Term Structure Forecasting: No-arbitrage Restrictions vs. Information set Expansion

Carlo A. Favero   Linlin Niu   Luca Sala [*]

This version: August 2006

Abstract: This paper addresses the issue of forecasting interest rate term structure by integrating no-arbitrage affine term structure model and macro information. A joint macro-finance model is expected to enhance forecasting performance by efficiently utilizing restrictions from well developed finance theory and information from large data set. We construct a unified framework to investigate the forecasting features of various specifications of macro-finance models.

Several groups of existing discrete-time yield curve models are analyzed in the structural framework along two dimensions, i.e. VAR in unrestricted form versus VAR with no-arbitrage restrictions, small information set (financial series only, or plus limited macro variables) versus large information set. Using the US yield curve data, we find that: 1) No-arbitrage restrictions help to produce balanced forecasting for the yield curve; in particular, yields with shorter maturities (up to 5 years) are explained relatively well at forecasting horizons from 1 to 12 months; 2) Under no-arbitrage restriction, the macro information has limited contribution upon the 3-latent factors model; 3) In unrestricted case where Nelson-Siegel factors are used to provide financial information, macro information contributes to the medium horizon (9-12 months ahead) forecasting; 4) In all cases, two to three financial factors are necessary to deliver precise short term forecasting for the yield curve; 5) The macro common factors do not provide clear evidence of contribution upon the single variables such as unemployment and inflation.

Keywords: Yield curve, term structure of interest rates, forecasting, large data set

JEL Classification: C33, C53, E43, E44

[*] Carlo A. Favero and Luca Sala: IGIER, Bocconi University, Via Salasco 5, 20136 Milan, Italy. E-mail: carlo.favero@unibocconi.it and luca.sala@unibocconi.it. Linlin Niu: Bocconi University, Via Salasco 5, 20136 Milan, Italy. E-mail: linlin.niu@phd.unibocconi.it.

1 Introduction

This paper addresses the issue of forecasting interest rate term structure by integrating no-arbitrage affine term structure model and macro information. A joint macro-finance model is expected to enhance forecasting performance on yield curve by efficiently utilizing restrictions from well developed finance theory and information from large data set. Recent research on some models in this line has shown advantages in both modelling and forecasting. However, little is known about the specific contribution to forecast of these two aspects: how do no-arbitrage restrictions and macro information affect the forecasting performance individually and jointly? What is the optimal strategy to exploit the advantages of both? In this paper we construct a unified framework to investigate the forecasting features of various macro-finance models. The results can be further utilised to develop a better forecasting strategy.

The term structure of interest rates is usually modelled with latent factors using different decomposition methods. It is shown that three latent factors explain most variances of the yield and the in-sample fit is usually good (Nelson and Siegel (1987), Dai and Singleton (2000)). Based on Nelson-Siegel (1987) factor interpolation, Diebold-Li (2005) further demonstrates the out-of-sample forecasting power of this factorization by modelling these factors as AR (1) or VAR(1) processes. The success of this method is attributed to its highly parsimonious setting, as Diebold-Li (2005) calls it the "KISS" principle – "keep it sophisticatedly simple".

On the other hand, we consider the well-developed finance theory as an alternative in maintaining parsimonious and efficient modelling. Since yields with longer maturities can be inferred from risk adjusted expectations of future short rates, which are driven by economic states, yields with different maturities are affined in their parameters. This relationship holds particularly in a relatively efficient market where arbitrage opportunities are ruled out at reasonable time horizon. No-arbitrage restrictions hence serve not only for reducing the parameter dimensions, but also for controlling consistency. Dai and Singleton (2000) and Piazzesi (2003) have surveyed the specification issues of affine term structure models in continuous time and discrete time respectively. Duffee (2002) proved the usefulness of forecasting with essentially affine $A_0(3)$ model in which risk varies independently with interest rate volatility.

Models mentioned above mostly utilize only the information contained in the financial series. However, financial market is by no means isolated with the states of macro economy. The short end of the yield curve is very much affected by monetary policy, while yields with longer maturities are closely related to real economic activities of firms and households. Both monetary authority and market participants make their decisions based on large information set, which may not be fully summarized by the financial time series (Bernanke and Boivin, 2003). Various empirical papers indicate that macroeconomic variables have strong effects on future movements of the yield curve (Ang and Piazzesi (2003), Diebold, Rudebusch, and Aruoba (2005), Rudebusch ad Wu (2003), etc). Ang and Piazzesi (2003) demonstrates that a mixed factor (latent with output and inflation) model performs better than yields only model for one-step ahead quarterly forecast.

Incorporating more macro information in yield curve modelling is likely to improve the in-sample fit as well as forecasting precision. But the gains are limited by the enlargement of parameter space, which leads to higher uncertainty in models and parameters. The recent advances on dynamic factor models provide a potential tool for tackling this dilemma. As shown in Stock and Watson (2002), Forni, Hallin, Lippi, and Reichlin (2003), by decomposing common and idiosyncratic components from a large panel of cross-sectional time series, the dimensionality problem is sufficiently reduced and forecasting efficiency is improved. Giannone et al (2004) shows that a two factor model produces forecasting accuracy of the federal funds rate similar to that of the market; Moench (2005) constructs a No-arbitrage Factor-Augmented VAR and finds better forecast for 6 to 12 month horizon compared to yields only model. We will examine the usefulness of this method and discuss the various issues in combining macro information with financial information.

In a unified state-space framework, we study the effects of incorporating common factor information and no-arbitrage restrictions on the forecasting performance of yield curve. Several groups of existing discrete-time yield curve models are analyzed in a structural framework along two dimensions, i.e. VAR in unrestricted form (Diebold-Li) versus VAR with no-arbitrage restrictions, small information set (financial series only, or plus limited macro variables) versus large information set. Hence, for each group of model, given number of state variables, four specifications emerge from combinations along these two dimensions. Among these models, special cases are Diebold-Li (2005), Dai-Singleton (2002) $A_0(3)$ three latent factor models, Ang-Piazzesi (2003) mixed factor models, and VAR of Taylor Rule variables, etc. Some of these special cases have been proved to be useful tools in modelling and forecasting yield curve. This paper aims at building a linkage among existing models, exploring the blank areas left between, and assisting deep understanding of optimal integration of information and restriction.

From this framework, we want to answer several questions.

Q1: Given a set of financial factors extracted from yield curve, to which extent the no-arbitrage restrictions affect the forecasting performance? A three-financial-factor model serves as the starting point. For comparison, we will take the Diebold-Li (2005) model with three Nelson-Siegel factors as a representative of the unrestricted case; then we construct a three latent factor model under no-arbitrage restriction, which can be regarded as a simple case of Dai-Singleton (2000) essentially affine $A_0(3)$ model.

Q2: How can macro information be effectively modelled jointly in the state equation and hence to improve out-of-sample forecasting? What are the gains by integrating large information set in addition to the financial factors? A comparison is made between augmenting macro common factors to the state equation and augmenting macro variables (e.g. proxies for output and inflation) to the state equation. We compare the difference both in unrestricted and restricted models.

And ***Q3***: Given number of state variables and estimated parameters, what are the optimal combination of financial factors and macro information? This trade-off is essential under the requirement of parsimony to forecasting models. Both financial and macro factors are informative for the yield curve dynamics, but the former is more indicative for short run dynamics and the latter for business cycle horizons. So change of the two ingredients in the state equation will affect the forecast performance on different parts of the yield curve along forecast horizons. Studying this pattern gives useful insight in tailoring the choice of states for specific forecast purpose.

By forecasting the US yield curve with various models in this framework, we find that: 1) No-arbitrage restrictions help to produce balanced forecasting for the yield curve; in particular, yields with shorter maturities (up to 5 years) are explained relatively well at forecasting horizons from 1 to 12 months; 2) Under no-arbitrage restrictions, the macro information has limited contribution upon the 3-latent factors model; 3) In unrestricted case where Nelson-Siegel factors are used to provide financial information, macro information contributes to the medium horizon (around 9-12 months ahead) forecasting; 4) In all cases, two to three financial factors are necessary to deliver precise short term forecasting of the yield curve dynamics; 5) The macro common factors do not provide clear evidence of contribution upon the single variables such as unemployment and inflation.

In the following, we will introduce the modelling framework in section 2. Data and econometric methodology are described in section 3. Section 4 reports results, and section 5 concludes.

