terms (see Cochrane, 2001, Chapter 19). Macro models, such as Fuhrer and Moore (1995), usually impose the

Expectations Hypothesis to infer long term yield dynamics from short rates. The dynamics of the term structure under the data-generating measure depend on the risk premia parameters λ_0 and λ_1 in Eq. (11). A non-zero vector λ_0 affects the long-run mean of yields because this parameter affects the constant term in the yield equation (18). A non-zero matrix λ_1 affects the time-variation of risk-premia, since it affects the slope coefficients in the yield equation (18). In a Vasicek (1977) model λ_0 is non-zero and λ_1 is zero, which allows the average yield curve to be upward sloping, but does not allow risk premia to be time-varying.

The number of parameters in λ to estimate is very large: λ_0 has $K_1 + K_2 = 5$ and λ_1 has $(K_1 + K_2)^2 = 25$ parameters in the case of the models with macro variables. To avoid over-fitting, we fix some of these parameters. In particular, we specify the λ_1 matrix to be block-diagonal, with zero restrictions on the upper-right 2×3 and lower-left 3×2 corner blocks. Time variation in the compensation for shocks to latent variables is thus only driven by the latent variables themselves. The analogous argument holds for the compensation for macro shocks. This parameterization assumption is in the spirit of the orthogonalization of macro and latent variables.

To summarize, we estimate 5 parameters in λ_0 and $4 + 9$ parameters in λ_1. The parameters in λ_0 correspond to the current macro variables and latent variables. Similarly, the parameters in λ_1 are contained in two non-zero matrices on the diagonal: an upper-left 2×2 matrix for current macro variables and a lower-right 3×3 matrix for the latent variables.

4.3. Estimation method

To estimate the model, we transform a system of yields and observables $(Y_t', X_t^{o'})$ into a system of observables and unobservables $X_t = (X_t^{o'}, X_t^{u'})$. The yields themselves are analytical functions of the state variables X_t, which allow us to infer the unobservable factors from the yields. We estimate using maximum likelihood, and derive the likelihood function in the appendices. In traditional VAR approaches, yields and macro variables are used directly as inputs into a VAR after specifying the autoregressive lag length. The likelihood for the VAR is a function of $(Y_t', X_t^{o'})$, and inferences about yield curve movements and macro shocks can be drawn from the parameters in the companion form coefficients and covariance terms. Our approach amounts to estimating a VAR of $(Y_t', X_t^{o'})$, with assumptions that identify an unobservable component orthogonal to macro shocks and guarantee no arbitrage.

We use a two-step consistent estimation procedure. In the first step, we estimate the macro dynamics (2) and the coefficients δ_0 and δ_{11} of the macro factors in the short rate dynamics equation (6). In a second step, we estimate the remaining parameters of the term structure model holding all pre-estimated parameters fixed. This two-step procedure avoids the difficulties associated with estimating a model with many factors using maximum likelihood when yields are highly persistent.[7] The

[7] We tried to estimate various versions of the model in a single step with maximum likelihood. These estimations typically produced explosive yield dynamics. Fixing the parameters that characterize the

procedure also avoids estimating a very large number of lag coefficients ($\rho_1, \ldots \rho_{12}$) in the bivariate VAR for the macro variables in the term structure model.

Both the macro dynamics (2) and the short rate coefficients of the macro variables in Eq. (6) are estimated by ordinary least squares, as reported in Sections 2 and 3. Since our constructed macro factors have zero mean, the constant δ_0 in the short rate equation represents the unconditional mean of the 1 month yield, which equals 5.10% on an annualized basis. This number has to be divided by 12 to obtain an estimate for δ_0 at a monthly frequency. The regression coefficients δ_{11} of the short rate equation give the maximal proportion of short rate movements explained by the macro factors, with all remaining orthogonal factors being unobservable. No-arbitrage assumptions identify the unexplained proportion.

In the second estimation procedure, we hold δ_0, δ_{11}, and the parameters entering the macro subsystem (6) fixed, and estimate all other parameters of the term structure model including the remaining coefficients in δ_{12} corresponding to the latent factors. We need to find good starting values to achieve convergence in this highly non-linear system. In particular, since unconditional means of persistent series are difficult to estimate, the likelihood surface is very flat in λ_0 which determines the mean of long yields. We therefore estimate the model in several iterative rounds.

We begin by obtaining starting values for ρ in Eq. (19) from estimating the model under the Expectations Hypothesis (with λ_0 and λ_1 equal to zero.) We then compute starting values for λ_1 holding λ_0 fixed at zero. Next, we estimate λ_0 setting any insignificant parameters in λ_1 at the 5% level equal to zero. Then we set insignificant λ_0 parameters to zero and re-estimate. This iterative procedure produces the zeros in the Φ and λ_1 matrices in Tables 5–7, which report the results. Most of the non-zero parameters in Φ and λ_1 are significant, and we expect these important effects to remain in other iterative estimation schemes. While our particular procedure may be path dependent, we could not find a feasible alternative which implies unconditional means for long yields close to those in the data.

Finally, our likelihood construction solves for the unobservable factors from the joint dynamics of the zero coupon bond yields and the macro factors. To do this, we follow Chen and Scott (1993) and assume that as many yields as unobservable factors are measured without error, and the remaining yields are measured with error. In particular, for our models we assume the 3 and 36-month yields are measured with error.

4.4. Summary of parameterization

To summarize, we estimate the following special case of the general model. The bivariate system of macro factors f_t^o follows the process:

$$f_t^o = \rho_1 f_{t-1}^o + \cdots + \rho_{12} f_{t-12}^o + \Omega u_t^o, \tag{20}$$

(*footnote continued*)
dependence of the short rate on the observable factors in a (consistent) first-step estimation is a tractable way to avoid the problem of nonstationary dynamics.

Table 5
Yields-only model estimates

Companion form Φ				
0.9924	0.0000	0.0000		
(0.0039)				
0.0000	0.9548	0.0000		
	(0.0062)			
0.0000	−0.0021	0.7646		
	(0.0001)	(0.0210)		
Short rate parameters δ_1 ($\times 100$)				
Unobs 1	Unobs 2	Unobs 3		
0.0136	−0.0451	0.0237		
(0.0020)	(0.0005)	(0.0015)		
Prices of risk λ_0 and λ_1				
		λ_1 matrix		
	λ_0	Unobs 1	Unobs 2	Unobs 3
Unobs 1	−0.0033	−0.0069	0.0000	0.0000
	(0.0004)	(0.0040)		
Unobs 2	0.0000	0.0445	0.0000	−0.2585
		(0.0050)		(0.0197)
Unobs 3	0.0000	−0.0490	0.0000	0.2412
		(0.0090)		(0.0256)
Measurement error ($\times 100$)				
3 month	36 month			
0.0203	0.0090			
(0.0003)	(0.0002)			

The table reports parameter estimates and standard errors in parenthesis for the 3-factor Yields-Only Model $X_t = \Phi X_{t-1} + \varepsilon_t$, with $\varepsilon_t \sim \mathrm{N}(0, I)$, Φ lower triangular and the short rate equation given by $r_t = \delta_0 + \delta_1' X_t$. All factors $X_t \equiv f_t^u$ are unobservable. The coefficient δ_0 is set to the sample unconditional mean of the short rate, 0.0513/12. Market prices of risk $\lambda_t = \lambda_0 + \lambda_1 X_t$ are restricted to be block diagonal. The sample period is 1952:06 to 2000:12.

with $u_t^o \sim \mathrm{IID\ N(0, I)}$. The 2×2 matrices $\rho_1, \ldots, \rho_{12}$, are unconstrained and Ω is lower-triangular.

The trivariate system of latent factors f_t^u follows the process:

$$f_t^u = \rho f_{t-1}^u + u_t^u \tag{21}$$

with $u_t^u \sim \mathrm{IID\ N(0, I)}$. The 3×3 matrix ρ is lower triangular to ensure identification. The shock processes u_t^o and u_t^u are independent.

