Predicting Future Bond Returns with Macro Variables: A Semi-parametric Approach*

Ai-ru (Meg) Cheng†

November 30, 2006

*I am grateful to Evan Anderson, Yin-Wong Cheung, Valentina Corradi, Ron Gallant, Eric Ghysels, and Raffaella Giacomini for their helpful comments. All remaining errors are my own.
†archeng@ucsc.edu, Department of Economics, University of California at Santa Cruz, CA 95064.

Abstract

This paper models the dynamic behavior of future bond returns. Using a no-arbitrage approach, we predict joint dynamics of future bond returns at all maturities using both term-structure and macro variables. The dimensionality reduction is achieved by composing index factors of the underlying variables. The modeling flexibility is preserved without specifying the functional form for the model dependence on the index factors. We develop a semi-parametric estimation scheme in Generalized Method of Moments (GMM) framework. We show that model based on both term-structure and macro variables is superior to model using only the term-structure variables in predicting future bond returns at all maturities. Moreover, it is shown that the index model can better forecast bond returns at longer time horizons.

JEL Classifications: C5, E4, G0.

Keywords: no-arbitrage, pricing kernel, semi-parametric. non-parametric, GMM.

1. Introduction

In recent years, a vast literature has modeled term structure dynamics as well as developed econometric tools for estimation. Most term-structure models have focused on fitting the yield curve, but relatively little attention has been paid on the forecast in yield changes over time. Although the expectations hypothesis (hereafter EH) implies that excess returns on zero-coupon bonds follow a white noise process, a series of classic regression analysis in Fama and Bliss (1987), Campbell and Shiller (1991), Cochrane and Piazzesi (2005), and many others have revealed evidences against EH and this result has motivated search for the underlying factors that forecast future movements in bond returns.

Cochrane and Piazessi (2005) found that the same linear combination of forward rates predicts 1-year-ahead excess returns at all maturities. This paper extends their work in investigating the return-forecasting ability of the common factors. Using no-arbitrage approach, we are able to model the joint dynamics of future bond returns at all maturities. Besides using yields and forward rates as the underlying variables, we include both real and macro variables to study their implications of the return forecasts.

This paper develops a semi-parametric estimation method to conduct our model estimation. Under the arbitrage-free assumption, the pricing kernel is projected onto a few index factors formed by the economic variables without specifying any functional form for the projection. One advantage of the proposed index factor model is that it allows us to predict all future bond returns based on only a few factors without imposing restrictive assumptions on the model specification. Our empirical study shows that the proposed index factor model is more flexible than a linear model in forecasting bond returns at all maturities.

We propose and estimate one- and two-index factor models to determine whether macro variables offer any incremental return-forecasting power in the presence of yields and forward rates. Our results show that by incorporating macro variables, the model produces better forecasts than its counterpart using only the information in term-structure variables. We also explore the index model's predictability for bond returns at long forecasting horizons. Similar to the finding in Diebold and Li (2006), our result indicates that the model has superior return-forecasting performance at longer horizons.

The rest of the paper is organized as follows. Section 2 provides a detailed description of the modeling framework and econometric method. Section 3 describes the data. Section 4 reports our empirical findings and relates it to preceding works. Section 5 concludes.

2. Model and econometric method

In this section, we develop a non-parametric kernel-based approximation to the pricing kernel using a no-arbitrage asset pricing approach. Section 2.1 describes the basic idea of the no-arbitrage condition. Section 2.2 introduces a kernel-based pricing kernel estimator in GMM framework. Section 2.3 gives details of a low-dimensional index modeling approach due to the curse of dimensionality problem in general non-parametric econometric methods.

2.1 A nonlinear arbitrage pricing model

The basic idea in the arbitrage pricing theory (hereafter, APT) by Ross (1976) is to assume that the markets are arbitrage-free and that the dynamics of asset payoffs are linear in a few state variables of the economy. The linearity assumption and the no-arbitrage restriction lead to a linear relationship between the pricing kernel (or state-price deflator) and the underlying state variables.

However, Bansal and Viswanathan (1993) and Chapman (1997) argued that the linearity assumption in APT can be relaxed by allowing the asset payoffs to be non-linear in the underlying state variables. They model the pricing kernel using polynomials of a small number of the state variables, so that the pricing kernel is capable of explaining nonlinear payoffs.

An an alternative approach, which we pursue in this paper, is to approximate the pricing kernel using a kernel-based non-parametric method. Under no-arbitrage opportunity assumption, we project the pricing kernel onto a set of state variables representing the current economic information set and estimate the pricing kernel based on the joint density of the underlying state variables. Let $\mathcal{F}_t$ be the information set up to time t, and denote m_{t+1} be a pricing kernel for a set of asset payoffs $\mathbf{P}_{t+1}$ to be realized at time $t+1$. The no-arbitrage assumption gives the following moment condition:

$$E\left[m_{t+1}\frac{\mathbf{P}_{t+1}}{\mathbf{P}_t}|\mathcal{F}_t\right] = 1 \tag{1}$$

One can also express the above restriction in terms of the holding-period returns $\mathbf{R}_{t+1} = \frac{\mathbf{P}_{t+1}}{\mathbf{P}_t}$:

$$E\left[m_{t+1}\mathbf{R}_{t+1} - 1|\mathcal{F}_t\right] = 0 \tag{2}$$

For illustration, consider the pricing kernel $m_{t+1} = m$ to be constant at the moment. In a general GMM settings, the sample moment condition $se(m)$ according to the

Euler condition (2) can be formd by the law of iterated expectations:

$$se(m) = \frac{1}{T}\sum_{t=1}^{T}\left[(m\mathbf{R}_{t+1} - 1)\otimes Z_t\right] \tag{3}$$

where Z_t is a set of instrumental variables that represent the information content of $\mathcal{F}_t$.

Denote X_t to be a d-dimensional vector of all the economic state variables underlying the pricing kernel m. One can obtain kernel-based non-parametric estimates of the pricing kernel by replacing the instrumental variables Z_t in equation (3) with a joint kernel density of X_t:

$$se(m) = \frac{1}{T}\sum_{t=1}^{T}\left[(m\mathbf{R}_{t+1} - 1)K\left(\frac{X_t - x}{h}\right)\right] \tag{4}$$

The kernel density function in equation (4) can be taken as a t-measurable $\mathcal{R}^d \mapsto \mathcal{R}$ instrumental variable that summarizes all the information in X_t. A series of dynamic estimates on the pricing kernel can be obtained by iterating the following GMM estimator through $t = 1, ..., T$:

$$\begin{aligned} x &= X_t, \quad t = 1, ..., T \\ \hat{m} &= \arg\min se(m)' W_T \, se(m), \\ & \quad s.t. \quad m \geq 0. \end{aligned}$$

where W_T is a weighting matrix discussed in Hansen and Singleton (1992).

2.1.2 Kernel density and bandwidth selection

The kernel density function of X_t in equation (4) is assumed to be equal to the product of the univariate kernel density of each variable in X_t. In general, as long as the kernel density function is bounded and symmetric, the choice of functional form is not as crucial as selecting the bandwidth h. In this paper, we use the Gaussian density function for the kernel density of X_t:

$$\begin{aligned} K\left(\frac{X_t - x}{h}\right) &= \prod_{j=1}^{d} k\left(\frac{X_{j,t} - x_j}{h_j}\right) \\ k(u) &= \frac{1}{\sqrt{2\pi}}\exp(-u^2/2) \end{aligned}$$

Selecting bandwidth includes a trade-off between bias and variance - a large h means local averaging over more data points, which reduces the variance of the estimates, but introduces bias in the estimator. On the other hand, a small h makes the estimator capable of capturing more local behaviors in the data, but it increases the variance. Following Silverman (1986) and Carroll, Fan, and Wand (1995), we use a simple plug-in rule for selecting h:

$$h_j = \lambda \sigma_j T^{-\frac{1}{(d+4)}}, \quad j = 1, ..., d. \tag{5}$$

Recall that d is the number of the state variables in X_t and the term $T^{-\frac{1}{(d+4)}}$ indicates that as d gets large, one needs to tune up the size of the window around x in order to include enough data points to analyze the local behaviors of X_t. σ_j is the standard deviation of the j-th series in X_t. λ is a tuning parameter, controlling the smoothness of the estimator $\hat{m}$. Since there is no conventional rule for selecting λ, particularly when our data is not serially independent, we use a trial and error approach suggested in Chaudhuri and Marron (1999) and settle with $\lambda = 1.0$ in our empirical analysis [1].

