WHAT DOES THE YIELD CURVE TELL US ABOUT GDP GROWTH?

Andrew Ang
Monika Piazzesi
Min Wei

What Does the Yield Curve Tell us about GDP Growth?
Andrew Ang, Monika Piazzesi, and Min Wei
NBER Working Paper No. 10672
August 2004
JEL No. E4, E5, G1

ABSTRACT

A lot, including a few things you may not expect. Previous studies find that the term spread forecasts GDP but these regressions are unconstrained and do not model regressor endogeneity. We build a dynamic model for GDP growth and yields that completely characterizes expectations of GDP. The model does not permit arbitrage. Contrary to previous findings, we predict that the short rate has more predictive power than any term spread. We confirm this finding by forecasting GDP out-of-sample. The model also recommends the use of lagged GDP and the longest maturity yield to measure slope. Greater efficiency enables the yield-curve model to produce superior out-of-sample GDP forecasts than unconstrained OLS regressions at all horizons.

Andrew Ang
Columbia Business School
3022 Broadway
805 Uris
New York, NY 10027
and NBER
aa610@columbia.edu

Monika Piazzesi
University of Chicago
Graduate School of Business
1101 East 58th Street
Chicago, IL 60637
and NBER
mpiazzesi@gsb.uchicago.edu

Min Wei
Board of Governors of the Federal Reserve
Division of Monetary Affairs
Washington, DC 20551
min.wei@frb.gov

1 Introduction

The behavior of the yield curve changes across the business cycle. In recessions, premia on long-term bonds tend to be high and yields on short bonds tend to be low. Recessions, therefore, have upward sloping yield curves. Premia on long bonds are countercyclical because investors do not like to take on risk in bad times. In contrast, yields on short bonds tend to be procyclical because the Federal Reserve lowers short yields in recessions in an effort to stimulate economic activity. For example, for every 2 percentage point decline in GDP growth, the Fed should lower the nominal yield by 1 percentage point according to the Taylor (1993) rule.

Inevitably, recessions are followed by expansions. During recessions, upward sloping yield curves not only indicate bad times today, but better times tomorrow. Guided from this intuition, many papers predict GDP growth in OLS regressions with the slope of the yield curve, usually measured as the difference between the longest yield in the dataset and the shortest maturity yield.[1] The higher the slope or term spread, the larger GDP growth is expected to be in the future. Related work by Fama (1990) and Mishkin (1990a and b) shows that the same measure of slope predicts real rates. The slope is also successful at predicting recessions with discrete choice models, where a recession is coded as a one and other times are coded as zeros (see Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1998). The term spread is also an important variable in the construction of Stock and Watson (1989)'s leading business cycle indicator index. Despite some evidence that parameter instability may weaken the performance of the yield curve in the future (see comments by Stock and Watson, 2001), it has been amazingly successful in these applications so far. For example, every recession after the mid-1960's was predicted by a negative slope - an inverted yield curve - within 6 quarters of the impending recession. Moreover, there has been only one "false positive" (an instance of an inverted yield curve that was not followed by a recession) during this time period.

Hence, the yield curve tells us something about future economic activity. We argue there is much more to learn from the yield curve when we explicitly model its joint dynamics with GDP growth. Our dynamic model also rules out arbitrage possibilities between bonds of different maturities and thus imposes more structure than the unrestricted OLS regression framework previously used in the literature. While OLS regressions show that the slope has predictive power for GDP, it is only an incomplete picture of the yield curve and GDP. For example, we would expect that the entire yield curve, not just the arbitrary maturity used in the construction of the term spread, would have predictive power. Using information across the whole yield curve, rather than just the long maturity segment, may lead to more efficient and more accurate forecasts of GDP. In an OLS

---

[1]See, among others, Harvey (1989 and 1993), Laurent (1988), Stock and Watson (1989), Chen (1991), Estrella and Hardouvelis (1991), Estrella and Mishkin (1998), Dotsey (1998), and Hamilton and Kim (2002), who regress GDP growth on term spreads on U.S. data. Jorion and Mishkin (1991), Harvey (1991), Estrella and Mishkin (1997), Plosser and Rouwenhorst (1994), Bernard and Gerlach (1998), among others, run a predictive GDP regression on international data. Other traditional GDP forecasting variables include Stock and Watson (1989)'s leading business cycle index and the consumption-output ratio in Cochrane (1994).

framework, since yields of different maturities are highly cross-correlated, it is difficult to use multiple yields as regressors because of collinearity problems. This collinearity suggests that we may be able to condense the information contained in many yields down to a parsimonious number of variables. We would also like a consistent way to characterize the forecasts of GDP across different horizons to different parts of the yield curve. With OLS, this can only be done with many sets of different regressions. These regressions are clearly related to each other, but there is no obvious way in an OLS framework to impose cross-equation restrictions to gain power and efficiency.

Our approach in this paper is to impose the absence of arbitrage in bond markets to model the dynamics of yields jointly with GDP growth. The assumption is reasonable in a world of hedge funds and large investment banks. Traders in these institutions take large bond positions that eliminate arbitrage opportunities arising from bond prices that are inconsistent with each other in the cross-section and with their expected movements over time. Based on the assumption of no-arbitrage, we build a model of the yield curve in which a few yields and GDP growth are observable state variables. This helps us to reduce the dimensionality of a large set of yields down to a few state variables. The dynamics of these state variables are estimated in a Vector Autoregression (VAR). Bond premia are linear in these variables and are thus cyclical, consistent with findings in Cochrane and Piazzesi (2002). Our yield-curve model leads to closed-form solutions for yields of any maturity which belong to the affine class of Duffie and Kan (1996).

We address two main issues regarding the predictability of GDP in the no-arbitrage framework. We first demonstrate that the yield curve model can capture the same amount of conditional predictability that is picked up by simple OLS regressions. However, OLS approaches and the forecasts implied by our model yield different predictions. Using the term structure model, we attribute the predictive power of the yield curve to risk premium and Expectations Hypothesis components and show how the model can generate the OLS coefficient patterns observed in data in small samples. The second question we investigate is how well GDP growth can be predicted out-of-sample using term structure information, where the coefficients in the prediction function are either estimated directly by OLS, or indirectly, by transforming the parameter estimates of our yield curve model.

We find that our yield-curve model has several main advantages over unrestricted OLS specifications. First, the theoretical framework advocates using long horizon forecasts implied by Vector Autoregressions (VARs). While some authors, like Dotsey (1998) and Stock and Watson (2001), use several lags of various instruments to predict GDP, to our knowledge the efficiency gains from long-horizon VAR forecasts have not previously been considered by the literature. Second, the estimated yield-curve model guides us in choosing the maturity of the yields that should be most informative about future GDP growth. Our results show that if we are to choose a single yield spread, the model recommends the use of the longest yield to measure the slope, regardless of the forecasting horizon. This result requires the framework of the term structure model to determine, in closed form, the predictive ability of any combination of yields of any maturity. Third, the model predicts that the nominal short rate contains more information about GDP growth than any yield spread. This finding stands in contrast to unconstrained OLS regressions

which find the slope to be more important. This prediction of the model is confirmed by forecasting GDP growth out-of-sample. Finally, our model is a better out-of-sample predictor of GDP than unrestricted OLS. This finding is independent of the forecasting horizon and of the choice of term structure regressor variables. The better out-of-sample performance from our yield-curve model is driven by the gain in estimation efficiency due to the reduction in dimensionality and imposing the cross-equation restrictions from the term structure model.

The rest of this paper is organized as follows. Section 2 documents the relationship between the yield curve, GDP growth and recessions. Section 3 describes the yield-curve model and the estimation method. Section 4 presents the empirical results. We begin by discussing the parameter estimates, perform specification tests, and then show how the model characterizes the GDP predictive regressions. We also present a series of small sample simulation experiments to help interpret the results of the model. We show that the model leads to better out-of-sample forecasts of GDP than unconstrained OLS regressions in Section 5. Section 6 concludes. We relegate all technical issues to the Appendix.

2   Data and Motivation

We use zero-coupon yield data for maturities 1, 4, 8, 12, 16 and 20 quarters from CRSP spanning 1952:Q2 to 2001:Q4. The 1-quarter rate is from the CRSP Fama risk-free rate file. All other bond yields are from the CRSP Fama-Bliss discount bond file. All yields are continuously compounded and expressed at a quarterly frequency. We denote the n-quarter yield as $y_t^{(n)}$. In their appendix, Fama and Bliss (1987) comment that data on long bonds before 1964 may be unreliable because there were few traded bonds with long maturities during the immediate post-war period 1952-1964. Fama and Bliss, and most of the papers in the literature looking at the predictive ability of the term spread for GDP, therefore choose to start their sample period in 1964. Our model will emphasize the role of the short rate in forecasting GDP growth, rather than the term spread. To ensure that we do not end up choose the sample period that favors our model, we report results for the sample where the term spread asserts its predictive ability most strongly, which is the post-1964 sample.[2]

We measure economic activity by real GDP growth rates. Data on real GDP is seasonally adjusted, in billions of chained 1996 dollars, from the FRED database (GDPC1). We denote log real GDP growth from t to $t + k$ expressed at a quarterly frequency as:

$$g_{t \to t+k} = 1/k \cdot (\log GDP_{t+k} - \log GDP_t).$$

For the special case of 1-quarter ahead GDP growth, we denote $g_{t \to t+1} \equiv g_{t+1}$. GDP

---

[2]The results for the sample 1952:Q2-2001:Q4 are qualitatively similar, but favor our model more. The results are available upon request. While Fama and Bliss advocate starting analysis with zero coupon bonds from 1964, others use all the yields available from the post-Treasury Accord period from 1952 onwards (see, for example, Campbell and Shiller, 1991).

numbers are subject to many revisions. We choose to use the revised figures rather than a real-time data set because we are forecasting what actually happens to the economy, not preliminary announcements of economic growth. GDP growth is significantly mean-reverting; one-quarter GDP growth, g_{t+1}, has an autocorrelation of 30%.

The extant literature focuses on using the longest available term spread to forecast GDP. In our sample, the longest 5-year term spread averages 0.99% over the sample period, reflecting the normal upward sloping pattern of the yield curve. Many authors, like Harvey (1991 and 1993) comment that recessions are often preceded by inverted yield curves. Table 1 shows that there are nine recessions during the post-1952 period. All except the first three recessions are preceded by negative term spreads. Table 1 shows that every recession since the 1964 start date, advocated by Fama and Bliss (1987), has been preceded by an inverted yield curve. There is one notable inversion from 1966:Q3-1966:Q4 which is not followed by an NBER recession, but this period is followed by a period of relatively slower GDP growth. While all the post-1964 recessions are preceded by inverted yield curves, the initial lead time between the onset of the inversion and the start of the NBER recession varies between two to six quarters. Between the time that the yield curve becomes inverted, it may stay inverted or return to its normal upward sloping shape before the onset of the recession. For example, the yield curve became inverted in 1973:Q2 and stayed inverted at the onset of the 1973:Q4-1975:Q1 recession, while the yield curve was upward sloping over the 2001:Q1-2002:Q1 recession (having been downward sloping from 2000:Q3-2000:Q4).

We can formalize the predictive power of the yield spread for economic activity in a predictive regression of the form:

$$g_{t \rightarrow t+k} = \alpha_k^{(n)} + \beta_k^{(n)} \left(y_t^{(n)} - y_t^{(1)} \right) + \varepsilon_{t+k,k}^{(n)}, \tag{1}$$

where future GDP growth for the next k quarters is regressed on the n-maturity term spread. Numerous authors have run similar regressions, but usually involving only very long spreads (5 or 10 years). Figure 1 plots 4-quarter GDP growth and the 5-year term spread lagged by 4 quarters together in its upper panel. For example, in 1990, we plot GDP growth from 1989-1990 together with the term spread at 1989. This corresponds to the two variables on the LHS and RHS of equation (1) taking values $k = 4$ and $n = 20$, respectively. The plot indicates that periods with high (low) spreads are typically followed by high (low) GDP growth. (For now, please ignore the lower panel of Figure 1, which shows another interesting correlation between 4-quarter GDP growth and lagged information from the yield-curve: the 1-quarter nominal rate).

Table 2 reports the results of regression (1) over 1964:Q1-2001:Q4. In regression (1), the long horizons of GDP growth on the left-hand side means that overlapping periods are used in the estimation, which induces moving average error terms in the residual. We use Hodrick (1992) standard errors to correct for heteroskedasticity and the moving average error terms, which Ang and Bekaert (2001) show have negligible size distortions. Inappropriate standard errors, like Newey-West (1987) standard errors, can vastly overstate the predictability of GDP growth from the term spread.

The literature concentrates on using long term spreads to predict GDP growth. Hence, in Table 2, the last two columns under the 5-year term spread list the known result that the long-term spread significantly predicts GDP growth. Estrella and Mishkin (1996) document that a large number of variables have some forecasting ability one quarter ahead, like the Stock-Watson (1989) index. But, in predicting recessions two or more quarters in the future, the term spread dominates all other variables and the dominance increases as the forecasting horizon increases. Since yields of different maturities are highly correlated, and movements of yields of different maturities are restricted by no-arbitrage, we expect that other yields might also have forecasting power.

Table 2 shows that the whole yield curve has significant predictive power for long-horizon GDP growth. In particular, the 16- and 20-quarter spreads significantly predict GDP growth 1-quarter ahead, and all the term spreads significantly predict GDP growth 4-quarters ahead. The predictability remains strong at 2 years out, but weakens at a 3-year forecasting horizon. The predictive power of the yield curve for GDP growth also varies with the maturity of the yields used as RHS variables. For example, while the 5-year term spread significantly predicts GDP growth at all horizons, the 1-year term spread significantly forecasts only GDP growth at 1- to 2-year horizons.

We use the regression in equation (1) as a useful starting point for showing the strong ability of the yield curve to predict future economic growth. However, Table 2 shows that the entire yield curve has predictive ability, but the predictive power differs across maturities and across forecasting horizons. Since yields are persistent, using the information from one particular forecasting horizon should give us information about the predictive ability at other forecasting horizons. Hence, we should be able to use the information for a 1-quarter forecasting horizon regression in our estimates of the slope coefficients from a 12-quarter forecasting horizon regression. Regression (1) only uses one term spread of an arbitrary maturity, but we may be able to improve forecasts by using combinations of spreads. However, the variation of yields relative to each other cannot be unrestricted, otherwise arbitrage is possible. We seek to incorporate these no-arbitrage restrictions using a yield-curve model to forecast GDP. This is a more efficient and powerful method than merely examining term spreads of arbitrary maturity as regressors in Table 2.

3   Model

Our yield-curve model is set in discrete time. The data is quarterly, so that we interpret one period to be one quarter. The nominal riskfree rate, $y^{(1)}$, is therefore the 1-quarter rate. We use a model with yield curve factors, augmented by including observable quarterly real GDP growth $g_t = \log GDP_t - \log GDP_{t-1}$ as the last factor. We collect these factors in a vector $(K+1) \times 1$ vector X_t, which has K term structure factors. The vector of state variables follows a Gaussian Vector Autoregression with one lag:

$$X_t = \mu + \Phi X_{t-1} + \Sigma \varepsilon_t, \qquad (2)$$

6

with $\varepsilon_t \sim$IID $\mathcal{N}(0, I)$, and μ is a $(K+1) \times 1$ vector and Φ is a $(K+1) \times (K+1)$ matrix.

Risk premia on bonds are linear in the state variables. More precisely, the pricing kernel is conditionally log-normal,

$$m_{t+1} = \exp\left(-y_t^{(1)} - \frac{1}{2}\lambda_t^\top \lambda_t - \lambda_t^\top \varepsilon_{t+1}\right), \tag{3}$$

where λ_t are the market prices of risk for the various shocks. The vector λ_t is a linear function of the state variables:

$$\lambda_t = \lambda_0 + \lambda_1 X_t,$$

for a $(K+1) \times 1$ vector λ_0 and a $(K+1) \times (K+1)$ matrix λ_1. We denote the parameters of the model by $\Theta = (\mu, \Phi, \Sigma, \lambda_0, \lambda_1)$.

Our specification has several advantages. First, we use a parsimonious and flexible factor model. This means that we do not need to specify a full general equilibrium model of the economy in order to impose no-arbitrage restrictions. Structural models, like Berardi and Torous (2002), allow the prices of risk to be interpreted as functions of investor preferences and production technologies. While this mapping is important for the economic interpretation of risk premia, a factor approach allows more flexibility in matching the behavior of the yield curve, especially in the absence of a general workhorse equilibrium model for asset pricing.

