What does the Yield Curve Tell us about GDP Growth?[¤]

Andrew Ang,[y] Monika Piazzesi[z] and Min Wei[x]

February 11, 2003

Abstract

A lot, including a few things you may not expect. Previous studies …nd 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 …ndings, we predict that the short rate has more predictive power than any term spread. We con…rm this …nding by forecasting GDP out-of-sample. The model also recommends the use of lagged GDP and the longest maturity yield to measure slope. Greater e¢ciency enables the yield-curve model to produce superior out-of-sample GDP forecasts than unconstrained OLS at all horizons.

[¤]We thank seminar participants at Columbia University, Northwestern University and an NBER EFG meeting. We thank Geert Bekaert, Charlie Calomiris, Mike Chernov, John Cochrane, Bob Hodrick, Urban Jermann, Lutz Kilian, Rick Mishkin and Martin Schneider for helpful discussions. Andrew Ang acknowledges funding from the NSF.

[y]Columbia Business School and NBER, 3022 Broadway 805 Uris, New York NY 10027; Email: aa610@columbia.edu; ph: 212 854 9154; WWW: http://www.columbia.edu/»aa610

[z]UCLA and NBER. Anderson School, 110 Westwood Plaza, Los Angeles CA 90095; Email: piazzesi@ucla.edu; ph: 310 825 9544; WWW: http://www.personal.anderson.ucla.edu/monika.piazzesi

[x]Columbia Business School, 3022 Broadway 311 Uris, New York NY 10027; Email: mw427@columbia.edu; ph: 212 864 8140; WWW: http://www.columbia.edu/»mw427

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. The lower demand for long bonds during recessions lowers their price and increases their yield. In contrast, yields on short bonds are procyclical because of monetary policy. The Federal Reserve lowers short yields in recessions in an e¤ort 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 with the slope of the yield curve in OLS regressions.[1] The higher the slope or term spread, the larger GDP growth is expected to be in the future. The slope is usually measured as the di¤erence between the longest yield in the dataset and the shortest maturity yield. 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; and Estrella and Mishkin, 1998). Finally, 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 impose more structure on the model 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, since bond yields are themselves dynamic, predictive regressions do not take into account the endogenous nature of the regressor variables. We would also expect that the entire yield curve, not just the arbitrary maturity used in the construction of the term spread, would have pre-

[1] See, among others, Harvey (1986, 1989 and 1993), Laurent (1988), Stock and Watson (1989), Chen (1991), Estrella and Hardouvelis (1991 and 1997), Estrella and Mishkin (1998), and Hamilton and Kim (2002), who regress GDP growth on term spreads on US data. Jorion and Mishkin (1991), Harvey (1991), Estrella and Mishkin (1997), Plosser and Rouwenhorst (1994), Bernard and Gerlach (1998), and Dotsey (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).

dictive power. Using information from 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 framework, since yields of different maturities are highly cross-correlated, it may be 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 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 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 which belong to the affine class of Duffie and Kan (1996).

The reduction in dimensionality and the cross-equation restrictions from no-arbitrage both help us to efficiently extract the business-cycle information contained in the yield curve. We find that a yield-curve model has four main advantages over unrestricted specifications. First, 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 the model recommends the use of the longest yield to measure the slope, regardless of the forecasting horizon. Second, 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. Third, the model recommends the use of lagged GDP growth as an additional regressor. GDP growth is autocorrelated and its mean-reversion especially helps in short (1-2 quarter horizon) forecasts of GDP. Finally, our arbitrage-free 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 regressor variables. The better out-of-sample performance from our yield-curve model is driven by the gain in estimation efficiency due to imposing the cross-equation restrictions implied by the absence of arbitrage.

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 and then showing how the model completely characterizes the predictive regressions. We show that imposing no-arbitrage restrictions leads to better out-of-sample forecasts of GDP than unconstrained OLS regressions or VAR's. Section 5 concludes. We relegate all technical issues to the Appendix.

2 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 we denote the yield for maturity n in quarters as $y^{(n)}$. 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 choose to therefore start their sample period in 1964. We follow their lead in this paper but we discuss the differences of including data from the immediate post-war period 1952-1964.[2]

Economic activity can be measured in different ways. We look at two alternative measures, real GDP growth rates and NBER recessions. Data on real GDP is seasonally adjusted, in billions of chained 1996 dollars, from the FRED database (GDPC1). We denote annualized log real GDP growth from t to $t+k$ as:

$$g_{t\to t+k} = 4/k \cdot (\log GDP_{t+k} - \log GDP_t).$$

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

We graph 4-quarter GDP growth, $g_{t-4\to t}$, and the 5-year term spread, $y_t^{(20)} - y_t^{(1)}$, in Figure 1 together with NBER recessions shown as shaded bars. In the top plot, negative GDP growth often coincides with NBER recessions. One conventional definition of a recession is two consecutive quarters of negative GDP growth, but the NBER takes into account other factors in defining recessions.[3] The figure shows that periods of negative 4-quarter GDP growth were always classified as recessions by the NBER, at least over the postwar sample. The correlation between quarterly GDP growth and an indicator variable for NBER recessions, that takes value 1 only during an NBER recession and zero elsewhere, is 63%. Hence, there is a strong correspondence between NBER recessions and negative economic growth. The top plot of Figure 1 shows that GDP growth has a strong

[2] In particular, our out-of-sample forecasting results for the term structure model are stronger when the model is estimated using data from 1952-1964. 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).