One distinct advantage of this non-parametric approach is that it allows a more flexible projection of the pricing kernel on the state variables and lends the model to capture more of the dynamics in asset payoffs. Another advantage of this method is that since the kernel density $K(\frac{X_t - x}{h})$ is a bounded function for any x given $h > 0$, it avoids some over-fitting problems that arise in some parametric approximation methods.

2.2 Index factor analysis

Due to the "curse of dimensionality" problem inherited in most non-parametric econometric methods, it is difficult to model the joint dynamics of future bond returns when the dimension of the underlying d is large. That is to say, as the number of state variables increases, the rate for the the pricing kernel estimator converging to its asymptotic distribution becomes exponentially low. To tackle this problem, we adopt a semi-parametric scheme by assuming that the pricing kernel m only depends on some linear combination of the underlying state variables and the functional form of this dependency stays unrestricted.

Modeling the pricing kernel based on composite indices of the state variables is motivated by the projection pursuit regression model by Friedman, et. al. (1981). The basic idea of projection pursuit is to find the projections (directions) from high- to low-dimensional space that reveal the most details about the structure of the data

[1] We verify that the estimated factor loadings are fairly close in a range of $\lambda = [0.7, 1.2]$.

set. Our index approach is also supported by Cochrane and Piazzesi (2004), which shows that a linear combination of the forward rates is a common factor for forecasting future excess returns on bonds at all maturities.

For conducting the model estimation, we extend the idea in Ait-Sahalia and Brandt (2001) and develop an estimation procedure that simultaneously estimates the index coefficients and the associated pricing kernel by iterating through the following estimators:

$$\begin{aligned} \hat{m} &= \arg\min \quad T\, se_1(m)\, W_{1T}\, se_1(m), \quad t = 1, ..., T. \qquad (6) \\ & \quad s.t. \quad m \geq 0. \\ \hat{\beta} &= \arg\min \quad T\, se_2(\beta)\, W_{2T}\, se_2(\beta) \qquad (7) \\ & \quad s.t. \quad \|\beta\| = 1. \qquad (8) \end{aligned}$$

Where

$$\begin{aligned} \mu &= \hat{I}_t, \quad t = 1, ..., T \\ se_1(m) &= \frac{1}{Th} \sum_{t=1}^{T} (m\mathbf{R}_{t+1} - 1) k(\frac{\hat{I}_t - \mu}{h}) \end{aligned}$$

and

$$se_2(\beta) = T^{-1} \sum_{t=1}^{T} [\hat{m}(X_t\beta)\mathbf{R}_{t+1} - 1] \otimes Z_t \qquad (9)$$

The discussion in Section 2.1 illustrates that for a given set of values on index coefficients β, one can obtain a series of dynamic estimates of the pricing kernel by equation (6) for each $t = 1, ..., T$. This procedure is then repeated at each GMM iteration of equations (7)-(8) until the convergence criterion[2] is met. To simplify the estimation procedure, we fix the weighting matrix W_{1T} in equation (6) to an identity matrix, so that all the sample moments in $se_1(m)$ are treated as equally important. For the estimation of β, we account for heteroskedastity and serial correlation of the sample moments in equation (7) by adopting the optimal weighting matrix W_{2T} as discussed in Hansen and Singleton (1992):

$$\begin{aligned} W_{2T} &= \hat{S}^{-1} \\ \hat{S} &= \hat{S}_o + \sum_{j=1}^{J} \omega(j) \left[\hat{S}_j + \hat{S}'_j\right] \end{aligned}$$

[2]We set our convergence tolerance to be 1.e-10 in our empirical work.

$$\hat{S}_j = \frac{1}{T-k}\sum_{t=j+1}^{T-1} se_{2t}(\hat{\beta})se_{2t}(\hat{\beta})'$$
$$se_{2t}(\hat{\beta}) = \left[\hat{m}(X_t\hat{\beta})\mathbf{R}_{t+1} - \mathbf{1}\right] \otimes Z_t$$

where $\omega(j)$ is a Parzen kernel weighting function (see Gallant, 1987) with J as the pre-determined lag truncation parameter.

3. Data

In this section, we discuss the data used in our empirical analysis.

3.1 Holding period returns, yields, and forward rates

According to Knez, Litterman, and Scheinkman (1994), the total variation in the interest rates can be decomposed by the "level", "spread", and "curvature", which are often regarded as some linear combinations of the yields themselves. Knez, Litterman, and Scheinkman (1994) also suggests that the "level" factor alone can explain up to 90 percent of the total variation. The regression analysis in Fama and Bliss (1987) and Campbell and Shiller (1991) show that a linear combination of the yields or forward rates can explain the dynamics of future bond returns quite well. Moreover, Cochrane and Piazzesi (2004) shows that forward rates are useful for predicting future bond returns. In light of all these previous studies, we include 1-year yield and forward rates as the term-structure variables for predicting future bond returns.

All the data on holding period returns, yields, and forward rates are constructed from the US government bond prices[3], which contains monthly-sampled bond prices with maturity varying from 1 to 5 years. The sample covers from January, 1960 to December, 2003, resulting in 528 monthly observations in each bond return series[4]. Figure 1 plots the standardized 1-year yield and 2-5 year forward rates over time.

Let $p_t^{(n)}$ denote the log price of n-year discount bond at time t, where n measures the time to maturity in years and t is measured in months corresponding to our sample frequency. Thus, the n-year yield $y_t^{(n)}$ and forward rate $f_t^{(n)}$ at t can be defined as follows:

$$y_t^{(n)} = -\frac{1}{n}p_t^{(n)}$$
$$f_t^{(n)} = p_t^{(n-1)} - p_t^{(n)}$$

[3] We use the unsmoothed Fama-Bliss data.
[4] The data is retrieved from CRSP.

Also let $r_{t+12}^{(n)}$ denote the 1-year holding period return, from buying a n-year bond at time t and selling it as a $n-1$-year bond a year later at time $t+12$:

$$r_{t+12}^{(n)} = p_{t+12}^{(n-1)} - p_t^{(n)} \tag{10}$$

Table 1 reports the sample statistics of the 1-year holding period returns r_{t+12} at all maturities. Cochrane and Piazzesi (2004) indicates that although both yields and forward rates should span the same space for bond prices, their empirical evidence suggests that a linear combination of the forward rates outperforms its counterpart using the yields in predicting future bond returns. Moreover, by looking directly at 1-year forecasting horizon, one can recover the return-forecasting ability of the forward rates. However, if looking at a higher frequency of monthly- or daily-horizon, this return-forecasting ability is likely to be concealed by measurement errors. We first adopt regression analysis to check if this common-factor feature for the expected return also exists in our data sample:

$$r_{t+12}^{(j)} = [1, y_t^{(1)}, \mathbf{f}_t]\gamma^{(j)} + \epsilon_{1,t+12}^{(j)}, \quad j = 2, 3, 4, 5. \tag{11}$$

where

$$\begin{aligned} \mathbf{f}_t &= [f_t^{(2)}, f_t^{(3)}, f_t^{(4)}, f_t^{(5)}], \\ \gamma^{(j)} &= [\gamma_0^{(j)}, ..., \gamma_5^{(j)}], \quad j = 2, 3, 4, 5. \end{aligned}$$

Figure 2 graphs the estimated regression coefficients $\hat{\gamma}$ across different terms to maturity. The sample $\mathbf{R}^2$ is 0.8075, 0.6183, 0.542, and 0.473 respectively for the maturity of 2-5 years. The tent-like pattern shown in Cochrane and Piazzesi (2004) is replicated here and it confirms that there is a common factor of the forward rates underlying the predictability of future bond returns at all maturities in our sample.