The short rate equation is

$$r_t = \delta_0 + \delta_{11}' X_t^o + \delta_{12}' X_t^u, \tag{22}$$

where the parameters δ_0 and δ_{11} are consistently estimated by least squares in a first-step procedure prior to maximizing the likelihood (since X_t^o and X_t^u are orthogonal

Table 6
Macro model estimates

Companion form Φ for latent factors		
0.9915	0.0000	0.0000
(0.0042)		
0.0000	0.9392	0.0000
	(0.0122)	
0.0000	0.0125	0.7728
	(0.0146)	(0.0217)

Short rate parameters δ_1 for latent factors ($\times 100$)		
Unobs 1	Unobs 2	Unobs 3
0.0138	−0.0487	0.0190
(0.0021)	(0.0007)	(0.0022)

Prices of risk λ_0 and λ_1

	λ_0	λ_1 matrix				
		Inflation	Real activity	Unobs 1	Unobs 2	Unobs 3
Inflation	0.0000	−0.4263	0.1616	0.0000	0.0000	0.0000
		(0.1331)	(0.0146)			
Real activity	0.0000	1.9322	−0.1015	0.0000	0.0000	0.0000
		(0.3893)	(0.0329)			
Unobs 1	−0.0039	0.0000	0.0000	−0.0047	0.0000	0.0000
	(0.0003)			(0.0043)		
Unobs 2	0.0000	0.0000	0.0000	0.0459	0.0000	−0.2921
				(0.0055)		(0.0205)
Unobs 3	0.0000	0.0000	0.0000	−0.0351	0.0000	0.1995
				(0.0087)		(0.0283)

Measurement error ($\times 100$)	
3 month	36 month
0.0207	0.0091
(0.0003)	(0.0002)

The table reports parameter estimates and standard errors in parenthesis for the Macro Model with the short rate equation specified with only current inflation and current real activity, as reported in Panel A of Table 4. The short rate equation is given by $r_t = \delta_0 + \delta_1' X_t$, where δ_1 only picks up current inflation, current real activity and the latent factors. The dynamics of inflation and real activity are given by a 12 lag VAR (not reported). The model is $X_t = \Phi X_{t-1} + \varepsilon_t$, with $\varepsilon_t \sim \mathrm{N}(0, I)$. X_t contains 12 lags of inflation and real activity and three latent variables, which are independent at all lags to the macro variables. In a pre-estimation we find the inflation and real activity VAR(12), and the coefficients on inflation and real activity in the short rate equation. The coefficient δ_0 is set to the sample unconditional mean of the short rate, 0.0513/12. Market prices of risk $\lambda_t = \lambda_0 + \lambda_1 X_t$ are restricted to be block diagonal. The sample period is 1952:06 to 2000:12.

and X_t^u has zero mean). The observable factors are $X_t^o = (f_t^{o\prime} f_{t-1}^{o\prime}, \ldots, f_{t-p-1}^{o\prime})$ and the latent factors are $X_t^u = f_t^u$. The full set of state variables is $X_t = (X_t^{o\prime}, X_t^{u\prime})'$.

Market prices of risk are affine in the state vector:

$$\lambda_t = \lambda_0 + \lambda_1 X_t. \tag{23}$$

The matrix λ_1 has an upper-left 2×2 matrix and a lower-right 3×3 matrix corresponding to f_t^o and f_t^u, while the remaining parameters are set to zero. The

Table 7
Macro lag model estimates

Companion form Φ for latent factors

0.9922 (0.0039)	0.0000	0.0000
0.0000	0.9431 (0.0118)	0.0000
0.0000	−0.0189 (0.0135)	0.8210 (0.0216)

Short rate parameters δ_1 for latent factors ($\times 100$)

Unobs 1	Unobs 2	Unobs 3
0.0130 (0.0020)	−0.0438 (0.0010)	0.0256 (0.0025)

Prices of risk λ_0 and λ_1

	λ_0	λ_1 matrix				
		Inflation	Real activity	Unobs 1	Unobs 2	Unobs 3
Inflation	0.0000	0.8442 (0.2397)	−0.0017 (0.0582)	0.0000	0.0000	0.0000
Real activity	0.0000	1.1209 (0.1375)	0.2102 (0.0275)	0.0000	0.0000	0.0000
Unobs 1	−0.0050 (0.0003)	0.0000	0.0000	−0.0048 (0.0040)	0.0000	0.0000
Unobs 2	0.0000	0.0000	0.0000	0.0483 (0.0068)	0.0000	−0.2713 (0.0195)
Unobs 3	0.0000	0.0000	0.0000	−0.0248 (0.0078)	0.0000	0.1624 (0.0292)

Measurement error ($\times 100$)

3 month	36 month
0.0251 (0.0005)	0.0107 (0.0003)

The table reports parameter estimates and standard errors in parenthesis for the Macro Lag Model with the short rate equation specified with 12 lags of inflation and current real activity, as reported in Panel B of Table 4. The short rate equation is given by $r_t = \delta_0 + \delta_1' X_t$, where δ_1 only picks up 12 lags of inflation and real activity and the latent factors. The dynamics of inflation and real activity are given by a 12 lag VAR (not reported). The model is $X_t = \Phi X_{t-1} + \varepsilon_t$, with $\varepsilon_t \sim \mathrm{N}(0, I)$. X_t contains 12 lags of inflation and real activity and three latent variables, which are independent at all lags to the macro variables. In a pre-estimation we find the inflation and real activity VAR(12), and the coefficients on inflation and real activity in the short rate equation. The coefficient δ_0 is set to the sample unconditional mean of the short rate, 0.0513/12. Market prices of risk $\lambda_t = \lambda_0 + \lambda_1 X_t$ are restricted to be block diagonal. The sample period is 1952:06 to 2000:12.

parameters in λ_0 corresponding to f_t^o and f_t^u are free, and all remaining parameters in λ_0 are restricted to be zero.

We estimate two versions of our most comprehensive model with two macro factors and three unobservable factors. The estimation of δ_{11} that restricts the parameters on lagged parameters to be zero as in Eq. (22) is denoted the

"Macro Model." The version with the full lagged Taylor rule is denoted the "Macro Lag Model." The estimation without any macro variables we call the "Yields-Only Model."

5. Estimation results

Section 5.1 interprets the parameter estimates of the Macro and Yields-Only Models. To determine the effect of the addition of macro factors into term structure models, we look at the IRs of each factor in Section 5.2. The variance decompositions in Section 5.3 allow us to attribute the forecast variance at a particular horizon to shocks in macro and latent factors. In Section 5.4, we find that imposing the cross-equation restrictions from no-arbitrage forecasts better than the unrestricted VARs common in the macro literature. Moreover, incorporating macro variables into a term structure model helps us obtain even better forecasts. We compare the latent factors from the different models in Section 5.5 and find that macro factors do account for some of the latent factors from the Yields-Only Model. Derivations for the IRs and variance decompositions are presented in the appendices.

5.1. Parameter estimates

5.1.1. Yields-only model

Table 5 presents the estimation results for the Yields-Only Model. The order of the latent factors in Table 5 is unspecified, but we present the estimation results by ordering the latent factors by decreasing autocorrelation. The model has one very persistent factor, one less persistent but still strongly persistent factor, and the last factor is strongly mean-reverting. This is consistent with previous multi-factor estimates in the literature.

Litterman and Scheinkman (1991) label these unobservable factors "level," "slope," and "curvature" respectively because of the effects of these factors on the yield curve. To show these effects, the first latent variable, Unobs 1, closely corresponds to a "level" transformation of the yield curve, which we define as $(y_t^1 + y_t^{12} + y_t^{60})/3$. The correlation between Unobs 1 and the level transformation is 92%. The second latent variable, Unobs 2, has a 58% correlation with a "spread" transformation, defined as $y_t^{60} - y_t^1$. The third latent variable, Unobs 3, has a 77% correlation with a "curvature" transformation $(y_t^1 - 2y_t^{12} + y_t^{60})$.

In Table 5, the estimated vector λ_0 has one significant parameter corresponding to the most highly autocorrelated factor. The parameter is negative, so that the unconditional mean of the short rate under the risk-neutral measure is higher than under the data-generating measure. Since bond prices are computed under the risk-neutral measure, negative parameters in λ_0 induce long yields to be on average higher than short yields. Time-variation in risk premia is mainly driven by the first and third unobservable factor. In other words, risk premia in bond yields mainly depend on the level and the curvature of the yield curve, and are not driven by the slope of the yield curve.

5.1.2. Models with yields and macro variables

Tables 6 and 7 contain estimation results of the Macro Model and the Macro Lag Model. The autocorrelation of the first latent factor is the same to three decimal places across the Yields-Only, Macro and Macro Lag Models (0.992). Hence, we would expect that this first factor has roughly similar very persistent effects in each model. There is more variation in the autocorrelations of Unobs 2 and Unobs 3 across the models. The risk premia estimates in Tables 6 and 7 corresponding to the latent factors have the same signs as in the Yields-Only Model in Table 5. Even though their magnitude differs somewhat across the three models, we expect that the latent factors behave in a similar fashion across the models.

The market price of risk coefficients corresponding to inflation and real activity are highly significant. This implies that observable macro factors drive time-variation in risk premia in both models! However, the estimates for λ_1 differ enormously across the Macro and Macro Lag Model. First, the λ_1 element corresponding to inflation ($\lambda_{1,11}$) is negative in the Macro Model (-0.4263) but positive in the Macro Lag Model (0.8442). Similarly, the real activity-term ($\lambda_{1,22}$) is also negative in the Macro Model (-0.1015) and positive in the Macro Model (0.2102). Finally, the inflation-real activity cross-terms ($\lambda_{1,12}$ and $\lambda_{1,21}$), where the additional two subscripts denote matrix elements, are much larger in absolute magnitude in the Macro Model than in the Macro Lag Model. Hence, we can expect inflation and real activity to play different roles in these two models. This also implies that estimates of the market price of risk are sensitive to the details of the model specification, particularly the parameterization chosen for the observable macro variables. Below, we show this sensitivity is important for economic inference.