2 Model

We study the dynamics of yields (y_t) in a state-space form, where yields are determined by a vector of variables X_t, endogenous or exogenous, observable or latent, and X_t follows a VAR process.

$$y_t = a + bX_t + \varepsilon_t,\ \varepsilon_t \text{ is i.i.d. N(0, } \sigma^2 \cdot I_n\text{)}.$$

$$X_t = \mu + \Phi X_{t-1} + v_t,\ v_t \text{ is i.i.d. N(0, } \Omega\text{)}.$$

In order to address the issue of integrating dynamic macro factors with no-arbitrage affine term structure model, we study the forecasting performance of models in this structural framework.

Several groups of existing discrete-time yield curve models are analyzed along two dimensions, i.e. VAR in unrestricted form (Diebold Li model with Nelson-Siegel factors as states) versus VAR with no-arbitrage restrictions, small information set (X_t includes financial series only, or plus limited macro variables) versus large information set (X_t includes financial series plus macro common factors) .

We specify the framework as following (Specifications are summarized in Table 1.):

1. With unrestricted reduced form VAR, we allow for arbitrage opportunities so that yields are modelled separately from each other. The coefficients in the *a, b* matrix for yields with different maturities are not restricted to rule out arbitrage restriction. The Diebold-Li (2005) model serves as a parsimonious representative of models in this group.

2. By No-arbitrage VAR, we mean that long yields are risk-adjusted expectations of average future short-rate, so that the coefficients of the state-space model are restricted in a certain form. Take measurement equation of the short rate as $r_t = \delta_0 + \delta_1^T X_t$, market price for risk associated with the state variable as $\Lambda_t = \lambda_0 + \lambda_1 X_t$, yield with maturity n as $y_{t,n} = a_n + b_n^{'} X_t = \frac{-1}{n}\left(A_n + B_n^{'} X_t\right)$ where $A_1 = -\delta_0$ and $B_1 = -\delta_1$, then the no-arbitrage restriction imposes the following structure on the coefficients of measurement equation (for n>1):

$$A_{n+1} = A_n + B_n'\left(\mu - \Omega\lambda_0\right) + \frac{1}{2}B_n'\Omega B_n - \delta_0$$

$$B_{n+1}' = B_n'\left(\Phi - \Omega\lambda_1\right) - \delta_1', \text{ equivalently, } B_{n+1} = \left(\Phi' - \lambda_1'\Omega\right)B_n - \delta_1$$

In terms of the direct coefficients on yields, the restrictions imply that once the coefficients on short rate equation are fixed, all other coefficients for longer maturity yields are determined by the parameters in the state equation and the risk pricing equation (Appendix 1).

$$b_{n+1} = \frac{1}{(n+1)}\left[\sum_{i=0}^{n}\left(\Phi' - \lambda_1'\Omega\right)^i\right]b_1 = \frac{1}{(n+1)}\left[\sum_{i=0}^{n}\left(\Phi' - \lambda_1'\Omega\right)^i\right]\delta_1$$

$$a_{n+1} = a_1 - \frac{1}{(n+1)}\sum_{i=1}^{n} B^{(i)}$$

$$B^{(i)} = B_i'\left(\mu - \Omega\lambda_0\right) + \frac{1}{2}B_i'\Omega B_i .$$

3. By small information set (small N), we mean that the state variables X_t are confined with financial time series, i.e. the term structure per se, plus at most a small set of macro variables such as inflation, output, etc.

Given a set of financial factors, we will augment macro variables to the state variables.

Firstly, with the three extracted financial factors either from Diebold-Li (2005) or essentially affine latent factor model A0(3), we can augment output and inflation as additional macro state variables.

Second, we use directly a set of Taylor Rule variables as the state vector, in which the only financial factor is either the short term interest rate in the unrestricted case, or it is extracted from the yield curve. As macro economic theory indicates, the short rate is a monetary policy tool which responds to the macro economic status. In a Taylor Rule setting, the short rate reacts to output gap and inflation. We use unemployment rate as a proxy of output gap. The state vector is written as: $X_t = [\ ffr_t \quad output_t \quad infl_t\ ]$.

4. By large information set (large N), we mean that a large set (in our case N = 171) of macroeconomic variables are exploited and common factors extracted from this set are used as the state variables. We extract the common factors using standard static principal components, as is suggested by Stock and Watson (2002). The first four common factors explain up to 50% of all variances in the macro panel. The first two are mostly correlated with the concepts of output and inflation (Table 2).

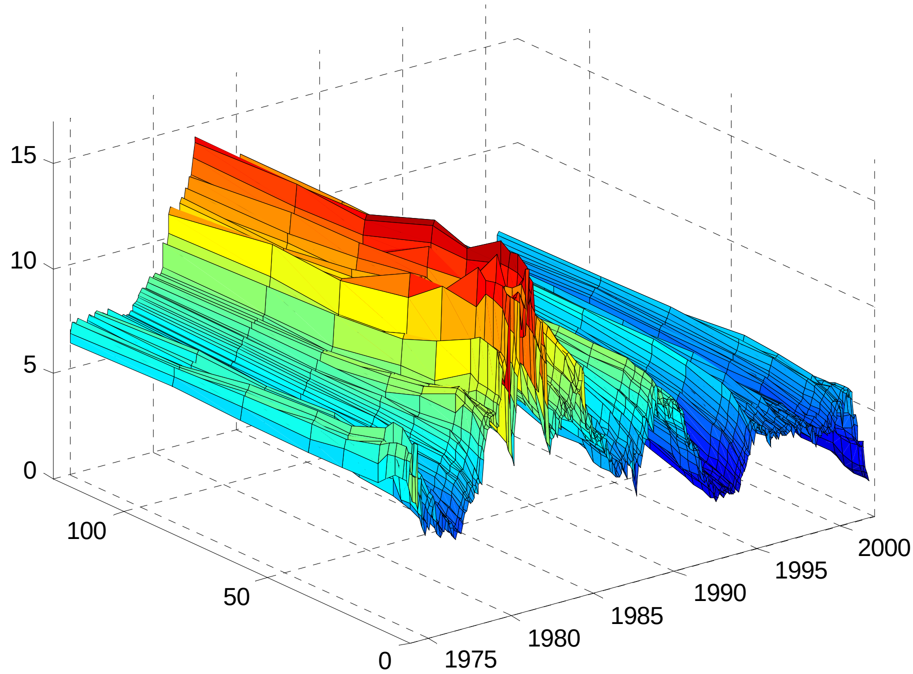
We will augment the first or first two common factors with the three extracted financial factors either from Diebold-Li (2005) or A0(3) essentially affine latent factor model. In comparison with the small information set case, we can see whether the common factors proxy the concept of output and inflation well and whether this delivers better forecasting.

Then, corresponding to the Taylor Rule where only limited macro variables are modelled, we introduce a general interest rate rule to the state equation. As in setting the short rate, central banks actually base their decisions on large macro economic information set (Bernanke and Boivin (2003)). The state vector is written as $X_t = [\ ffr_t \quad mf_{1t} \ldots mf_{nt}\ ]$. In previous work (Moench (2005)), it is shown to be a better forecasting tool with respect to several popular models. We will introduce two to four macro factors in this setting, to see how the augmentation of macro factors is likely to contribute to the forecasting performance.

3 Data and Econometric Methodology

Our basic data set consists of a set of zero-coupon equivalent US yields (1974:02-2001:12), provided by Brousseau, V. and B. Sahel (1999). We consider zero-coupon equivalent yields for US data measured at the following 11 maturities: 1-

month, 2-month, 3-month, 6-month, 9-month, 1-year, 2-year, 3-year, 5-year, 7-year, and 10-year.



The macro panel which contains 171 series of US macro time series (1974:2-2002:12) is the same as used in Giannone, Reichlin and Sala (2004). We transform the data set in order to obtain stationarity (Stock and Watson (2002)). Then macro common factors are extracted by applying standard principal component method. The first 4 common factors are used in our analysis. We rank the factors according to their explanation power to the whole macro panel, instead of their correlation with the yield curve. This is different from Ng and Ludvigson (2006), where they construct a composite factor by combining several common factors according to their in sample significance in explaining the bond risk premia. However, by running experiments on ranking the factors according to their contribution to R-squares of yields, we didn't find clear pattern that such a combination may improve out-of-sample forecasting. This may be due to two reasons: 1) the explaining power of factor to yields changes over time, to single yields as well as to the whole yield curve. Hence no gains could be made by sticking to a set of pre-selected factors according to their in-sample behaviour, and little gains from tracing the most correlated factors with yields either. 2) There is a trade off between using more significantly correlated factor (which could be an unimportant factor with respect to the macro panel, hence small in magnitude) and more significantly estimated factors (which could be less correlated with yields, but more robustly estimated). So we augment factors according to their original usefulness in explaining the variances in the macro panel.