Second, parameterizing prices of risk to be time-varying allows the model to match many stylized facts about yield curve dynamics. For example, Dai and Singleton (2001) show that (3) allows the yield-curve model to match deviations from the Expectations Hypothesis, but they do not relate yields to movements in macro variables. Finally, we jointly model yield curve factors and economic growth. Ang and Piazzesi (2003) and Piazzesi (2003) demonstrate that directly incorporating both observed macro factors and traditional yield-curve factors is important for capturing the dynamics of yields. For example, Ang and Piazzesi (2003) find that incorporating macro factors allows for better out-of-sample forecasts of yields than only using yield-curve factors. However, the Ang and Piazzesi (2003) model does not allow feed-back between macro factors and term structure factors, so they cannot address the question of forecasting macro factors using term structure information.[3]

The only macroeconomic variable in our model is GDP growth. Of course, we may be omitting other macro variables that contain additional information about future real activity. We consider a variety of other forecasting variables in Section 5. An obvious candidate is inflation (see, for example, King, Stock and Watson, 1995). We do not include inflation in our basic model because we seek to make our results directly comparable to the large literature using only term structure information (mostly the term

---

[3]Recent approaches integrating term structure approaches with macro models have been developed by Wu (2002), Hordahl, Tristani, and Vestin (2003), and Rudebusch and Wu (2003). Only Hordahl, Tristani, and Vestin report output forecasts from a general equilibrium model.

spread) to forecast GDP. Following this literature, our model only uses term structure information and GDP, but our model yields dissimilar predictions to the extant literature. We show that this is due to imposing the structure from a no-arbitrage pricing model, rather than from the confounding effects of including inflation as an additional state variable. Nevertheless, the factor structure (2) allows additional factors to be included by augmenting the VAR. In Section 5, we include inflation in the factor VAR, but this does not provide significantly better out-of-sample forecasts for GDP. This result may be due to the fact that the short rate and the term spread already contain expectations about future inflation. For example, under the null of the Fisher Hypothesis, the short rate is the sum of the real rate and short-term expected inflation. Moreover, Ang and Bekaert (2003) show that inflation is responsible for a large part of the movements in the term spread.

We can solve for the price $p_t^{(n)}$ of an n-period nominal bond at time t by recursively solving the relation:

$$p_t^{(n)} = E_t \left(m_{t+1} p_{t+1}^{(n-1)} \right),$$

with the terminal condition $p_t^{(0)} = 1$. The resulting bond prices are exponential linear functions of the state vector:

$$p_t^{(n)} = \exp \left(A_n + B_n^\top X_t \right),$$

for a scalar A_n and a 3×1 vector B_n of coefficients that are functions of time-to-maturity n. The absence of arbitrage is imposed by computing these coefficients from the following difference equations (see the derivations in Ang and Piazzesi, 2003):

$$A_{n+1} = A_n + B_n^\top \left(\mu - \Sigma \lambda_0 \right) + \frac{1}{2} B_n^\top \Sigma \Sigma^\top B_n \tag{4}$$

$$B_{n+1} = \left(\Phi - \Sigma \lambda_1 \right)^\top B_n - e_1,$$

where $e_1 = [1 \ 0 \ 0]^\top$. The initial conditions are given by $A_1 = 0$ and $B_1 = -e_1$. Bond yields are then affine functions of the state vector:

$$y_t^{(n)} = -\frac{\log p_t^{(n)}}{n} \tag{5}$$
$$= a_n + b_n^\top X_t,$$

for coefficients $a_n = -A_n/n$ and $b_n = -B_n/n$.

Equation (5) enables the entire yield curve to be modelled. This allows us to characterize predictive relationships for yields of any maturity. If there are no risk premia, $\lambda_0 = 0$ and $\lambda_1 = 0$, a local version of the Expectations Hypothesis holds. In this case, yields are simply expected values of future average short rates (apart from some Jensen's inequality terms). From the difference equations (4), we can see that the constant risk premia parameter λ_0 only affects the constant yield coefficient a_n while the parameter λ_1 also affects the factor loading b_n. The parameter λ_0 therefore only impacts average term

8

spreads and average expected bond returns, while λ_1 controls the time variation in term spreads and expected returns.

One caveat is that we assume that the factor dynamics are stable over time. There is evidence for regime shifts in the dynamics of interest rates (see Ang and Bekaert, 2002). Stock and Watson (2002) find that the differences in coefficients across regimes for GDP relationships with other economic time series mostly reflect time-varying volatility and find that the coefficients in the conditional mean of GDP growth are much more stable. Nevertheless, we compare our forecasts to a Hamilton (1989) regime-switching model of GDP growth in Section 5 below.[4]

How Many Yield Curve Factors?

While the model in equations (2) and (3) allows term structure factors and GDP to feed back on each other, the remaining question is how many K yield curve factors should be included? At a quarterly frequency, the first principal component of yields accounts for 97.2% of the variation of yields. Adding the second principal component brings the percentage of yield curve variation to 99.7%. The first two principal components have almost one to one correspondences with the short rate and term spread. The first (second) principal component has a -95.6% (-86.5%) correlation with the short rate (5-year term spread). Hence, even if the first two true yield-curve factors are unknown, the observable short rate and spread are good proxies. Hence, we need at most two factors to capture almost all of the variation of yields at a quarterly frequency. Studies like Knez, Litterman, and Scheinkman (1994) find that the third principal component accounts for approximately 2% of the movements of yields, whereas we find it accounts for less than 0.3%. The reason is that we work at a quarterly frequency, whereas the third principal component is more important at daily and weekly frequencies. Diebold, Rudebusch, and Aruoba (2003) also find that a third curvature factor is unimportant at a monthly frequency. The third principal component is related to heteroskedasticity, and yields exhibit little heteroskedasticity at monthly or quarterly frequencies.

While two yield curve factors are sufficient to model the dynamics of yields, are two yield curve factors sufficient to forecast GDP? A simple way to answer this question is to look at simple OLS regressions of GDP growth onto principal components. Although these unconstrained regressions do not impose any economic structure, they can give us an idea of how many term structure factors are important for predicting GDP growth. Table 3 reports predictive GDP regressions with principal components of yields. The third principal component (PC3) is never significant at any forecasting horizon, but the first two principal components (PC1 and PC2) are significant predictors of GDP.[5] This confirms Hamilton and Kim (2002), who directly use interest rate volatility and find it has little ability to forecast GDP growth. Hence, two yield curve factors are sufficient not only

---

[4]Including regime-switches in this model is computationally intensive, and difficult to implement in out-of-sample forecasts. While recent term structure models allow for regime changes, bond prices can only be computed in closed form under restrictive assumptions (see Ang and Bekaert, 2003; Dai, Singleton, and Yang, 2003).

[5]There is little difference in the R^2s from using principal components versus the short rate, term spread and a curvature transformation $(y^{(1)} - 2y^{(12)} + y^{(20)})$ of yields.

to model the dynamics of yields, but they should be sufficient to capture relationships between the yield curve and future GDP. Note that Table 2 shows that lagged GDP is also important, at least for horizons up to one year. Hence, we specify X_t to be a 3×1 vector, with $K = 2$ term structure factors. While our basic set-up uses 2 term structure factors, we also consider the effect of a higher number of term structure factors for out-of-sample forecasting in Section 5.

Estimating the Model

The most efficient way to estimate the model is to assume that the term structure factors are latent and to use a one-step maximum likelihood estimation. However, this procedure is computationally intensive which makes out-of-sample forecasting difficult. Instead, we develop a two-step procedure that is fast and consistent. The key assumption is that our term structure factors are observable.

We use two factors from the yield curve, the short rate expressed at a quarterly frequency, $y^{(1)}$, to proxy for the level of the yield curve, and the 5-year term spread expressed at a quarterly frequency, $y^{(20)} - y^{(1)}$, to proxy for the slope of the yield curve. With quarterly real GDP growth $g_t = \log GDP_t - \log GDP_{t-1}$ as the last factor, the vector of state-variables can be expressed as:

$$X_t = \left[y_t^{(1)}, \; \left(y_t^{(20)} - y_t^{(1)} \right), \; g_t \right]^\top, \tag{6}$$

and is thus entirely *observable*.[6]

Observable yield-curve factors enable us to use a consistent two-step procedure to estimate the model. The parameters Θ can be partitioned into the parameters μ, Φ and Σ governing the factor dynamics (2) and the risk premia λ_0 and λ_1. In the first step, we estimate the VAR parameters μ, Φ and Σ using standard SUR. In the second step, we estimate λ_0 and λ_1 given the estimates of the VAR parameters from the first step. This is done by minimizing the sum of squared fitting errors of the model subject to the constraint. We compute standard errors for our parameter estimates using the procedure described in the Appendix, which takes into account the two-stage estimation. While the two-step procedure is not as efficient as a one-step maximum likelihood procedure with latent term structure factors, the loss in efficiency turns out to have negligible effects for forecasting GDP growth. We present the estimation details, and a comparison to a latent factor maximum likelihood approach, in the Appendix.

An important detail in the estimation is how we treat the $n = 20$ quarter yield. For this yield, we want to ensure that the following two implications of the model are consistent with each other. First, the model implies that we can price the 5-year bond and compute its yield using equation (5). Second, the model implies that the 5-year yield is the sum of the first two factors in (6). The 5-year yield is thus $e_1^\top X_t + e_2^\top X_t$, where e_i is a 3×1 vector of zeros with a 1 in the i-th element. To ensure that these two implications

---

[6]Using the 2-year, or 4-year, term spread in place of the 5-year term spread as a term structure factor produces almost identical results.

are consistent with each other, we impose the constraints

$$a_{20} = 0 \tag{7}$$
$$b_{20} = e_1 + e_2$$

on the estimation of the market prices of risk. These constraints ensure that our no-arbitrage model and the trivariate VAR in equation (2) have identical implications for the 5-year yield.

Our two-step procedure is consistent, robust and fast, because many parameters are estimated in a VAR. Speed is crucial for out-of-sample forecasting. Moreover, the estimation procedure allows us to cleanly separate several sources of improvement over OLS. We can improve on the OLS predictive regressions for GDP by moving to a dynamic system describing the endogenous evolution of yields and GDP, parsimoniously captured by a finite number of factors. Hence, the first source of the efficiency gains over unconstrained OLS comes from moving from a large VAR with many yields to a low-dimensional factor model. This is the first part of the estimation procedure. The second source of efficiency gain is to describe the movement of yields in an internally consistent way that rules out arbitrage across yields. In the term structure model, risk-adjusted expectations of the short rate are consistent with the cross-section of yields. This is accomplished by estimating the prices of risk in the second part of the estimation procedure.

Forecasting GDP Growth from the Model

The yield equation (5) allows the model to infer the dynamics of every yield and term spread. Once the parameters Θ are estimated, the coefficients a_n and b_n are known functions of Θ. All the yield dynamics are therefore known functions of X. The yield-curve model completely characterizes the coefficients in the regression (1) for any spread maturity and for any forecasting horizon. For example, the predictive coefficient in the regression (1) is given by:

$$\beta_k^{(n)} = \frac{\mathrm{cov}\left(g_{t \to t+k}, y_t^{(n)} - y_t^{(1)}\right)}{\mathrm{var}\left(y_t^{(n)} - y_t^{(1)}\right)}.$$

We can use equation (5) together with the implied long-run forecast for GDP growth from equation (2) to write:

$$E_t[g_{t \to t+k}] = c + 1/k \cdot e_3^\top \Phi (I - \Phi)^{-1}(I - \Phi^k) X_t, \tag{8}$$

where c is a constant term. Hence, we can solve for $\beta_k^{(n)}$ as:

$$\beta_k^{(n)} = \frac{1/k \cdot e_3^T \Phi (I - \Phi)^{-1}(I - \Phi^k)\Sigma_X \Sigma_X^\top (b_n - b_1)}{(b_n - b_1)^\top \Sigma_X \Sigma_X^\top (b_n - b_1)}, \tag{9}$$

where $\Sigma_X \Sigma_X^\top$ is the unconditional covariance matrix of the factors X, which is given by $\mathrm{vec}(\Sigma_X \Sigma_X^\top) = (I - \Phi \otimes \Phi)^{-1}\mathrm{vec}(\Sigma\Sigma^\top)$. We compute the standard errors for $\beta_k^{(n)}$ using the delta method.

The yield-coefficients b_n together with the VAR parameters Φ and Σ (through Σ_X) completely determine the predictive regression coefficients in (9). Note that only the factor loadings b_n, and not the constants a_n, affect the slope coefficients. The reason is that a_n only determines average term spreads and therefore is absorbed into constant term of the forecasting regression (1). By contrast, the factor loadings b_n determine the dynamic response of yields to GDP growth and vice versa. The loadings therefore impact the slope coefficient of the predictability regressions. The loadings are in turn affected by the risk premia parameter λ_1.

Equation (9) provides some intuition on how the no-arbitrage model provides efficiency gains for forecasts. First, the term $e_3^T \Phi (I - \Phi)^{-1} (I - \Phi^k)$ comes from a VAR. It is the same expression for a k-period ahead forecast from a VAR, which comes from assuming the VAR factor structure in (2) for the 1-quarter yield, the n-quarter spread and GDP growth. For longer forecasting horizons k, the regression coefficient (9) exploits the long-horizon forecasts of the VAR, just as in Campbell and Shiller (1988) and Hodrick (1992), who look at predictive regressions that forecast excess returns of the market return using the dividend yield. However, long-horizon VAR forecasts have not previously been exploited in GDP predictive relations.[7]

Second, the yield-coefficients b_n enter equation (9). These coefficients tie down the predictive coefficients to the movements of the factors and reflect the imposition of no-arbitrage restrictions. In our empirical work, we determine the efficiency gains from each of these two components separately. Importantly, the regression coefficients (9) allow us to characterize the relation of future GDP growth to a term spread of any maturity n, and also characterize the response of future GDP at any forecasting horizon k. An important special case arises when $n = 20$. In this case, the yield maturity in (9) happens to be the maturity of the long yield used as factor in our no-arbitrage model. In this case, our two-step estimation procedure implies that the GDP forecasts computed from our yield-curve model are identical to those computed directly from the trivariate VAR in equation (2). For all other term spread maturities, $n \neq 20$, forecasts from the yield-curve model differ from (unconstrained) VAR predictions. The yield-curve model uses the same factor dynamics under no-arbitrage restrictions to compute equation (9) for various maturities. In contrast, an unconstrained VAR has to be re-estimated including these other term spreads.

We can go further than simply forecasting GDP growth with only the term spread. Since GDP growth is autocorrelated, we can also control for the autoregressive nature of GDP growth and augment the regression (1) by including lagged GDP growth. Whereas the simple regression (1) also only uses one yield-curve factor (the term spread) to predict

---

[7]The implied iterated VAR forecasts derived from a term structure model is also convenient because many macro studies of the relationships between monetary policy, yields and real activity work with VARs. Christiano, Eichenbaum, and Evans (1999) provide an excellent survey of this literature. In a contemporaneous paper, Marcellino, Stock and Watson (2004) investigate iterated multi-step forecasts from autoregressive models for a large dataset of macro time series. They do not motivate this approach from a term structure model and so cannot investigate the role of risk premia, or investigate the underlying sources of discrepancy between OLS forecasts and long-horizon VAR forecasts. Schorfheide (2004) examines multi-step VAR forecasts with potentially misspecified VAR models.

GDP and ignores other yield-curve factors (like the first short rate level factor), the yield-curve model can fully characterize each predictive coefficient in the regression:

$$g_{t \to t+k} = \alpha_k^{(n)} + \beta_{k,1}^{(n)} \cdot y_t^{(1)} + \beta_{k,2}^{(n)} \cdot \left(y_t^{(n)} - y_t^{(1)} \right) + \beta_{k,3}^{(n)} \cdot g_t + \varepsilon_{t+k,k}^{(n)}, \tag{10}$$

of which equation (1) is a special case. In equation (10), the subscript k notation on each coefficient denotes the dependence of the coefficients on the forecast horizon of GDP. The superscript (n) notation denotes the dependence of the coefficients on the choice of the n quarter maturity term spread.