[3] The NBER Business Cycle Dating Committee actually places little weight on real GDP for recession dating. See http://www.nber.org/cycles/recessions.html

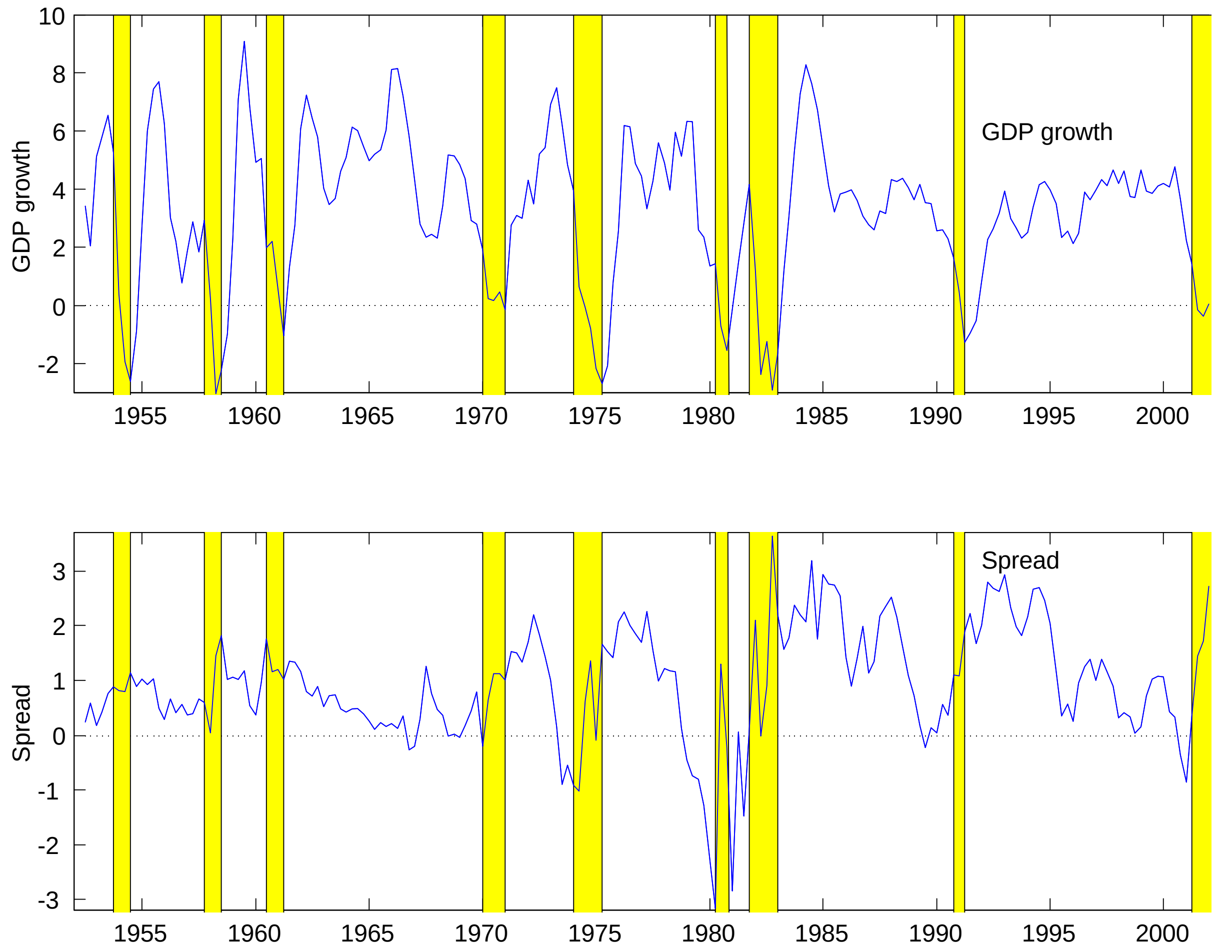


Figure 1: Term spread and GDP growth with shaded NBER recessions

cyclical pattern. GDP growth is significantly mean-reverting; one-quarter GDP growth, g_{t+1}, has an autocorrelation of 30%.

In the bottom plot of Figure 1, we graph the term spread between the 5-year and 1-quarter zero-coupon bonds. The term spread averages 0.99% over the sample period, reflecting the normal upward sloping pattern of the yield curve. To document the relationship between the term spread and GDP growth, we focus on the behavior of the spread before and during the periods of NBER recessions marked by shaded bars. There are nine recessions during the post-1952 period. All except the first three recessions are preceded by negative term spreads. However, the figure shows that there is a difference in the time interval when the term spread becomes negative and when a recession is declared. Moreover, there are times like 1966 where the term spread is negative but are not followed by NBER recessions. To examine this in detail, Table 1 lists periods of negative term spreads and recessions.