When we estimate yields with macro variables in the state equation, we particularly pick up unemployment as a proxy for output, and annual change of CPI as inflation.

In order to maintain parsimony, we first start modelling all state equations as VAR(1) process. In the case of three latent factors, we model all state variables as AR(1) processes assuming all latent factors are orthogonal, as a comparison with Diebold-Li AR(1) model. In order to make the estimates comparable, we set observation involved in all estimation fixed as 212 periods. By moving the window forward, we make rolling estimation. In the restricted estimation, instead of time-consuming initial condition searching, we rather use a training period of 46 months for the algorithm to stabilize. The first period's initial values on risk loadings are set to zeros as in the risk neutral case. All subsequent initial values of parameters are set to the previous maximization results, but if this period's likelihood is much worse than last period, i.e. exceeding certain criteria, we make additional search for desirable likelihood value. In the end, we make multiple step ahead forecast (1 month, 3 months, 6 months, 9 months, 12 months, 18 months, 24 months) for a total of 74 periods (1995:11-2001:12).

Forecasts for different specifications are computed as follows:

A. Unrestricted models.

1. Diebold-Li (2005)

The Diebold-Li (2005) model is a parsimonious modelling of the yield curve with Nelson-Siegel (1987) factor interpolation. At each point of time, the yield curve is fitted with three factors (Level, Slope, and Curvature) from a constant plus Laguerre function.

$$y_t^{(n)} = \beta_{1,t} + \beta_{2,t} \cdot \left(\frac{1-e^{-\lambda n}}{\lambda n} \right) + \beta_{3,t} \cdot \left(\frac{1-e^{-\lambda n}}{\lambda n} - e^{-\lambda n} \right)$$

In Diebold-Li model, the factors, once extracted, are modelled as either univariate AR(1) or VAR(1) process. The yield forecasts are then based on this underlying factor specification. We use Diebold-Li AR(1) three factor model as a benchmark for comparison for the restricted case.

With this model, we follow two step procedures for the regression: *(i),* extract Nelson-Siegel factors by OLS; *(ii),* estimate the factor dynamics with AR(1). For forecasting beyond one month, $\hat{c}$, $\hat{\Gamma}$, are obtained by regressing the factors onto a constant and their h-month lag.

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

where $\hat{\beta}_t = \lfloor \hat{\beta}_{1,t} \quad \hat{\beta}_{2,t} \quad \hat{\beta}_{3,t} \rfloor'$, and $\hat{\Gamma}$ is diagonal for AR(1) process. Then the forecasting of yields is carried out with the same parameterization:

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

2. Diebold-Li (2005) financial factors with macro variables or factors in unrestricted VAR form

The Nelson-Siegel factors are extracted as before, then these factors together with macro variables or factors are modeled as a VAR(1). The state vector is: X_t = [$\hat{\beta}_t$; z_t], where z contains the macro information. For forecasting beyond one month, state equation coefficients are obtained by regressing the factors onto a constant and their h-month lag.

$$\hat{y}_{t+h|t} = \hat{a} + \hat{b}\hat{X}_{t+h|t}$$
$$\hat{X}_{t+h|t} = \hat{\mu} + \hat{\Phi} X_t$$

3. Interest rate rule type VAR as the state equation in the unrestricted form.

In this setting, the only one financial factor is extracted in the state vector, the forecasting procedure is the same as the three financial factor model

B. Models with no-arbitrage restriction.

The difference in forecasting with no-arbitrage restriction is not only that parameters on the yield equations are restricted according to the recursive difference equation, but the VAR for the state equation is also restricted, in the sense that, for forecasting beyond one month, state equation coefficients are obtained by recursively multiplying forward the VAR coefficients for one-month ahead forecasting.

One month ahead: $\hat{X}_{t+1|t} = \hat{\mu} + \hat{\Phi} X_t$

h-month ahead: $\hat{X}_{t+h|t} = \sum_{i=0}^{h} \hat{\Phi}^i \hat{\mu} + \hat{\Phi}^h X_t$

4 Results (incomplete)

In the stated results, we make the following specifications:

1) Measure of forecasting performance is the ratio of forecasting root mean square error (RMSE) from specified models with respect to the forecasting RMSE from Random Walk model. In the Random Walk model, we assume no change in yields.

2) Forecasting is made for the following horizons (denoted by *h*): 1 month, 3 months, 6 months, 9 months, 12 months, 18 months, and 24 months.
3) For each forecasting horizon, forecasting is conducted for the period of 1995:11 - 2001:12, in total, 74 periods. For example, yield values at 1995:11 are compared with their one-month ahead forecast made at period 1995:10, three-month ahead forecast made at 1995:8, etc., until 24-month ahead forecast made at 1993:11.
4) Better forecasts with respect to Random Walk are highlighted with colours: blue and bold characters for ratios in the range of [0.9, 1); orange, bold and italic for ratios within [0.8, 0,9); pink, bold and underlined for ratios within (0, 0.8).

From Table 3 to Table 6, we show comparisons of forecasting results from various groups of models.

First, three financial factor models are estimated. In the unrestricted estimation, we follow Diebold-Li (2005) and model the Nelson-Siegel factors as AR(1) processes. In the restricted one, we first estimate all three latent factors by using Kalman Filter. Then we follow Ang-Piazzessi (2003), use the method proposed by Chen-Scott (1993). The latter is convenient because it is faster.

Table 3: three financial factors only, unrestricted versus restricted.

Then, three financial factors are augmented with small macro information – output and inflation in the small N case, and large macro information – macro common factors in the large N case. Given number of state variables, the following comparisons are made along two dimensions:

	Unrestricted	Restricted
Small N	3financial factors +output, inflation	3financial factors +output, inflation
Large N	3financial factors + macro factors	3financial factors + macro factors

Table 4: three financial factors + 1 macro variable/factor models.

Table 5: three financial factors + 2 macro variable/factor models.

> There are other cases where one or two financial factors are augmented with up to two macro variables and four macro factors. The above analytical framework is reserved. So we can compare how the weights of financial factors and macro factor matters in the forecasting precision. Table 6 shows the unrestricted case of 5 state variables. And the restricted results show similar pattern.

The results indicate that:

1) Compared with Dai-Singleton three-factor $A_0(3)$ model estimated by Kalman Filter, the Chen-Scott approach provides competitive results (Table 3). Due to the efficiency of Chen-Scott approach in estimation, we further explore our interested issues by this method.

2) No-arbitrage restriction helps to produce balanced forecasting for the yield curve; in particular, yields with shorter maturities (up to 5 years) are explained relatively well at forecasting horizons from 1 to 12 months. We present here a simplest $A_0(3)$ case where risks are assumed to be time invariant with respect to the latent states. This case actually forecast better than that with time varying risk.

3) With no-arbitrage restrictions, the macro information has limited contribution upon the 3-latent factors model.

4) In unrestricted case where Nelson-Siegel factors are used to provide financial information, macro information contributes to the medium horizon forecasting (9-12 months ahead).

5) In all cases, more financial factors (2 or 3) are necessary to deliver precise short term forecasting for the yield curve. (Table 6).

6) The macro common factors do not provide clear evidence of contribution upon the single variables like unemployment and inflation.

Further results need to be shown: forecast from models of two financial factors…

5 Conclusions

In this paper we address the issue of integrating dynamic macro factors with no-arbitrage affine term structure model. A joint macro-finance model is expected to enhance forecasting performance by efficiently utilizing information from large data set and restrictions from well developed finance theory. However, we need to decompose the specific contribution to forecast of these two aspects: how do no-arbitrage restrictions and macro information affect the forecasting performance individually and jointly? What is the optimal strategy to exploit the advantages of both? In our unified framework, we can investigate the forecasting features of various macro-finance models.