To compute the regression coefficients in equation (10), we can write k-quarter GDP growth as $g_{t \to t+k} = \frac{1}{k} e_3^T \sum_{i=1}^{k} X_{t+i}$. We decompose GDP growth into the part that is in the time-t information set and future shocks:

$$
\begin{aligned}
g_{t \to t+k} &= \frac{1}{k} e_3^T \sum_{i=1}^{k} X_{t+i} \\
&= \frac{1}{k} e_3^T \sum_{i=1}^{k} \left(\sum_{j=0}^{i-1} \Phi^j \mu + \Phi^i X_t + \sum_{j=1}^{i} \Phi^{i-j} \Sigma \varepsilon_{t+j} \right) \\
&= c + e_3^T \Phi \widetilde{\Phi}_k X_t + \frac{1}{k} e_3^T \sum_{i=1}^{k} \sum_{j=1}^{i} \Phi^{i-j} \Sigma \varepsilon_{t+j}
\end{aligned}
\tag{11}
$$

where c and $\widetilde{\Phi}_k$ are given by

$$\widetilde{\Phi}_k = \frac{1}{k} \sum_{j=0}^{k-1} \Phi^j = \frac{1}{k} (I - \Phi)^{-1} \left(I - \Phi^k \right)$$

$$c = e_3^T \sum_{i=1}^{k} \widetilde{\Phi}_i \mu = e_3^T (I - \Phi)^{-1} \left(I - \Phi \widetilde{\Phi}_k \right) \mu$$

are constants. Hence the time-t expectation of the regressor is:

$$E_t \left[g_{t \to t+k} \right] = c + e_3^T \Phi \widetilde{\Phi}_k X_t. \tag{12}$$

The regressors in regression (10) are linear combinations of the state variables and can be represented as $A + B X_t$ with $A = \begin{bmatrix} 0 & a_n - a_1 & 0 \end{bmatrix}^T$ and $B = \begin{bmatrix} e_1 & b_n - b_1 & e_3 \end{bmatrix}^T$ where a_n and b_n are the yield coefficients defined in equation (5). Hence, the regression coefficients $\beta_k^{(n)} = [\beta_{k,1}^{(n)}, \beta_{k,2}^{(n)}, \beta_{k,3}^{(n)}]^\top$ implied by the model can be computed as:

$$\beta_k^{(n)} = \left(B \Sigma_X \Sigma_X^T B^T \right)^{-1} \left[B \Sigma_X \Sigma_X^T \widetilde{\Phi}_k^T \Phi^T e_3 \right],$$

where $\Sigma_X \Sigma_X^T$ is the unconditional covariance matrix of the factors X_t with $\text{vec}(\Sigma_X \Sigma_X^T) = (I - \Phi \otimes \Phi)^{-1} \text{vec}(\Sigma \Sigma^T)$. Importantly, when $n = 20$, the forecasts of the yield curve model reduce to long-horizon forecasts implied by a 1-step VAR, due to the factor VAR structure of the model in equation (2) and the constraints on the risk prices imposed in the estimation in equation (7).

Our model also allows us to compute R^2s of each regression specification in closed-form. We detail the computations of R^2s and their standard errors in the Appendix.

4 Empirical Results

4.1 Parameter Estimates

Table 4 reports estimates of the parameters of the observable yield-curve model with GDP growth. High short rates Granger-cause low GDP growth. The significantly negative coefficient (-0.269) in the last row of the Φ matrix is consistent with the Fed raising (lowering) rates to cool (stimulate) economic growth. The term spread does not Granger-cause GDP growth. This finding foreshadows the observation that term spreads behave differently in multivariate regressions compared to univariate regressions (equation (1)) in Table 3. GDP growth remains significantly autocorrelated (0.258), when we control for the short rate and spreads, which are themselves highly persistent. (The quarterly autocorrelation of the short rate is 88%.)

There is little evidence of Granger-causality of GDP growth to short rates or spreads, but shocks to all three factors are significantly correlated. In particular, shocks to the short rate and GDP growth are 22% correlated, and shocks to the spread and GDP growth are -5% correlated. (These numbers are computed using the covariance matrix $\Sigma\Sigma^\top$ of shocks obtained from Table 4.) This implies that the conditional covariance structure plays an important role in bond pricing in equation (4).

Some of the parameter estimates of the risk premia in Table 4 have large standard errors, as is common in many yield curve studies. Average risk premia (λ_0) are estimated imprecisely because yields are very persistent, which makes it hard to pin down unconditional means in small samples. Most of the time variation in the time-varying risk premia λ_1 parameters are, however, significant. Some of this time-variation is captured by the slope of the yield curve. Many term structure studies also find this effect because it reflects the fact that term spreads have predictive power for future holding period returns on bonds (see Fama and Bliss, 1987). When the yield curve is upward sloping, expected returns on long bonds are higher than on short bonds. From equation (4), we can see that the more negative the λ_1 terms, the more positive the loading b_n on long bonds. Hence, more negative λ_1 terms lead to larger positive responses to conditional short rate shocks. Since the time-variation in risk premia impacts the yield-coefficients b_n, they also significantly affect the predictive coefficients $\beta_k^{(n)}$ in equation (9).

4.2 Specification Tests

Matching Unconditional Moments

Table 5 lists means and standard deviations of the short rate, term spreads and GDP growth implied by the yield-curve model, against moments computed from data. (To make the moments easily comparable with other empirical studies, we report annualized numbers.) The standard errors implied by the model are given in parentheses. All the data moments are well within one standard error bound. Thus, our model matches the

unconditional moments of term spreads. Later, we will investigate the ability of our model to capture the conditional moments implied by predictive GDP regressions using term spreads.

Since the model, by construction, matches the dynamics of the 1- and 20-quarter yields, we can perform χ^2 over-identification tests on the various moments on the yields not included in the VAR. Table 6 reports p-values of tests separately matching sample means, variances and first-order autocorrelations. We use a robust Newey-West (1987) covariance estimated from the data by GMM. The last column of Table 6 reports results of these jointly tests for the mean, variance, or first-order autocorrelation of the 4 yields jointly. Table 6 shows that we match the moments of the over-identified yields very closely (with p-values all larger than 50%) when each of the yields are considered separately. The only difficulty the model has is matching the variance of the four yields jointly. Whereas we fail to reject that the implied means of the 4 term spreads not used as factors are jointly equal to their sample means at the 5% level, we reject jointly for the variance test (with a p-value of 0.007). The reason is that the data estimate of the covariance matrix of these moments is nearly singular, because yields are highly cross-correlated.

Matching the R^2s of GDP Regressions

To characterize the explanatory power of the predictive GDP regressions, we begin by computing theoretical R^2s implied by the model in Figure 2. The figure shows the model R^2s for three regressions in each panel. In the first panel, GDP growth is regressed on the term spread; in the middle panel, the regressors are the term spread and lagged GDP growth; and in the last panel, we have a trivariate regression on the short rate, term spread and lagged GDP growth. On the x-axis, we plot the term spread maturity. For the maturity corresponding to one quarter, we plot the R^2 from a regression involving the short rate. The R^2s from OLS regressions are superimposed as empty squares, with one OLS standard error bound given by the black squares. The OLS regressions can be run only for a number of selected yields, but our model derives the predictive power for any spread. Finally, we plot the theoretical R^2s implied from the model. For the short rate, the R^2 is shown as a gray cross and one standard error bounds are shown as black crosses. For 2- to 20-quarter spreads, the theoretical R^2s are shown as solid lines with one standard error bounds as dotted lines. While our model enables the computation of the regression R^2s from placing any horizon GDP growth on the left-hand side, the figure shows only a 4-quarter horizon. We view this horizon as representative; the patterns are qualitatively the same for other horizons.

The OLS one standard error bounds for the OLS regression R^2s easily encompass the R^2s implied by the yield-curve model. The OLS R^2s also lie very near the model-implied R^2s. This is good news for the model – the model does not seem to be at odds with the amount of predictability of GDP observed in the data. The model-implied R^2 in the last panel is flat at 27%. This is because our model has three factors (level, slope, and GDP growth), and two yields together with GDP growth is sufficient to capture exactly the same information as the three factors. The remainder (73%) of the GDP forecast variance cannot be attributed to predictable factor dynamics according to our yield-curve

model. The model R^2s provide us with a guide as to what to expect from the numerous OLS specifications. While running many different OLS regressions with various term spread maturities, together with other predictive variables, may give us some a rough picture of how the whole yield curve may predict GDP, the OLS standard errors are large enough from Figure 2 that only loose characterizations are possible. In contrast, our model provides very clear predictions, in addition to producing tighter standard error bounds.

First, the model-implied R^2s are highest for the short rate specifications. That is, we would expect the greatest predictive power from using a short rate in the regression, and if we were to choose between using a term spread and a short rate, we would prefer to choose the short rate. The theoretical R^2s do not imply that we should only use short rates instead of spreads but instead suggest that we should at least augment the term spread with the short rate for better explanatory power. This is not what Plosser and Rouwenhorst (1994) suggest. They find that in multiple regressions that include the short rate and spread, the short rate, given the spread, has little predictive ability to forecast output for the U.S. We will examine this conflicting finding in detail.

Second, if we were to use a spread to forecast GDP, the theoretical R^2s give us a guide as to which term spread we would choose. Term spreads of very long maturities have greater predictive content for forecasting GDP in predictive regressions. The model R^2s increase with the maturity of the term spread. Interestingly, this implication of the model is independent of the forecasting horizon that we consider and is consistent with the OLS evidence in Table 2, where coefficients on 5-year term spreads have the highest significance levels. This is also consistent with the previous literature using the longest maturity spreads available, instead of spreads of intermediate maturity bonds. Also, this result is not simply due to the fact that the longest term spread is one of the factors in our model. To check this, we re-estimated the model with the 2-year term spread or the 4-year term spread as factors and obtained the same graph as Figure 2.

Third, controlling for lagged GDP is not as important as controlling for the short rate. Moving from the first panel of Figure 2 to the second panel only slightly increases the R^2 by adding lagged GDP. This is because GDP is only slightly autocorrelated (30%), and the autoregressive effect is only important for forecasting GDP at short (1- and 2-quarter) horizons. However, we obtain a large jump in the R^2 moving from the second to the last panel, where the short rate is included. In the first two panels, the R^2 plotted as a cross for the one quarter maturity is also much higher than the term spread regressors.

4.3   Predictive GDP Regression Coefficients

To see how our model completely characterizes the GDP predictive regressions, we plot the model and OLS regression coefficients for regression (1) in Figure 3. Each panel of Figure 3 shows the term spread coefficients for a different forecasting horizon (1-, 4-, 8- or 12-quarters ahead), with the same scale. On the x-axis, we show the maturity of the term spread from two to twenty quarters. The solid line plots the coefficients implied from the

yield-curve model with two standard errors bounds as dotted lines. We can compute a coefficient for every horizon, even those spreads not readily available from data sources. The cross corresponding to $x = 1$, represents the model-implied coefficient for regressing GDP growth onto the 1-quarter short rate. The OLS coefficients are shown as empty squares, with two standard error bounds denoted by solid squares. All the model-implied coefficients lie within two OLS standard error bands except for the 4-quarter spread coefficient in the 1-quarter GDP growth horizon regression, which is borderline.

The model-implied term-spread coefficients in Figure 3 have a strong downward sloping pattern, and the largest coefficients occur at the shortest maturity spreads (the unobserved 2-quarter spread). The regression coefficients rapidly decrease and then level off. The horizontal asymptote coincides almost exactly with the OLS estimates at the 1-quarter forecasting horizon but lies slightly below the OLS estimates for the other horizons. Looking only at the point estimates, the OLS coefficients of term spreads are slightly biased upwards, relative to the yield-curve model coefficients for the 4-, 8- and 12-quarter forecasting horizons. However, the difference between the coefficients is not statistically significant for using only term spreads as regressors.

However, there are large differences in the point estimates between OLS coefficients and the predictive coefficients implied by our model for other choices of regressors. Table 7 shows some selected regression coefficients for different forecasting horizons. Here, we list coefficients for two specifications: the short rate alone and a trivariate regression with short rates, term spreads and lagged GDP. We compare model-implied and OLS coefficients across different forecasting horizons given in rows. While we can compute the regression coefficients for any term spread, we choose the longest term spread in our data to be comparable to the literature. Note that since the term spread is an observable factor, the model-implied coefficients in Table 7 are just coefficients implied from a long-horizon VAR.

In Table 7, the short rate alone is always significantly negative from our model at the 5% level, and also significant in OLS regressions, except at the 12-quarter horizon. Turning to the trivariate regression, the 1-quarter forecasting horizon coefficients are identical because this regression represents the last row of the Φ matrix in the state dynamics (2). There is a difference in the standard errors across the model-implied and OLS columns because the model-implied standard errors are computed by the delta-method, and take into account the sampling error of all the parameters. We use Hodrick (1992) standard errors for the OLS regressions. As the forecasting horizon increases from one quarter to twelve quarters, the autocorrelation of GDP becomes less important.

There are two major differences between the model-implied coefficients and the OLS coefficients. First, at long forecasting horizons, the point estimate of the short rate coefficient implied by the model is significant in univariate regressions. The short rate model-implied coefficient retains its magnitude and significance in the trivariate regression. For example, in forecasting GDP growth 12-quarters out, the model-implied coefficient on $y_t^{(1)}$ is -0.24 in the univariate regression and -0.22 in the trivariate regression. Both coefficients are significantly negative. This pattern does not occur in the OLS results.

17

The OLS short rate coefficients, like the model-implied coefficients, are also significantly negative when GDP is regressed solely on the short rate at the 1-, 4- and 8-quarter forecasting horizons. However, at the 12-quarter horizon, the OLS coefficient on $y_t^{(1)}$ is -0.14 and insignificant in the univariate regression. Moreover, the OLS short rate coefficient is small and insignificant (-0.07) in the trivariate regression.

Second, the model-implied coefficients assign more predictive power to the short rate than to term spreads, while the opposite is true for OLS. This is consistent with the theoretical R^2 patterns observed in Figure 2. In the trivariate regressions, the model-implied coefficients on $y_t^{(20)} - y_t^{(1)}$ are all insignificant. For the OLS regressions, it is only at the 1-quarter horizon where the spread coefficient is insignificant. At the 12-quarter forecasting horizon, the model-implied (OLS) coefficient on $y_t^{(20)} - y_t^{(1)}$ is 0.08 (0.42). According to the yield-curve model, the only significant forecaster of GDP growth is the short rate, not the spread, at long horizons. Moreover, the largest effect should still come from the short rate, even after including the spread. According to OLS, the short rate has little predictive power controlling for the spread, in line with Plosser and Rouwenhorst (1994).

The differences in the $y_t^{(1)}$ and $y_t^{(20)} - y_t^{(1)}$ coefficients are not statistically significant – otherwise the model would be rejected. The OLS coefficients in Table 7 lie within two standard error bounds of the model-implied coefficients, where the standard errors are due solely to sampling error of the OLS predictive coefficients. The model-implied coefficients also lie within two standard error bounds of the OLS coefficients, where the standard errors are computed using the delta-method. Hence, statistically there may not be much difference, but the two sets of coefficients have vastly different implications for the role of the short rate and spread in forecasting GDP growth. Note that while some authors, like Dueker (1997) and Dotsey (1998) demonstrate the weak performance of the spread over specific sample periods (particularly over the 1990's), the unconstrained OLS regressions still firmly make the case for the predictive power of the spread. However, our model, estimated over the whole sample, implies that the short rate – not the spread – should be the most powerful predictor of GDP growth. The strong model prediction that the short rate should have the most explanatory power for future GDP movements is not the product of looking at a particular subsample, but is due to estimating the predictive relationship more efficiently.

The source of discrepancy between the model-implied and the OLS coefficients is sampling error, both from estimating the model and from estimating the OLS coefficients.[8] When we estimate the yield-curve model, we use information from across the entire curve and impose restrictions derived from no-arbitrage factor models that yield greater efficiency. In contrast, the OLS regressions treat the regressors as exogenous variables. To better characterize the important differences between the OLS and model-implied pre-

---

[8] An alternative explanation for the difference is small-sample endogenous regressor bias (see Stambaugh, 1999). We performed small sample simulations which indicate that this is not the source of the discrepancy between the OLS and model-implied coefficients. Of course, another explanation – which can never be ruled out – is model misspecification. However, Monte Carlo simulations in Section 4.5 show that our model can give rise to the OLS patterns observed in the data in small samples.

dictions for GDP, we next examine the role of risk premia in Section 4.4 and describe the small sample distributions of the OLS regression coefficients implied by our model in Section 4.5.