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 Figure 1 shows that it is followed by a period of relatively slower GDP growth. While all the post-1964 recessions are predicted by inverted yield curves, the initial lead time between the onset of the inversion and the start of the NBER recession varies between 2 to 6 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 before the 1973:Q4-1975:Q1 recession, while the yield curve was upward sloping when the 2001:Q1-2002:Q1 recession began (having been downward sloping from 2000:Q3-2000:Q4).

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

The predictive power of the yield spread for economic activity, documented in Figure 1 and Table 1, can be formalized in a predictive regression of the form:

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

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). 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. Ang and Bekaert (2001) show that Hodrick (1992) standard errors have negligible size distortions, unlike standard OLS or Newey-West (1987) standard errors. The correct choice is important, because inappropriate standard errors can vastly overstate the predictability of GDP growth from the term spread. Throughout this paper, we use Hodrick (1992) standard errors for OLS regressions. Table 2 reports the results of regression (1) over 1964:Q1-2001:Q4.

Table 2: Forecasts of GDP Growth from Term Spreads

	Term Spread Maturity									
Horizon	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 $\beta_k^{(n)}$ and R^2 for equation (1). Sample period 1964:Q1-2001:Q4.

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 1-quarter ahead, like the Stock-Watson (1989) index. But, in predicting recessions 2 or more quarters into 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 would 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. This 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 differs across maturity. 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 regression (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 from 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 regressor variables 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 two factors from the yield curve, the short rate expressed at a quarterly frequency, $y^{(1)}/4$, to proxy for the level of the yield curve, and the 5–year term spread expressed at a quarterly frequency, $\left(y^{(20)} - y^{(1)}\right)/4$, to proxy for the slope of the yield curve. We augment these yield-curve factors by including observable quarterly real GDP growth $g_t/4 = \log GDP_t - \log GDP_{t-1}$ as the last factor. The vector of state-variables

$$X_t = 1/4 \left[y_t^{(1)}, \left(y_t^{(20)} - y_t^{(1)} \right), g_t \right]^\top ; \tag{2}$$

is thus entirely observable.

The 3 factors in X_t follow a Gaussian Vector Autoregression with one lag:

$$X_t = \mu + \Phi X_{t-1} + \Sigma \varepsilon_t ; \tag{3}$$

with $\varepsilon_t \sim IID\ N(0, I)$; and μ is a 3×1 vector and Φ is a 3×3 matrix.

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

$$m_{t+1} = \exp\left(-y_t^{(1)} - \frac{1}{2} \lambda_t^\top \lambda_t - \lambda_t^\top \varepsilon_{t+1} \right) ; \tag{4}$$

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 3×1 vector λ_0 and a 3×3 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 (2001), 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 (4) 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. Piazzesi (2001) and Ang and Piazzesi (2002) 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 find (2002) that incorporating macro factors allows for better out-of-sample forecasts of yields than only using yield-curve factors.

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 Ang and Piazzesi, 2002):

$$\begin{aligned} A_{n+1} &= A_n + B_n^\top (\mu - \Sigma \lambda_0) + \frac{1}{2} B_n^\top \Sigma \Sigma^\top B_n \\ B_{n+1} &= (\Phi - \Sigma \lambda_1)^\top B_n - e_1; \end{aligned} \tag{5}$$

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:

$$\begin{aligned} y_t^{(n)} &= -\frac{\log p_t^{(n)}}{n} \\ &= a_n + b_n^\top X_t; \end{aligned} \tag{6}$$

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

If there are no risk premia, $\lambda_0 = 0$ and $\lambda_1 = 0$; a local version of the Expectations Hypothesis (EH) 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 (5), we can see that the 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 spreads and average expected bond returns, while λ_1 controls the time variation in term spreads and expected returns.

Estimation of the Model

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 (3) 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. We compute standard errors for our parameter estimates using GMM, adjusting for the two-stage estimation procedure. Details are in the Appendix.

The 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. The OLS predictive regressions for GDP can be improved by moving to a dynamic system describing the evolution yields and GDP over time. Hence, the first source of efficiency gains over unconstrained OLS is to use information about the evolution of yields over time. The second source of efficiency gain comes from moving from a large VAR with many yields to a low-dimensional factor model. Finally, no arbitrage enables us to specify a low-dimensional system in a consistent way. Risk-adjusted expectations of the short rate in this system are consistent with the cross-section of yields.

Parameter estimates can be obtained with somewhat greater efficiency using a one-step maximum likelihood estimation. The loss in efficiency of the two-step procedure has two origins. First, the estimation of the VAR parameters μ, Φ and Σ only uses information contained in the state vector X instead of the entire cross-sectional information on yields. Second, the moment conditions are not chosen with efficiency in mind. This is especially true for the moments used to estimate the risk premia parameters λ_0 and λ_1: The latter loss in efficiency turns out to have negligible effects for forecasting GDP growth, as we show in the Appendix.