5.2. Impulse responses

5.2.1. Factor weights across the yield curve

From Eq. (18), the effect of each factor on the yield curve is determined by the weights B_n that the term structure model assigns on each yield of maturity n. These weights B_n also represent the initial response of yields to shocks from the various factors. Fig. 4 plots these weights as a function of yield maturity for the Macro Model in the upper graph, and the Macro Lag Model in the lower graph. For the Macro Lag Model, we only plot the B_n coefficients corresponding to the contemporaneous macro variables. The B_n coefficients have been scaled to correspond to movements of one standard deviation of the factors, and have been annualized by multiplying by 1200.

Fig. 4 shows that the latent factors act in almost the same way in both the Macro and Macro Lag Models. The weight on the most persistent factor (Unobs 1) is almost horizontal. This means that it affects yields of all maturities the same way, hence the name "level" factor. The coefficient of the second factor (Unobs 2) is upward sloping. It mainly moves the short end of the yield curve relative to the long

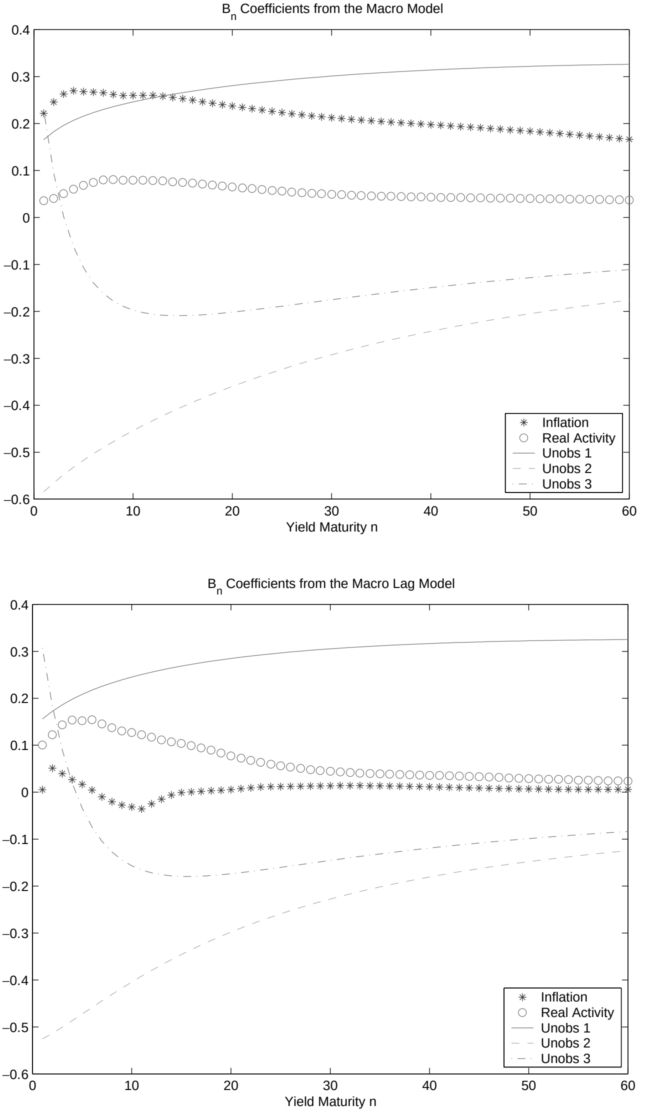


Fig. 4. B_n yield weights for the macro and macro lag model. The figure displays B_n yield weights as a function of maturity n for the Macro (Macro Lag) model in the top (bottom) plot. The plots show only the B_n yield weights corresponding to contemporaneous state variables in each system. The weights are scaled to correspond to one standard deviation movements in the factors and are annualized by multiplying by 1200.

end, so Unobs 2 is therefore a "slope" factor. The coefficient on the least persistent factor (Unobs 3) is hump-shaped. Movements in this factor affect yields at the short-end of the yield curve and middle and long-end of the yield curve with different signs. Hence, the B_n weights corresponding to Unobs 3 have a twisting effect, so Unobs 3 is thus a "curvature" factor. The inverse hump in the coefficient of this factor cannot be accommodated in a model with independent factors and constant risk premia, where yield coefficients are monotonic functions of maturity.

We now turn to the B_n coefficients corresponding to inflation and real activity. These coefficients differ across the Macro and Macro Lag Models. In the top plot of Fig. 4, the effect of inflation is hump-shaped but mostly affects short yields and less so long yields. The magnitude of the inflation weights are higher than the level factor weights at short maturities, and about half the magnitude of the slope factor weights. Initial shocks to real activity have a much weaker effect across the yield curve. In contrast, the bottom plot of Fig. 4 shows that the effects of shocks to inflation and real activity in the Macro Lag Model are weaker than the Macro Model, and shocks to real activity impact the yield more than shocks to inflation.

There are several reasons for the differences in the B_n coefficients for macro factors across the Macro and Macro Lag Models. First, in the estimates of the Taylor rules in Table 4, the Macro Model gives inflation a very strong effect on the short rate (coefficient = 0.1535). In the Macro Lag Model, initial shocks to inflation have little impact (coefficient = 0.0037), and it is only after 11 lags of inflation where inflation begins to have a large impact (coefficient on the 11th lag of inflation = 0.1427). Given that both models estimate the same standard deviation of inflation shocks (both rely on the same VAR(12) for inflation and real activity), we get a much stronger initial effect of inflation on yields in the Macro Model. Second, in the Macro Model, real activity has little initial impact (coefficient = 0.0143) while the effect in the Macro Lag Model is larger (coefficient = 0.0398). Given that the standard deviation of real activity shocks is also the same across models, the initial effect of real activity is stronger in the Macro Lag Model than in the Macro Model.

The time-varying prices of risk for inflation and real activity vary across the Macro and the Macro Lag Models in Tables 6 and 7. The prices of risk control how the variation of longer yields respond relative to the short rate. In the Macro Model, the λ_1 time-varying prices of risk for inflation and real activity are both negative ($\lambda_{1,11} = -0.4263$ and $\lambda_{1,22} = -0.1015$) , while in the Macro Lag Model these are both positive ($\lambda_{1,11} = 0.8442$ and $\lambda_{1,22} = 0.2102$). The more negative the terms on the λ_1 diagonal, the more positively longer yields react to positive factor shocks. Since the time-varying prices of risk for inflation and real activity are more negative for the Macro Model, the initial shocks are larger across the yield curve in this model.

While Fig. 4 shows only the initial effect of shocks as a function of yield maturity, we are also interested in how the initial shocks propagate through time. To trace out the long-term responses of the yield curve from shocks to the macro variables after the yield curve's initial response, we now compute IRs.

5.2.2. *Impulse responses from macro shocks*

We look at IRs to yields of maturities 1, 12 and 60 months. Our term structure model allows us to obtain the movements of the yield curve in response to driving shocks at all horizons, including maturities omitted in estimation. The IRs for *all* maturities are known analytical functions of the parameters. This is in contrast to

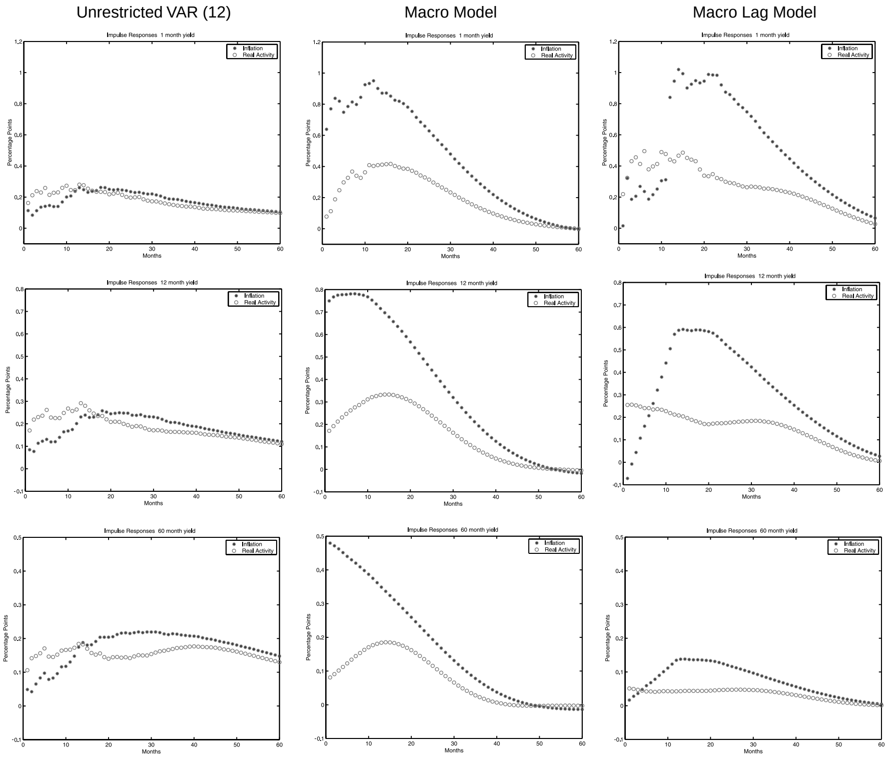


Fig. 5. Impulse response functions. Impulse Responses (IR's) for 1 month (top row), 12 month (middle row) and 60 month (bottom row) yields. The first column presents IR's from an unrestricted VAR(12) fitted to macro variables and yields ; the middle column presents IR's from the Macro model; and the last column presents IR's from the Macro Lag model. The IR's from the latent factors are drawn as lines, while the IR's from inflation (real activity) are drawn as stars (circles). All IR's are from a one standard deviation shock.

estimations with VARs where IRs can only be calculated for yields included in the VAR. Our estimation also guarantees that the movements of yields are arbitrage-free.