By forecasting the US yield curve with various models in this framework, we found with available results that: 1) No-arbitrage restriction helps to produce balanced forecasting for the yield curve; in particular, yields with shorter maturities (up to 5 years) are explained relatively well at forecasting horizons from 1 to 12 months. 2) With no-arbitrage restriction, the macro information has limited contribution upon the 3-latent factors model; 3) In unrestricted case where Nelson-Siegel factors are used to provide financial information, macro information contributes to the medium horizon forecasting (9-12 months ahead); 4) In all cases, two or three financial factors are necessary to deliver precise short term forecasting for the yield curve; 5) The macro common factors do not provide clear evidence of contribution upon the single variables like unemployment and inflation.

The results can be further utilised to develop a better forecasting strategy. It seems that no single best model exists to provide precise forecasting in wide time horizons for the whole curve. Instead, different combinations have their own merits and shortcomings. How to combine these features of various specifications to deliver better forecast remains to be an interesting research task for the future.

References:

1. Ang, A. and M. Piazzesi, 2003, “A No-Arbitrage Vectorautoregression of Term Structure Dynamics with Macroeconomic and Latent Variables.” *Journal of Monetary Economics* 50, 4, 745-787.

2. Ang, A., M. Piazzesi, and M. Wei ,2004, "What does the Yield Curve Tell us about GDP Growth?" *Journal of Econometrics 131*, 359-403.

3. Bernanke, Ben S., and Jean Boivin, 2003, “Monetary Policy in a Data-Rich Environment.” *Journal of Monetary Eocnomics*,

4. Brousseau, V. and B. Sahel, 1999, “The advantages of using bootstrapped yield curves for monetary policy analyses.” Paper presented at the *ECB Workshop on yield curve modelling*, Frankfurt.

5. Carriero, A., Favero C., Kaminska,I., 2005, “Financial Factors, Macroeconomic Information and the Expectation Theory of the Term Structure of the Interest Rates.” *Journal of Econometrics 131*, 339-358.

6. Chen, R. and L. Scott, 1993, “Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates.” Journal of Fixed Income 3, 14-31.

7. Dai, Q. and K. Singleton, 2000, “Specification Analysis of Affine Term Structure Models.” *Journal of Finance* 55, 1943-78.

8. Dai, Q. and K. Singleton, 2002, “Expectation Puzzles, Time-varying Risk Premia, and Affine Models of the Term Structure.” *Journal of Financial Economics* 63, 415-41.

9. Diebold, F.X. and Li., C. ,2005, "Forecasting the Term Structure of Government Bond Yields." *Journal of Econometrics*, forthcoming.

10. Diebold, F.X., Glenn D. Rudebusch, and S. Boragan Aruoba, 2006, “The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach.” *Journal of Econometrics 131*, 309-338.

11. Duffee, G.R., 2002, “Term premia and interest rate forecasts in affine models.” *Journal of Finance*, 57, 405-443.

12. Duffie, D. and R. Kan, 1996, “A Yield-Factor Model of Interest Rates.” *Mathematical Finance* 6, 379-406.

13. Forni, M., M. Hallin, M. Lippi, and L. Reichlin, 2003, “The Generalized Dynamic Factor Model. One-Sided Estimation and Forecasting.” *LEM Working Paper Series*.

14. Giannone, Dominico, Lucrezia Reichlin, and Luca Sala, 2004, “Monetary Policy in Real-Time.” *NBER Macroeconomics Annual*.

15. Hördahl, P., O. Tristani and D. Vestin, 2004, “A Joint Econometric Model of Macroeconomic and Term Structure Dynamics.” *Working Paper, European Central Bank*.

16. Moench, Emanuel, 2005, "Forecasting the Yield Curve in a Data-Rich Environment: A No-Arbitrage Factor-Augumented VAR Approach." *ECB Working Paper No. 544*.

17. Nelson and Siegel, 1987, "Parsimonious modelling of yield curves." *Journal of Business*, 60, 473-89

18. Ng, Serena, and S. Ludvigson, 2006, "Macro Factors in Bond Risk Primia." Working Paper, New York University.

19. Piazzesi, Monika, 2003, "Affine Term Structure Models." prepared for *Handbook of Financial Econometrics*, March 2003.

20. Stock, James H., and Mark W. Watson, 2002, "Macroeconomic Forecasting Using Diffusion Indexes." *Journal of Business and Economic Statistics* 20, 147-162.

Appendix 1. No-Arbitrage Restriction on Bond Pricing Parameters

1. State variable dynamics

Transition equation for X_t follows VAR(1):

$$X_t = \mu + \Phi X_{t-1} + v_t$$

v_t is i.i.d. N(0, Ω).

2. Short rate equation:

$$r_t = \delta_0 + \delta_1^T X_t$$

δ_0: a scalar.

δ_1: K × 1 vector.

3. Time-varying prices of risk (associated with the sources of uncertainty v_t)

$$\Lambda_t = \lambda_0 + \lambda_1 X_t$$

Λ_t: K × 1 vector.

λ_0: K × 1 vector.

λ_1: K × K matrix.

If investors are risk-neutral, λ_0 =0 and λ_1 =0, hence Λ_t =0, no risk adjustment.

4. Pricing kernel

No arbitrage opportunity between bonds with different maturities implies that there is a discount factor m linking the price of yield of maturity n this month with the yield of maturity n-1 next month.

$$P_t^{(n)} = E_t\left[m_{t+1} P_{t+1}^{(n-1)}\right]$$

The stochastic discount factor is related to the short rate and risk perceived by the market,

$$m_{t+1} = \exp\left(-r_t - \frac{1}{2}\Lambda_t^T \Omega \Lambda_t - \Lambda_t^T v_{t+1}\right)$$

No-arbitrage recursive relation can be derived from the above equations as:

$$P_t^{(n)} = E_t\left[m_{t+1}P_{t+1}^{(n-1)}\right] = E_t\left[m_{t+1}m_{t+2}P_{t+2}^{(n-2)}\right] = E_t\left[m_{t+1}m_{t+2}...m_{t+n}P_{t+n}^{(0)}\right]$$
$$= E_t\left[m_{t+1}m_{t+2}...m_{t+n}\cdot 1\right]$$

$$= E_t\left[\exp\left(-\sum_{i=0}^{n-1}\left(r_{t+i} + \frac{1}{2}\Lambda_{t+i}{}^T\Omega\Lambda_{t+i} + \Lambda_{t+i}{}^T v_{t+1+i}\right)\right)\right]$$

$$= E_t\ \left[\exp\left(A_n + B_n^{'}X_t\right)\right] = E_t\ \left[\exp\left(-ny_{t,n}\right)\right]$$

$$= E_t^{\,Q}\left[\exp\left(-\sum_{i=0}^{n-1}\left(r_{t+i}\right)\right)\right]$$

$E_t^{\,Q}$ denotes the expectation under the risk-neutral probability measure, under which the dynamics of the state vector X_t are characterized by the risk-neutral constant and autocorrelation matrix:

$$\mu^Q = \mu - \Omega\lambda_0$$

$$\Phi^Q = \Phi - \Omega\lambda_1$$

Affine functions of the state variables for yields are:

$$p_{t,n} = A_n + B_n^{'}X_t$$

$$y_{t,n} = a_n + b_n^{'}X_t = \frac{-1}{n}\left(A_n + B_n^{'}X_t\right)$$

where the coefficients A_n and B_n follow the difference equation:

$$A_{n+1} = A_n + B_n{}'\left(\mu - \Omega\lambda_0\right) + \frac{1}{2}B_n{}'\Omega B_n - \delta_0$$

$$B_{n+1}{}' = B_n{}'\left(\Phi - \Omega\lambda_1\right) - \delta_1{}'$$

with $a_1 = \delta_0$ and $b_1 = \delta_1$.