Forecasting GDP growth from the Model

The yield equation (6) 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 (1) is given by:

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

Then we can use (6) together with the implied long-run forecast for GDP growth from (3):

$$E_t[g_{t\to t+k}] = c + 4/k \cdot e_3^\top \Phi (I - \Phi)^{-1} (I - \Phi^k) X_t, \tag{7}$$

where c is a constant term and e_3 is a 3×1 vector of zero's with a 1 in the last element. Hence, we can solve for $\beta_k^{(n)}$ as:

$$\beta_k^{(n)} = \frac{4/k \cdot e_3^T \Phi (I - \Phi)^{-1} (I - \Phi^k) \Sigma_X \Sigma_X^\top (b_n - b_1)}{(b_n - b_1)^\top \Sigma_X \Sigma_X^\top (b_n - b_1)}, \tag{8}$$

where $\Sigma_X \Sigma_X^\top$ is the unconditional covariance matrix of the factors X, which is given by $\mathrm{vec}(\Sigma_X \Sigma_X^\top) = (I - \Phi \otimes \Phi)^{-1} \mathrm{vec}(\Sigma \Sigma^\top)$.

The yield-coefficients b_n together with the VAR parameters Φ and Σ (through Σ_X) completely determine the predictive regression coefficients in (8). 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. This makes it clear that in order for a yield-curve model to capture the forecastability inherent in the predictability regressions, we must model time-variation in the risk premia.

The regression coefficients (8) 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. For longer forecasting horizons k, the regression coefficient (8) exploits the long-horizon forecasts of the VAR, just as in Campbell and Shiller (1988) and Hodrick (1992).[4] Our two-step estimation procedure implies that GDP forecasts based on the $n = 20$ quarter term spread from the yield-curve model are simply the ones from a VAR. 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 and no-arbitrage to compute (8). 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. The simple regression (1) also only uses one yield-curve factor (the term spread) to predict GDP and ignores other yield-curve factors (like the first short rate level factor). For example, the yield-curve model can fully characterize each predictive coefficient in the regression:

$$g_{t \to t+k} = \alpha_k^{(n)} + \beta_{k,1}^{(n)} \cdot y_t^{(1)} + \beta_{k,2}^{(n)} \cdot \left(y_t^{(n)} - y_t^{(1)} \right) + \beta_{k,3}^{(n)} \cdot g_t + \varepsilon_{t+k,k}^{(n)}, \qquad (9)$$

of which (1) is a special case. In (9), 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. Our model also allows us to compute R^2's of each regression specification in closed-form. We detail the computations of predictive coefficients and R^2's in the Appendix.

Choice of Factors and Caveats

[4] This is also convenient as many macro studies of the relationships between monetary policy, yields and real activity work with VAR's. See, for example, Christiano, Eichenbaum and Evans (1999) and Clarida, Gali and Gertler (2000).

The choice of our factors is motivated by three considerations. First, the model needs to capture the dynamics of yields. These dynamics can be explained with a very small number of factors, because the first two principal components of yields already explain 99.7% of the variation of our yield data, consistent with Knez, Litterman and Scheinkman (1994). These 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.

Most yield-curve models use unobservable factors to capture the dynamics of yields. These factors are inferred from yield data alone. Thus, traditional latent models do not address the forecastability of GDP growth. The Appendix discusses a model that combines GDP growth with latent yield factors. The augmented latent model is equivalent to our observable yield-curve model for a particular assumption on measurement errors. The latent factor model augmented with GDP can be estimated in one step, using maximum likelihood. GDP forecasts from this model are more efficient than from our observable yield-curve model, but they turn out to be essentially identical, and take much longer to compute.

A second consideration is that the model needs to accurately describe the dynamics of GDP growth. We check whether other combinations of yields forecast GDP growth in OLS regressions (not reported). The third principal component does not enter significantly once we control for the first two principal components. Similarly, we find that a curvature transformation $(y^{(1)} - 2y^{(12)} + y^{(20)})$, does not enter significantly, once we control for the short rate and the 5–year term spread.[5] Moreover, there is little difference in the R^2 from using principal components versus our factor definitions.

One caveat of our approach is that yields are slightly skewed and heteroskedastic. We assume homoskedasticity as a first-order modeling approach and the ease in directly estimating (3) as a SUR system. However, we can incorporate heteroskedasticity by making Σ a linear function of X along the lines of Duffie and Kan (1996). However, there is potentially little to be gained by modelling heteroskedasticity for the purposes of forecasting GDP. The third principal component, or the curvature transformation, is related to interest rate volatility which has no forecasting power for GDP. This confirms Hamilton and Kim (2002), who directly use interest rate volatility and find it has little ability to forecast GDP growth.

Another caveat is that we assume that the factor dynamics are stable over time. There is substantial evidence for regime shifts in the dynamics of interest rates (see Ang and Bekaert, 2002). However, the largest difference in the dynamics of yields across regimes is time-varying volatility, which has little predictive power for GDP. Moreover, Stock and Watson (2002) find that the differences in coefficients across regimes for GDP

[5] Cochrane and Piazzesi (2002) find that higher principal components than the third enter significantly for longer forecasting horizons. These are not included in our model. Hence, our model understates the predictability of the yield curve for GDP growth.

relationships with other economic time series mostly reflect time-varying volatility. They find that the coefficients in the conditional mean of GDP growth are much more stable. Hence, we might expect that regime-switching would matter little for our GDP forecasts.