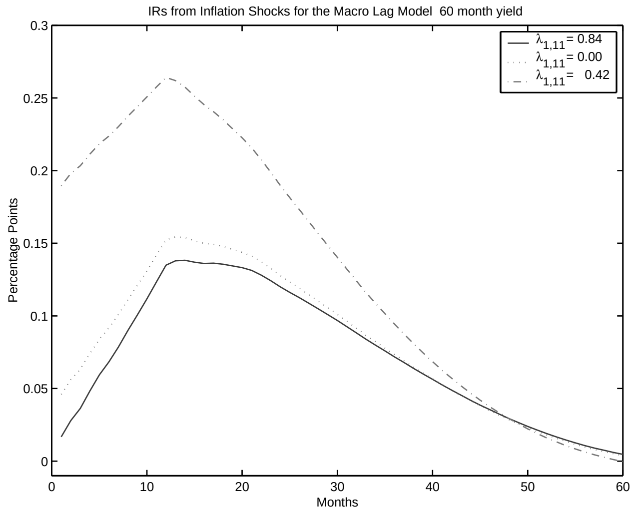
Fig. 5 shows IRs of 1, 12 and 60 month yields from the Macro Model and the Macro Lag Model. The initial IR (corresponding to 1 month on the x-axis) for each factor correspond to the 1, 12 and 60 month maturity in Fig. 4. In addition, we compute the IRs from a simple unrestricted VAR(12), with macro factors and 5 yields similar to Campbell and Ammer (1993). We order the variables with macro factors first, and then yields with increasing maturities, but the effect is robust to the ordering of variables in the VAR. The x-axis on each plot is in months and the IRs are given in terms of annualized percentages for a shock of one standard deviation.

Fig. 5 shows that the IRs for the Macro and the Macro Lag Models are much larger than the IRs from the unrestricted VAR(12) (except for the 60 month yield for the Macro Lag Model). The maximum magnitude of the responses for the 1 month and 60 month yields is up to five times larger than the VAR(12). Turning first to the IRs of the unrestricted VAR in the first column of Fig. 5, a one-standard deviation shock to inflation initially raises the 1-month yield about 10 basis points. The response peaks after about two years at 30 basis points and then slowly levels off. The response of longer yields has the same overall shape. The initial response of the 1-year yield (5-year yield) is only 8 basis points (5 basis points). The response increases to around 25 basis points (22 basis points) after 2 years, and then dies off slowly. The response of yields to real activity shocks in the unrestricted VAR is slightly smaller than the response to inflation shocks. The real activity response is also hump-shaped with the hump occurring after one year.

The second column of Fig. 5 plots IRs for the Macro Model. The hump-shape of the IRs are similar to the shape of the IRs from the unrestricted VAR, but the IRs are much larger. For example, the initial response of the 1-month yield to a 1 standard deviation inflation shock is around 60 basis points, peaking after 12 months to slightly under 1%. This is over six times the effect as the unrestricted VAR(12). For the 5 year yield, the initial response to inflation is around 50 basis points, compared to a less than 5 basis point move for the VAR(12). However, the effect of real activity is about the same order of magnitude as the VAR(12) and is much smaller than the IRs from inflation shocks. This is due primarily to the small loading on real activity (0.0143) in the Taylor rule, compared to the much larger loading on inflation (0.1535).

We plot IRs for the Macro Lag Model in the final column of Fig. 5. For inflation, there are much longer lagged effects, after 12 months, than in the Macro Model. This is because the Taylor rule with lags has a significant weight on the 11th lag of inflation, which has its highest impact after 12 months (see Table 4). In contrast, the weights in the Taylor rule for real activity are largely flat, so there is little hump-shape and also less impact from shocks to real activity. The Macro Lag Model IRs for inflation reach almost the same magnitude as the IRs for the Macro Model for the 1 and 12-month yields, but are much smaller for the 60-month yield. This is in contrast to the Macro Model, where inflation shocks have much bigger impacts across the yield curve.

The reason for the different effects across the Macro and Macro Lag Model at longer maturities is due to the estimates of the time-varying price of risk λ_1 for each model in Tables 6 and 7. The diagonal elements of λ_1 in the Macro Lag Model are negative, while they are positive in the Macro Model. Lower (more negative) prices of risk have higher positive impacts from the macro factors to long yields. Fig. 6 focuses on IRs for the 60-month yield in the Macro Lag Model. The top (bottom) plot traces IRs for three different values of $\lambda_{1,11}$ ($\lambda_{1,22}$), starting from the parameter estimates 0.84 (0.21). The negative parameter choice for $\lambda_{1,11}$ ($\lambda_{1,22}$) is the corresponding parameter estimate for the Macro Model, -0.42 (-0.10). In each case, decreasing the diagonal prices of risk increases the magnitude of the IRs. Note that for IRs from real activity shocks, as $\lambda_{1,22}$ decreases, there is higher exposure to the oscillatory effects from the lagged Taylor rule combined with the VAR(12) fitted to inflation and real activity.



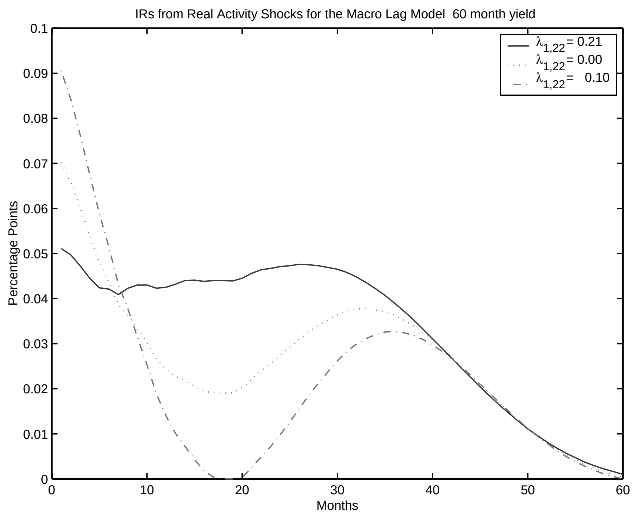


Fig. 6. Impulse responses for the 60-month yield from the macro lag model. We show IR's for the 60-month yield for various time-varying diagonal prices of risk $\lambda_{1,11}$ and $\lambda_{1,22}$. All other parameters are held fixed at their estimated values in Table 7. The top (bottom) plot displays the IR's for inflation (real activity) for the Macro Lag Model.

5.3. Variance decompositions

To gauge the relative contributions of the macro and latent factors to forecast variances we construct variance decompositions. These show the proportion of the forecast variance attributable to each factor, and are closely related to the IRs of the previous section. Table 8 summarizes our results for the Macro and Macro Lag Models.

Table 8
Proportion of variance explained by macro factors

	Forecast horizon h			
	1 mth	12 mth	60 mth	∞
Macro model				
Short end	50%	78%	85%	83%
Middle	67%	79%	78%	73%
Long end	61%	63%	48%	38%
Macro lag model				
Short end	11%	57%	87%	85%
Middle	23%	52%	71%	64%
Long end	2%	8%	11%	7%

We list the contribution of the macro factors to the h-step ahead forecast variance of the 1 month yield (short end), 12 month yield (middle) and 60 month yield (long end) for the Macro and Macro Lag Models. These are the sum of the variance decompositions from the macro factors in Table 9.

The proportion of unconditional variance accounted for by macro factors is decreasing with the maturity of yields: highest at the short and middle-ends of the yield curve, and smallest for the long-end. The largest effect is for the 1-month yield where macro factors account for 83% (85%) of the unconditional variance (where the forecasting horizon is infinite) for the Macro (Macro Lag) Model. The proportion of the unconditional variance for the 60-month yield is much smaller for the Macro Lag Model (only 7%) versus the Macro Model (38%). This is because of the more negative prices of risk for the Macro Model compared to the Macro Lag Model, allowing the response of the longer yields to be more larger to real activity shocks in the Macro Model.

That the macro factors explain so much of the variance decomposition for the 1-month yield at a 1-month forecasting horizon is no surprise, since this result is by construction. The latent factors in the Macro and Macro Lag Models now explain the residuals after taking out the effects of inflation and real activity in the Taylor rule (see Table 4). In contrast, the latent factors in the Yields-Only Model account for interest rates themselves, and are merely transformations of yields. What is more interesting is the behavior of the longer yields compared to the short rate, since these are driven by the no-arbitrage restrictions on the VAR.