These can be derived from the pricing kernel equation.

$$P_t^{(n+1)} = E_t\left[m_{t+1}P_{t+1}^{(n)}\right]$$

$$= E_t\left[\exp\left\{-r_t - \frac{1}{2}\Lambda_t'\Omega\Lambda_t - \Lambda_t'v_{t+1}\right\}\exp\left\{A_n + B_n'X_{t+1}\right\}\right]$$

$$= \exp\left\{-r_t - \frac{1}{2}\Lambda_t'\Omega\Lambda_t + A_n\right\}E_t\left[\exp\left\{-\Lambda_t'v_{t+1} + B_n'X_{t+1}\right\}\right]$$

$$= \exp\left\{-\delta_0 - \delta_1'X_t - \frac{1}{2}\Lambda_t'\Omega\Lambda_t + A_n\right\}E_t\left[\exp\left\{-\Lambda_t'v_{t+1} + B_n'\left(\mu + \Phi X_t + v_{t+1}\right)\right\}\right]$$

$$= \exp\left\{-\delta_0 - \delta_1'X_t - \frac{1}{2}\Lambda_t'\Omega\Lambda_t + A_n + B_n'\left(\mu + \Phi X_t\right)\right\}E_t\left[\exp\left\{-\Lambda_t'v_{t+1} + B_n'v_{t+1}\right\}\right]$$

$$= \exp\left\{-\delta_0 + A_n + B_n'\mu + \left(B_n'\Phi - \delta_1'\right)X_t - \frac{1}{2}\Lambda_t'\Omega\Lambda_t\right\}E_t\left[\exp\left\{\left(-\Lambda_t' + B_n'\right)v_{t+1}\right\}\right]$$

$$= \exp\left\{-\delta_0 + A_n + B_n'\mu + \left(B_n'\Phi - \delta_1'\right)X_t - \frac{1}{2}\Lambda_t'\Omega\Lambda_t\right\}\exp\left\{E_t\left[\left(-\Lambda_t' + B_n'\right)v_{t+1}\right] + \frac{1}{2}\mathrm{var}\left[\left(-\Lambda_t' + B_n'\right)\right.\right.$$

$$= \exp\left\{-\delta_0 + A_n + B_n'\mu + \left(B_n'\Phi - \delta_1'\right)X_t - \frac{1}{2}\Lambda_t'\Omega\Lambda_t\right\}\exp\left\{\frac{1}{2}\mathrm{var}\left[\left(-\Lambda_t' + B_n'\right)v_{t+1}\right]\right\}$$

$$= \exp\left\{-\delta_0 + A_n + B_n'\mu + \left(B_n'\Phi - \delta_1'\right)X_t - \frac{1}{2}\Lambda_t'\Omega\Lambda_t\right\}\exp\left\{\frac{1}{2}E_t\left[\left(-\Lambda_t' + B_n'\right)v_{t+1}v_{t+1}'\left(-\Lambda_t + B_n\right)\right]\right\}$$

$$= \exp\left\{-\delta_0 + A_n + B_n'\mu + \left(B_n'\Phi - \delta_1'\right)X_t - \frac{1}{2}\Lambda_t'\Omega\Lambda_t\right\}\exp\left\{\frac{1}{2}\left[\Lambda_t'\Omega\Lambda_t - 2B_n'\Omega\Lambda_t + B_n'\Omega B_n\right]\right\}$$

$$= \exp\left\{ -\delta_0 + A_n + B_n'\mu + \left(B_n'\Phi - \delta_1'\right)X_t - \frac{1}{2}\Lambda_t'\Omega\Lambda_t + \frac{1}{2}\Lambda_t'\Omega\Lambda_t - B_n'\Omega\Lambda_t + \frac{1}{2}B_n'\Omega B_n \right\}$$

$$= \exp\left\{ -\delta_0 + A_n + B_n'\mu + \left(B_n'\Phi - \delta_1'\right)X_t - B_n'\Omega\Lambda_t + \frac{1}{2}B_n'\Omega B_n \right\}$$

$$= \exp\left\{ -\delta_0 + A_n + B_n'\mu + \left(B_n'\Phi - \delta_1'\right)X_t - B_n'\Omega\left(\lambda_0 + \lambda_1 X_t\right) + \frac{1}{2}B_n'\Omega B_n \right\}$$

$$= \exp\left\{ -\delta_0 + A_n + B_n'\left(\mu - \Omega\lambda_0\right) + \frac{1}{2}B_n'\Omega B_n + \left(B_n'\Phi - B_n'\Omega\lambda_1 - \delta_1'\right)X_t \right\}$$

$$= \exp\left\{ \left[-\delta_0 + A_n + B_n'\left(\mu - \Omega\lambda_0\right) + \frac{1}{2}B_n'\Omega B_n \right] + \left[B_n'\left(\Phi - \Omega\lambda_1\right) - \delta_1'\right]X_t \right\}$$

5. No-arbitrage coefficients in alternative presentation:

In order to understand intuitively how these restrictions are imposed directly on the coefficients a_n, b_n in the yield equation, we can write them in the following affined form,

Given that

$$p_{t,n} = A_n + B_n^{'} X_t$$

$$y_{t,n} = a_n + b_n^{'} X_t = \frac{-1}{n}\left(A_n + B_n^{'} X_t\right)$$

$$b_{n+1} = \frac{1}{(n+1)}\left[\sum_{i=0}^{n}\left(\Phi' - \lambda_1'\Omega\right)^i\right] b_1$$

$$= \frac{1}{(n+1)}\left[\sum_{i=0}^{n}\left(\Phi' - \lambda_1'\Omega\right)^i\right]\delta_1;$$

$$a_{n+1} = a_1 - \frac{1}{(n+1)}\sum_{i=1}^{n} B^{(i)}$$

where $B^{(i)} = B_i'\left(\mu - \Omega\lambda_0\right) + \frac{1}{2} B_i'\Omega B_i$

Appendix 2. The likelihood function with latent variables

1. Latent factors only, extracted with Kalman Filter.

This method is straightforward, but computationally costly. With Dai-Singleton $A_0(3)$ model (three factors, but no factor driving the conditional variances of yields), the Kalman Filter is illustrated in the following

Model:

$$y_t = A + BX_t + e_t, \quad e_t \text{~i.i.d. N}(0, \sigma^2 \cdot I_n)$$
$$X_t = \Phi X_{t-1} + D\varepsilon_t, \quad \varepsilon_t \text{~i.i.d. N}(0, \Omega_x)$$

1) Initial values (at steady state): $X_{0|0}$, $\Sigma_{0|0}$

   (1) $X_{0|0}$

   For specification where all factors are assumed to be mean zero, $X_{0|0} = 0$;

   (2) $\Sigma_{0|0}$. Unconditional covariance matrix of stationary X_t is described as:

   Cov(X_t) = $\Phi \cdot Cov(X_t) \cdot \Phi' + DCov(\varepsilon_t)D'$

   $\Sigma_{0|0} = \Phi \cdot \Sigma_{0|0} \cdot \Phi' + D\Omega_x D'$

   $vec(\Sigma_{0|0}) = vec(\Phi \cdot \Sigma_{0|0} \cdot \Phi') + vec(D\Omega_x D')$

   $vec(\Sigma_{0|0}) = (\Phi \otimes \Phi')vec(\Sigma_{0|0}) + vec(D\Omega_x D')$

   $vec(\Sigma_{0|0}) = \left(I - (\Phi \otimes \Phi')\right)^{-1} vec(D\Omega_x D')$

2) Prediction.

   For t = 1, …, T:

   $X_{t|t-1} = \Phi X_{t-1|t-1}$

$$\Sigma_{t|t-1} = \Phi\Sigma_{t-1|t-1}\Phi' + D\Omega_x D'$$
$$\eta_{t|t-1} = y_t - y_{t|t-1} = y_t - A - BX_{t|t-1}$$
$$f_{t|t-1} = B\Sigma_{t|t-1}B' + \sigma^2 I$$

Likelihood function:

$$l(\Theta) = l(\Theta) - \frac{1}{2}\ln\left((2\pi)^n \left| f_{t|t-1} \right|\right) - \frac{1}{2}\eta'_{t|t-1} f_{t|t-1}^{-1}\eta_{t|t-1}$$
$$= l(\Theta) - \frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln\left| f_{t|t-1} \right| - \frac{1}{2}\eta'_{t|t-1} f_{t|t-1}^{-1}\eta_{t|t-1}$$

3) Updating.

Kalman Gains: $K_t = \Sigma_{t|t-1}B\ f_{t|t-1}{}^{-1}$

$$X_{t|t} = X_{t|t-1} + K_t\eta_{t|t-1}$$
$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t B\Sigma_{t|t-1}$$

4) Likelihood function.

By repeating 2-3, the likelihood function for the regression is obtained as:

$$l(\Theta) = -\frac{1}{2}\sum\ln\left((2\pi)^n \left| f_{t|t-1} \right|\right) - \frac{1}{2}\sum\eta'_{t|t-1} f_{t|t-1}^{-1}\eta_{t|t-1}$$
$$= -\frac{nT}{2}\ln(2\pi) - \frac{1}{2}\sum\ln\left| f_{t|t-1} \right| - \frac{1}{2}\sum\eta'_{t|t-1} f_{t|t-1}^{-1}\eta_{t|t-1}$$

2. Latent factors only or mixed with observable factors, with Chen-Scott (1993) method

(This part follows Ang, A., Piazzesi, M., 2003, which applies the method proposed by Chen and Scott (1993). Details are discussed for the extraction of unobservable variables.)