4.4 The Role of Risk Premia in Forecasting

To help interpret the finding that the model assigns little role to the term spread in forecasting GDP growth in the OLS regression, but assigns more importance to the short rate, we can write the spread as the sum of two components:

$$
y_t^{(n)} - y_t^{(1)} = \underbrace{E_t\left[\frac{1}{n}\sum_{i=0}^{n-1}y_{t+i}^{(1)}\right] - y_t^{(1)}}_{\text{EH-spread}} + \underbrace{y_t^{(n)} - E_t\left[\frac{1}{n}\sum_{i=0}^{n-1}y_{t+i}^{(1)}\right]}_{\text{RP}} \tag{13}
$$

The first component is the spread computed under the Expectations Hypothesis (EH). The n-quarter yield under the EH is just the expected average future short rate. The second component is the risk premium (RP) – the deviation of the actual yield from the yield implied by the EH.

The OLS predictive regressions (1) are often motivated by the idea that the term spread captures expectations of the future path of GDP. The regressions, however, do not differentiate between the risk premium and expectations of future interest rates. There is thus no distinction between correlations between future GDP growth and expectations of future short rates and the premium that investors demand for holding long-term bonds. Based on equation (13), we can disentangle expected future short rates and risk premia in term spreads. We compute the EH-spread from the VAR dynamics (2) and regress GDP growth on both components in (13) separately for the 4- and 20-quarter term spreads in Table 8.[9]

The results in Table 8 are striking. For comparison, we repeat the results with only the spread regressor in the first set of columns for different spread maturities from Table 2. The other columns report the EH and RP components of the spreads, treated as separate regressors. The R^2s from the bivariate regression with the two separate EH and RP components are much higher than the R^2s from the univariate regressions using the term spread. Moreover, the EH-spread coefficient is much larger in magnitude than the term spread coefficient for all forecasting horizons. Finally, the coefficient on the RP component is always insignificant, while the EH-spread is highly significant. These

---

[9]Strictly speaking, the interpretation of the EH-spread computed using the VAR (2) is inaccurate. The VAR dynamics in equation (2) do not impose the null of the EH, because the 20-year spread contains risk premia that is already captured by the VAR. To strictly impose the EH on the VAR requires complex estimation procedures (see, for example, Bekaert and Hodrick, 2001). The EH component also does not take into account Jensen's inequality terms. To check whether these two approximations matter for our results, we re-ran the results of Table 8 using the EH-spread directly computed from the latent factor model in the Appendix by imposing the prices of risk to be zero ($\lambda_1 = \lambda_2 = 0$). However, the results in Table 8 did not change.

results suggest that to predict GDP growth, we should subtract the RP component from the spread. Otherwise the expectations contained in the term spread are contaminated by the RP component, which blurs the GDP forecasts.

The EH-spread is highly correlated with the current short rate, because the short rate is slowly mean-reverting and, thus, highly predictable. The short rate instruments for the EH-spread in multivariate regressions based on the short rate and the term spread. This may explain why our model finds that the short rate does better at forecasting GDP than term spreads. An advantage of our model is that it effectively allows for a decomposition of term spreads into expected short rate changes and RP according to equation (13) using prices of risk.

4.5   Small Sample Simulations

The model coefficients that we estimate and the implied predictions for GDP predictive regressions that favor the short rate (rather than the term spread as the unconstrained OLS regressions suggest) are statements about population coefficients. It is conceivable that even though the OLS spread regression does not deliver an optimal population forecast, OLS may have less sampling variability than the model-implied forecast and is preferable in finite samples. To examine this possibility, and the implied small sample behavior of the forecasting coefficients, we use the estimated yield-curve model to generate 10,000 samples of the same length as our sample period (152 observations). In each simulated sample, we re-estimate our factor dynamics (a VAR), and compute an implied long-run forecast, which is a long-horizon VAR forecast, and compare them to a forecast from a simple OLS regression. For this exercise, we focus on the 5-year term spread. In this case, the GDP forecasts computed from our no-arbitrage model and those computed from the VAR in equation (2) are identical.[10] Table 9 presents the results.

Table 9 lists three rows for each forecasting horizon. First, the row labelled "Truth" reports the population forecasting coefficients from the estimated yield-curve model. The row labelled "Re-estimated Model" reports the average coefficients across all the simulations from re-estimating the model VAR, and computing long-horizon VAR forecasts. The numbers in parentheses represent the standard deviations of the coefficients across the simulations. Similar numbers are reported for the row labelled "OLS," which refer to the simple unconstrained OLS specification. The last column "RMSE" in Table 9 reports the Root Mean Squared Errors an out-of-sample forecasting period of 44 observations that corresponds to the same length of the out-of-sample forecasting period we use below in Section 5. The RMSE are relative to an AR(1) which we estimate for each simulated sample. The RMSE numbers for "Truth" use the true populations coefficients for forecasting in each simulated sample. In the out-of-sample forecasting period, each model is re-estimated in a rolling procedure, corresponding to the out-of-sample exercise using actual U.S. data in Section 5.

---

[10]Unfortunately, we cannot re-estimate prices of risk in each simulation due to computational complexity. Note that when using the 5-year yield as a factor, the model reduces to a VAR.

Table 9 shows that the small sample distributions of the implied forecasting coefficients from the re-estimated model come much closer to the truth than the OLS regressions. In fact, the averages are often exactly aligned with the population coefficients. The small sample distributions of the OLS regressions are very interesting. The OLS regressions place a large weight on the spread, even though the population places little weight on the spread. For example, for the 12-quarter horizon, the spread coefficient in population is only 0.08 (which the re-estimated model nails at 0.09), whereas the spread coefficient is a very large 0.61. This is exactly the result we see from the OLS coefficients estimated from data reported in Table 7! Hence, OLS has a tendency to favor the spread in small samples, even though the population distribution places little weight on the spread. In contrast, the greater efficiency of our yield-curve model recovers the true population forecasting coefficients that emphasize the short rate. Finally, consistent with the small sample distributions of the forecasting coefficients, the model RMSE are lower than OLS in the simulated out-of-sample.

5   Out-of-Sample Forecasting

5.1   Out-of-Sample Forecasting Results

We perform out-of-sample forecasts over the period 1990:Q1-2001:Q4. Our choice of the out-of-sample period must balance the need for a long in-sample time-series of data needed to estimate the yield-curve model, and a sufficiently long out-of-sample period to conduct the forecasting analysis. Our out-of-sample period covers 11 years and encompasses two recessions (one from 1990-1991 and one beginning in 2001). The forecasting exercise we conduct is rolling. That is, we first estimate the model parameters $\Theta = \{\mu, \Phi, \Sigma, \lambda_0, \lambda_1\}$ starting in 1964:Q1, and forecast k-quarter GDP growth from 1990:Q1 ($k > 1$). Then at each later point in time t, we re-estimate Θ again using the same number of observations up to time t, and forecast for horizons $t + k$.

Table 10 focuses on the forecasting performance using the short rate as opposed to spreads, while Table 11 examines the performance of different spread maturities. Our benchmark is the RMSE from an AR(1) specification, which is a reasonable benchmark that is also used by Stock and Watson (2001) in the context of forecasting macro-series. We first focus on the ability of the yield-curve model's factor structure to beat unconstrained OLS regressions. In out-of-sample forecasts, very often it is the most parsimonious statistical models that produce the best forecasts even if they are not based on economic theory. Over-parameterized models usually perform very well on in-sample tests but perform poorly out-of-sample. Our yield-curve model has many more parameters than an unconstrained OLS approach. This puts the performance of our model at a potential disadvantage relative to the standard OLS regressions.

Table 10 reports the out-of-sample forecasts results for several specifications to illustrate the efficiency gains from moving from OLS to a parsimonious factor VAR representation. The second last column refers to a system that includes all yields and GDP. The

last column refers to a 4-variable system with the short rate, 5-year term spread, GDP growth and the CPI inflation rate. The remaining columns compare the results from the yield-curve model, which here reduces to a long-horizon VAR because the short rate and term spread are factors. The rows labeled "OLS" refer to unconstrained OLS regressions. For comparison, a random walk performs atrociously, producing RMSE ratios as large as 1.496 for the 12-quarter forecasting horizon. Similarly, a simple OLS regression with all yields and GDP produces RMSE ratios all larger than the RMSE ratios in Table 10, reaching 1.378 for the 8-quarter horizon.

The first striking result of Table 10 is that the best performing models use short rates rather than term spreads. This is in line with the theoretical R^2 results in Section 4.2, which advocates using short rates as predictors. The RMSE ratios are particularly low for long-horizon out-of-sample forecasts at 8- and 12-quarters for the short rate only and short rate plus lagged GDP specifications. While the OLS in-sample results (Table 7) advocate the use of the term spread to forecast GDP, the out-of-sample results confirm the predictions from the yield-curve model that stress the use of the short rate. Not surprisingly, adding GDP helps to lower the RMSE ratios, particularly at short horizons.

As we add more predictors, the out-of-sample forecasts implied by the yield-curve model tend to produce lower RMSE ratios than the corresponding unconstrained OLS forecasts. For the full specification of the short rate, spread and GDP, the OLS and model forecasts are identical at the 1-quarter horizon by construction. As we increase the forecasting horizon, the model RMSE ratios become markedly lower than the OLS forecasts. This result is not unexpected, since the yield-curve model is able to incorporate information from the yield curve more efficiently than OLS, where the regressors are exogenous. This demonstrates that the yield-curve model can improve forecasts of GDP by exploiting the factor structure.

In our framework, it is easy to accommodate more factors than the 3 factors (short rate, spread and GDP) of our basic model. Because of the factor structure in equation (2), including additional factors simply augments the VAR. Table 10 shows that using two term structure factors is not only more parsimonious, but also produces more accurate forecasts than employing all the yields as factors. In particular the column labelled "Short Rate, All Spreads and GDP" reports higher RMSE ratios than the columns lying on its left, partly due to the added noise from having many more parameters from the additional state variables. Hence, there are valuable forecasting gains by reducing the dimensionality of the system. When we include inflation in the system, the forecasts are similar to our chosen 3-factor specification (slightly worse for short horizon and slightly better at long horizons).

The first potential gain from forecasting from the yield-curve model is the reduction in dimensionality of the system relative to a large VAR with all yields. The second potential gain is that the yield-curve model uses the risk premia to ensure the yields not used as factors move in a manner consistent with no-arbitrage. To investigate which gain is more important, Table 11 reports results for 3-variable systems with the short rate, a term spread of an arbitrary maturity, and GDP growth. To compute the "Model" RMSE

ratios, we compute the predictive coefficients using the model with relations similar to those in equation (9). (Note that when we use the 5-year term spread, as in Table 10, the model implies long-horizon VAR forecasts.) The row labelled "OLS" in Table 11 reports the RMSE ratio of an unconstrained OLS regression using the short rate, the term spread of arbitrary maturity, and GDP growth. Finally, the "VAR" refers to long-horizon forecasts from a 3-variable VAR with the same regressors as the OLS regressions, which is why the OLS and VAR RMSE ratios coincide for the 1-quarter forecasting horizon. The VAR forecasts differ from the forecasts implied by the model, because the model uses the prices of risk to pin down the implied forecasting coefficients, rather than just using the arbitrary term spread from data.

Table 11 shows that the risk premia not already captured by the factor structure does not contribute much to improving the GDP forecasts. The RMSE of the last column are much larger than those in the other columns. Table 11 also reports that the forecasting gains come from moving from OLS to the model or from OLS to the VAR. However, there is little difference between the RMSE ratios of the model and the VAR, and the model RMSE ratios are slightly higher for the model forecasts at for longer term spread maturities. Hence, the main advantage of the yield-curve model comes from using the VAR factor dynamics (equation (2)). The VAR structure places larger weight on the short rate (with a negative sign) and less weight on the spread than OLS (see Table 7), especially at long horizons. The better performance of the model is not due to the lack of inverted yield curves during the sample: the yield curve was inverted before both the 1991 and 2001 recessions for substantial periods of time (see Table 1).

There are two important messages in these out-of-sample forecasting results. First, while the point estimates of the predictive coefficients in the various regression specifications are well within OLS confidence bounds (see Table 7 and Figure 3), the magnitudes of the model-implied coefficients are quite different from the OLS estimates. In out-of-sample forecasting, this difference becomes important, particularly for the weight assigned to the short rate. By imposing restrictions implied by a term structure pricing framework, our estimates are more efficient than unrestricted OLS estimates. The stark implications of the superior forecasting power from short rates in the model versus term spreads in OLS show that *these efficiency gains matter*.

Second, the biggest gain comes from imposing a low-dimensional factor structure. Estimating a full VAR with all yields and GDP results in an over-parameterized model. A select choice of factors captures the common information in the yield curve, allowing us to reduce the sampling variability of the prediction functions. In this sense, the factor structure works like the shrinkage approaches on large VAR systems (see, for example, Schorfheide, 2003; Del Nelgro and Schorfheide, 2004). While our model has, for simplicity, used observable short rates and term spreads as factors, we expect that other methods of constructing factors will have similar efficiency gains. Some alternative factors include principal components (which we also consider below) or Nelson and Siegel (1987) splines (as used by Diebold and Li, 2003).

Our forecasting results are subject to a number of important qualifications. For exam-

ple, the out-of-sample period covers only two recessions and during the mid-1990's, there was a large decline in interest rates. While the full sample shows strong predictability (see Tables 2 and 7), there may be issues of the structural stability of these relations, particularly since monetary policy has undergone several different regimes. Another concern is that the superior forecasting ability of adding short rates or using a yield-curve model is small. None of the forecasts from the yield-curve model are strongly significant relative to their OLS counterparts using Diebold and Mariano (1995) tests. However, the number of observations in the out-sample is small and out-of-sample forecasting tests have very low power in these cases. Another observation is that there is one prediction from the model R^2s that is not borne out in the out-of-sample forecasting. The model predicts that greater explanatory power should come from longer term spreads. In Table 11, the RMSE ratios for longer term spread maturities should be lower, but they are not always so.

5.2   What is Driving the Short Rate Predictability?

Are the 1990's Special?

The high predictability of GDP growth by the short rate during the out-of-sample forecasting exercise may be due to the 1990's. The best performing models in the out-of-sample forecasting exercise used short rates. Over the out-sample (1990:Q1 to 2001:Q4), this particular sample may favor the short rate over the spread. Looking at Figure 1, we can see that during the 1990's, particularly, post 1994, the lagged term spread, if anything, is slightly negatively correlated with GDP growth. In contrast, the short rate moves clearly in the opposite direction to future GDP growth. Hence, the 1990's does seem to be a period where the spread does not seem to positively predict GDP growth but high short rates forecast negative GDP growth. However, the top panel of Figure 1 shows that prior to 1970, the term spread is almost flat and has little to do with GDP fluctuations. It is only between 1970 and 1990 that the lagged spread is positively correlated with GDP growth. In contrast, in the bottom panel of Figure 1, lagged short rates and GDP are negatively correlated throughout the whole sample period. Table 12 makes this clear.

Table 12 reports unconstrained OLS coefficients on the short rate, 5-year term spread and lagged GDP growth, forecasting future GDP growth k-quarters ahead across various sample periods. Over the out-of-sample period, post-1990, the OLS regressions place little weight on the spread but significantly large negative weights on the short rate. Prior to 1970, we see that the 5-year spread is also insignificant (but also has negative point estimates) and the coefficients on short rates are large in magnitude, highly statistically significant, and negative. It is only from 1971-1989 that the term spread significantly predicts future GDP growth with a positive sign and drives out the predictive power of the short rate. It is this period that dominates in the unconstrained OLS regressions over the full sample (listed in Table 7). Hence, although the 1990's favor the short rate, this period is by no means special. Over the full sample, OLS favors the term spread rather than the short rate. In contrast, the forecasting coefficients implied by our model always

24

favor the short rate over the term spread. Our model estimations do not let the 1970's and 1980's dominate the results.

Is it Inflation or the Real Rate?

To better understand why the nominal short rate predicts GDP growth, we decompose the nominal rate into the real rate and inflation. The actual real rate is unobservable, so we use an ex-post real rate that we construct by using the nominal short rate less realized CPI inflation over the last quarter. Since we use the ex-post real rate, the two sum up to the nominal short rate. To decompose the effect of the nominal rate into the real rate and inflation, we fit a 4-variable VAR of inflation, the real rate, the 5-year term spread and GDP. We compute long-horizon coefficients of each variable for forecasting inflation k-quarters out.