The proportion of forecast variance explained by macro factors has an interesting hump-shaped pattern for short and intermediate maturities. For example for the Macro Model, macro factors account for 50% of the 1-step ahead forecast variance of the 1-month yield. This percentage rises to 85% at a 60-month horizon, but then converges to 83% for extremely long forecast horizons. Generally, as the yield maturity increases, the proportion of the forecast variance attributable to the latent factors increases. For the 60-month yield, the latent factors account for 62% and 93% of the unconditional variance in the Macro and Macro Lag Models, respectively. The low variance decomposition of long yields is due to the dominance of persistent unobserved factors (the near unit-root factor). As the level latent factor has the highest autocorrelation, the weights on this factor are highest for long maturities (see the recursion in Eq. (17)).

More detailed variance decompositions are listed in Table 9 for 1, 12 and 60 month maturities. To interpret the top row of Table 9, for the Yields-Only model, 14% of the 1-step ahead forecast variance of the 1-month yield is explained by the first unobserved factor, 33% by the second unobserved factor and 53% by the third unobserved factor. In the row labeled $h = 1$ of the Macro Model in the first panel corresponding to the 1-month yield, 49% of the 1-step ahead forecast variance is attributable to inflation, 1% to real activity and the remainder to the latent factors.[8]

Focusing on the Macro Model, inflation has more explanatory power for forecast variances than real activity at all points of the yield curve and for all forecast horizons. The explanatory power of real activity generally rises with the forecast interval h. At the long end of the yield curve, the explanatory power of inflation decreases with h. At short horizons, very little of the forecast variance can be attributed to real activity across the yield curve, but as the horizon increases, the proportion due to real activity shocks increases to 13% of the 1-month and 11% of the 12-month yield. The effect at the long end of the yield curve is much smaller (less than 6% of the unconditional variance for the 60-month yield). This pattern is due to the large weight on inflation in the simple Taylor rule and the much smaller weight on real activity. At long yields, the higher persistence of the latent factors dominates, which decreases the explanatory power for the macro factors.

In the Macro Lag Model, inflation and real activity explain roughly the same proportion of the unconditional and long-horizon variances for the short and medium segments of the yield curve as the Macro Model. However, the initial variance decompositions at short forecasting horizons are much smaller than the Macro Model. For example, for the 1-month yield, the initial Macro (Macro Lag) inflation variance decomposition is 49% (0%). This is because the full effect of the macro variables, particularly inflation, does not kick in until the 11th lag of inflation in the Taylor rule with lags. At the long end of the yield curve, the Macro Lag Model has very little role for macro factors. Here, the Macro Lag Model's positive (diagonal) prices of risk for the macro factors do not allow the long end of the yield curve to share the same positive short end exposure to macro shocks as the Macro Model's negative (diagonal) prices of risk.

Turning now to the latent factors in Table 9, Unobs 1, is the most persistent latent factor corresponding to a level effect. For the Yields-Only model, this factor dominates the variance decompositions at long horizons across the yield curve. The variance decomposition of Unobs 1 is markedly reduced for the 1-month and 12-month yields for the Macro and Macro Lag Models. Here, the persistence of inflation plays a major role in absorbing most of the mean forecast error due to the large effect inflation has in the Taylor rule. For the 60-month yield, Unobs 1 has a

[8] A Cholesky orthogonalization was used for inflation and real activity, which are correlated. Changing the ordering of inflation and real activity has very little effect on the results. Note that the macro factors and latent factors are orthogonal.

Table 9
Variance decompositions

	h	Macro factors		Latent factors		
		Inflation	Real activity	Unobs 1	Unobs 2	Unobs 3
1 month yield						
Yields-Only	1			0.14	0.33	0.53
	12			0.31	0.43	0.26
	60			0.56	0.31	0.13
	∞			0.68	0.23	0.09
Macro	1	0.49	0.01	0.03	0.41	0.06
	12	0.69	0.09	0.03	0.18	0.01
	60	0.71	0.14	0.04	0.11	0.00
	∞	0.70	0.13	0.06	0.10	0.00
Macro	1	0.00	0.11	0.05	0.62	0.21
Lag	12	0.22	0.35	0.04	0.34	0.05
	60	0.69	0.18	0.03	0.09	0.01
	∞	0.67	0.18	0.05	0.09	0.01
12 month yield						
Yields-Only	1			0.60	0.35	0.05
	12			0.71	0.28	0.01
	60			0.86	0.14	0.00
	∞			0.91	0.09	0.00
Macro	1	0.63	0.03	0.07	0.21	0.05
	12	0.71	0.08	0.07	0.13	0.01
	60	0.66	0.12	0.13	0.09	0.00
	∞	0.62	0.11	0.19	0.08	0.00
Macro	1	0.02	0.22	0.21	0.47	0.10
Lag	12	0.33	0.19	0.20	0.26	0.02
	60	0.59	0.12	0.19	0.09	0.01
	∞	0.52	0.11	0.28	0.08	0.01
60 month yield						
Yields-Only	1			0.75	0.20	0.05
	12			0.84	0.14	0.01
	60			0.93	0.06	0.00
	∞			0.96	0.04	0.00
Macro	1	0.60	0.02	0.28	0.08	0.03
	12	0.57	0.06	0.31	0.06	0.01
	60	0.40	0.08	0.49	0.03	0.00
	∞	0.32	0.06	0.60	0.02	0.00
Macro	1	0.00	0.02	0.81	0.12	0.05
Lag	12	0.06	0.02	0.84	0.07	0.02
	60	0.09	0.02	0.86	0.03	0.00
	∞	0.06	0.01	0.91	0.02	0.00

The table lists the contribution of factor i to the h-step ahead forecast of the 1 month yield. To interpret the top row, for the Yields-Only model, 14% of the 1-step ahead forecast variance is explained by the first unobserved factor, 33% by the second unobserved factor and 53% by the third unobserved factor. The Yields-Only Model only has three latent factors. The macro models have inflation, real activity and three latent factors. The Macro Model has no lags of inflation and real activity in the short rate equation, while the Macro Lag Model does.

much larger impact, explaining 60% (91%) of the unconditional variance for the Macro (Macro Lag) Model. As we move across maturities, the initial macro shocks to the short rate become more muted, since the weights to the macro factors become smaller as maturity increases, as Fig. 4 shows. Making the time-varying prices of risk more negative increases the positive responses at the long-end of the yield curve to shocks in inflation and real activity.

5.4. Forecasts

The variance decompositions hint that term structure models with observable macro variables may help in forecasting future movements in yields. However, these are statements based on assuming a particular model as the true model after estimation, and may not hold in a practical setting where more parsimonious data representations often out-perform sophisticated models. To determine if this is actually the case we conduct an out-of-sample forecasting experiment.

Our procedure for examining out-of-sample forecasts over the last 5 years of our sample is as follows. We examine forecasts for all the five yields used in estimation. At each date t, we estimate the models using data up to and including time t, and then forecast the next month's yields at time $t+1$. The macro factor data is formed using the principal components of the macro data up to time t, and we estimate the short rate equation and the bivariate VAR of the macro dynamics only using data up to time t for the Macro and Macro Lag Models. Hence, we only use data available in the information set at time t in making the forecast at time $t+1$.

We perform a comparison of out-of-sample forecasts for six models. First, we use a simple random walk. Second, we investigate out-of-sample forecasts for the corresponding VAR(12)'s which do not impose cross-equation restrictions. Our first VAR uses only yields, and we use a second VAR which incorporates yields and macro variables. Our last three models are the Yields-Only, Macro and the Macro Lag Models. We use two criteria to compare our forecasts across the models. The first is the Root Mean Squared Error, RMSE, of actual and forecasted yields, and the second is the Mean Absolute Deviation, MAD.

Table 10 lists the results of the out-of-sample comparisons. Lower RMSE and MAD values denote better forecasts. The best model RMSE or MAD is listed in bold. We forecast over the last 60 months of the sample, where interest rate volatility is much lower than over the full sample, which includes the very volatile late 1970s and early 1980s. We note the following points regarding the forecasting performance of the models. First, a random walk easily beats an unconstrained VAR. The result holds independently of whether the VARs only contain yields, or are augmented with macro variables. In fact, the forecasts are worse adding macro factors into the unconstrained VARs. The bad performance is due to the high persistence of yields and small sample biases in the estimation of autoregressive coefficients in over-parameterized VARs.