4 Empirical Results

4.1 Parameter Estimates

Table 3 contains 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 already foreshadows that term spreads may behave completely different in multivariate regressions than in the univariate regressions (equation (1)) from Table 2. GDP growth is remains significantly autocorrelated (0.258), when we control for the short rate and spreads, which are themselves highly persistent. (The 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, conditional shocks to the short rate and GDP growth are 36% correlated, and shocks to the spread and GDP growth are -16% correlated. This implies that the conditional covariance structure plays an important role in bond pricing from (5). Simple predictive OLS regressions of future economic growth assume that the term spread is exogenous. However, the short rate and spread processes with GDP growth are very much jointly endogenous.

Some of the parameter estimates of the risk premiums in Table 3 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 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 as 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 (5), 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 therefore significantly affect the predictive coefficients $\beta_k^{(n)}$ in (8).

Table 3: Parameter estimates for Yield-Curve Model

State dynamics X_t = ¹ + ©$X_{t ¡ 1}$ + §"$_t$

	¹	©			§		
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 ¸$_t$ = ¸$_0$ + ¸$_1 X_t$				Std dev of errors -	
	¸$_0$	¸$_1$			1-yr yield	2-yr yield
Short Rate	0.29	-34.18	-52.61	1.60	0.336	0.234
	(0.22)	(0.27)	(0.17)	(10.88)	(0.035)	(0.025)
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	0.164	0.126
	(4.89)	(0.09)	(0.01)	(0.38)	(0.049)	(0.034)

NOTE: Parameters ¹, § and standard deviations of measurement errors are multiplied by 100. Sample period: 1964:Q1-2001:Q4.

Table 4: The Yield Curve Implied by the Model

Maturity	Short Rate 1-qtr	Term Spreads 4-qtr	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. Sample period: 1964:Q1-2001:Q4.

The performance of the estimated model can be seen in Table 4, which 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. Standard errors implied by the model are given in parentheses. All the data moments are well within one standard error bound. This shows that our model is at least able to match the unconditional behavior of term spreads. We now investigate the ability of our model to capture the conditional moments implied by predictive GDP regressions using term spreads. We already know from the outset that the model matches time-variation in conditional expected bond premia. The reason is that the model is forced to match the short rate and 5–year yield by construction. The model thus exactly replicates the behavior of the 5-year term spread shown in Figure 1.

4.2 Characterizing the R^2 of GDP Regressions

To characterize the explanatory power of the predictive GDP regressions, we begin by computing theoretical R^2's implied by the model in Figure 2. The …gure shows the model R^2's for three regressions in each panel. In the …rst 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 1 quarter, we plot the R^2 from a regression involving the short rate as a cross. The solid lines from maturities 2 to 20 quarters are the theoretical R^2's implied by the model. The OLS regressions can be run only for a number of selected yields, but our model derives the predictive explanatory power for any spread. We superimpose R^2's from OLS regressions in empty squares, with 1 OLS standard error bound given by the black squares. While our model enables the computation of the regression R^2's from placing any horizon GDP growth on the left-hand side, the …gure 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^2's easily encompass the R^2's implied by the yield-curve model. The OLS R^2's also lie very near the model-implied R^2's. The model-implied R^2 in the last panel is ‡at at 27%. This is because our model has three factors (level, slope, and GDP growth), and two yields together with GDP growth is su¢ cient 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^2's provide us with a guide as to what to expect from the numerous OLS speci…cations. While running many di¤erent 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.

First, the model-implied R^2's are highest for the short rate speci…cations. 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 the short rate. The theoretical R^2's do not imply that we should only use short rates instead