Table 13 reports the GDP forecasting coefficients where we distinguish between the real rate and inflation. We compare the 4-variable inflation-real rate system with the last 3 columns, which report the results from the yield-curve model that uses the nominal rate. We can see that the forecasting power from the nominal rate is mostly due to inflation. While both inflation and the real rate have negative coefficients, the inflation coefficients are larger in magnitude than the real rate coefficients. Moreover, the real rate coefficients are all insignificant while the inflation coefficients are all significant at the 5% level. The magnitude of the inflation coefficient is also roughly comparable to the coefficient on the nominal rate in the yield-curve model. We can conclude that it is the inflation content of the nominal rate that has predictive power for future economic growth. These results are consistent with Stock and Watson (1999), who find that the nominal rate is a leading business-cycle indicator, while the ex-post real rate is less cyclical. Bernanke and Blinder (1992) argue that the forecasting power of the nominal rate may be due to monetary policy. However, a full analysis of how monetary policy influences GDP through the short rate is beyond the scope of this paper, and should be approached with a structural model with no-arbitrage assumptions, building on the models of Hordahl, Tristani and Vestin (2003) and Rudebusch and Wu (2003).

5.3   Comparison with Other Predictive Variables

Best practice in forecasting involves the use of combinations of forecasts from a variety of models (see Granger and Newbold, 1987; Diebold, 2004). Our results show the benefit of using term structure approaches for forecasting GDP, but the yield curve is only one set of predictive instruments available to forecast economic growth. While finding optimal combinations of term structure forecasts with other variables is beyond the scope of this paper, we can compare the out-of-sample forecasts of the term structure model with a laundry list of predictive instruments. We compare the following variables as predictors: consensus GDP forecasts, the Stock-Watson experimental leading index XLI and the alternative non-financial experimental leading index XLI-2, default spreads, commercial paper spreads, lagged stock returns, housing starts, residential investments, growth in help-wanted ads, the first three principal components of yields and predictions from

Hamilton (1989)'s regime-switching model. A full description of these variables is given in the Appendix.

Table 14 compares the term structure model forecasts along with forecasts with other predictive variables, concentrating on results only with the short rate and GDP, and with the full 3-variable specification of the short rate, spread and GDP. Table 14 lists ratios of RMSE's relative to an AR(1) over the same out-of-sample period as Tables 10 and 11, 1990:Q1 to 2001:Q4. The results are very interesting. At all horizons, many variables perform much more poorly than an AR(1). The strongest competitor of the yield-curve model is the Stock-Watson XLI index. Note that the index employs term structure information along with other economic variables. For 1-quarter ahead forecasts, the consensus forecasts and both Stock-Watson indices do well. At the 12-quarter horizon, the term structure model outperforms all alternative predictors in the table, including the Stock-Watson XLI index.

6   Conclusion

We present a model of yields and GDP growth for forecasting GDP. Our approach is motivated from term structure approaches for pricing bonds in a no-arbitrage framework. The model is easily estimated and gives us a number of advantages to forecasting future economic growth. First, the model advocates using a select number of factors to summarize the information in the whole yield curve. These factors follow a Vector Autoregression (VAR), and long-term forecasts for these factors and GDP are simply long-horizon forecasts implied by the VAR. Second, the yield-curve model guides us in choosing the right spread maturity in forecasting GDP growth. We find that the maximal maturity difference is the best measure of slope in this context. Third, the nominal short rate dominates the slope of the yield curve in forecasting GDP growth both in- and out-of-sample. We find that the factor structure is largely responsible for most of the efficiency gains resulting in better out-of-sample forecasts. In contrast, our term structure approach allows us to show that risk premia not captured by the factor dynamics matter less in forecasting GDP. However, an unanswered question is whether we can improve on these yield curve forecasts by combining both term structure information and other macro variables. Furthermore, a better out-of-sample test than just using U.S. data is to use international data to test the efficiency gains of factor approaches implied by a term structure model.

Appendix

Estimation procedure

We use a two-step procedure to estimate the model. The parameters Θ can be partitioned into the parameters μ, Φ and Σ governing the factor dynamics (2) and the risk premia λ_0 and λ_1. In the first step, we estimate the VAR parameters μ, Φ and Σ using standard SUR. In the second step, we estimate λ_0 and λ_1 given the estimates of the parameters μ, Φ and Σ estimated in the first step. This is done by minimizing the sum of squared fitting errors of the model. More precisely, we compute model-implied yields $\hat{y}_t^{(n)} = a_n + b_n^\top X_t$ for given values of the state vector at time t and then solve:

$$\min_{\{\lambda_0,\lambda_1\}} \sum_{t=1}^{T} \sum_{n=1}^{N} \left(\hat{y}_t^{(n)} - y_t^{(n)} \right)^2 \tag{14}$$

for the N yields used to estimate the model.

The observed factors, the short rate and spread, make direct use of the yields $y_t^{(1)}$ and $y_t^{(20)}$. Hence these are yields to be considered to be measured without any observation error. The other yields (with maturities 4, 8, 12, and 16 quarters) are then functions of $y_t^{(1)}$ and $y_t^{(20)} - y_t^{(1)}$ and GDP growth, according to the model pricing equation (5). Naturally, this stochastic singularity means that the model generates yields $\hat{y}_t^{(n)}$ slightly different from the observed yields for the intermediate maturities. We therefore place a small sampling error on these yields not included as factors. We assume that the sampling errors have mean zero and estimate their standard deviation Ω in the second stage. Since $y_t^{(1)}$ and $y_t^{(20)} - y_t^{(1)}$ are generated by the VAR, and the model also determines $y_t^{(20)}$, we ensure that the no-arbitrage restrictions are satisfied by $y_t^{(20)}$ by imposing this constraint on the estimation. Effectively, the price of risk parameters adjust in our estimation so that this constraint is satisfied. We can also view the minimization (14) as a GMM estimation, subject to a constraint (see Newey and McFadden, 1994; Bekaert and Hodrick, 2001).

We compute standard errors for our parameter estimates using GMM, with moments from each stage of our two-step procedure. The moments are the first order conditions of ordinary least squares for μ and Φ, the covariance conditions of Σ, and the first-order conditions of λ_0 and λ_1 to satisfy (14). The standard errors we compute adjust for the two-stage estimation process. This is done as follows.

Let $\theta_1 = \{\mu, \Phi, \Sigma\}$ and $\theta_2 = \{\lambda_0, \lambda_1, \Omega\}$ be the partitioned parameter space and denote the corresponding sample estimates by overlined letters. We estimate $\Theta = \{\theta_1, \theta_2\}$ following a consistent two-step procedure. In the first step, we fit a VAR(1) process to the observed state variable series and estimate θ_1 by solving the system of equations:

$$\sqrt{T} g_1 \left(\overline{\theta}_1 \right) = 0, \tag{15}$$

where the function $g_1 (.)$ represents the usual SUR moment conditions.

In the second step we choose θ_2 so as to minimize the sum of squared fitting errors of the yields:

$$\min_{\theta_2} \frac{1}{2} g_2 \left(\theta_2; \overline{\theta}_1\right)^\top g_2 \left(\theta_2; \overline{\theta}_1\right),$$

where $g_2 \left(\theta_2; \overline{\theta}_1\right) = \left\{\hat{y}_t^{(n)} - y_t^{(n)}\right\}_{n=1}^N$ is a column vector of fitting errors of the yields evaluated at any given θ_2 and a fixed $\overline{\theta}_1$ from the first step. The first-order conditions for this minimization problem are:

$$G_{2,2}'\sqrt{T}g_2 \left(\overline{\theta}_2; \overline{\theta}_1\right) = 0, \tag{16}$$

where $G_{i,j} = \partial g_i/\partial \theta_j'$, $i,j = 1,2$.

We can expand the function g_1 and g_2 around the true parameter values using the first-order Taylor approximation:

$$\sqrt{T}g_1 \left(\overline{\theta}_1\right) = \sqrt{T}g_1 \left(\theta_1\right) + G_{1,1}\sqrt{T} \left(\overline{\theta}_1 - \theta_1\right),$$
$$\sqrt{T}g_2 \left(\overline{\theta}_2, \overline{\theta}_1\right) = \sqrt{T}g_2 \left(\theta_2, \theta_1\right) + G_{2,1}\sqrt{T} \left(\overline{\theta}_1 - \theta_1\right) + G_{2,2}\sqrt{T} \left(\overline{\theta}_2 - \theta_2\right),$$

which can then be substituted back into the moment conditions (15) and (16) to yield:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -\sqrt{T}g_1 \left(\theta_1\right) \\ -G_{2,2}'\sqrt{T}g_2 \left(\theta_2, \theta_1\right) \end{bmatrix} - \begin{bmatrix} G_{1,1} & 0 \\ G_{2,2}'G_{2,1} & G_{2,2}'G_{2,2} \end{bmatrix} \begin{bmatrix} \sqrt{T} \left(\overline{\theta}_1 - \theta_1\right) \\ \sqrt{T} \left(\overline{\theta}_2 - \theta_2\right) \end{bmatrix} \tag{17}$$

Invoking the rule of partitioned matrix inversion we obtain:

$$\begin{bmatrix} G_{1,1} & 0 \\ G_{2,2}'G_{2,1} & G_{2,2}'G_{2,2} \end{bmatrix}^{-1} = \begin{bmatrix} G_{1,1}^{-1} & 0 \\ -\left(G_{2,2}'G_{2,2}\right)^{-1}\left(G_{2,2}'G_{2,1}\right)G_{1,1}^{-1} & \left(G_{2,2}'G_{2,2}\right)^{-1} \end{bmatrix},$$

which allows us to rewrite (17) as:

$$\begin{bmatrix} \sqrt{T} \left(\overline{\theta}_1 - \theta_1\right) \\ \sqrt{T} \left(\overline{\theta}_2 - \theta_2\right) \end{bmatrix} = \begin{bmatrix} G_{1,1} & 0 \\ G_{2,2}'G_{2,1} & G_{2,2}'G_{2,2} \end{bmatrix}^{-1} \begin{bmatrix} -\sqrt{T}g_1 \left(\theta_1\right) \\ -G_{2,2}'\sqrt{T}g_2 \left(\theta_2, \theta_1\right) \end{bmatrix}$$
$$= \begin{bmatrix} G_{1,1}^{-1} & 0 \\ -\left(G_{2,2}'G_{2,2}\right)^{-1}\left(G_{2,2}'G_{2,1}\right)G_{1,1}^{-1} & \left(G_{2,2}'G_{2,2}\right)^{-1}G_{2,2}' \end{bmatrix} \begin{bmatrix} -\sqrt{T}g_1 \left(\theta_1\right) \\ -\sqrt{T}g_2 \left(\theta_2, \theta_1\right) \end{bmatrix}.$$

The asymptotic variances of the parameter estimates are thus:

$$\text{var}\left[\sqrt{T} \left(\overline{\theta}_1 - \theta_1\right)\right] = G_{1,1}^{-1}\Omega_{11}G_{1,1}'^{-1}$$
$$\text{var}\left[\sqrt{T} \left(\overline{\theta}_2 - \theta_2\right)\right] = \begin{bmatrix} -\left(G_{2,2}'G_{2,2}\right)^{-1}\left(G_{2,2}'G_{2,1}\right)G_{1,1}^{-1} & \left(G_{2,2}'G_{2,2}\right)^{-1}G_{2,2}' \end{bmatrix} \Omega$$
$$\begin{bmatrix} -\left(G_{2,2}'G_{2,2}\right)^{-1}\left(G_{2,2}'G_{2,1}\right)G_{1,1}^{-1} & \left(G_{2,2}'G_{2,2}\right)^{-1}G_{2,2}' \end{bmatrix}'$$

where $\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}$ is the variance of the sample mean of the moment conditions $\{g_1, g_2\}$.

Computing Regression R^2s

We detail the computation of the R^2 statistic for the most general regression (10), which regresses the k-quarter GDP growth $g_{t \to t+k}$ onto the short rate, the n-quarter term spread and GDP growth. By switching the order of summation, the last term in (11) can be rewritten as:

$$\frac{1}{k} e_3^T \sum_{j=1}^{k} \sum_{i=j}^{k} \Phi^{i-j} \Sigma \varepsilon_{t+j} = \frac{1}{k} e_3^T \sum_{j=1}^{k} \sum_{i=0}^{k-j} \Phi^i \Sigma \varepsilon_{t+j}$$

$$= \frac{1}{k} e_3^T \sum_{j=1}^{k} \sum_{i=0}^{j-1} \Phi^i \Sigma \varepsilon_{t+k-j+1}$$

$$= e_3^T \sum_{j=1}^{k} \widetilde{\Phi}_j \Sigma \varepsilon_{t+k-j+1} \tag{18}$$

Combining (11) and (18), we can compute the unconditional variance of the k-quarter ahead GDP growth as

$$\mathrm{var}\left(g_{t \to t+k}\right) = e_3^T \Phi \widetilde{\Phi}_k (\Sigma_X \Sigma_X^\top) \widetilde{\Phi}_k^T \Phi^T e_3 + e_3^T \sum_{j=1}^{k} \widetilde{\Phi}_j \Sigma \Sigma^T \widetilde{\Phi}_j^T e_3$$

and the R^2 statistic as

$$R^2 = \frac{\left(\beta_k^{(n)}\right)^T B(\Sigma_X \Sigma_X^\top) B^T \beta_k^{(n)}}{e_3^T \Phi \widetilde{\Phi}_k (\Sigma_X \Sigma_X^\top) \widetilde{\Phi}_k^T \Phi^T e_3 + e_3^T \sum_{j=0}^{k-1} \widetilde{\Phi}_j (\Sigma \Sigma^T) \widetilde{\Phi}_j^T e_3}.$$

We compute the standard errors for R^2 using the delta method.

Comparison with a Latent Factor Model

To compare the performance of our model with a more traditional model with unobservable factors, we estimate a system with 2 latent yield curve factors and GDP. Latent factors leads to identification issues of parameters, as rotations and linear transformations can be applied to the latent factors that result in observationally equivalent systems (see Dai and Singleton, 2000). We estimate the most general identified model.[11] We use maximum likelihood following Chen and Scott (1993) by inverting the latent factors from the 1 and 20-quarter yields.

While the estimation of the observable yield-curve model is not as efficient as the 1-step estimation of the latent yield-curve model, the results from the two approaches

---

[11]The first two elements of μ are set to zero, Φ is lower triangular, Σ can be partitioned into as $\Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{bmatrix}$. The short rate is $r_t = \delta_0 + \delta_1^\top X$ for a scalar δ_0 and a 3-dimensional vector δ_1. The first element in λ_0 is constrained to be zero.

are almost identical. The extracted latent factors are almost exact transformations of the first two principal components, or level and slope. The first latent factor has a 97% correlation with the first principal component and a -100% correlation with the short rate level. The second latent factor has a 98% correlation with the second principal component and a -99% correlation with the 20-quarter term spread. The correlations of the two latent factors with any other principal components are very small.

The second, more convincing, evidence is that the latent and observable models have near-identical implications for forecasting GDP. Figure 4 graphs the implied coefficients from a predictive GDP regression with short rate, term spread and lagged GDP regressors, as given by equation (10). In the left (right) hand column we show the coefficients for a forecast horizon of $k = 1$ quarter (4 quarters). In each plot, the x-axis denotes the maturity of the term spread used in the regression. For example, the coefficients $\beta_{1,1}^{(10)}$ for the short rate, $\beta_{1,2}^{(10)}$ for the 10-quarter term spread and $\beta_{1,3}^{(10)}$ for GDP shown in the top, middle and bottom panels on the left column, respectively, correspond to using a term spread maturity of $n = 10$ quarters for a forecasting horizon of $k = 1$ quarter. The thin solid line in Figure 4 represents the implied predictive coefficients from the latent yield-curve model. The implied coefficients from the observable yield-curve model are shown as diamonds. The two lines are almost identical. The dotted lines represent two standard deviation bounds computed using the delta-method from the latent yield-curve model. Figure 4 clearly shows that the predictive implications for forecasting GDP from the latent and observable yield-curve models are the same. Hence, we focus on the more easily interpretable observable yield-curve model that is more tractable, especially for out-of-sample forecasting.