In order to extract one unobserved factor f_t^u, we need one yield measured without error. (Likewise, if there are K_2 unobserved factors, we will need K_2 yields measured without error.) Then the observed factors f_t^u will be solved from the yields and the observed variables which includes observed macro variables f_t^o and possibly lagged terms of both observed latent variables. $X_t^o = \left[f_t^o \quad f_t^u \right]$.

When number of yields exceeds number of unobserved factors, following Chen and Scott (1993), and others, we assume that some of the yields are observed with measurement error. Then K_2 yields are set to infer the latent factors, N - K_2 yields are measured with error. Denote these yields by $Y_t^E = (y_{t,n_{K_2+1}}, ..., y_{t,n_N})'$.

Yields measured with errors:

Assume that the measurement error is IID, and uncorrelated across yields measured with error. Let $b^{E,m}$ denote a $(N-K_2)\times(N-K_2)$ measurement matrix (identity matrix) and u_t^m be an (N - K_2)-dimensional Gaussian white noise with a diagonal covariance matrix independent of X_t. For the yields measured with error, we can write:

$$Y_t^E = a^E + b^{E,o} X_t^o + b^{E,u} X_t^u + b^{E,m} u_t^m$$

Yields measured without errors:

Denote yield(s) measured without error by $Y_t^{NE} = (y_{t,n_1}, ..., y_{t,n_{K_2}})'$,

$$Y_t^{NE} = a^{NE} + b^{NE} X_t$$

where a^{NE} is $K_2 \times 1$ and b^{NE} is $K_2 \times K$. Partition the matrix b^{NE} into $b^{NE} = \left[b^{NE,o} \; b^{NE,u} \right]$ where $b^{NE,o}$ is $K_2 \times (K - K_2)$ matrix which contains the loadings on observable factors and $b^{NE,u}$ is a $K_2 \times K_2$ invertible matrix that picks up the unobservable factor(s).

$$Y_t^{NE} = a^{NE} + b^{NE,o} X_t^o + b^{NE,u} X_t^u$$

Latent factors:

Then the unobservable factors in $X_t^u \equiv f_t^u$ can be inferred from Y_t^{NE} and the pricing coefficient matrices a^{NE} and b^{NE} using the inversion of the equation:

$$f_t^u \equiv X_t^u = (b^{NE,u})^{-1}\left[Y_t^{NE} - a^{NE} - b^{NE,o} X_t^o\right]$$

Measurement equation:

With N yields, the system can be written as a compact equation. 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.

$$Y_t = a + b^o X_t^o + b^u X_t^u + b^m u_t^m$$

where: $Y_t = \begin{bmatrix} Y_t^{NE} \\ Y_t^E \end{bmatrix}$, $a = \begin{bmatrix} a^{NE} \\ a^E \end{bmatrix}$, $b^o = \begin{bmatrix} b^{NE,o} \\ b^{E,o} \end{bmatrix}$, $b^u = \begin{bmatrix} b^{NE,u} \\ b^{E,u} \end{bmatrix}$, $b^m = \begin{bmatrix} \underline{0}_{(K_2 \times (N-K_2))} \\ b^{E,m} \end{bmatrix}$

For a given parameter vector $\theta = (\mu, \Phi, \Sigma, \delta_0, \delta_1, \lambda_0, \lambda_1)$, we can invert the above equation to solve for X_t and u_t^m.

Likelihood function:

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

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

$$\log\left(\lambda(\theta)\right) = \sum_{t=2}^{T} \log|\det(J^{-1})| + \log f_X\left(X_t^o,\ X_t^u \mid X_{t-1}^o, X_{t-1}^u\right) + \log f_{u^u}\left(u_t^m\right)$$

$$= -(T-1)\log|\det(J)| - \frac{(T-1)}{2}\log\left(\det(\Omega)\right) - \frac{1}{2}\sum_{t=2}^{T}\left(X_t - \mu - \Phi X_{t-1}\right)'\Omega^{-1}\left(X_t - \mu - \Phi X_{t-1}\right)$$

$$-\frac{(T-1)}{2}\log\sum_{i=1}^{N-K_2}\sigma_i^2 - \frac{1}{2}\sum_{t=2}^{T}\sum_{i=1}^{N-K_2}\frac{\left(u_{t,i}^m\right)^2}{\sigma_i^2}$$

(The constant terms like $\frac{(T-1)}{2}\log(2\pi)$ are ignored)

The Jacobian term is:

$$J = \begin{pmatrix} I_{K-K_2} & 0_{(K-K_2)\times K_2} & 0_{(K-K_2)\times(N-K_2)} \\ B^o & B^u & B^m \end{pmatrix}$$

Tabel 1. Modeling Framework

State variables		Unrestricted model		Restricted Model (No-arbitrage)	
# Financial factors	# Macro variables (factors)	Small information set: {Yield curve, output, inflation}	Large information set: {Macro factors}	Small information set: {Yield curve, output, inflation}	Large information set: {Macro factors} (mf)
3	0	3 Nelson Siegel (NS) factors (Diebold-Li, 2003)		3latent factors 1) By Chen-Scott (1993) method; 2) By Kalman Filter	
	1	3 NS' + output	3 NS' + mf1	3 CS' + output	3 CS' + mf1
	2	3 NS' + output, inflation	3 NS' + mf1, mf2	3 CS' + output, inflation	3 CS' + mf1, mf2
2	1	2 NS' + output	2 NS' + mf1	2 CS' + output	2 CS' + mf1
	2	2 NS' + output, inflation	2 NS'+ mf1, mf2	2 CS' + output, inflation	2 CS' + mf1, mf2
	3		2 NS'+ mf1, mf2, mf3		2 CS' + mf1, mf2, mf3
1	2	1 NS + output, inflation	1 NS + mf1, mf2	1 CS + output, inflation	1 CS + mf1, mf2
	3		1 NS + mf1, mf2, mf3		1 CS + mf1, mf2, mf3
	4		1 NS + mf1, mf2, mf3, mf4		1 CS + mf1, mf2, mf3, mf4

NS: Nelson-Siegel factors used in Diebold-Li model; **CS**: Latent factor extracted by using Chen-Scott method; **mf**: macro common factors.

Tabel 2. Factor loadings

Factors are extracted from a panel with 171 macro variables (1974:2-2002:12). First four factors with the eight variables that it is most highly correlated. The first four factors together explain 49.58% of the total variation in the macro panel. The number of series is their original order in the panel used in Giannone, Reichlin and Sala (2004).

Factor 1	**Total variance explained: 28.57%**	R^2
1	Index of IP: Total	0.8826
44	Employment on nonag. payrolls: Goods-producing	0.8794
17	Index of IP: Non-energy, total	0.8730
11	Index of IP: Mfg	0.8704
21	Index of IP: Non-energy excl CCS and MVP	0.8668
22	Capacity of Utilization: Total	0.8636
20	Index of IP: Non-energy excl CCS	0.8605
43	Employment on nonag. payrolls: Total private	0.8577
Factor 2	**Total variance explained: 12.67%**	
133	CPI: all items less food	0.6118
126	CPI: housing	0.5798
136	CPI: all items less food and energy	0.5603
124	CPI: all items (urban)	0.5287
135	CPI: all items less medical care	0.5018
122	PPI: finished goods excl food	0.4805
132	CPI: service	0.4802
82	Inventories: Mfg and Trade: Mfg (Mil of chained 96$)	0.4669
Factor 3	**Total variance explained: 4.55%**	
130	CPI: commodities	0.4198
141	PCE prices: non durables	0.3797
134	CPI: all items less shelter	0.3431
128	CPI: transportation	0.3412
119	PPI: finished consumer goods	0.3214
118	PPI: finished goods (1982=100 for all PPI data)	0.2930
138	PCE chain weight price index: Total	0.2663
122	PPI: finished goods excl food	0.2418
Factor 4	**Total variance explained: 3.78%**	
99	Nominal effective exchange rate	0.3533
101	Spot SZ/US	0.3307
100	Spot Euro/US	0.3278
102	Spot Japan/US	0.2767
105	M1 (in bil. of current $)	0.2477
137	Price of gold ($/oz) on the London market (recorded in the p.m.)	0.1793
103	Spot UK/US	0.1757
54	Employment on nonag. payrolls: Financial activities	0.1654