3.2 Macro variables

Although it is quite common to use factors like yields or forward rates to explain term structure dynamics, these factors do not lend themselves to understanding the economic conditions, if any, that interacts with the term structure dynamics. To better relate economic conditions to term structure, recent studies used macro variables to explain the movements in bond prices and found that some macro variables do have influences on the dynamic behavior of the term structure. Examples include, (a) consumer pricing index for all urban consumers (CPI, Hardouvelis, 1988); (b) producer price index for finished goods (PPI, Dwyer and Hafer, 1989; McQueen and Roley, 1993; Piazzesi, 2000); (c) M2 money stock (M2, Ghysels and Ng, 1998); (d) industrial production index (IP, McQuenn and Roley, 1993; Roley et al., 1983, Ghysels and Ng, 1998); (e) non-farm payroll employment (EMPLOY, McQueen, and Roley, 1993; Piazzesi, 2000); (f) real personal consumption (PCE, Hardouvelis, 1983); and (g) housing starts (HS, Ghysels and Ng, 1998). As suggested in these preceding literature, we use CPI, PPI, and M2 as our nominal macro variables and IP, EMPLOY, PCE, and HS as the real macro variables for predicting future movements in bond prices. All the macro series used in our analysis are seasonally adjusted and retrieved from Federal Reserve Economic Data (FRED). We take each macro variable as the annual growth rate of the macro series, where the growth rate is measured as the difference in logs of the series levels at time t and $t-12$. We then standardize each constructed macro data series and use them as the underlying predictors in our empirical study. Figure 3 plots the time series of the macro variables.

Let $S = [CPI, PPI, M2, IP, EMPLOY, PCE, HS]$ be the collection of all the macro variables mentioned above. To study the causal relation between the current economic conditions and future bond returns, we project $\mathbf{r}_{t+12}$ onto S_t:

$$r_{t+12}^{(j)} = [1, S_t]\alpha^{(j)} + \epsilon_{1,t+12}^{(j)}, \quad j = 2, 3, 4, 5. \tag{12}$$

where

$$\alpha^{(j)} = [\alpha_0^{(j)}, ..., \alpha_7^{(j)}], \quad j = 2, 3, 4, 5. \tag{13}$$

Figure 4 plots the coefficients estimates $\hat{\alpha}$ with the maturity. The sample $\mathbf{R}^2$ is 0.3712, 0.2463, 0.196, and 0.175 respectively for 2-5 year bonds. Although the pattern is no longer tent-shaped, it suggests that a linear combination of the macro variables forecasts bond returns at all maturities. The above interpretation of the plot in Figure 4 is based on observing that the coefficient estimates $\hat{\alpha}_i^{(j)}$ associated with each macro variable has the same sign and is monotonic in j - bond returns with longer maturities have greater loadings on the same linear combination of the macro variables. A comparison of the sample $\mathbf{R}^2$ illustrates that the regression model with the macro variables is not as nearly impressive as the model with the term-structure variables for explaining the total variations in future bond returns.

For a clarification of some notations used in the next section, let X_t contain all the observations on the selected state variables at time t; namely,

$$X_t = [y_t^{(1)}, f_t^{(2)}, f_t^{(3)}, f_t^{(4)}, f_t^{(5)}, S_t] \tag{14}$$

Also, we set the instrumental variables Z_t equal to X_t augmented by a constant term in all the following GMM estimations.

4. Empirical Results

This section discusses our empirical findings. Section 4.1 shows the forecasting performance based on each variable in X_t. Section 4.2 compares the return-forecasting abilities of a one- and two-index factor models. Section 4.3 revisits the issue of the linearity assumption in APT. Section 4.4 applies the index factor model to forecast long-horizon holding period returns.

4.1 Forecasting performance with univariate time series

We first evaluate the return-forecasting performance for 1-year-ahead bond returns. Figure 5 plots the return-forecasting dynamics of the pricing kernel based on each series in X_t. Table 2 presents the prediction errors (in root mean square errors, hereafter RMSE) associated with the pricing kernels plotted in Figure 5. Each prediction error is computed as follows:

$$e_{jt} = \hat{m}(X_{jt})\mathbf{R}_{t+12} - 1, \quad j = 1, ..., 12. \quad t = 1, ..., 516. \tag{15}$$

Not to our surprise, the model based on the term-structure variables is superior to the one with the macro variables in predicting 1-year-ahead bond returns at all maturities. A comparison to the standard deviations reported in Table 1 illustrates the point that any term-structure variable alone can predict about 50, 40, 30, and 25 percent of the total variations of 1-year-ahead bond returns at respectively 2, 3, 4, and 5 years to the maturity. However, the prediction power of any given macro variable is much less impressive: each macro variable can only contribute, on average, no more than 20 percent of the total variation of each return series.

In a search for an "adequate" bandwidth h required in our model estimation, we detect a trade-off between the forecasting abilities for the short- and long-term bond returns. This trade-off is due to the bandwidth size h in equation (4). Specifically, within a certain range, a large h generates a smooth dynamic estimates on the pricing kernel, hence, it reduces the prediction errors of the short-term bond returns while inducing prediction errors for the long-term bonds. On the other hand, a small h reduces the prediction errors for the long-term bonds by incorporating more dynamics in the pricing kernel, but simultaneously creating more errors in predicting short-term bond returns. This observation not only reflects the statistical fact that holding-period returns on long-term bonds are more volatile than the short-term bonds, but it also suggests that a single common factor is not good enough to simultaneously capture the dynamics of the bond returns at all maturities.

4.2 Forecasting performance with index factor

To explore the return-forecasting ability of the index factor model, we follow the estimation procedures described in Section 2 to obtain GMM estimates on the index loadings β. Since the estimator $\hat{\beta}$ is dependent on another non-parametric estimator $\hat{m}$, when computing the standard error of the estimates on β, the uncertainty in $\hat{m}$ should be taken in account. Details of the asymptotic results are provided in Appendix A. Despite the existence of a non-parametric component $\hat{m}$ in the estimator of β, $\hat{\beta}$ is shown to be consistent and achieves its asymptotic normal distribution at the convergence rate of $\sqrt{T}$.

Table 3 reports our GMM estimates on index loadings of a two-index model. One can judge the importance of each index component's from its standard error reported in the parenthesis. In the first composite index, both 2- and 5-year forward rates are shown to be significant for predicting 1-year-ahead bond returns. For the second composite index, only non-farm payroll employment (EMPLOY) appears to be a significant predictor among all the selected macro variables. The contemporary correlation of the indices is -0.33222 and Figure 6 shows the time series plot of each index factor. Figure 7 plots the estimated pricing kernel over time based on both composite indices (in solid line).[5]

To examine the return-forecasting ability of the two-factor model with both term-structure and macro variables, we estimate a one-index model based on the term-structure variables alone. The prediction errors (in RMSE) in Table 4 indicate that the pricing kernel implied by the two-index model outperforms its one-index counterpart. Moreover, Table 4 shows that the two-index model can reduce, on average, 50 percent of the total variations in bond returns at all maturities. To further contrast the forecasting abilities of the models, we adopt the forecast accuracy test proposed by Diebold and Mariano (1995). Let $d_{jt} = e^2_{jt,1} - e^2_{jt,2}, \quad j = 2, 3, 4, 5$ denote the difference in squared prediction errors at time t across different maturities, where $e_{jt,1}$ and $e_{jt,2}$ are each attributed to the one- and two-index models, respectively. According to Diebold and Mariano (1995):

$$\frac{\bar{d}}{\sqrt{\frac{\hat{V}(\bar{d})}{T}}} \sim \mathcal{N}(0,1),$$

$$\hat{V}(\bar{d}) = \hat{V}_0 + \sum_{j=1}^{J} \omega(j) \left[\hat{V}_j + \hat{V}_j' \right],$$

[5] Due to the property of the kernel density function, the positivity restriction imposed on the pricing kernel $m \geq 0$ is never bounded in our estimation. Therefore, we regard it as an unconstrained optimization problem when we estimate m in this paper.

$$\hat{V}_j = \frac{1}{T}\sum_{t=j+1}^{T}(d_t - \bar{d})(d_{t-j} - \bar{d})'$$

where $\omega(.)$ is a decaying function in j with J as the truncation parameter[6].

The test result is reported in Table 4 and it favors the two-index model in predicting the 1-year-ahead bond returns at all maturities.