Data Appendix of Other Predictive Instruments

Consensus forecasts on next-quarter real GDP growth are computed using consensus forecasts on current and next-quarter nominal GDP numbers together with current and next-quarter GDP deflators obtained from the Federal Reserve Bank of Philadelphia. The resulting series starts in 1968:Q4. We use the Stock-Watson Experimental Leading Index (*XLI*) and Alternative Nonfinancial Experimental Leading Index-2 (*XLI-2*) from Jim Stock's website. The *default spread* is computed as the difference between Moody's AAA corporate bond yield and Moody's BAA corporate bond yield, obtained from the Federal Reserve Bank of St. Louis FREDII database. The *commercial paper spread* is computed as the difference between the 3-month AA nonfinancial commercial paper rate, obtained from the Board of Governors of the Federal Reserve System, and the 3-month T-Bill rate. The 3-month non-financial commercial paper rate is only available back to 1997:Q1 and is supplemented using the 3-month commercial paper rate from 1971:Q2 to 1996:Q4. The *stock returns* are NYSE value-weighted returns from CRSP. The *housing starts* are the level of "new privately owned housing units started" available from the Census Bureau, while *residential investment growth* is the quarterly log growth rate of "real private residential fixed investment" from the Bureau of Economic Analysis. The index of *help-wanted advertising* in newspapers is in levels and is obtained from

the Conference Board. All data start in 1964:Q1, except those marked with special start dates. The out-of-sample forecast period is the same as in Table 10: 1990:Q1 to 2001:Q4.

The two remaining forecasts are based on the first *3 principal components of yields* and *Hamilton (1989)'s regime-switching model*. The regime-switching model is specified as:

$$g_{t+1} = \mu\left(s_{t+1}\right) + \rho \cdot g_t + \sigma(s_{t+1})\varepsilon_{t+1},$$

where $\varepsilon_{t+1} \sim N(0,1)$ and the regime variable $s_t = \{1,2\}$ follows a Markov chain with the constant transition probability matrix:

$$\Pi = \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix}.$$

The model is estimated using maximum likelihood. The time t forecast is computed as

$$E_t\left[g_{t \to t+k}\right] = E_t \sum_{i=1}^{k} \rho^{k-i} \mu\left(s_{t+i}\right) + \rho^k g_t$$

$$= \sum_{i=1}^{k} \rho^{k-i} \left\{ \sum_{S_t} \sum_{S_{t+i}} \mu\left(s_{t+i}\right) P\left[s_{t+i} | s_t\right] \right\} + \rho^k g_t.$$

Both the principal components and the regime-switching model are estimated recursively through the sample.

References

[1] Ang, A. and G. Bekaert, 2001, Stock Return Predictability: Is it There?, Working paper, Columbia University.

[2] Ang, A. and G. Bekaert, 2002, Regime Switches in Interest Rates, *Journal of Business and Economic Statistics* 20, 2, 163-189.

[3] Ang, A. and G. Bekaert, 2003, The Term Structure of Real Rates and Expected Inflation, Working paper, Columbia University.

[4] 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.

[5] Bekaert, G. and R.J. Hodrick, 2001, Expectations Hypotheses Tests, *Journal of Finance* 56, 1357-1394.

[6] Bernanke, B.S. and A.S. Blinder, 1992, The Federal Funds Rate and the Channels of Monetary Transmission, *American Economic Review* 82, 4, 901-921.

[7] Bernard, H. and S. Gerlach, 1998, Does the Term Structure Predict Recessions? The International Evidence, *International Journal of Finance and Economics* 3, 195-215.

[8] Berardi, A. and W. Torous, 2002, Does the Term Structure Forecast Consumption Growth?, Working paper, UCLA.

[9] Campbell, J.Y. and R.J. Shiller, 1988, The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, *Review of Financial Studies* 1, 195-228.

[10] Campbell, J.Y. and R.J. Shiller, 1991, Yield Spreads and Interest Rate Movements: A Bird's Eye View, *Review of Economic Studies* 58, 495-514.

[11] Chen, N.F., 1991, Financial Investment Opportunities and the Macroeconomy, *Journal of Finance* 46, 529-54.

[12] Chen, R.R. and L. Scott, 1993, Maximum Likelihood Estimation for a Multi-factor Equilibrium Model of the Term Structure of Interest Rates, *Journal of Fixed Income* 3, 14-31.

[13] Christiano, L.J., M. Eichenbaum and C. Evans, 1999, Monetary Policy Shocks: What Have we Learned and to What End? in: *Handbook of Macroeconomics*, edited by M. Woodford and J.B. Taylor, North Holland, Amsterdam.

[14] Cochrane, J.Y., 1994, Permanent and Transitory Components of GNP and Stock Prices, *Quarterly Journal of Economics* 109, 241-66.

[15] Cochrane, J.Y. and M. Piazzesi, 2002, Bond Risk Premia, NBER Working paper 9178, September.

[16] Dai, Q. and K. Singleton, 2000, Specification Analysis of Affine Term Structure Models, *Journal of Finance* 55, 1943-78.

[17] 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.

[18] Dai, Q., K. Singleton, and W. Yang, 2003, Regime Shifts in a Dynamic Term Structure Model of the U.S. Treasury Yields, Working paper, NYU.

[19] Del Negro, M. and F. Schorfheide, 2004, Priors from General Equilibrium Models for VARs, Forthcoming *International Economic Review*.

[20] Diebold, F.X. 2004. *Elements of Forecasting*. South-Western College Publishing.

[21] Diebold, F.X. and C. Li, 2003, Forecasting the Term Structure of Government Bond Yields, Working Paper, University of Pennsylvania.

[22] Diebold, F.X. and R.S. Mariano, 1995, Comparing Predictive Accuracy, *Journal of Business and Economic Statistics* 13, 3, 253-63.

[23] Diebold, F.X., G.D. Rudebusch and B. Aruoba, 2003, The Macroeconomy and the Yield Curve, Working paper, University of Pennsylvania.

[24] Dotsey, M., 1998, The Predictive Content of the Interest Rate Term Spread for Future Economic Growth, *Federal Reserve Bank of Richmond Economic Quarterly* 84, 3, 31-51.

[25] Dueker, M.J., 1997, Strengthening the Case for the Yield Curve as a Predictor of U.S. Recessions, *Federal Reserve Bank of St. Louis Review*, 79, March/April, 41-51.

[26] Duffie, D. and R. Kan, 1996, A Yield-Factor Model of Interest Rates, *Mathematical Finance* 6, 379-406.

[27] Estrella, A. and G.A. Hardouvelis, 1991, The Term Structure as Predictor of Real Economic Activity, *Journal of Finance* 46, 555-76.

[28] Estrella, A. and F.S. Mishkin, 1996, The Yield Curve as a Predictor of U.S. Recessions, *Current Issues in Economics and Finance*, Federal Reserve Bank of New York, 2, 1-6.

[29] Estrella, A. and F.S. Mishkin, 1997, The Predictive Power of the Term Structure of Interest Rates in Europe and the United States: Implications for the European Central Bank, *European Economic Review*, 41, 1375-401.

[30] Estrella, A. and F.S. Mishkin, 1998, Predicting U.S. Recessions: Financial Variables as Leading Indicators, *Review of Economics and Statistics* 1, 45-61.

[31] Evans, C. and D. Marshall, 1998, Monetary Policy and the Term Structure of Nominal Interest Rates: Evidence and Theory, *Carnegie-Rochester Conference Series on Public Policy* 49, 53-111.

[32] Fama, E.F., 1990, Term Structure Forecasts of Interest Rates, Inflation and Real Returns, *Journal of Monetary Economics* 25, 56-76.

[33] Fama, E.F. and R.R. Bliss, 1987, The Information in Long-Maturity Forward Rates, *American Economic Review* 77, 680-692.

[34] Granger, C.W.J. and P. Newbold, 1986, *Forecasting Economic Time Series*. Academic Press.

[35] Hamilton, J.D., 1989, A New Approach To The Economic Analysis Of Nonstationary Time Series And The Business Cycle, *Econometrica* 57, 2, 357-384.

[36] Hamilton, J.D. and D.H. Kim, 2002, A Re-Examination of the Predictability of the Yield Spread for Real Economic Activity, *Journal of Money, Credit, and Banking* 34, 340-60.

[37] Harvey, C.R., 1989, Forecasts of Economic Growth from the Bond and Stock Markets, *Financial Analysts Journal* 45, 38-45.

[38] Harvey, C.R., 1991, The Term Structure and World Economic Growth, *Journal of Fixed Income* 1, 7-19.

[39] Harvey, C.R., 1993, The Term Structure Forecasts Economic Growth, *Financial Analyst Journal* 49, 6-8.

[40] Hodrick, R.J., 1992, Dividend Yields And Expected Stock Returns: Alternative Procedures For Influence And Measurement, *Review of Financial Studies* 5, 357-86.

[41] Hordahl, P., O. Tristani and D. Vestin, 2003, A Joint Econometric Model of Macroeconomic and Term Structure Dynamics, Working Paper, European Central Bank.

[42] Jorion, P. and F.S. Mishkin, 1991, A Multi-Country Comparison of Term-Structure Forecasts at Long Horizons, *Journal of Financial Economics* 29, 59-80.

[43] King, R.G., J.H. Stock and M.W. Watson, 1995, Temporal Instability of the Unemployment-Inflation Relationship, *FRB Chicago - Economic Perspectives* 19, 3, 2-12.

[44] Knez, P.J., R. Litterman and J. Scheinkman, 1994, Explorations Into Factors Explaining Money Market Returns, *Journal of Finance* 49, 5, 1861-882.

[45] Laurent, R., 1988, An Interest Rate-Based Indicator of Monetary Policy, *Economic Perspectives*, Federal Reserve Bank of Chicago 12, 3-14.

[46] Mishkin, F.S., 1990a, The Information in the Longer Maturity Term Structure about Future Inflation, *Quarterly Journal of Economics* 55, 815-28.

[47] Mishkin, F.S., 1990b, A Multi-country study of the Information in the Term Structure about Future Inflation, *Journal of International Money and Finance* 19, 2-22.

[48] Mody, A., and M.P. Taylor, 2003, The High Yield Spread as a Predictor of Real Economic Activity: Evidence of a Financial Accelerator for the United States, Working paper, IMF.

[49] Nelson, C.R. and A.F. Siegel, 1987, Parsimonious Modeling of Yield Curves, *Journal of Business* 60, 4, 473-490.

[50] Newey, W.K. and D.L. McFadden, 1994, Large Sample Estimation and Hypothesis Testing, in: *Handbook of Econometrics Vol. 4*, edited by R.F. Engle and D.L. McFadden, 2111-2245, Elsevier, Amsterdam.

[51] Newey, W.K. and K.D. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703-08.

[52] Piazzesi, M., 2003, Bond Yields and the Federal Reserve, Forthcoming *Journal of Political Economy.*

[53] Plosser, C.I. and K.G. Rouwenhorst, 1994, International Term Structures and Real Economic Growth, *Journal of Monetary Economics* 33, 133-55.

[54] Rudebusch, G.D., and T. Wu, 2003, A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy, Working Paper, Federal Reserve Bank of San Francisco.

[55] Schorfheide, F., 2003, Learning and Monetary Policy Shifts, Working paper, University of Pennsylvania.

[56] Schorfheide, F., 2004, VAR Forecasting under Misspecification, Forthcoming *Journal of Econometrics.*

[57] Stambaugh, R.F., 1999, Predictive Regressions, *Journal of Financial Economics* 54, 375-421.

[58] Stock, J.H. and M.W. Watson, 1989, New Indexes of Coincident and Leading Indicators, in: *1989 NBER Macroeconomics Annual*, edited by O. Blanchard and S. Fischer, MIT Press, Cambridge MA.

[59] Stock, J.H. and M.W. Watson, 1999, Business Cycle Fluctuationsin U.S. Macroeconomic Time Series, in: *Handbook of Macroeconomics Vol. 1*, edited by J.B. Taylor and M. Woodford, 3-63, Elsvier, Amsterdam.

[60] Stock, J.H. and M.W. Watson, 2001, Forecasting Output and Inflation: The Role of Asset Prices, NBER Working Paper 8180.

[61] Stock, J.H. and M.W. Watson, 2002, Has the Business Cycle Changed and Why? in *NBER Macroeconomics Annual 2002*, edited by M. Gertler and K. Rogoff.

[62] Taylor, J.B., 1993, Discretion versus Policy Rules in Practice, *Carnegie-Rochester Conference Series on Public Policy* 39, 195-214.

[63] Wu, T., 2002, Monetary Policy and the Slope Factors in Empirical Term Structure Estimations, Federal Reserve Bank of San Francisco Working Paper 02-07.

Table 1: NBER Recession Forecasts from Term Spreads

NBER recession	Inversion	Lead Time
1953:Q3-1954:Q2		
1957:Q3-1958:Q2		
1960:Q2-1961:Q1		
- - - - - - - - Fama-Bliss sample starts 1964:Q1 - - - - - - - -		
	1966:Q3-1966:Q4	
1969:Q4-1970:Q4	1968:Q2,1968:Q4,1969:Q4	6 qtrs
1973:Q4-1975:Q1	1973:Q2-1974:Q1,1974:Q4	2 qtrs
1980:Q1-1980:Q3	1978:Q4-1980:Q1	5 qtrs
1981:Q3-1982:Q4	1980:Q3-1980:Q4,1981:Q2,1982:Q1	4 qtrs
1990:Q3-1991:Q1	1989:Q2	5 qtrs
2001:Q1-2002:Q1	2000:Q3-2000:Q4	2 qtrs

NOTE: The first column lists periods of recessions as defined by the NBER. The second column reports the periods of inverted yield curves (a negative difference between the 5-year zero coupon yield and the 3-month T-bill yield) prior to the onset of a recession. The final column reports the time from an initial inverted yield curve and the onset of a recession.

Table 2: Forecasts of GDP Growth from Term Spreads

| Horizon | Term Spread Maturity | | | | | | | | | |
| | 4-qtr | | 8-qtr | | 12-qtr | | 16-qtr | | 20-qtr | |
k-qtrs	$\beta_k^{(4)}$	R^2	$\beta_k^{(8)}$	R^2	$\beta_k^{(12)}$	R^2	$\beta_k^{(16)}$	R^2	$\beta_k^{(20)}$	R^2
1	0.31	0.00	0.78	0.03	0.72	0.04	0.66	0.04	0.65	0.04
	(0.73)		(0.49)		(0.39)		(0.33)		(0.29)	
4	1.18	0.06	1.23	0.16	1.06	0.18	0.90	0.17	0.89	0.20
	(0.49)		(0.38)		(0.32)		(0.28)		(0.26)	
8	1.06	0.10	1.04	0.20	0.91	0.25	0.78	0.24	0.73	0.24
	(0.41)		(0.33)		(0.29)		(0.26)		(0.24)	
12	0.56	0.05	0.67	0.16	0.59	0.19	0.53	0.20	0.48	0.20
	(0.32)		(0.27)		(0.24)		(0.21)		(0.20)	

NOTE: The table reports the slope coefficient $\beta_k^{(n)}$ and R^2 for the regression $g_{t \to t+k} = \alpha_k^{(n)} + \beta_k^{(n)} \left(y_t^{(n)} - y_t^{(1)} \right) + \varepsilon_{t+k,k}^{(n)}$, where $g_{t \to t+k}$ is annualized GDP growth over the next k quarters and $y_t^{(n)}$ denotes a zero-coupon yield of maturity n quarters. Hodrick (1992) standard errors are reported in parentheses.

Table 3: Forecasting GDP with Principal Components

Horizon	Principal Components			Lagged	
k-qtrs	PC1	PC2	PC3	GDP	R^2
1	0.10	−0.68	1.48	0.27	0.18
	(0.05)	(0.28)	(1.00)	(0.08)	
4	0.11	−0.91	0.05	0.12	0.32
	(0.06)	(0.30)	0.72)	(0.06)	
8	0.07	−0.81	−0.10	−0.02	0.32
	(0.05)	(0.27)	(0.54)	(0.04)	
12	0.02	−0.59	0.37	−0.02	0.24
	(0.04)	(0.22)	(0.43)	(0.03)	

NOTE: The table reports the slope estimates and R^2 from regressing GDP growth on a constant and the indicated right-hand side variables. Hodrick (1992) standard errors are reported in parentheses.