Table 3. Three financial factor models

Forecasting RMSE ratio with respect to Random Walk

Unrestricted model

h	1	3	6	9	12	18	24
m01	**0,911**	**1,000**	**0,997**	**0,940**	**0,955**	1,134	1,277
m02	1,007	1,046	1,010	**0,947**	**0,949**	1,116	1,267
m03	1,122	1,099	1,032	**0,961**	**0,949**	1,091	1,241
m06	1,172	1,129	1,049	**0,969**	**0,942**	1,064	1,274
m09	1,139	1,102	1,040	**0,978**	**0,950**	1,076	1,310
y01	***<u>0,794</u>***	1,014	1,012	**0,950**	**0,940**	1,077	1,314
y02	1,186	1,086	1,037	1,012	1,009	1,108	1,427
y03	***<u>0,776</u>***	**0,960**	**0,979**	**0,970**	**0,990**	1,147	1,516
y05	1,046	1,031	1,033	1,073	1,116	1,283	1,731
y07	**0,997**	1,002	1,011	1,059	1,123	1,308	1,749
y10	1,001	**0,998**	1,050	1,117	1,195	1,412	1,849

Restricted model

3 latent factors with Chen-Scott(1993) method, AR(1) in state variables.
(m01, y2, y7 are assumed measured without error.)

h	1	3	6	9	12	18	24
m01	**0.913**	***0.807***	***0.808***	***0.814***	**0.961**	1.242	1.274
m02	1.016	***0.817***	***0.805***	***0.819***	**0.952**	1.215	1.250
m03	1.148	***0.844***	***0.831***	***0.840***	**0.948**	1.183	1.221
m06	1.022	***0.828***	***0.847***	***0.852***	**0.926**	1.103	1.160
m09	**0.964**	***0.838***	***0.870***	***0.873***	**0.928**	1.071	1.112
y01	***<u>0.791</u>***	***0.823***	***0.885***	***0.869***	**0.921**	1.056	1.091
y02	**0.988**	**0.976**	**0.966**	**0.945**	**0.943**	**0.979**	1.009
y03	***<u>0.665</u>***	***0.895***	**0.936**	**0.908**	**0.903**	**0.962**	1.009
y05	1.008	**0.985**	**0.983**	**0.972**	**0.958**	**0.978**	1.016
y07	**0.998**	**0.993**	**0.988**	**0.979**	**0.974**	**0.997**	1.031
y10	1.013	**0.969**	**1.000**	**1.000**	1.004	1.036	1.075

3 latent factors with Kalman Filter, AR(1) in state variables

h	1	3	6	9	12	18	24
m01	***0,841***	***0,852***	***0,806***	***<u>0,711</u>***	***0,858***	1,279	1,622
m02	**0,926**	**0,914**	***0,846***	***<u>0,750</u>***	***0,888***	1,305	1,657
m03	1,048	**0,967**	***0,881***	***<u>0,781</u>***	**0,902**	1,296	1,646
m06	1,254	1,055	**0,946**	***0,834***	**0,932**	1,269	1,662
m09	1,215	1,057	**0,966**	***0,863***	**0,955**	1,259	1,655
y01	***0,802***	**0,942**	**0,929**	***0,827***	**0,927**	1,231	1,602
y02	1,174	1,066	1,000	**0,927**	1,009	1,189	1,559
y03	***<u>0,789</u>***	**0,943**	**0,949**	***0,884***	**0,968**	1,166	1,521
y05	1,148	1,011	**0,971**	**0,938**	1,033	1,153	1,452
y07	1,090	1,045	1,012	**0,989**	1,098	1,218	1,501
y10	1,424	1,234	1,205	1,170	1,258	1,365	1,635

Table 4. Three financial factors + 1 macro variable/factor models
Forecasting RMSE ratio with respect to Random Walk

Small N

Unrestricted

3Nelson-Siegle factors +1ma, AR(1) in state variables

h	1	3	6	9	12	18	24
m01	**0,939**	1,020	1,007	**0,938**	**0,946**	1,103	1,227
m02	1,028	1,054	1,015	**0,950**	**0,948**	1,109	1,257
m03	1,100	1,090	1,030	**0,966**	**0,956**	1,108	1,267
m06	1,170	1,127	1,048	**0,973**	**0,950**	1,085	1,306
m09	1,153	1,101	1,036	**0,977**	**0,953**	1,086	1,323
y01	***0,812***	1,002	1,012	**0,967**	**0,969**	1,125	1,388
y02	1,198	1,095	1,032	**0,990**	**0,976**	1,059	1,347
y03	***0,787***	**0,956**	**0,980**	**0,973**	**0,990**	1,140	1,511
y05	1,024	1,001	**0,997**	1,026	1,063	1,223	1,660
y07	1,002	1,013	1,022	1,069	1,133	1,319	1,757
y10	1,091	1,050	1,099	1,173	1,254	1,474	1,920

Restricted

3latent factors (Chen-Scott)+1ma, AR(1)

h	1	3	6	9	12	18	24
m01	**0.941**	***0.865***	***0.873***	***0.869***	1.005	1.302	1.376
m02	1.055	***0.877***	***0.865***	***0.868***	**0.990**	1.268	1.343
m03	1.198	**0.906**	***0.886***	***0.885***	**0.983**	1.230	1.306
m06	1.070	***0.870***	***0.882***	***0.881***	**0.947**	1.134	1.224
m09	**0.992**	***0.862***	***0.892***	***0.892***	**0.942**	1.093	1.165
y01	***0.821***	***0.848***	**0.906**	***0.888***	**0.936**	1.079	1.145
y02	**0.991**	**0.981**	**0.973**	**0.952**	**0.948**	**0.990**	1.046
y03	***0.665***	***0.896***	**0.941**	**0.915**	**0.908**	**0.974**	1.048
y05	1.010	**0.982**	**0.983**	**0.974**	**0.961**	**0.988**	1.050
y07	**0.999**	**0.996**	**0.996**	**0.990**	**0.983**	1.015	1.075
y10	1.085	1.001	1.027	1.020	1.013	1.047	1.102

Large N

Unrestricted

3Nelson-Siegle factors +1mf, AR(1) in state variables

h	1	3	6	9	12	18	24
m01	**0,931**	1,011	1,002	**0,937**	**0,946**	1,103	1,227
m02	1,033	1,057	1,018	**0,952**	**0,949**	1,110	1,257
m03	1,116	1,099	1,034	**0,966**	**0,956**	1,108	1,268
m06	1,166	1,125	1,046	**0,971**	**0,949**	1,084	1,306
m09	1,123	1,089	1,031	**0,974**	**0,951**	1,083	1,322
y01	***0,869***	1,029	1,024	**0,970**	**0,971**	1,128	1,390
y02	1,159	1,070	1,019	**0,986**	**0,973**	1,055	1,345
y03	***0,813***	**0,974**	**0,989**	**0,976**	**0,992**	1,144	1,513
y05	1,009	1,002	1,004	1,036	1,072	1,227	1,653
y07	**0,996**	1,004	1,015	1,063	1,128	1,316	1,759
y10	1,081	1,045	1,094	1,168	1,251	1,474	1,925

Restricted

3latent factors (Chen-Scott)+1mf, AR(1)

h	1	3	6	9	12	18	24
m01	***0.882***	***0.822***	***0.818***	***0.818***	**0.914**	1.184	1.278
m02	**0.907**	***0.857***	***0.852***	***0.857***	**0.944**	1.218	1.341
m03	**0.923**	***0.883***	***0.879***	***0.884***	**0.956**	1.214	1.353
m06	**0.986**	**0.945**	**0.947**	**0.939**	**0.984**	1.213	1.433
m09	1.001	**0.962**	**0.982**	**0.981**	1.011	1.227	1.469
y01	***0.658***	***0.872***	**0.953**	**0.949**	**0.992**	1.205	1.444
y02	1.013	1.031	1.047	1.045	1.045	1.180	1.491
y03	***0.644***	**0.936**	1.012	1.015	1.022	1.194	1.550
y05	1.122	1.101	1.134	1.166	1.173	1.339	1.778
y07	1.014	1.043	1.093	1.140	1.179	1.370	1.807
y10	1.022	**0.973**	1.069	1.137	1.198	1.417	1.834