Second, imposing the cross-equation restrictions from no-arbitrage helps in forecasting. The improvement in forecasting performance is substantial, generally about 25% of the RMSE and 30% of the MAD for all yields. These constrained

Table 10
Forecast comparisons

Yield (mths)	RW	Unconstrained VARs		VARs with cross-equation restrictions		
		VAR Yields Only	VAR with Macro	Yields Only	Macro model	Macro lag model
RMSE criteria						
1	0.3160	0.3905	0.3990	0.3012	**0.2889**	0.3906
3	**0.1523**	0.2495	0.2540	0.1860	0.2167	0.2876
12	0.1991	0.2776	0.2722	0.1914	**0.1851**	0.2274
36	0.2493	0.3730	0.3644	0.2489	**0.2092**	0.2665
60	0.2546	0.3793	0.3725	0.2497	**0.2333**	0.2530
MAD criteria						
1	0.2252	0.3076	0.3242	0.2155	**0.2039**	0.2981
3	**0.1159**	0.1987	0.2056	0.1442	0.1693	0.2344
12	0.1639	0.2176	0.2204	0.1616	**0.1559**	0.1870
36	0.1997	0.2991	0.2924	0.1974	**0.1604**	0.2111
60	0.2054	0.2957	0.2930	0.2017	**0.1883**	0.2064

We forecast over the last 60 months (the out-sample) of our sample and record the root mean square error (RMSE) and the mean absolute deviation (MAD) of the forecast versus the actual values. Lower RMSE and MAD values denote better forecasts, with the best statistics highlighted in bold. Forecasts are 1-step ahead. We first estimate models on the in-sample, and update the estimations at each observation in the out-sample. RW denotes a random walk forecast, VAR Yields Only denotes a VAR(12) only with 5 yields, VAR with Macro denotes a VAR(12) fitted to the macro factors and all 5 yields, Yields-Only denotes the 3 factor latent variable model without macro variables, the Macro model has only contemporaneous inflation and real activity in the short rate equation, and the Macro Lag model has contemporaneous and 12 lags of inflation and real activity in the short rate equation. The first three of these models are thus unconstrained estimations, while the last three impose the cross-equation restrictions derived from the absence of arbitrage.

VAR's perform in line with, and slightly better, than a random walk (except for the 3-month yield). Duffee (2002) remarks that beating a random walk with a traditional affine term structure model is difficult. From forecasting exercises without risk premia (not reported here), we know that this result crucially depends on the type of risk adjustment. Linear risk premia, not considered by Duffee (2002), seem to do well in this regard.

Third, the forecasts of the Macro Model are far better than those of the Macro Lag Model. While the forecasts of the Macro Lag Model are comparable to those of unconstrained VARs, the Macro model slightly outperforms a random walk (except for the 3-month yield). Both the Macro Model and the Macro Lag Model impose cross-equation restrictions on a VAR with yields and macro variables. The Macro Lag Model, however, has a large number of insignificant coefficients entering the short rate equation. This over-parameterization may cause its poor out-of-sample performance.

Finally, incorporating macro variables helps in forecasting. More precisely, the forecasts of the Macro Model are uniformly better than the Yields-Only Model (except for the 3 month yield). Hence, we can conclude that (i) adding term structure restrictions improves forecasts relative to unconstrained VARs, even beating a random walk, and (ii) forecasts can be further improved by including macro variables. Note, however, that we have shown this improvement is only in incrementally adding macro factors to a given number of latent factors.

5.5. Comparison of factors

We now finally address the issue of how adding macro factors changes the original latent factors of the Yields-Only model in Table 11. In this table we regress the latent

Table 11
Comparison of Yields-Only and macro factors

Dependent variable	Independent variables					
	Inflation	Real activity	Unobs 1	Unobs 2	Unobs 3	Adj R^2
Panel A: Regressions on macro factors						
Unobs 1	0.4625	−0.0726				0.2180
"level"	(0.0735)	(0.0860)				
Unobs 2	−0.6707	−0.1890				0.4902
"spread"	(0.0716)	(0.0611)				
Unobs 3	0.0498	−0.1794				0.0343
"curvature"	(0.0629)	(0.0714)				
Panel B: Regressions on factors from macro model						
Unobs 1	0.1118	0.0307	0.9507	−0.0174	0.0038	0.9971
	(0.0054)	(0.0056)	(0.0055)	(0.0056)	(0.0047)	
Unobs 2	−0.9364	−0.1026	0.0199	0.7624	0.0279	0.9981
	(0.0037)	(0.0037)	(0.0042)	(0.0032)	(0.0029)	
Unobs 3	0.0427	−0.1238	0.1656	−0.1455	0.9071	0.9256
	(0.0262)	(0.0260)	(0.0289)	(0.0241)	(0.0233)	
Panel C: Regressions on factors from macro lag model						
Unobs 1	−0.0580	−0.0207	1.0248	0.0035	0.0058	0.9979
	(0.0049)	(0.0040)	(0.0044)	(0.0047)	(0.0036)	
Unobs 2	−0.7069	−0.1132	−0.2955	0.5700	0.1306	0.8715
	(0.0393)	(0.0313)	(0.0356)	(0.0376)	(0.0315)	
Unobs 3	0.1112	−0.0081	0.2059	0.0228	0.8119	0.7470
	(0.0458)	(0.0386)	(0.0507)	(0.0365)	(0.0424)	

Regressions of the latent factors from the Yields-Only model with only latent factors (dependent variables) onto the macro factors and latent factors from the Macro and Macro Lag model (independent variables). All factors are normalized, and standard errors, produced using 3 Newey–West (1987) lags, are in parentheses. Panel A lists coefficients from a regression of the Yields-Only latent factors onto only macro factors. Panel B lists coefficients from a regression of Yields-Only latent factors on the macro and latent factors from the Macro model with only contemporaneous inflation and real activity in the short rate equation. Panel C lists coefficients from a regression of Yields-Only latent factors on the macro and latent factors from the Macro Lag model with contemporaneous inflation and real activity and 11 lags of inflation and real activity in the short rate equation.

factors from the Yields-Only model onto the macro and latent factors from the Macro and Macro Lag Models. We run three series of regressions, first only on the macro variables (Panel A), and then onto the macro and latent variables of the Macro Model (Panel B), and then onto the macro and latent variables of the Macro Lag Model (Panel C). All the variables in the regressions are normalized.

Turning first to Panel A of Table 11, the traditional level factor loads significantly onto inflation and real activity, with an adjusted R^2 of 22%. In particular, the loading on inflation is positive and large (0.46). This suggests that the traditional level factor captures a strong inflation effect. When the second latent factor, labeled "slope," is regressed onto the macro factors, we obtain a high R^2 of 49%, with significant negative loadings particularly on inflation (-0.67). Hence, much of the traditional slope factor is also related to the dynamics of inflation. Finally, the third latent factor ("curvature") is poorly accounted by macro factors $R^2 = 3\%$. However, the traditional curvature factor does load significantly on real activity.

Panel B of Table 11 reports the regression from the traditional Yields-Only factors onto the macro and latent factors implied by the Macro Model. The level factor from the Yields-Only model translates almost one for one with the level factor of the Macro Model. The magnitude of the coefficient on Unobs 1 of the Macro model is close to 1 (0.95), showing that there is some qualitative similarity. However, we reject that the coefficient is equal to 1 at the 1% level, showing that the two latent factors are statistically different. The loadings on the macro factors remain significant suggesting that macro variables do account for some of the level factor.

The reason why the level factor survives largely intact when macro factors are introduced is because the level factor proxies for the first principal component of the yield curve. Fig. 3 shows that the residuals from the Taylor rules largely mimic the level of the short rate. Since the unobservable factors are linear combinations of the yields, the best linear combination of yields which explains term structure movements is the first principal component, or the level of the short rate. When macro factors are added, these factors still do not resemble the level of the yield curve, and so this factor is still necessary to explain the movements across the term structure.

When we regress the Yields-Only slope factor (Unobs 2) onto the Macro Model factors, the loading of the Unobs 2 factor from the Macro Model is much smaller than 1 (0.76), while the coefficient on inflation is very large and negative, and the coefficient on real activity is also significant. The loading on the Unobs 2 factor from the Macro Lag Model is even smaller (0.57). This means that a large part of the traditional slope factor can be attributed to macro factors, in particular, inflation movements. When inflation is high, the slope narrows because the short rate increases relative to the long rate. Turning finally to the regression of the Yields-Only curvature factor (Unobs 3), this regression still has a significant negative coefficient on real activity, but most of the correspondence is with the Unobs 3 factor from the Macro Model (the coefficient is 0.91). Nevertheless, this coefficient is also statistically different from 1 at the 1% level.

Panel C of Table 11 reports the regression coefficients of the latent factors from the Yields-Only Model onto the macro and latent factors of the Macro Lag Model.

We see that here the level effect survives almost one for one and there is still a large loading on the inflation factor by the Yields-Only model's Unobs 2. However, the R^2's of the Unobs 2 and Unobs 3 regressions are much smaller than the Macro Model regressions in Panel B.

In summary, Table 11 shows that the traditional level and slope factors are markedly associated with and accounted by observable macro factors. In particular, inflation accounts for large amounts of the dynamics of the traditional slope factor. However, the level effect qualitatively survives largely intact when macro factors are added to a term structure model.

6. Conclusion

This paper presents a Gaussian model of the yield curve with observable macroeconomic variables and traditional latent yield variables. The model takes a first step towards understanding the joint dynamics of macro variables and bond prices in a factor model of the term structure. Risk premia are time-varying; they depend on both observable macro variables and unobservable factors. The approach extends existing VAR studies of yields and macro variables by imposing no-arbitrage assumptions.

We find that macro factors explain a significant portion (up to 85%) of movements in the short and middle parts of the yield curve, but explain only around 40% of movements at the long end of the yield curve. The effects of inflation shocks are strongest at the short end of the yield curve. Comparing the latent factors from traditional three latent factor models of term structure, the "level" factor survives almost intact when macro factors are incorporated, but a significant proportion of the "level" and "slope" factors are attributed to macro factors, particularly to inflation. Moreover, we find that imposing the cross-equation restrictions from no arbitrage helps in out-of-sample forecasts. Incorporating macro factors in a term structure model further improves forecasts.