Table 4: Parameter estimates for Yield-Curve Model

State dynamics $X_t = \mu + \Phi X_{t-1} + \Sigma \varepsilon_t$

	μ	Φ			Σ		
Short Rate	0.086	0.923	0.096	0.012	0.286		
	(0.087)	(0.044)	(0.094)	(0.040)	(0.062)		
Spread	0.025	0.026	0.768	−0.008	−0.150	0.124	
	(0.056)	(0.029)	(0.072)	(0.024)	(0.038)	(0.009)	
GDP growth	0.935	−0.269	0.320	0.258	0.177	0.147	0.773
	(0.303)	(0.132)	(0.325)	(0.090)	(0.061)	(0.075)	(0.068)

Risk premia $\lambda_t = \lambda_0 + \lambda_1 X_t$ | Std dev of errors Ω

	λ_0	λ_1			1-yr yield	2-yr yield
Short Rate	0.29	−34.18	−52.61	1.60	0.084	0.058
	(0.22)	(0.27)	(0.17)	(10.88)	(0.009)	(0.006)
Spread	0.41	2.04	−102.90	0.85		
	(0.67)	(0.11)	(0.08)	(1.64)	3-yr yield	4-yr yield
GDP Growth	6.57	2.84	−2.56	8.87	0.041	0.031
	(4.89)	(0.09)	(0.01)	(0.38)	(0.012)	(0.008)

NOTE: Parameters μ, Σ and standard deviations of measurement errors are multiplied by 100. Standard errors (in parentheses) are computed following the method in the Appendix.

Table 5: Unconditional Moments from Model and Data

Maturity	Short Rate 1-qtr	4-qtr	Term Spreads 8-qtr	12-qtr	16-qtr	20-qtr	GDP growth
			Model Implied Moments				
Mean	6.28	0.44	0.70	0.85	0.96	1.03	3.21
	(1.06)	(0.07)	(0.12)	(0.16)	(0.18)	(0.20)	(0.42)
Std.Dev.	2.66	0.36	0.70	0.91	1.06	1.16	3.52
	(0.72)	(0.15)	(0.24)	(0.29)	(0.31)	(0.33)	(0.30)
			Data				
Mean	6.34	0.46	0.66	0.81	0.93	0.99	3.20
Std.Dev	2.61	0.50	0.74	0.93	1.06	1.15	3.54

NOTE: The table compares model-implied means and standard deviations of short rates, term spreads and GDP growth with data. Standard errors (in parentheses) are computed using the delta method.

Table 6: P-values of χ^2 Specification Tests

	4-qtr	8-qtr	12-qtr	16-qtr	Joint
Mean	0.845	0.947	0.958	0.932	0.071
Variance	0.908	0.967	0.909	0.971	0.007
Autocorrelation	0.510	0.544	0.591	0.673	0.905

NOTE: The table reports p-values for GMM χ^2 tests for matching means, variances and autocorrelations of yields not used as factors in the estimation of the model. The last column lists p-values for χ^2 tests performed jointly across the 4-, 8-, 12-, and 16-quarter yields.

Table 7: Model-Implied and OLS Coefficients

| Horizon qtrs | Model Implied Coefficients | | | OLS Coefficients | | |
	Short Rate	5-Year Term Spread	GDP Growth	Short Rate	5-Year Term Spread	GDP Growth
1	−0.39			−0.39		
	(0.14)			(0.14)		
	−0.27	0.32	0.26	−0.27	0.32	0.26
	(0.13)	(0.33)	(0.09)	(0.14)	(0.26)	(0.09)
4	−0.35			−0.42		
	(0.13)			(0.15)		
	−0.28	0.25	0.08	−0.28	0.60	0.11
	(0.14)	(0.31)	(0.04)	(0.15)	(0.25)	(0.06)
8	−0.28			−0.28		
	(0.11)			(0.13)		
	−0.25	0.14	0.04	−0.19	0.56	−0.02
	(0.14)	(0.25)	(0.03)	(0.13)	(0.21)	(0.03)
12	−0.24			−0.14		
	(0.10)			(0.12)		
	−0.22	0.08	0.02	−0.07	0.42	−0.02
	(0.13)	(0.20)	(0.02)	(0.11)	(0.17)	(0.02)

NOTE: Slope estimates for predictive GDP regressions using only short rates as regressors, and short rates, term spreads and GDP growth as regressors. The table reports predictive coefficients implied by the yield-curve model (left panel) versus unrestricted OLS coefficients (right panel). Standard errors (in parentheses) are computed by the delta-method for the model-implied coefficients, and following the method of Hodrick (1992) for the OLS coefficients.

Table 8: The Role of Risk Premia

Horizon	4-qtr Term Spread					20-qtr Term Spread				
k-qtrs	Spread	R^2	EH	RP	R^2	Spread	R^2	EH	RP	R^2
1	0.31	0.00	2.75	-1.00	0.12	0.65	0.04	0.92	0.20	0.09
	(0.73)		(1.02)	(0.73)		(0.29)		(0.34)	(0.31)	
4	1.18	0.06	3.41	-0.08	0.30	0.89	0.20	1.15	0.47	0.30
	(0.49)		(0.65)	(0.45)		(0.26)		(0.23)	(0.30)	
8	1.06	0.10	2.46	0.26	0.27	0.73	0.24	0.87	0.50	0.31
	(0.41)		(0.44)	(0.41)		(0.24)		(0.15)	(0.27)	
12	0.56	0.05	1.33	0.12	0.15	0.48	0.20	0.53	0.40	0.21
	(0.32)		(0.50)	(0.35)		(0.20)		(0.14)	(0.21)	

NOTE: This table reports predictive GDP regression results from equation (1) and results from regressing GDP growth on the Expectations Hypothesis (EH) and risk premium (RP) components in equation (13) separately. Hodrick (1992) standard errors are in parentheses.

Table 9: Small Sample Predictive Coefficients

Horizon k-qtrs		Short Rate	5-Year Term Spread	GDP Growth	RMSE
4	Truth	−0.28	0.25	0.08	0.896
	Re-estimated Model	−0.30	0.20	0.07	0.951
		(0.07)	(0.07)	(0.10)	(0.014)
	OLS	−0.16	0.59	0.10	0.960
		(0.09)	(0.16)	(0.14)	(0.018)
8	Truth	−0.25	0.14	0.04	0.887
	Re-estimated Model	−0.22	0.14	0.03	0.961
		(0.07)	(0.07)	(0.10)	(0.014)
	OLS	−0.07	0.65	−0.02	0.981
		(0.09)	(0.16)	(0.14)	(0.018)
12	Truth	−0.22	0.08	0.02	0.888
	Re-estimated Model	−0.18	0.09	0.02	0.974
		(0.07)	(0.07)	(0.10)	(0.014)
	OLS	0.02	0.61	−0.01	1.004
		(0.09)	(0.16)	(0.14)	(0.018)

NOTE: The row labelled "Truth" reports population coefficients. In each simulation of 152 observations, we re-estimate the model, compute predictive coefficients implied by the model, and estimate OLS coefficients. The numbers in brackets are the standard deviations of these coefficient estimates across simulated sample paths. The column labelled "RMSE" reports the RMSE relative to an AR(1) that corresponds to a series of rolling out-of-sample estimations over the last 44 observations in each simulation. We perform 10,000 simulations.

Table 10: Out-of-Sample Forecasts – Effect of Factor Dynamics

Horizon k-qtrs		Short Rate Only	Spread Only	Short Rate and GDP	Spread and GDP	Short Rate, Spread and GDP	Short Rate, All Spreads and GDP	Short Rate, Spread, GDP and Inflation
1	OLS	1.133	1.130	1.043	1.033	1.059	1.160	1.077
	Model	1.134	1.135	1.043	1.038	1.059	1.160	1.077
4	OLS	1.041	1.169	1.011	1.121	1.151	1.151	1.117
	Model	1.010	1.064	0.980	1.018	1.014	1.068	1.024
8	OLS	0.851	1.190	0.881	1.225	1.281	1.271	1.233
	Model	0.831	0.972	0.827	0.947	0.888	1.007	0.886
12	OLS	0.742	0.955	0.762	1.002	1.116	1.179	1.054
	Model	0.720	0.863	0.710	0.836	0.750	0.831	0.740

NOTE: Table entries are RMSE ratios relative to an AR(1). The spread refers to the 5-year term spread. The out-of-sample period is 1990:Q1-2001:Q4.

Table 11: Out-of-Sample GDP Forecasts – Effect of Risk Premia

Horizon		Term Spread Maturity			
$k$qtrs		4-qtrs	8-qtrs	12-qtrs	16-qtrs
1	Model	1.021	1.041	1.054	1.063
	OLS	1.080	1.044	1.048	1.054
	VAR	1.080	1.044	1.048	1.054
4	Model	0.993	1.003	1.013	1.023
	OLS	0.997	1.060	1.091	1.117
	VAR	1.029	1.000	1.002	1.009
8	Model	0.862	0.878	0.885	0.899
	OLS	0.908	1.057	1.148	1.212
	VAR	0.899	0.857	0.861	0.874
12	Model	0.709	0.726	0.741	0.755
	OLS	0.798	0.918	0.991	1.048
	VAR	0.783	0.727	0.728	0.736

NOTE: Table entries are RMSE ratios relative to an AR(1) for systems with the short rate, a term spread (of varying maturity), and GDP growth. The out-of-sample period is 1990:Q1-2001:Q4.

Table 12: Subsample Analysis

Horizon	1964:Q1 - 1970:Q4			1971:Q1 - 1989:Q4			1990:Q1 - 2001:Q4		
k-qtrs	Short Rate	5-year Spread	GDP Growth	Short Rate	5-year Spread	GDP Growth	Short Rate	5-year Spread	GDP Growth
1	-1.37	-2.35	0.04	-0.27	0.53	0.25	-0.75	-0.27	0.34
	(0.56)	(1.48)	(0.22)	(0.21)	(0.39)	(0.11)	(0.14)	(0.41)	(0.12)
4	-1.69	-0.05	-0.01	-0.22	0.93	0.06	-1.00	-0.43	0.08
	(0.64)	(1.23)	(0.12)	(0.22)	(0.40)	(0.07)	(0.30)	(0.33)	(0.07)
8	-2.16	-0.94	-0.07	-0.06	0.97	-0.06	-0.55	-0.51	-0.02
	(0.45)	(0.73)	(0.04)	(0.19)	(0.35)	(0.04)	(0.13)	(0.28)	(0.03)
12	-1.56	-1.42	0.03	0.11	0.83	-0.05	-0.25	-0.12	0.03
	(0.23)	(0.54)	(0.02)	(0.17)	(0.29)	(0.03)	(0.09)	(0.19)	(0.04)

NOTE: Unconstrained OLS predictive regressions over various sample periods. Hodrick (1992) standard errors are in parentheses.

Table 13: Forecasting GDP Growth

Horizon	VAR with Inflation and Real Rate				Yield-Curve Model		
k-qtrs	Inflation	Real Rate	Spread	GDP	Nom. rate	Spread	GDP
1	−0.33	−0.17	0.21	0.24	−0.27	0.32	0.26
	(0.14)	(0.15)	(0.34)	(0.08)	(0.13)	(0.33)	(0.09)
4	−0.32	−0.20	0.24	0.07	−0.28	0.25	0.08
	(0.13)	(0.15)	(0.29)	(0.04)	(0.14)	(0.31)	(0.04)
8	−0.26	−0.18	0.18	0.03	−0.25	0.14	0.04
	(0.12)	(0.13)	(0.22)	(0.02)	(0.13)	(0.24)	(0.03)
12	−0.21	−0.16	0.13	0.02	−0.22	0.08	0.02
	(0.11)	(0.12)	(0.18)	(0.02)	(0.12)	(0.19)	(0.02)

NOTE: The table reports predictive coefficients from a 4-variable VAR (left panel) and from the yield-curve model (right panel) for forecasting GDP over various horizons. Standard errors (in parentheses) are computed by the delta-method for the model-implied coefficients. We compute standard errors for the long-horizon VAR coefficients using GMM on the 1-step ahead VAR moment conditions and then applying the delta method.

Table 14: Comparison with Other Predictors

| | Forecasting Horizon | | | |
	1-qtr	4-qtr	8-qtr	12-qtr
Short Rate and GDP	1.043	0.980	0.827	0.710
Short Rate, Spread and GDP	1.059	1.014	0.888	0.750
Consensus Forecasts	0.992	1.035	—	—
Stock-Watson XLI	0.966	0.907	0.795	0.803
Stock-Watson XLI-2	0.938	0.851	0.828	0.866
Default Spread	1.092	1.083	1.175	1.382
Commercial Paper Spread	1.095	1.026	0.859	0.828
Stock Returns	1.124	0.998	0.850	0.812
Housing Starts	1.060	1.015	0.930	0.922
Residential Investment Growth	1.003	0.920	0.836	0.817
Help-Wanted Ads	1.093	1.049	1.029	1.015
3 Principal Components of Yields	1.283	1.225	1.243	1.185
Hamilton Regime-Switching Model	1.022	0.997	0.892	0.859

NOTE: Table entries are RMSE ratios relative to an AR(1). The in-sample period starts at 1964:Q1, unless otherwise indicated in the Appendix. The out-of-sample period is 1990:Q1-2001:Q4.

Table A-1: Parameter estimates for the Latent-Factor Model

State dynamics $X_t = \mu + \Phi X_{t-1} + \Sigma \varepsilon_t$

	μ	Φ			Σ		
factor 1		0.936			100.000		
		(0.035)			(0.000)		
factor 2		0.156	0.744			100.000	
		(0.072)	(0.071)			(0.000)	
GDP growth	0.585	−0.001	0.001	0.264	0.238	0.035	0.770
	(0.113)	(0.000)	(0.001)	(0.110)	(0.137)	(0.163)	(0.049)

	Risk premia $\lambda_t = \lambda_0 + \lambda_1 X_t$				Std dev of errors Ω	
	λ_0	λ_1			1-yr yield	2-yr yield
factor 1		0.00	-0.10	2.12	0.084	0.059
		(0.04)	(0.05)	(4.96)	(0.010)	(0.007)
factor 2	−0.14	0.10	-0.16	−3.95		
	(2.54)	(0.12)	(0.13)	(50.85)	3-yr yield	4-yr yield
GDP Growth	13.14	0.28	0.53	150.56	0.041	0.031
	(49.64)	(2.54)	(3.76)	(1069.78)	(0.004)	(0.002)

NOTE: Parameters μ, Σ and standard deviations of measurement errors are multiplied by 100. Standard errors (in parentheses) are computed using the Hessian.
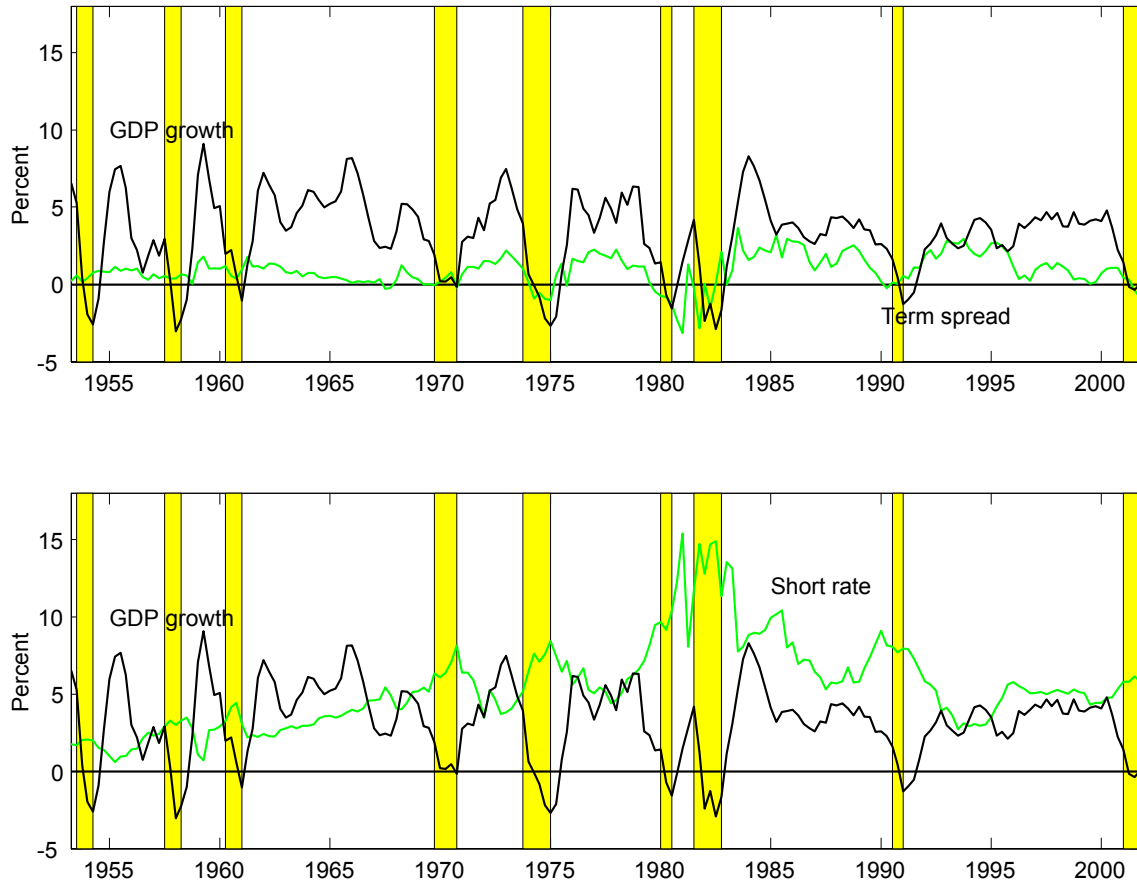
Figure 1: Four-Quarter GDP Growth.