4.3 Linearity assumption in no-arbitrage pricing model

Despite the fact that the main focus of this paper is to explore the return-forecasting power of a nonlinear no-arbitrage pricing model, it is still an open question whether linearity is a valid assumption in the application of forecasting in bond returns. For the answer to this question, we re-model the relationship between the pricing kernel and the underlying X_t in accordance with a linearity constraint and solve the following optimization problem[7]:

$$\hat{\delta} = se_3(\delta)' W_{3T}\, se_3(\delta), \tag{16}$$

$$se_3(\delta) = T^{-1}\sum_{t=1}^{T}[(Z_t'\delta)\ \mathbf{R}_{t+12} - 1] \otimes Z_t \tag{17}$$

The pricing kernel in equation (15) is modeled as a linear function of the state variables X_t added by a constant. Table 3 indicates that all the term-structure and real macro variables have significant contributions to the return-forecasting model. This result implies that the term-structure and the nominal macro variables may share similar information contents regarding the future movements in bond returns. Table 4 shows that the forecast accuracy test rejects the linear model in favor of the two-index nonlinear model, and the test results are even more significant with the long-term bonds (maturity of 3-5 years). Figure 7 plots the pricing kernel estimated over time based on the linear model (in dotted line).

4.4 Forecasting long-horizon bond returns

We discuss forecasting ability of a nonlinear no-arbitrage model in Section 4.2. In this section, we turn to explore the index factor model's return-forecasting ability at longer horizons.

[6] In our empirical study, we use Parzen kernel in Gallant (1987) and set $J = 12$

[7] This is an unconstrained optimization problem, since the positivity constraint for the pricing kernel is never bounded in our estimation procedure.

Let $\mathbf{R}_{t+24}$ denote a 3 by 1 vector of 2-year holding period returns on 3-5 year bonds. Similarly, $\mathbf{R}_{t+36}$ is denoted as a 2 by 1 vector containing 3-year holding period returns on 4-5 year bonds:

$$\begin{aligned}\mathbf{R}_{t+24} &= [r^{(3)}_{t+24}, r^{(4)}_{t+24}, r^{(5)}_{t+24}],\\ \mathbf{R}_{t+36} &= [r^{(4)}_{t+36}, r^{(5)}_{t+36}]\end{aligned}$$

where

$$\begin{aligned}r^{(j)}_{t+24} &= (p^{(j-2)}_{t+24} - p^{(j)}_{t})/2, \quad j = 3, 4, 5\\ r^{(j)}_{t+36} &= (p^{(j-3)}_{t+36} - p^{(j)}_{t})/3, \quad j = 4, 5\end{aligned}$$

Similar to the steps taken in Section 4.2, here we compare the forecasting abilities of both one- and two-index models for 2- and 3-year-ahead bond returns. Table 4 reports the prediction errors (in RMSE) associated with each index model. The ratios of the prediction errors and the corresponding sample statistics (standard deviations are reported in Table 1) indicate that at 2-year forecasting horizon, the two-index model predicts respectively 68, 65, and 56 percent of the total variations of 3-, 4-, and 5-year bond returns. At the horizon of 3 years, the two-index model can explain up to 76 and 67 percent of the total variations of the returns on long-term bonds. Therefore, we find that the two-index model can better forecast bond returns at longer horizons. This suggestion is in line with Cochrane and Piazzesi (2004) and Diebold and Li (2006), who find that term-structure factors have better forecasting power for longer horizon yields.

Since the no-arbitrage model is generally used to fit a cross section of asset payoffs at any point in time, which, by construction, constrains its ability to explain the dynamics of each payoff over time, we worry that in a no-arbitrage setting, the superior forecasting performance for the long-horizon returns is mainly due to the fact that the dimensions of $\mathbf{R}_{t+24}$ or $\mathbf{R}_{t+36}$ are lower than that of $\mathbf{R}_{t+12}$, and makes the model less restricted to track the dynamics of each return over time. To justify this concern, we run an experiment by re-estimating the two-index model at a 1-year forecasting horizon, but only for the long-term bonds (4- and 5-year bonds). The estimation result (not reported) shows that the two-index model accounts for less than 50 percent of the total variations of the long-term bonds at 1-year forecasting horizon. The result of this experiment suggests that the model does offer a better fit for the long forecasting horizons.[8].

[8]The estimated return-forecasting pricing kernel can only capture 49 and 43 percent of the total variations in 4- and 5-year bonds.

Despite being insignificant for the 2-year forecast of the 3-year bond return, the test result in Table 4 shows that the two-index model is superior to its one-index counterpart in forecasting long-horizon bond returns. This implies that macro variables do help in predicting future dynamics of the holding period returns.

5 Conclusion

In this paper, we explored the in-sample forecasting performance of a no-arbitrage model for holding period returns on zero-coupon bonds at different horizons. We take yields, forward rates, and macro variables as the underlying predictors and evaluate their return-forecasting power. The dimensionality reduction is achieved by assuming that the return-forecasting pricing kernel is dependent on a few index factors formed by the underlying economic predictors, and yet the functional form of the dependency is left unspecified.

In methodology, we propose a semi-parametric econometric method to simultaneously identify the index factors and the associated dynamic pricing kernel in an iterative GMM estimation framework. The proposed estimation method is shown to be able to enhance modeling flexibility for capturing the dynamic behavior of future bond returns at all maturities.

Despite detecting a rather unimpressive performance in fitting future dynamics of the bond returns based on macro variables alone, we show that an index factor composed of the macro variables improves the term structure model's forecasting ability at different time horizons. This index-factor model performs even better in forecasting bond returns at longer horizons.

One possible future research direction is to extend this work and investigate the model's return-forecasting ability over the entire yield curve, as the current return data used in this paper only covers a portion of the whole term structure dynamics.

Appendix A

This section provides a technical support on deriving the asymptotic distribution of the index coefficient estimates in equations (6)-(8).

Take $X_t = [X_{1t}, X_{2t}]$ and $\hat{\beta} = [\hat{\beta}_1, \hat{\beta}_2]$. The asymptotics of $\hat{\beta} = [\hat{\beta}_1, \hat{\beta}_2]$ follow from a series of Taylor expansions of the GMM moments (6) and (7):

For $t = 1, ..., T$

$$\begin{aligned} \frac{\partial se_1(\hat{m}(X_t\hat{\beta};\hat{\beta}))'}{\partial m}\hat{W}_1 se_1(\hat{m}(X_t\hat{\beta};\hat{\beta})) &= 0 \quad (18) \\ \left[\frac{\partial se_2(\hat{m}(\hat{\beta}))}{\partial m}\frac{\partial \hat{m}(\hat{\beta})}{\partial \beta}\right]' \hat{W}_2 se_2(\hat{m}(\hat{\beta})) &= 0 \quad (19) \end{aligned}$$

First, we expand the term $se_2(\hat{m}\hat{\beta})$ in equation (7) around β_0:

$$\left[\frac{\partial se_2(\hat{m}(\hat{\beta}))}{\partial m}\frac{\partial \hat{m}(\hat{\beta})}{\partial \beta}\right]' \hat{W}_2 \left[se_2(\hat{m}(\beta_0)) + \frac{se_2(\hat{m}(\beta^*))}{\partial m}\frac{\partial \hat{m}(\beta^*)}{\partial \beta}(\hat{\beta}-\beta_0)\right] = 0 \quad (20)$$

where $\beta^* \in [\beta_0, \hat{\beta}]$.
Let

$$Q = E\left[\left[\frac{\partial se_2(m_0(\beta_0))}{\partial m}\frac{\partial m_0(\beta_0)}{\partial \beta}\right]' W_2 \frac{\partial se_2(m_0(\beta_0))}{\partial m}\frac{\partial m_0(\beta_0)}{\partial \beta}\right] \quad (21)$$

and expand the term $se_2(\hat{m}(\beta_0))$ in equation (19) around $m_0(X_t\beta_0;\beta_0)$, we get

$$\begin{aligned} \hat{\beta}-\beta_0 &= -Q^{-1}\left[\frac{\partial se_2(m_0(\beta_0))}{\partial m}\frac{\partial m_0(\beta_0)}{\partial \beta}\right]' W_2 \Bigg[se_2(m_0(\beta_0)) + \\ &\quad T^{-1}\sum_{t=1}^{T}[\mathbf{R}_{t+12}\otimes Z_t](\hat{m}(X_t\beta_0;\beta_0) - m_0(X_t\beta_0;\beta_0))\Bigg](1+o_p(1)) \quad (22) \end{aligned}$$