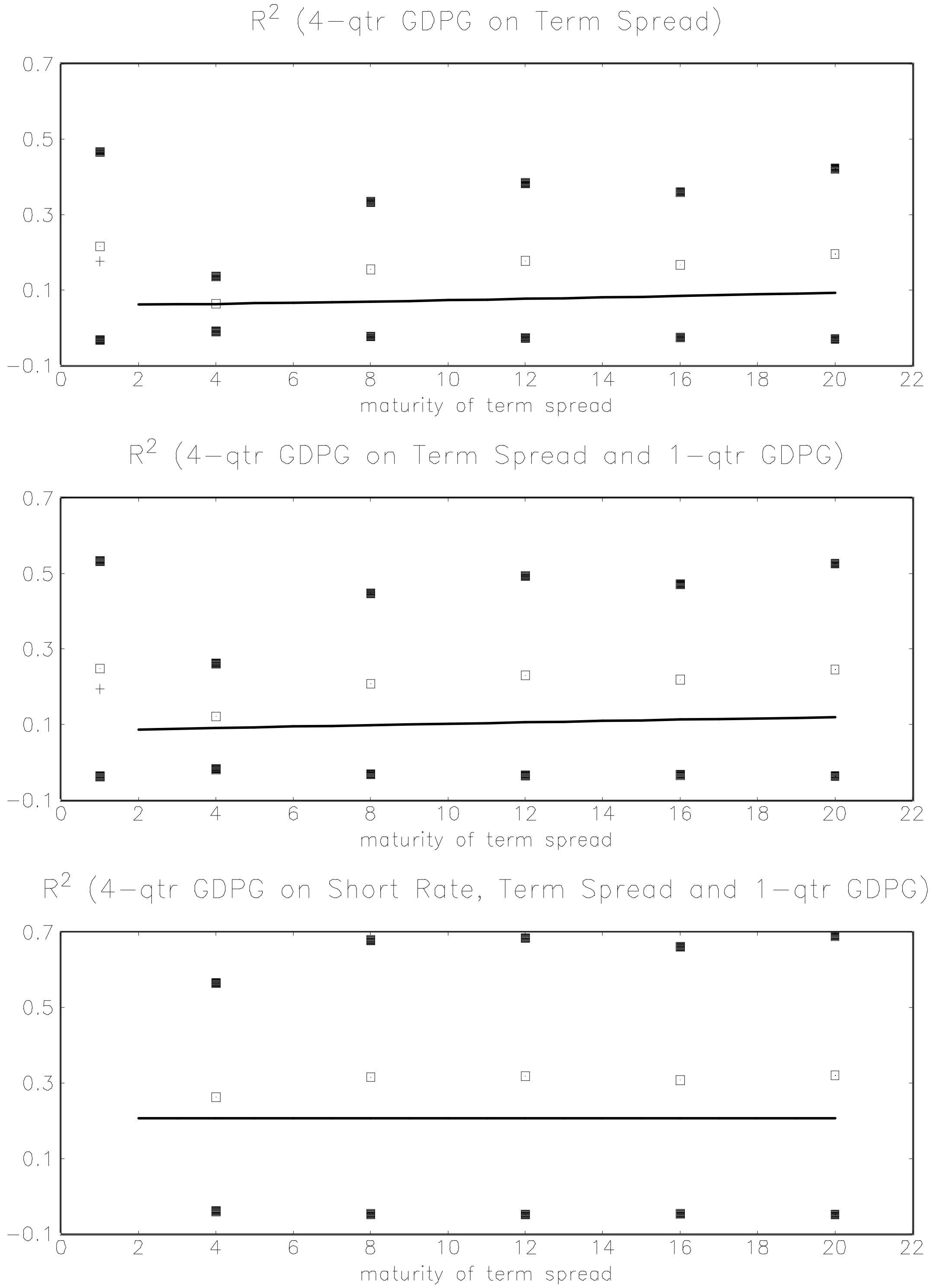


Figure 2: Model and OLS R^2.

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^2's 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^2's increase with the maturity of the term spread. We can see some OLS evidence in Table 2, where 5–year spreads have the highest significance levels. This is consistent with the previous literature using the longest maturity spreads available, instead of spreads of intermediate maturity bonds.

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 (1-2 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 maturity 1-quarter is also much higher than the term spread regressors.

To test these three predictions, we next compute the regression coefficients and then examine the out-of-sample forecasting power of the various regression specifications in Section 4.4.

4.3 GDP Regression Coefficients

To see how our model completely characterizes the GDP predictive regressions, we plot the model and OLS regression coefficients for the 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), with the same scale. On the x-axis, we show the maturity of the term spread from 2 to 20 quarters. The solid line plots the coefficients implied from the yield-curve model. 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 in empty squares, with 2 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