FIGURE 1 NOTE: The figure plots 4-quarter GDP growth together with the 20-quarter term spread (upper panel) and the 1-quarter short rate (lower panel) lagged four quarters. For example, in 1990 we plot GDP growth from 1989-1990 together with the term spread (short rate) at 1989.
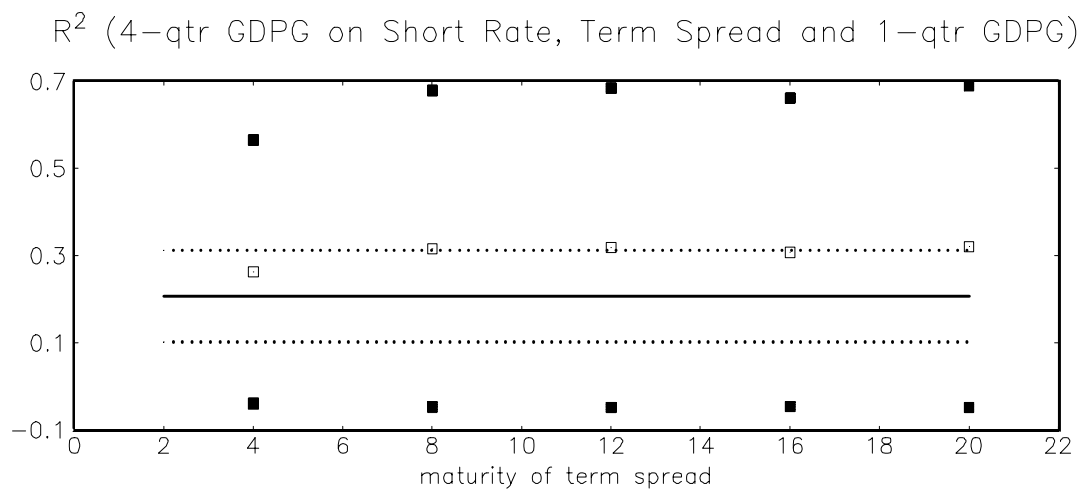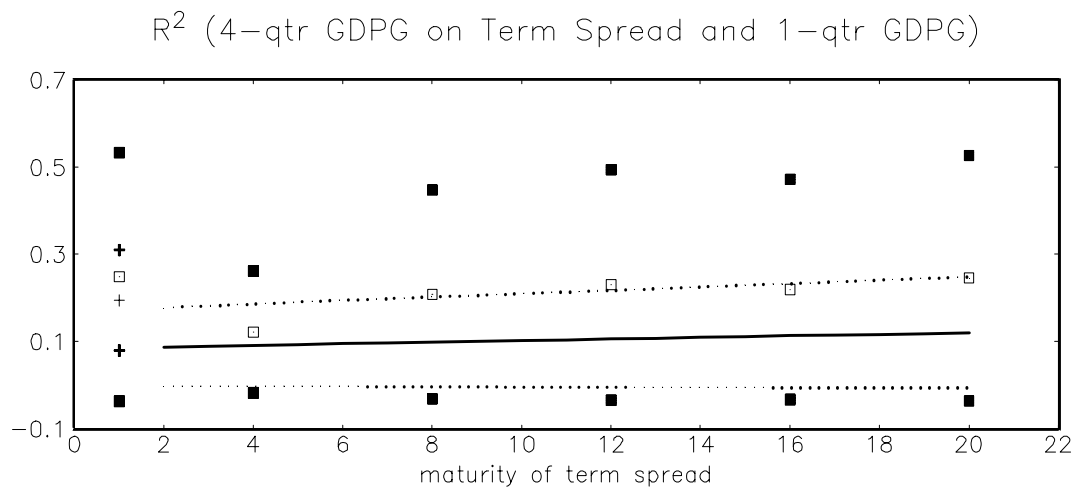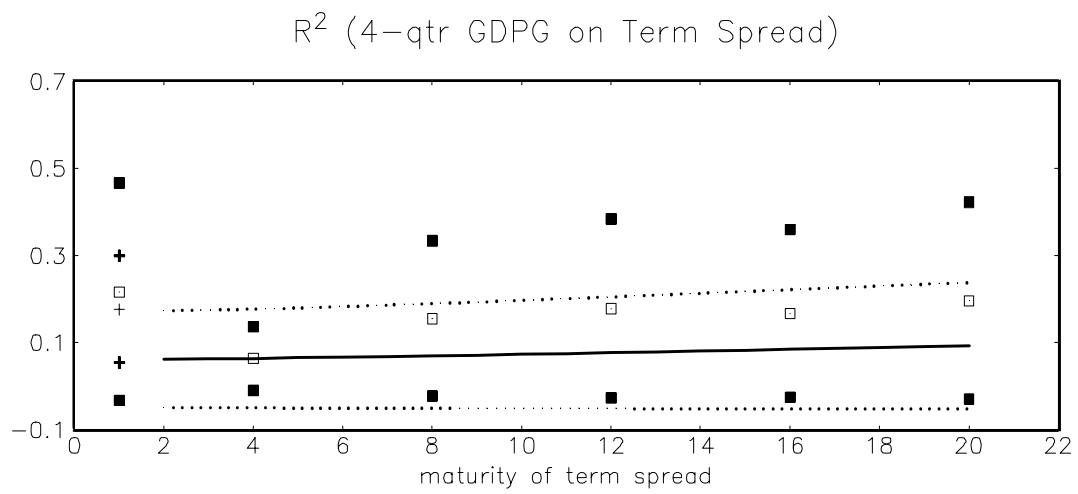
Figure 2: Model and OLS R^2.

53

FIGURE 2 NOTE: In the top panel, GDP growth is regressed on the term spread; in the middle panel, the regressors are the term spread and lagged GDP growth; and in the last panel, we have a trivariate regression on the short rate, term spread and lagged GDP growth. On the x-axis, we plot the term spread maturity. For the maturity corresponding to one quarter, we plot the R^2 from a regression involving the short rate. The R^2s from OLS regressions are superimposed as empty squares, with one OLS standard error bound given by the black squares. For the regressions with the short rate, we plot the theoretical R^2s implied from the model as a gray cross and one standard error bounds are shown as black crosses. For 2- to 20-quarter spreads, the theoretical R^2s are shown as solid lines with one standard error bounds plotted as dotted lines.
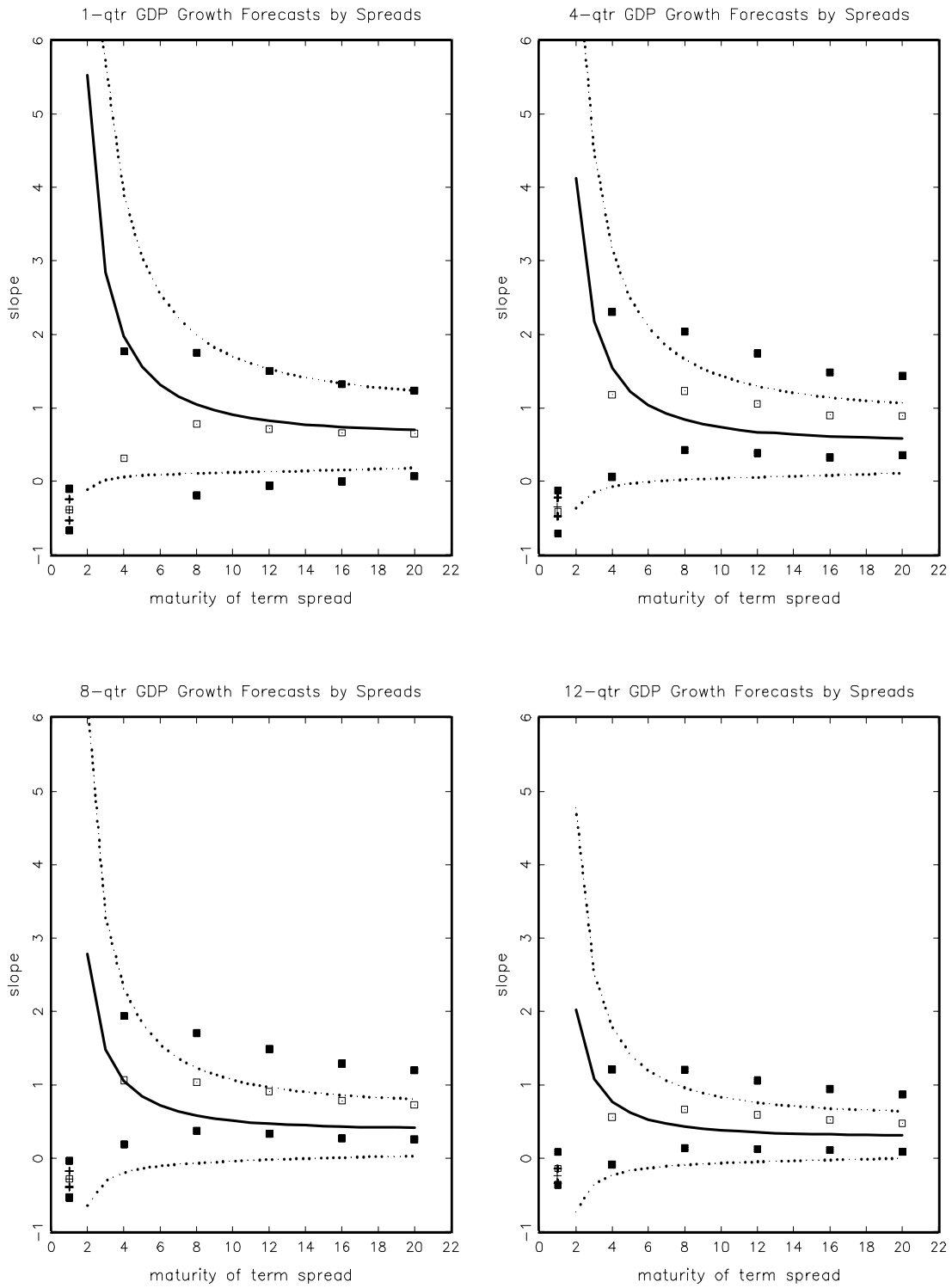
Figure 3: Model and OLS Regression Coefficients.

FIGURE 3 NOTE: Each panel of shows the term spread coefficients for a different forecasting horizon (1-, 4-, 8- or 12-quarters ahead). On the x-axis, we show the maturity of the term spread from two to twenty quarters. The solid line plots the coefficients implied from the yield-curve model with two standard errors bounds as dotted lines. The cross corresponding to $x = 1$, represents the model-implied coefficient for regressing GDP growth onto the 1-quarter short rate. The OLS coefficients are shown as empty squares, with two standard error bounds denoted by solid squares.
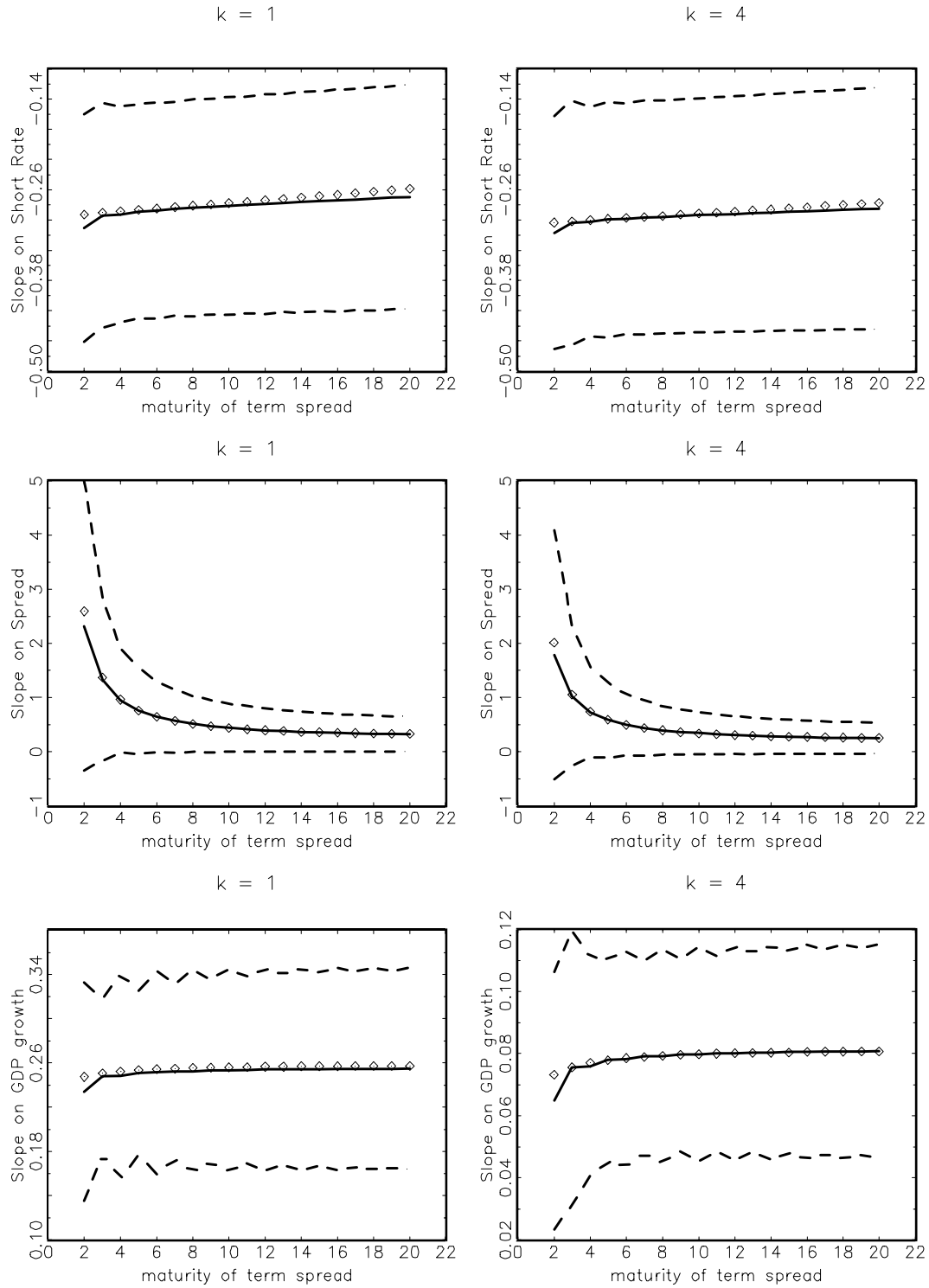
Figure 4: Latent and Observable Yield Curve Model Predictive Coefficients

FIGURE 4 NOTE: We graph the implied coefficients from a predictive GDP regression with short rate, term spread and lagged GDP regressors, as given by equation (10). In the left (right) hand column we show the coefficients for a forecast horizon of $k = 1$ quarter (4 quarters). In each plot, the x-axis denotes the maturity of the term spread used in the regression. For example, the coefficients $\beta_{1,1}^{(10)}$ for the short rate, $\beta_{1,2}^{(10)}$ for the 10-quarter term spread and $\beta_{1,3}^{(10)}$ for GDP shown in the top, middle and bottom panels on the left column, respectively, correspond to using a term spread maturity of $n = 10$ quarters for a forecasting horizon of $k = 1$ quarter. The thin solid line represents the implied predictive coefficients from the latent yield-curve model. The implied coefficients from the observable yield-curve model are shown as diamonds. The dotted lines represent two standard deviation bounds computed using the delta-method from the latent yield-curve model.