Given β_0, expand $se_1(\hat{m}(X_t\beta_0;\beta_0))$ in equation (6) around $m_0(X_t\beta_0;\beta_0)$:

$$\begin{aligned} &\frac{\partial se_1(\hat{m}(X_t\beta_0;\beta_0))'}{\partial m}\hat{W}_1\Bigg[se_1(m_0(X_t\beta_0)) + \\ &\frac{\partial se_1(m^*(X_t\beta_0;\beta_0))}{\partial m}[\hat{m}(X_t\beta_0;\beta_0) - m_0(X_t\beta_0;\beta_0)]\Bigg] = 0 \quad (23) \end{aligned}$$

and now denote

$$A = E\Big[\frac{\partial se_1(m_0(X_t\beta_0;\beta_0))'}{\partial m}W_1\frac{\partial se_1(m_0(X_t\beta_0;\beta_0))}{\partial m}\Big] \tag{24}$$

$$\begin{aligned}\hat{m}(X_t\beta_0;\beta_0) - m_0(X_t\beta_0;\beta_0) &= -A^{-1}\frac{\partial se_1(m_0(X_t\beta_0;\beta_0))'}{\partial m}W_1 \times \\ &\quad se_1(m_0(X_t\beta_0;\beta_0))(1+o_p(1))\end{aligned} \tag{25}$$

According to Section 2.2:

$$\begin{aligned}se_1(m_0(X_t\beta_0;\beta_0)) &= \frac{1}{Th_1h_2}\sum_{s=1}^{T}(m_0\mathbf{R}_{s+12}-1)k(\frac{X_{1s}\beta_{10}-X_{1t}\beta_{10}}{h_1}) \times \\ &\quad k(\frac{X_{2s}\beta_{20}-X_{2t}\beta_{20}}{h_2})\end{aligned} \tag{26}$$

Substitute equation (24) into the second term of equation (21), and define

$$\begin{aligned}p_{t+12,s+12}(X_t,X_s) &= -\frac{1}{h_1h_2}[\mathbf{R}_{t+12}\otimes Z_t]A^{-1}\frac{\partial se_1(m_0(X_t\beta_0;\beta_0))'}{\partial m}W_1 \times \\ &\quad (m_0\mathbf{R}_{s+12}-1)k(\frac{X_{1s}\beta_{10}-X_{1t}\beta_{10}}{h_1}) \times \\ &\quad k(\frac{X_{2s}\beta_{20}-X_{2t}\beta_{20}}{h_2})\end{aligned} \tag{27}$$

Notice that $p_{t+12,s+12}(X_t,X_s) \neq p_{s+12,t+12}(X_s,X_t)$, and in order to write the second term in equation (21) using U-statistics representation, we define a function $q(.)$ which is symmetric in t and s:

$$q_{t+12,s+12}(X_t,X_s) = \frac{p_{t+12,s+12}(X_t,X_s)+p_{s+12,t+12}(X_s,X_t)}{2} \tag{28}$$

Thus, the second term in equation (21) can be rewritten as:

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T}[\mathbf{R}_{t+12}\otimes Z_t](\hat{m}(X_t\beta_0;\beta_0)-m_0(X_t\beta_0;\beta_0)) \\ = &\Big[\frac{2}{T^2}\sum_{t=1}^{T}\sum_{s=t+1}^{T}q_{t+12,s+12}(X_t,X_s)\Big](1+o_p(1))\end{aligned} \tag{29}$$

The bracket term in equation (28) is a representation of U-statistics, and by following the asymptotic behavior of U-statistics in Powell, Stock, and Stoker (1989), one can show that $\sqrt{T}(\hat{\beta} - \beta_0)$ has a limiting normal distribution.

Moreover, let $r(X_t) = E[q_{t+12,s+12}(X_t, X_s)|X_t]$,

$$\begin{aligned}\sqrt{T}(\hat{\beta} - \beta_0) &= -Q^{-1}\left[\frac{\partial se_2(m_0(\beta_0))}{\partial m}\frac{\partial m_0(\beta_0)}{\partial \beta}\right]' W_2\left[\sqrt{T}se_2(m_0(\beta_0)) + \right. \\ &\left. \frac{1}{\sqrt{T}}\sum_{t=1}^{T} r(X_t)\right] + o_p(1) \qquad (30)\end{aligned}$$

Denote

$$C = \frac{\partial se_2(m_0(\beta_0))}{\partial m}\frac{\partial m_0(\beta_0)}{\partial \beta} \qquad (31)$$

and since

$$se_2(m_0(\beta_0)) = \frac{1}{T}\sum_{t=1}^{T}(m_0(X_{1t}\beta_{10}, X_{2t}\beta_{20})\mathbf{R}_{t+12} - 1) \otimes Z_t \qquad (32)$$

and one can verify that:

$$E\left[C'W_2(m_0(X_{1t}\beta_{10}, X_{2t}\beta_{20})\mathbf{R}_{t+12} - 1) \otimes Z_t\right] = 0; \qquad (33)$$

$$E\left[C'W_2 r(X_t)\right] = 0 \qquad (34)$$

Finally, we show that
$\sqrt{T}(\hat{\beta} - \beta_0) \sim \mathcal{N}(0, Q^{-1}C'W_2\Sigma W_2 C Q^{-1})$;
where

$$\Sigma = Cov\left[\sqrt{T}se_2(m_0(\beta_0)) + \frac{1}{\sqrt{T}}\sum_{t=1}^{T} r(X_t)\right] \qquad (35)$$

References

Ait-Sahalia, Y.and Michael Brandt, 2001,"Variable Selection for Portfolio Choice," Journal of Finance 56, 1297-1351.

Andrews, W.K., 1999, "Consistent Moment Selection Procedures for Generalized Method of Moments Estimations," Econometrica 67, 543-564.

Ang, A. and Monika Piazzesi, 2003, "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," Journal of Monetary Economics 50, 745-787.

Balduzzi, P., Elton, E., and C. Green, 2001, "Economic News and Bond Prices: Evidence from the U.S. Treasury Market," Journal of Financial and Quantitative Analysis 36, 523-543.

Bansal, R. and S. Viswanathan, 1993, "No-Arbitrage and Arbitrage Pricing: A New Approach," Journal of Finance 48, 1231-1262.

Brandt, M., 1999, "Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach," Journal of Finance 54, 1609-1646.

Campbell, J.Y., 1987, "Stock Returns and Term Structure," Journal of Financial Economics 18, 373-399.

Campbell, J.Y. and Robert J. Shiller, 1991, "Yield Spreads and Interest Rates Movements: A Bird's Eye View," Review of Economic Studies 58, 495-514.

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

Carroll, R.J., Fan, J., and M.P. Wand, 1995, "Generalized Partially Linear Single-index Models," Working paper.

Chapman, D., 1997, "Approximating the Asset Pricing Kernel," Journal of Finance 52(4), 1383.

Chaudhuri, P. and J.S. Marron, 1999, "SiZer for Exploration of structures in Curves," Journal of the American Statistical Association 94, 807-823.

Cochrane, J.H. and Monika Piazzesi, 2005, "Bond Risk Premia," American Economic Review 95, 138-160.

Cox, J.C., Ingersoll, J.E., and S.A. Ross, 1985, "A Theory of the Term Structure of Interest Rates," Econometrica 53, 385-407.

Dai, Q. and Kenneth Singleton, 2000, "Specification Analysis of Affine Term Structure Models," Journal of Finance 55, 1943-1978.

Dewachter, H., Lyrio, M., and K. Maes, 2001, "Estimation of a Joint Model for the Term Structure of Interest Rates and the Macroeconomy," Working paper.

Dewyer, G.P. and R.W. Hafer, 1989, "Interest Rates and Economic Announcements," Federal Reserve Bank of St. Louis Review 71, 34-46.

Diebold, F.X. and Canlin Li, 2006, "Forecasting the Term Structure of Government Bond Yields, " Journal of Econometrics 130, 337-364.

Diebold, F.X., Rudebusch, G.D., and S.B. Aruoba, 2004, "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach," Journal of Econometrics, forthcoming.

Donkers, B. and M. Schafgans, 2004, "A Derivative Based Estimator For Semiparametric Index Models," Working paper.