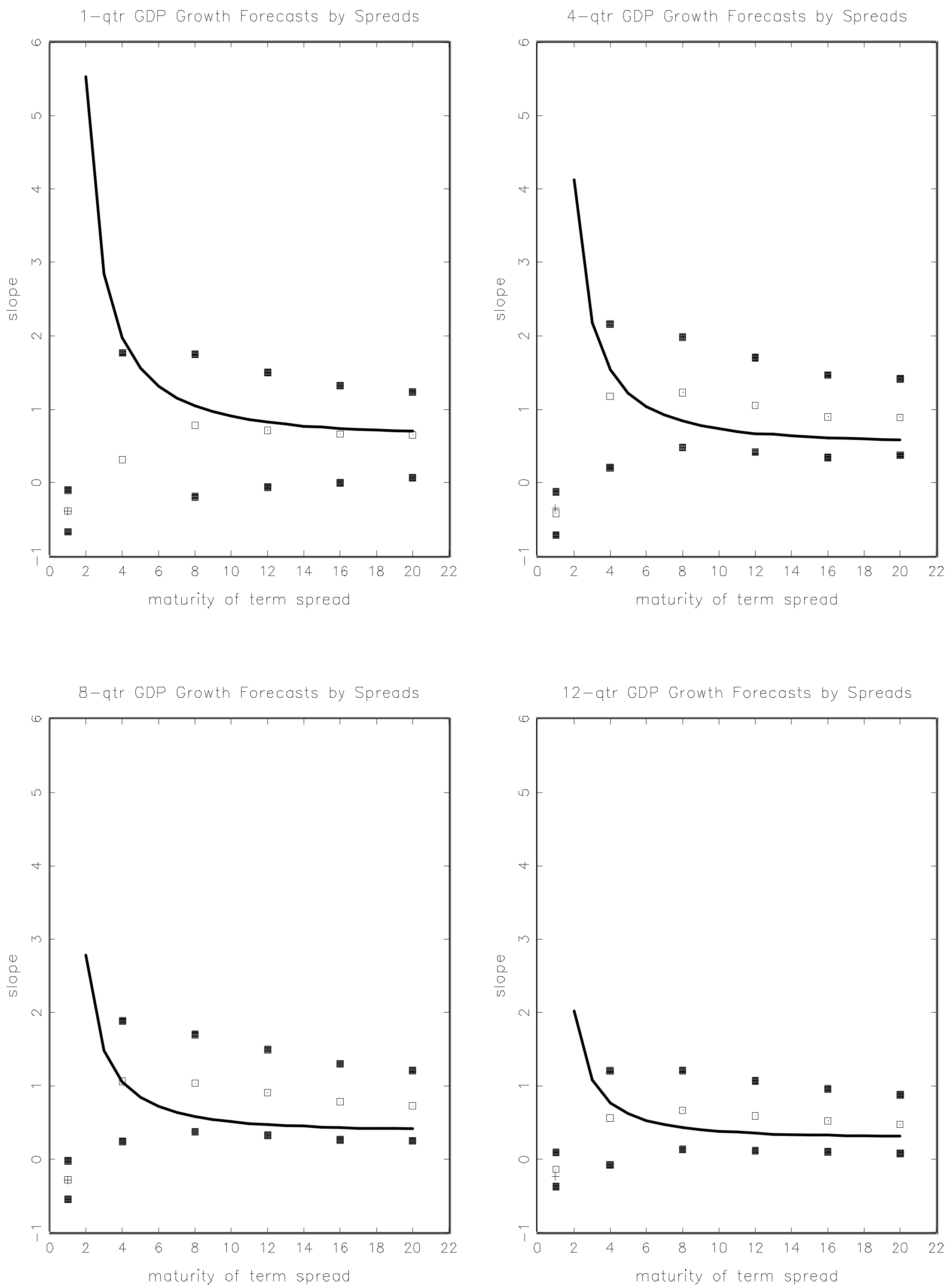


Figure 3: Model and OLS Regression Coe¢cients.

1-quarter forecasting horizon but lie 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.

The difference between the OLS coefficients and the regression coefficients implied from our yield-curve model may come from two sources. The first source may be just sampling error, which we can see from the large OLS standard error bounds in Figure 3. Second, the simple regression (1) treats the spread as an exogenous variable. However, the spread is very much an endogenous regressor and so the OLS coefficient suffers from endogenous variable bias. We capture the dynamics of the spread, jointly with GDP growth in the factor process (3), and model the relationship of yields to the factors via the prices of risk. The difference between OLS and our model-implied coefficients can be negative (such as for the 1-quarter forecasting horizon) or positive, but for long forecasting horizons the OLS coefficient for the term spread is higher than the model-implied coefficient. Although this bias is not statistically significant, it may impact out-of-sample forecasting performance, which we examine below.

We focus on some selected regression coefficients in Table 5. 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 5-year term spread to be comparable to the literature. Note that since the 5-year term spread is an observable factor, the model-implied coefficients in Table 5 are just coefficients implied from long-horizon VAR's.

In Table 5, 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 (3). There is a difference in the standard errors across the model-implied and OLS columns because the model also takes into account the standard errors of all the other parameters in computing the standard errors for the regression coefficients and we use Hodrick (1992) standard errors for the OLS regressions. As the forecasting horizon increases from 1-quarter to 12-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 significantly negative 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. The OLS short rate coefficients, like the model-implied coefficients, are also significantly negative when GDP is regressed solely on the short rate at forecasting horizons 1-8 quarters. However, at the 12-quarter horizon, the OLS coefficient on $y_t^{(1)}$ is -0.14 and insignificant in the univariate regression. The OLS short rate coefficient is also small and insignificant (-0.07) in the trivariate regression.

Table 5: Model-Implied and OLS Coefficients

Horizon	Model Implied Coefficients			OLS Coefficients		
qtrs	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)	(0.31)	(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)

Second, the model-implied coefficients assign more predictive power to the short rate than OLS does. 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 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, however, not statistically significant. The OLS coefficients in Table 5 lie within two standard error bounds of the model-implied coefficients, where the standard errors are due solely to sampling error of the OLS coefficients. The model-implied coefficients also lie within two standard error bounds of the OLS coefficients, where the standard errors are computed using the covariance matrix of the model coefficients. 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. The model implies that the short rate, not the spread, should be the most powerful predictor of GDP growth while OLS implies that the spread has the highest explanatory power. Before testing these implications with out-of-sample forecasting, we first try to distinguish what drives the differences in the coefficients.

The sources of the model-implied and OLS discrepancy are either small sample error or endogenous regressor bias from the omission of modelling the process for the regressor variables in the OLS regression. With endogenous regressors in univariate regressions, the bias can be signed and the magnitude computed. The bias depends on the autocorrelation of the regressor variable and the correlation of shocks to the regressor and regressand (see Stambaugh, 1999). However, in a multivariate regression, characterizing the bias in closed-form is not possible. We examine the small sample distribution of the coefficients implied by the model in Table 6 for the long horizons, k = 8 and 12 quarters, where the differences between the model-implied coefficients and OLS coefficients are pronounced.