In future research, we plan to extend our empirical specification to allow for non-diagonal terms in the companion form for the factors which introduces feedback from latent factors to macro variables. Yields can then forecast macro variables along the lines of Estrella and Hardouvelis (1991) but with the dynamics of the yield curve modeled in a no-arbitrage pricing approach.

Appendix A. Recursive bond prices

To derive the recursions in Eq. (17), we first note that for a one-period bond, $n = 1$, we have

$$\begin{aligned} p_t^1 &= \mathrm{E}_t[m_{t+1}] = \exp\{-r_t\} \\ &= \exp\{-\delta_0 - \delta_1' X_t\}. \end{aligned} \tag{A.1}$$

Matching coefficients leads to $\bar{A}_1 = -\delta_0$ and $\bar{B}_1 = -\delta_1$. Suppose that the price of an n-period bond is given by $p_t^n = \exp(\bar{A}_n + \bar{B}_n' X_t)$. Now we show that the exponential form also applies to the price of the $n+1$ period bond:

$$\begin{aligned} p_t^{n+1} &= \mathrm{E}_t[m_{t+1} p_{t+1}^n] \\ &= \mathrm{E}_t\left[\exp\left\{-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\varepsilon_{t+1} + \bar{A}_n + \bar{B}_n' X_{t+1}\right\}\right] \\ &= \exp\left\{-r_t - \frac{1}{2}\lambda_t'\lambda_t + \bar{A}_n\right\}\mathrm{E}_t[\exp\{-\lambda_t'\varepsilon_{t+1} + \bar{B}_n' X_{t+1}\}] \\ &= \exp\left\{-r_t - \frac{1}{2}\lambda_t'\lambda_t + \bar{A}_n\right\}\mathrm{E}_t[\exp\{-\lambda_t'\varepsilon_{t+1} + \bar{B}_n'(\mu + \phi X_t + \Sigma\varepsilon_{t+1})\}] \\ &= \exp\left\{-\delta_0 + \bar{A}_n + \bar{B}_n'\mu + (\bar{B}_n'\phi - \delta_1')X_t - \frac{1}{2}\lambda_t'\lambda_t\right\} \\ &\quad \times \mathrm{E}_t[\exp\{-(\lambda_t' + \bar{B}_n'\Sigma)\varepsilon_{t+1}\}] \\ &= \exp\left\{-\delta_0 + \bar{A}_n + \bar{B}_n'(\mu - \Sigma\lambda_0) + \frac{1}{2}\bar{B}_n'\Sigma\Sigma'\bar{B}_n - \delta_1' X_t \right. \\ &\quad \left. + \bar{B}_n'\phi X_t - \bar{B}_n'\Sigma\lambda_1 x_t\right\}. \end{aligned} \tag{A.2}$$

The last equality relies on ε_t being IID normal with $\mathrm{E}[\varepsilon_t] = 0$ and a degenerate variance-covariance matrix $\mathrm{var}[\varepsilon_t]$ which contains many zeros (see Eq. (8)). Also, λ_t contains zero submatrices (see Eq. (11)). Taken together, these assumptions imply that $\lambda_t'\lambda_t = \lambda_t'\mathrm{var}(\varepsilon_t)\lambda_t$. Matching coefficients results in the recursive relations in Eqs. (17).

Appendix B. Likelihood function

We have data on an N vector of zero coupon yields Y_t. Our approach to estimation is to solve for the unobserved factors f_t^u from the yields Y_t and the observed variables X_t^o, which includes observed macro variables f_t^o and latent variables f_t^u and lagged terms of the driving factors.

Suppose first that we have $N = K_2$ yields of different maturity $n_1, \ldots, n_{K_2}$, as many yields as we have unobserved factors, f_t^u. Stacking the equations for the K_2 yields, with $Y_t = (y_t^{n_1} \ldots y_t^{n_{K_2}})'$, we can write

$$Y_t = A + BX_t, \tag{B.1}$$

where A is $K_2 \times 1$ and B is $K_2 \times K$. Partition the matrix B into $B = [B^0 \;\; B^u]$ where B^o is a $K_2 \times (K - K_2)$ matrix which picks up the observable factors and B^u is a $K_2 \times K_2$ invertible matrix that picks up the unobservable factors. Then we can infer the unobservable factors in $X_t^u \equiv f_t^u$ from Y_t and the pricing matrices A and B using an inversion from the equation:

$$Y_t = A + B^o X_t^o + B^u X_t^u. \tag{B.2}$$

The term structure model only prices exactly the yields used to invert the latent factors. To increase the number of yields to $N > K_2$ in the estimation, we follow Chen and Scott (1993), and others, in assuming that some of the yields are observed with measurement error. There will be K_2 yields from which we invert to obtain the latent variables, and the other $N - K_2$ yields are measured with error. We assume this measurement error is IID, and the measurement error is uncorrelated across the yields measured with error. Let B^m denote a $N \times (N - K_2)$ measurement matrix and u_t^m be an $(N - K_2)$-dimensional Gaussian white noise with a diagonal covariance matrix independent of X_t. With N yields, the matrix $[B^o B^u]$ of yield coefficients now has dimension $N \times K$. (B^o is $N \times (K - K_2)$, while B^u is $N \times K_2$.) We can then write

$$Y_t = A + B^o X_t^o + B^u X_t^u + B^m u_t^m. \tag{B.3}$$

In Eq. (B.3) the yields measured without error will be used to solve for X_t^u, and the yields measured with error have non-zero u_t^m. For a given parameter vector $\theta = (\mu, \Phi, \Sigma, \delta_0, \delta_1, \lambda_0, \lambda_1)$, we can invert Eq. (B.3) to obtain X_t^u and u_t^m.

Denoting the normal density functions of the state variables X_t and the errors u_t^m as f_X and f_{u^m} respectively, the joint likelihood $\mathfrak{L}(\theta)$ of the observed data on zero coupon yields Y_t and the observable factors X_t^o is given by

$$\mathfrak{L}(\theta)) = \prod_{t=2}^{T} f(Y_t, X_t^o \mid Y_{t-1}, X_{t-1}^o)$$

$$\begin{aligned}
\log(\mathfrak{L}(\theta)) = \sum_{t=2}^{T} & - \log|\det(J)| + \log f_x(X_t^o, X_t^u \mid X_{t-1}^o, X_{t-1}^u) + \log f_{u^m}(u_t^m) \\
= & - (T-1)\log|\det(J)| - (T-1)\frac{1}{2}\log(\det(\Sigma\Sigma')) \\
& - \frac{1}{2}\sum_{t=2}^{T} (X_t - \mu - \Phi X_{t-1})'(\Sigma\Sigma')^{-1}(X_t - \mu - \Phi X_{t-1}) \\
& - \frac{T-1}{2}\log \sum_{i=1}^{N-K_2} \sigma_i^2 - \frac{1}{2}\sum_{t=2}^{T}\sum_{i=1}^{N-K_2} \frac{(u_{t,i}^m)^2}{\sigma_i^2},
\end{aligned} \tag{B.4}$$

where σ_i^2 is the variance of the ith measurement error and the Jacobian term is given by

$$J = \begin{pmatrix} I & 0 & 0 \\ B^o & B^u & B^m \end{pmatrix}.$$

Note that the Jacobian terms of the likelihood in equation (B.4) do not involve A_n, and hence the constant prices of risk λ_0 but do involve the linear prices of risk λ_1.

Appendix C. Impulse responses

To derive the IR's of the yields from shocks to the macro variables and latent yield factors $F_t = (f_t^{o'}, f_t^{u'})'$ consider the VAR(12) form of F_t in Eq. (7), repeated here:

$$F_t = \Phi_0 + \Phi_1 F_{t-1} + \cdots + \Phi_{12} F_{t-12} + \theta_0 u_t. \quad \text{(C.1)}$$

The Φ_i coefficients take the following form in our parameterization:

$$\Phi_0 = 0 \quad \Phi_1 = \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho \end{pmatrix}, \quad \Phi_i = \begin{pmatrix} \rho_i & 0 \\ 0 & 0 \end{pmatrix} \quad \text{for } i = 2, \ldots, 12.$$

We write this as an implied Wold MA(∞) representation:

$$F_t = \sum_{i=0}^{\infty} P_i u_{t-i}, \quad \text{(C.2)}$$

where $u_t = (u_t^{o'} u_t^{u'})'$ are the shocks to F_t. Note that a Choleski adjustment is needed to take into account the contemporaneous correlation of the shocks.

The yield on an n-period zero coupon bond y_t^n is a linear combination of current and lagged values of u_t from Eq. (18), which we can write as

$$y_t^n = A_n + \sum_{i=0}^{\infty} \psi_i^n u_{t-i}, \quad \text{(C.3)}$$

where the row vectors ψ_i^n are functions of B_n. Note that this is just a linear transformation of the original MA(∞) form, and the B_n are closed-form from Eq. (17).