Duffie, D. and R. Kan, 1996, "A Yield-factor Model of Interest Rates," Mathematical Finance 6(4), 379-406.

El Karoui, Nichole, Myneni, R., and R. Viswanathan, 1993, "Arbitrage Pricing and Hedging of Interest Rates Claims with State Variables," working paper.

Estrella, A.and F.S. Mishkin, "The Term Structure of Interest Rates and its Role in Monetary Policy for the European Central Bank," NBER working paper 5279.

Evans, C.L., D.A. Marshall, 1998, "Monetary Policy and the Term Structure of Nominal Interest Rates: Evidence and Theory," Working paper.

Evans, C.L. and D. Marshall, 2001, "Economic Determinants of the Nominal Treasury Yield Curve," working paper.

Fama, E. and Robert Bliss, 1987, "The Information in Long-Maturity Forward Rates," American Economic Review 77, 680-692.

Fama, E. and Kenneth French, 1993, "Common Risk Factors in the Returns on Stocks and Bonds," Journal of Financial Economics 33, 3-56.

Friedman, J.H. and W. Stuetzle, 1981, "Projection Pursuit Regression," J. Amer. Statist. Assoc. 76, 817.

Gallant, A.R., 1987, Nonlinear statistical models. John Wiley and Sons, New York.

Ghysels, E.and Serena Ng, 1998, "A Semi-parametric Factor Model of Interest Rates and Tests of the Affine Term Structure," The Review of Economics and Statistics 80, 535-548.

Harvey, C. and Wayne Ferson, 1991, "The Variation of Economic Risk Premiums," Journal of Political Economy 99, 285-315.

Knez, P.J., Litterman, R., Jose Scheinkman, 1994, "Explorations into Factors Explaining Money Market Returns," Journal of Finance 49, 1861-1882.

Litterman, R., and Jose Scheinkman, 1991, "Common Factors Affecting Bond Returns," Journal of Fixed Income 1(1) , 54-61.

McQueen, G. and V.V. Roley, 1993, "Stock Prices, News, and Business Conditions," Review of Financial Studies 6, 683-707.

Newey, W.K. and Daniel McFadden, 1994, Handbook of Econometrics (edited by R.F. Engle and D.L. McFadden) 6, 2111-2245.

Powell, J.L., Stock, J.H., and T. Stoker, 1989, "Semiparametric Estimation of Index Coefficients," Econometrica 57, 1403-1430.

Ross, S., 1976, "The Arbitrage Theory of Capital Asset Pricing," Journal of Economic Theory 13, 341-360.

Silverman, B.W., 1986, Density Estimation, London, Chapman and Hall.

Stoker, T., 1986, "Cocsistent Estimation of Scaled Coefficients," Econometrica 54, 1461-1481.

Stoker, T., 1992, Lectures on Semi-parametric Econometrics, CORE Lecture Series, CORE Foundation, Louvain-la-Neuve, Belgium.

Vasicek, O., 1977, "An Equilibrium Characterization of the Term Structure," Journal of Financial Economics 5, 177-188.

Table 1: Sample Statistics of US Zero-Coupon Bond Returns

Maturity(yr.)	Mean	Std.Dev.	Skewness	Kurtosis
		1-year Horizon		
2	1.0681	0.0349	1.2223	4.8539
3	1.0714	0.0449	1.1131	4.7272
4	1.0733	0.0555	0.9665	4.4198
5	1.0730	0.0647	0.8730	4.2530
		2-year Horizon		
3	1.0702	0.0317	1.0752	3.9077
4	1.0726	0.0384	1.0342	3.7456
5	1.0733	0.0449	0.9501	3.6334
		3-year Horizon		
4	1.0716	0.0300	0.9792	3.5500
5	1.0729	0.0349	0.8913	3.1764

Values shown are the sample statistics for the holding-period returns on 2-5 year bonds at 1-3 year forecasting horizons. The holding-period returns are computed from the unsmoothed Fama-Bliss zero-coupon prices.

Table 2: Prediction Errors for 1-year-ahead Bond Returns

	2-year	3-year	4-year	5-year
	Term-structure Variables			
1-year yield	0.0166	0.0289	0.0405	0.0500
2-year forward rate	0.0148	0.0266	0.0378	0.0471
3-year forward rate	0.0164	0.0264	0.0362	0.0454
4-year forward rate	0.0180	0.0290	0.0398	0.0485
5-year forward rate	0.0153	0.0257	0.0367	0.0453
	Nominal Macro Variables			
CPI	0.0271	0.0370	0.0471	0.0560
PPI	0.0304	0.0397	0.0498	0.0586
M2	0.0309	0.0393	0.0498	0.0575
	Real Macro Variables			
IP	0.0306	0.0390	0.0488	0.0574
EMPLOY	0.0322	0.0405	0.0501	0.0586
PCE	0.0316	0.0405	0.0504	0.0589
HS	0.0303	0.0396	0.0499	0.0587

Shown are the prediction errors in RMSE for 1-year-ahead bond returns at 2-5 years of maturity. Each pricing error is computed as the difference between the actual and the model-based bond return by each of the underlying state variables. The details of the computation is given in Section 4.1.

Table 3: GMM Estimates on Index Coefficients for Return-forecasting Models

	Two-index Factor Model			Linear Model
Forecasting Horizon	1-yr.	2-yr.	3-yr.	1-yr.
Const.				0.9367
				(0.0008)
1-year yield	0.0694	-0.1615	-0.0614	0.0180
	(0.6747)	(0.3241)	(1.3558)	(0.0039)
2-year forward	0.0787	0.7427	0.5111	-0.0393
	(1.1769)	(0.9747)	(3.8458)	(0.0060)
3-year forward	0.8090	0.3628	0.7166	-0.034
	(0.2665)	(0.7973)	(3.3188)	(0.0042)
4-year forward	0.2595	0.3523	0.24	-0.0106
	(0.7087)	(1.3216)	(1.6461)	(0.0032)
5-year forward	-0.5168	-0.4082	-0.4048	0.0319
	(0.2241)	(0.6575)	(0.7647)	(0.0025)
CPI	-0.4837	-0.5192	-0.4826	0.0019
	(0.6138)	(0.3552)	(2.3902)	(0.0034)
PPI	0.2661	0.1987	0.1838	0.0025
	(0.5514)	(0.4936)	(1.8274)	(0.0031)
M2	-0.2379	-0.1624	-0.1124	0.0015
	(0.3644)	(0.2785)	(0.9842)	(0.0009)
IP	0.1824	0.1626	0.1618	-0.0046
	(0.5143)	(0.3201)	(2.7580)	(0.0017)
EMPLOY	-0.7633	-0.7679	-0.799	0.0089
	(0.4109)	(0.2722)	(1.2581)	(0.0014)
PCE	-0.0738	-0.0774	-0.0944	0.0041
	(0.5915)	(0.4941)	(0.4502)	(0.0016)
HS	-0.1318	-0.2063	-0.2171	0.0044
	(0.1479)	(0.2048)	(0.7172)	(0.0011)

Column 1-3 report the coefficient estimates on a two-index factor model for predicting 1-, 2-, and 3-year-ahead bond returns. The last column shows the coefficient estimates on a linear model. All models are estimated using GMM optimal weighting matrix for the heteroskedasticity and serial correlation in the sample moments. The truncation lag parameter $J = 12$ in the weighting matrix. A detailed description of the estimation procedures is given in Section 2.2.

Table 4: Forecasting Performances for 1-year-ahead Bond Returns

	Model				
Maturity	1	2	3	t-stat (1-2)	t-stat (2-3)
2-year	0.0161	0.0157	0.0173	0.0508	1.1096
3-year	0.0288	0.0208	0.0234	4.1478*	2.5311*
4-year	0.0405	0.0299	0.0324	4.3054*	2.4213*
5-year	0.0500	0.0391	0.0416	4.3096*	2.3816*

Column 2-4 show the prediction errors (in RMSE) of 3 models: a one-index model(Model 1), a two-index model (Model 2), and a linear (Model 3) model. The pricing kernel in one-index model is assumed only dependent on the term-structure variables. While the two-index and linear models include both term-structure and macro variables. Column 5 shows the Diebold-Mariano test statistics on the forecasting powers of Model 1 and 2. The last column reports the test statistics for Model 2 and 3. The DM t statistics are computed with the adjustment for heteroskedasticity and serial correlation in the pricing errors.