Table 5. Three financial factors + 2 macro variable/factor models

Forecasting RMSE ratio with respect to Random Walk

Small N

Unrestricted

3Nelson-Siegle factors +2ma, AR(1) in state variables

h	1	3	6	9	12	18	24
m01	0,932	1,016	1,006	0,941	0,954	1,116	1,240
m02	1,030	1,054	1,017	0,953	0,952	1,114	1,261
m03	1,098	1,092	1,031	0,966	0,955	1,106	1,266
m06	1,161	1,125	1,045	0,967	0,942	1,074	1,295
m09	1,141	1,097	1,031	0,969	0,945	1,074	1,312
y01	0,803	0,999	1,010	0,962	0,962	1,117	1,380
y02	1,192	1,094	1,032	0,993	0,982	1,066	1,353
y03	0,797	0,963	0,988	0,983	1,000	1,153	1,525
y05	1,020	1,004	1,003	1,037	1,076	1,237	1,672
y07	1,001	1,013	1,023	1,070	1,135	1,320	1,758
y10	1,050	1,028	1,081	1,150	1,232	1,453	1,903

Restricted

3latent factors (Chen-Scott)+2ma, AR(1) in state variables

h	1	3	6	9	12	18	24
m01	0.938	0.866	0.880	0.876	1.011	1.308	1.392
m02	1.034	0.869	0.865	0.870	0.992	1.269	1.353
m03	1.162	0.893	0.884	0.885	0.984	1.229	1.313
m06	1.049	0.863	0.880	0.880	0.947	1.130	1.227
m09	0.984	0.861	0.892	0.891	0.941	1.087	1.166
y01	0.817	0.851	0.909	0.890	0.936	1.075	1.148
y02	0.993	0.982	0.973	0.951	0.946	0.983	1.043
y03	0.633	0.890	0.936	0.911	0.905	0.967	1.043
y05	1.055	0.991	0.983	0.973	0.962	0.986	1.048
y07	1.000	0.997	0.995	0.987	0.978	1.003	1.057
y10	1.067	0.996	1.022	1.014	1.008	1.032	1.076

Large N

Unrestricted

3Nelson-Siegle factors +2mf, AR(1) in state variables

h	1	3	6	9	12	18	24
m01	0,929	1,010	1,002	0,936	0,946	1,103	1,227
m02	1,032	1,058	1,018	0,951	0,949	1,109	1,257
m03	1,116	1,099	1,034	0,966	0,957	1,108	1,268
m06	1,169	1,126	1,047	0,972	0,950	1,085	1,306
m09	1,124	1,090	1,032	0,975	0,952	1,083	1,322
y01	0,867	1,028	1,022	0,969	0,969	1,126	1,389
y02	1,159	1,070	1,020	0,986	0,973	1,055	1,345
y03	0,813	0,973	0,988	0,976	0,992	1,144	1,514
y05	1,009	1,002	1,004	1,036	1,072	1,228	1,653
y07	0,995	1,004	1,014	1,063	1,127	1,315	1,758
y10	1,080	1,045	1,095	1,168	1,251	1,474	1,925

Restricted

3latent factors (Chen-Scott)+2mf, AR(1) in state variables

h	1	3	6	9	12	18	24
m01	0.856	0.777	0.786	0.786	0.897	1.186	1.297
m02	0.856	0.801	0.815	0.823	0.925	1.217	1.357
m03	0.877	0.835	0.850	0.857	0.941	1.217	1.373
m06	0.978	0.926	0.933	0.924	0.976	1.219	1.454
m09	0.983	0.944	0.969	0.964	1.000	1.229	1.483
y01	0.652	0.852	0.938	0.930	0.978	1.202	1.450
y02	1.007	1.030	1.048	1.041	1.045	1.188	1.507
y03	0.690	0.966	1.039	1.037	1.042	1.220	1.591
y05	1.191	1.143	1.173	1.198	1.201	1.370	1.820
y07	1.013	1.049	1.106	1.149	1.188	1.380	1.823
y10	1.021	0.991	1.097	1.162	1.222	1.440	1.867

Table 6. Five factors unrestricted models

Forecasting RMSE ratio with respect to Random Walk

3Nelson-Siegel factors +2ma, VAR(1) in state variables

h	1	3	6	9	12	18	24
m01	***0.844***	***0.798***	***0.794***	***0.734***	***0.867***	1.177	1.309
m02	**0.923**	***0.845***	***0.814***	***0.774***	***0.886***	1.200	1.350
m03	**0.981**	***0.898***	***0.839***	***0.801***	***0.895***	1.171	1.303
m06	1.059	**0.984**	***0.881***	***0.840***	***0.899***	1.079	1.235
m09	1.085	1.003	***0.891***	***0.864***	**0.906**	1.044	1.204
y01	***0.771***	**0.944**	***0.879***	***0.858***	**0.907**	1.021	1.164
y02	1.194	1.056	**0.908**	***0.890***	***0.893***	***0.885***	1.072
y03	***0.813***	**0.953**	***0.880***	***0.872***	***0.873***	***0.866***	1.089
y05	1.032	**0.975**	***0.892***	***0.893***	***0.889***	***0.886***	1.196
y07	1.002	**0.964**	***0.899***	**0.915**	**0.931**	**0.970**	1.331
y10	**0.995**	**0.944**	**0.946**	**0.988**	1.019	1.115	1.513

3Nelson-Siegel factors +2mf, VAR(1) in state variables

h	1	3	6	9	12	18	24
m01	***0.839***	***0.717***	***0.804***	**0.952**	1.203	1.930	2.184
m02	**0.942**	***0.765***	***0.850***	1.013	1.241	1.951	2.236
m03	**0.977**	***0.782***	***0.871***	1.043	1.255	1.913	2.211
m06	**0.932**	***0.812***	***0.898***	1.059	1.239	1.771	2.158
m09	**0.921**	***0.825***	**0.917**	1.090	1.255	1.739	2.140
y01	***0.712***	***0.791***	**0.942**	1.119	1.306	1.798	2.235
y02	1.104	**0.933**	**0.959**	1.116	1.247	1.533	1.988
y03	***0.801***	***0.887***	**0.985**	1.151	1.290	1.618	2.174
y05	1.050	**0.977**	1.032	1.213	1.347	1.635	2.217
y07	1.030	**0.998**	1.062	1.249	1.406	1.715	2.308
y10	1.084	1.059	1.195	1.400	1.556	1.896	2.479

First two Nelson-Siegel factors +3mf, VAR(1) in state variables

h	1	3	6	9	12	18	24
m01	***0.865***	***0.803***	***0.856***	**0.945**	1.087	1.496	1.782
m02	1.029	***0.866***	***0.896***	**0.996**	1.116	1.519	1.833
m03	1.197	**0.905**	**0.915**	1.017	1.123	1.492	1.821
m06	1.438	**0.963**	**0.940**	1.027	1.099	1.393	1.796
m09	1.470	**0.963**	**0.944**	1.044	1.101	1.371	1.788
y01	1.470	**0.953**	**0.956**	1.060	1.122	1.417	1.866
y02	1.699	1.025	**0.963**	1.063	1.091	1.236	1.688
y03	1.541	**0.982**	**0.968**	1.083	1.119	1.328	1.860
y05	1.408	1.003	1.007	1.154	1.208	1.392	1.940
y07	1.250	**1.000**	1.037	1.200	1.280	1.490	2.038
y10	1.103	1.040	1.165	1.353	1.443	1.687	2.219

First Nelson-Siegel factor +4mf, VAR(1) in state variables

h	1	3	6	9	12	18	24
m01	3.908	2.100	1.329	**0.936**	***0.850***	1.078	1.223
m02	4.770	2.215	1.321	**0.932**	***0.849***	1.088	1.258
m03	5.677	2.353	1.359	**0.965**	***0.870***	1.089	1.263
m06	5.668	2.326	1.360	**0.989**	***0.883***	1.066	1.288
m09	5.268	2.186	1.317	**0.986**	***0.882***	1.051	1.278
y01	5.382	2.272	1.364	1.006	***0.897***	1.076	1.315
y02	3.891	1.814	1.216	**1.000**	**0.908**	**0.991**	1.238
y03	3.667	1.714	1.181	**0.988**	**0.907**	1.041	1.345
y05	2.448	1.339	1.064	1.009	**0.981**	1.104	1.455
y07	1.958	1.194	1.029	1.025	1.033	1.176	1.546
y10	1.358	1.008	1.020	1.083	1.130	1.306	1.684