For example, for the Macro Model, the state-space X_t is given by

$$X_t = (f_t^{o'} \; f_{t-1}^{o'} \; \ldots \; f_{t-11}^{o'} \; f_t^{u'})',$$

where f_t^o are the two macro factors, and f_t^u are the three unobservable factors. The yields for maturity n, y_t^n, can be written as:

$$\begin{aligned} y_t^n &= A_n + B_n' X_t \\ &= A_n + B_{n0}^{o'} f_t^o + \cdots + B_{n11}^{o'} f_{t-11}^o + B_n^{u'} f_t^u \\ &= A_n + \bar{B}_{n0}' F_t + \cdots + \bar{B}_{n11}' F_{t-11}, \end{aligned} \quad \text{(C.4)}$$

where we partition as $B_n = [B_{n0}^o \ldots B_{n11}^o B_n^u]$, where B_{ni}^o corresponds to f_{t-i}^o for $i = 0, \ldots, 11$ and B_n^u corresponds to f_t^u, and $\bar{B}_{n0} = [B_{n0}^o \; B_n^u]$, and $\bar{B}_{ni} = [B_{ni}^o \; 0]$ for $i = 1, \ldots, 11$.

$$\begin{aligned} \psi_0^n &= \bar{B}_{n0}' P_0 \\ \psi_1^n &= \bar{B}_{n0}' P_1 + \bar{B}_{n1}' P_0 \\ &\vdots \\ \psi_i^n &= \bar{B}_{n0}' P_i + \cdots + \bar{B}_{n11}' P_{i-11} \quad \text{for } i \geqslant 11 \end{aligned} \quad \text{(C.5)}$$

and so on.

The vector ψ_i^n is the IR for the n-period yield at horizon i for shocks to the driving variables F_t at time 0. For k yields of maturities $n_1, \ldots, n_k$, we can stack the coefficients of each yield to write:

$$Y_t = A + \sum_{i=0}^{\infty} \Psi_i u_{t-i}, \tag{C.6}$$

where $Y_t = (y_t^{n_1} \dots y_t^{n_k})'$ and the jth row of Ψ_i is ψ_i^n.

Appendix D. Variance decompositions

Working with the MA(∞) representation of the yields in Eq. (C.6), the error of the optimal h-step ahead forecast at time t, $\hat{Y}_{t+h\,|\,t}$ is

$$\hat{Y}_{t+h\,|\,t} - Y_{t+h} = \sum_{i=0}^{h-1} \Psi_i u_{t+h-i}. \tag{D.1}$$

Let the jth component of a vector be denoted by a superscript j and $\Psi_{jk,i}$ denote the element in row j, column k of Ψ_i. Then

$$\hat{Y}^j_{t+h\,|\,t} - Y^j_{t+h} = \sum_{k=1}^{K} (\Psi_{jk,0} u^k_{t+h} + \dots \Psi_{jk,h-1} u^k_{t+1}). \tag{D.2}$$

Denote the mean squared error of $\hat{Y}^j_{t+h\,|\,t}$ as $\text{MSE}(\hat{Y}_{t+h\,|\,t})$. Then

$$\text{MSE}(\hat{Y}_{t+h\,|\,t}) = \sum_{k=1}^{K} (\Psi^2_{jk,0} + \cdots + \Psi^2_{jk,h-1}). \tag{D.3}$$

The contribution $\Omega_{jk,h}$ of the kth factor to the MSE of the h-step ahead forecast of the jth yield is

$$\Omega_{jk,h} = \frac{\sum_{i=0}^{h-1} \Psi^2_{jk,i}}{\text{MSE}(\hat{Y}_{t+h\,|\,t})}, \tag{D.4}$$

which decomposes the forecast variance at horizon h of the jth yield to the various factors.

References

Ang, A., Bekaert, G., 2002. Regime switches in interest rates. Journal of Business and Economic Statistics 20, 163–182.

Bekaert, G., Hodrick, R.J., 2001. Expectations hypotheses tests. Journal of Finance 56, 1357–1394.

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

Campbell, J.Y., Ammer, J., 1993. What moves the stock and bond markets? A variance decomposition for long-term asset returns. Journal of Finance 48, 3–37.

Campbell, J.Y., Shiller, R.J., 1991. Yield spreads and interest rate movements: a bird's eye view. Review of Economic Studies 58, 495–514.

Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton.

Chen, R.R., Scott, L., 1993. Maximum likelihood estimation for a multi-factor equilibrium model of the term structure of interest rates. Journal of Fixed Income 3, 14–31.
Christiano, L., Eichenbaum, M., Evans, C., 1996. The effects of monetary policy shocks: evidence from the flow of funds. Review of Economics and Statistics 78, 16–34.
Christiano, L., Eichenbaum, M., Evans, C., 1999. Monetary policy shocks: what have we learned and to what end? In: Woodford, M., Taylor, J. (Eds.), Handbook of Macroeconomics. North Holland, Amsterdam.
Clarida, R., Gali, J., Gertler, M., 2000. Monetary policy rules and macroeconomic stability: evidence and some theory. Quarterly Journal of Economics 65, 147–180.
Cochrane, J.H., 1998. What do the VAR's mean? Measuring the output effects of monetary policy. Journal of Monetary Economics 41, 277–300.
Cochrane, J.H., 2001. Asset Pricing. Princeton University Press, Princeton.
Constantinides, G.M., 1992. A theory of the nominal term structure of interest rates. Review of Financial Studies 5, 531–552.
Dai, Q., Singleton, K., 2000. Specification analysis of affine term structure models. Journal of Finance 55, 1943–1978.
Dai, Q., Singleton, K., 2002. Expectation puzzles, time-varying risk premia, and affine models of the term structure. Journal of Financial Economics 63, 415–441.
Duffee, G.R., 2002. Term premia and interest rate forecasts in affine models. Journal of Finance 57, 405–443.
Duffie, D., Kan, R., 1996. A yield-factor model of interest rates. Mathematical Finance 6, 379–406.
El Karoui, N., Myneni, R., Viswanathan, R., 1992. Arbitrage pricing and hedging of interest rate claims with state variables I: theory. Working paper, Université de Paris VI, Laboratoire de Probabilité.
Estrella, A., Hardouvelis, G.A., 1991. The term structure as a predictor of real economic activity. Journal of Finance 46, 555–576.
Estrella, A., Mishkin, F.S., 1997. The predictive power of the term structure of interest rates in Europe and the United States: implications for the European central bank. European Economic Review 41, 1375–1401.
Evans, C.L., Marshall, D.A., 1998. Monetary policy and the term structure of nominal interest rates: evidence and theory. Carnegie-Rochester Conference Series on Public Policy 49, 53–111.
Evans, C.L., Marshall, D.A., 2001. Economic determinants of the term structure of nominal interest rates. Working paper, Federal Reserve Bank of Chicago, Chicago, IL.
Fisher, M., 1998. A simple model of the failure of the expectations hypothesis. Working paper, Federal Reserve Bank of Atlanta, Atlanta, GE.
Fuhrer, J.C., Moore, G.R., 1995. Monetary policy trade-offs and the correlation between nominal interest rates and real output. American Economic Review 85, 219–239.
Gali, J., 1992. How Well does the ISLM model fit post war data? Quarterly Journal of Economics 107, 709–738.
Hansen, L.P., Jagannathan, R., 1991. Implications of security market data for models of dynamic economies. Journal of Political Economy 99, 225–262.
Harrison, J.M., Kreps, D.M., 1979. Martingales and arbitrage in multiperiod securities markets. Journal of Economic Theory 2, 381–408.
Harvey, C.R., 1988. The real term structure and consumption growth. Journal of Financial Economics 22, 305–333.
Knez, P.K., Litterman, R., Scheinkman, J.A., 1994. Explorations into factors explaining money market returns. Journal of Finance 49, 1861–1882.
Litterman, R., Scheinkman, J., 1991. Common factors affecting bond returns. Journal of Fixed Income 1, 51–61.
Liu, J., 1999. Portfolio selection in stochastic environments. Working paper, UCLA, Los Angeles, CA.
Longstaff, F.A., Schwartz, E.S., 1992. Interest rate volatility and the term structure: a two factor general equilibrium model. Journal of Finance 47, 1252–1282.
Newey, W.K., West, K., 1987. A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Pearson, N., Sun, T.S., 1994. Exploiting the conditional density in estimating the term structure: an application to the Cox, Ingersoll and Ross Model. Journal of Finance 54, 1279–1304.

Piazzesi, M., 2001. An econometric model of the yield curve with macroeconomic jump effects. National Bureau of Economic Research, Working paper no. 8246.

Sargent, T.J., 1979. A note on maximum likelihood estimation of the rational expectations model of the term structure. Journal of Monetary Economics 35, 245–274.

Sims, C., Zha, T., 1995. Does monetary policy cause recessions? Working paper, Princeton University, Princeton, NJ.

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

Uhlig, H., 2001. What are the effects of monetary policy? Results from an agnostic identification procedure. Working paper, Humboldt University, Berlin.

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