Table 5: Forecasting Performances for 2-year-ahead Bond Returns

Maturity	One-index Model	Two-index Model	t-stat
3-year	0.0117	0.0103	1.2407
4-year	0.0201	0.0135	3.0755*
5-year	0.0274	0.0196	3.1385*

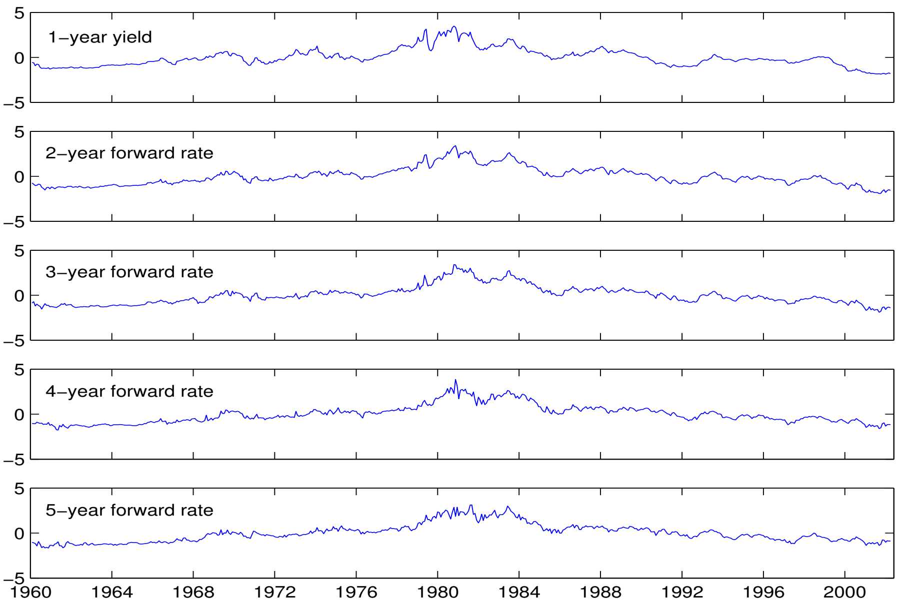
Column 1-2 show the prediction errors (in RMSE) based on one- and two-index models, respectively. The one-index model is only dependent on the term-structure variables, and the two-index model uses both term-structure and macro variables. Column 3 shows the Diebold-Mariano test statistics on the forecasting powers of the models. The t statistics is computed with the adjustment for heteroskedasticity and serial correlation in the pricing errors.

Table 6: Forecasting Performances for 3-year-ahead Bond Returns

Maturity	One-index Model	Two-index Model	t-stat
4-year	0.0088	0.0056	2.1212*
5-year	0.0145	0.0107	2.045*

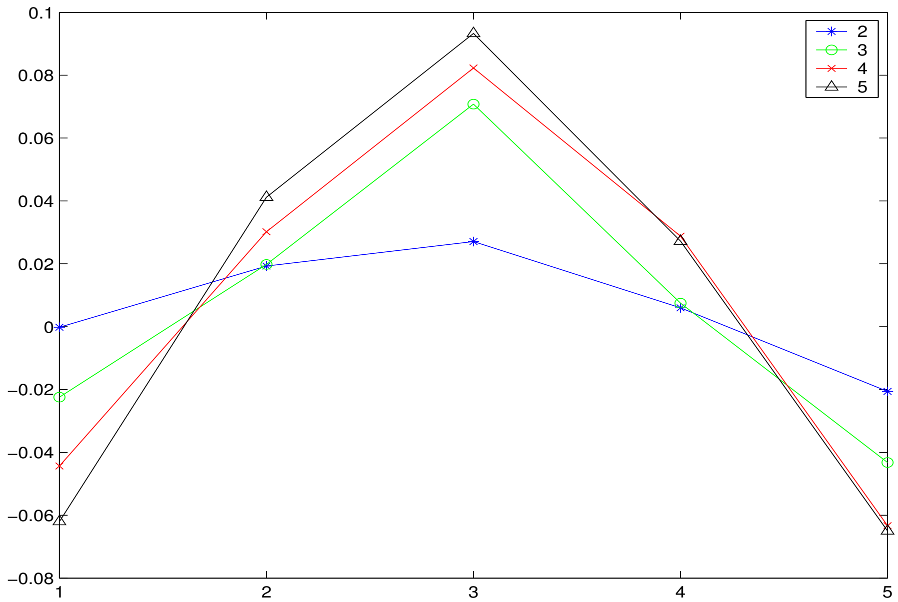
Column 1-2 show the prediction errors (in RMSE) based on one- and two-index models, respectively. The one-index model is only dependent on the term-structure variables, and the two-index model uses both term-structure and macro variables. Column 3 shows the Diebold-Mariano test statistics on the forecasting powers of the models. The t statistics is computed with the adjustment for heteroskedasticity and serial correlation in the pricing errors.

Figure 1: Times Series of 1-year Yield and 2-5 Year Forward Rates



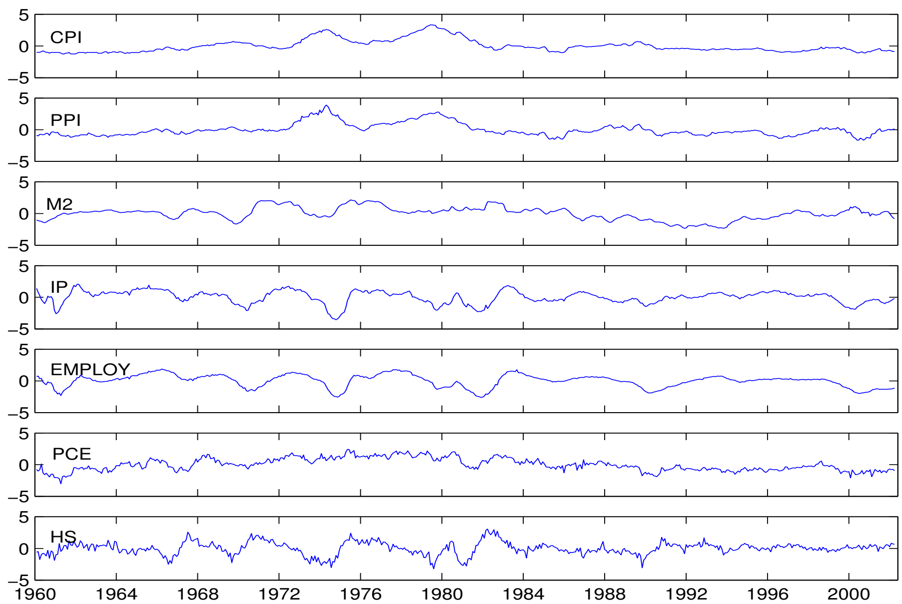
Shown are the time series of the standardized 1-year yield and 2-5 forward rates during the sample period of January 1960 to December 2003.

Figure 2: Factor Loadings of the Term-structure Variables



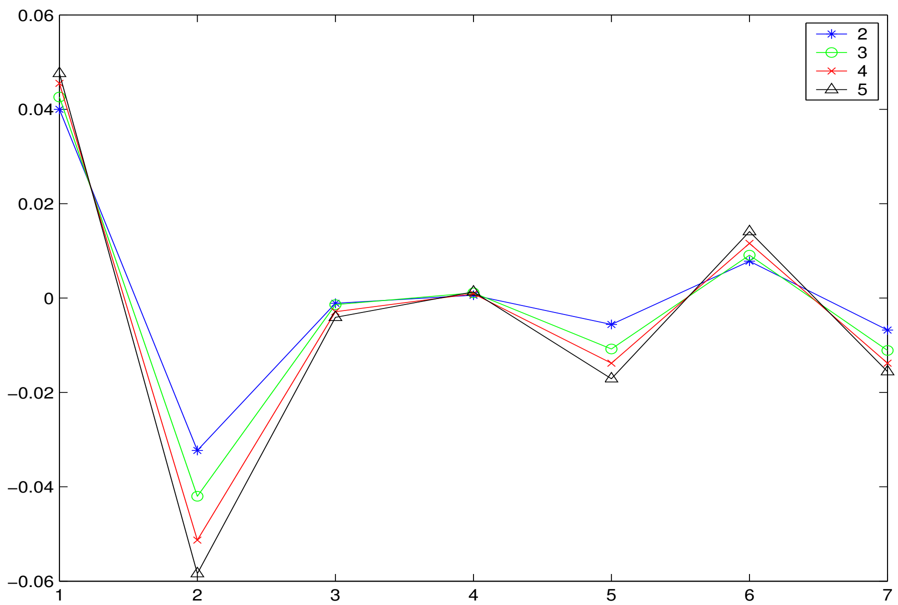
Shown are the OLS coefficients from the projections 2-5 year bond returns on 1-year yield and 2-5 year forward rates.

Figure 3: Times Series of Macro variables



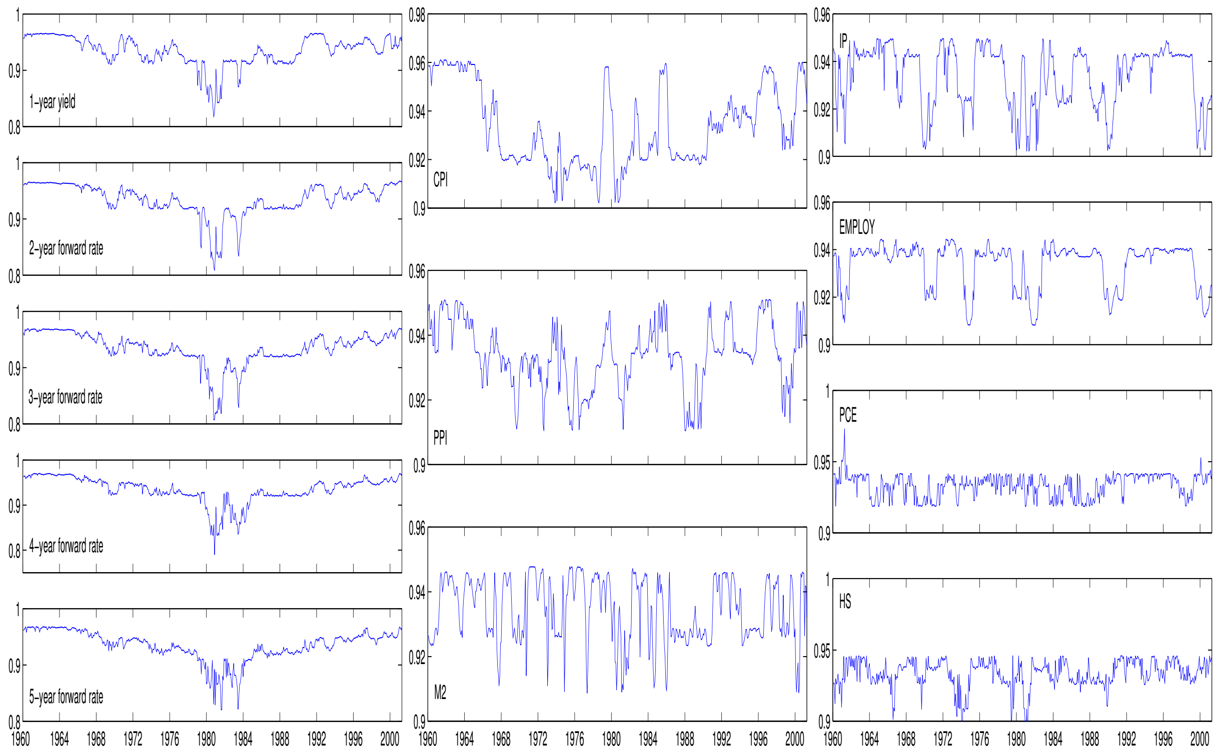
Plotted are the time series of the nominal and real macro variables during the sample period of January 1960 to December 2003.

Figure 4: Factor Loadings of the Macro Variables



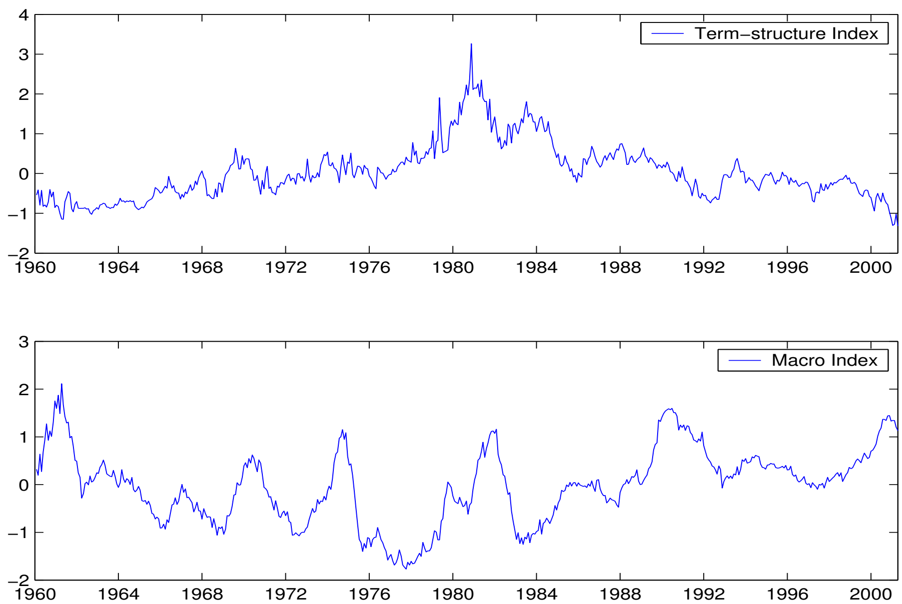
Shown are the OLS coefficients from the projections 2-5 year bond returns on the selected macro variables.

Figure 5: Pricing Kernels from Single-factor Nonlinear Pricing Models



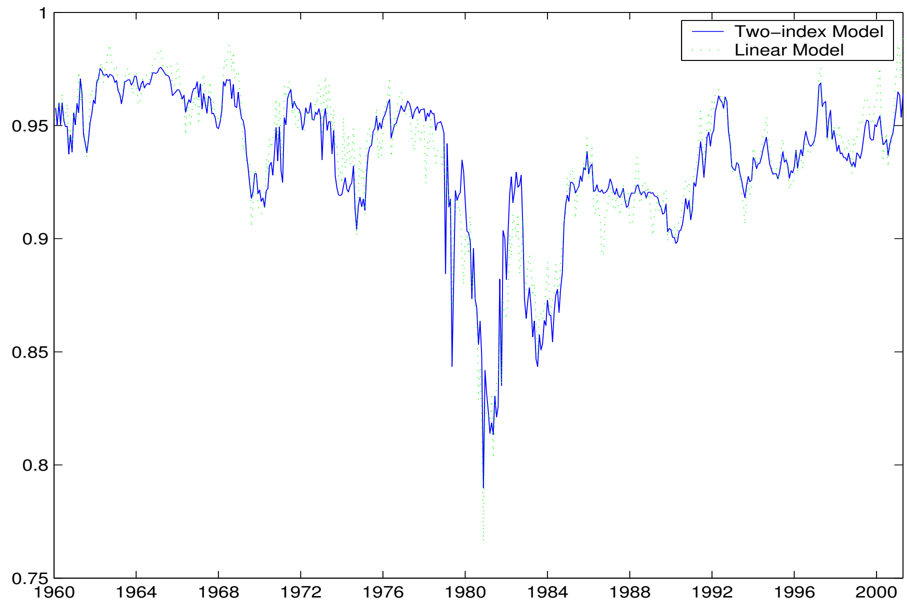
Plotted are the time series of the pricing kernels, estimated based on each of the underlying variables: 1-year yield, 2-5 forward rates, nominal, and real macro variables.

Figure 6: Times Series of the Index Factors



The first panel shows the term-structure index dynamics, and the second panel plots the macro index dynamics. Both indices are simultaneously estimated from a two-index nonlinear model for predicting future bond returns at all maturities.

Figure 7: Pricing Kernels from Two-index and Linear Models



The pricing kernel according to the two-index model is plotted in solid line. The dotted line is for the pricing kernel from the model with the linearity constraint.