Table 6: Endogenous Regressor Bias

Horizon k-qtrs		Model Implied Coefficients: Short Rate	5-Year Term Spread	GDP Growth
8	Model-Implied Coefficient	-0.25	0.14	0.04
	Small Sample Coefficient	-0.25	0.10	0.02
12	Model-Implied Coefficient	-0.22	0.08	0.02
	Small Sample Coefficient	-0.21	0.05	0.01

Table 6 repeats the model-implied coefficients from Table 5 for comparison. For each forecasting horizon, we also report the small sample coefficient computed as follows. From the model (3), we simulate out a small sample of exactly the same length as our data sample 1964:Q1-2001:Q4. Then, we compute yields using (6), run the trivariate regression (9) on the small sample, and record the coefficients. We repeat this 10,000 times to build a distribution of the small sample coefficients. The mean of this distribution is the expected small sample coefficient. The small sample coefficients for both $y_t^{(1)}$ and $(y_t^{(20)} - y_t^{(1)})$ show a slight downward bias. In contrast, the OLS coefficients in Table 5 are much higher than the model-implied coefficients. For example, for k = 8 quarters, the OLS coefficients for the short rate and spread are -0.19 and 0.56, respectively. Hence, endogenous regressor bias is not the source of the discrepancies between the OLS and model-implied coefficients.

The remaining explanation of the model-implied and OLS difference is sampling error, both from estimating the model and from estimating the OLS coefficients. When we estimate the yield-curve model, we use information across the entire curve and impose no-arbitrage restrictions that bring greater efficiency. In contrast, the OLS regressions use only yields from particular maturities, that are treated as exogenous variables. To gauge the model and OLS predictions, we now conduct an out-of-sample forecasting exercise.

4.4 Out-of-Sample Forecasts

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 su¢ ciently 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 …rst estimate the model using 105 data points from 1964:Q1 to 1990:Q1 and forecast k-quarter GDP growth from 1990:Q1 ($k > 1$). Then at each later point in time t, we re-estimate the model again using 105 observations up to time t, and forecast for horizons $t + k$.

In our forecasting exercise, we run a horse race between forecasts implied by the yield-curve model (denoted by "M") and forecasts from simple OLS regressions (denoted by "OLS"). We consider three regression speci…cations. First, we use only the spread as in the standard spread regression (1) (M1 and OLS1). Second, we use both the spread and lagged GDP growth (M2 and OLS2). Finally, we consider a trivariate regression (9) of the short rate, spread and lagged GDP growth (M3 and OLS3). Each of these regressions can use spreads of di¤erent maturities. We also consider the k-horizon forecast implied from an unconstrained tri-variate VAR with 1 lag, comprising the short rate, term spread and GDP.[6] If the term spread used is the 5–year spread, then our model simpli…es to a VAR. We do not consider predictors other than yields in this horse race because Estrella and Mishkin (1996) document that in forecasting GDP, forecasts from the yield curve dominate many other predictors including stock market indices, default spreads, monetary aggregates and other macro variables.

In out-of-sample forecasts, very often it is the most parsimonious statistical models that yield the best forecasts (see Meese and Rogo¤, 1983) 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 7 reports the out-of-sample forecast results. We report RMSE as a percentage of the RMSE from a benchmark AR(1) forecast (so the AR(1) forecast corresponds to 1.000). Since GDP growth is slightly autocorrelated (30%), this is a reasonable benchmark that is also used by Stock and Watson (2001) in the context of forecasting GDP and other macro-series. Table 7 groups the RMSE ratios by each regression speci…cation (1, 2 and 3). The 1-quarter maturity column in Table 7 corresponds to a short rate regressor. Hence, there are blank entries in Table 7 corresponding to MLD3 and OLS3 because this regression speci…cation can only be performed for term spreads.[7] We mark the lowest RMSE ratios for each forecast horizon with asterisks.

[6] We also considered a random walk forecast, but this performs atrociously because GDP growth is strongly mean reverting. It is beaten by an AR(1) and all the speci…cations in Table 7.

[7] The blank entry for the VAR line is because the special case of using the M3 speci…cation with the 5–year term spread, the yield-curve model reduces to a VAR.

Table 7: Out-of-Sample GDP Forecasts

	Term Spread Maturity											
	1	4	8	12	16	20	1	4	8	12	16	20
	1-qtr Horizon						4-qtr Horizon					
M1	1.018	1.109	1.080	1.090	1.104	1.121	0.884	1.089	1.030	1.026	1.035	1.052
OLS1	1.133	1.087	1.084	1.101	1.116	1.130	1.041	1.039	1.088	1.109	1.130	1.169
M2	0.953*	1.027	1.114	1.009	1.018	1.030	0.861*	1.048	0.991	0.986	0.993	1.007
OLS2	1.043	1.006	1.004	1.013	1.023	1.033	1.012	1.001	1.050	1.066	1.083	1.121
M3	¡	0.975	0.973	0.982	0.990	1.000	¡	0.925	0.912	0.918	0.927	0.941
OLS3	¡	1.080	1.044	1.048	1.054	1.059	¡	0.997	1.060	1.091	1.117	1.151
VAR	1.043	1.080	1.044	1.048	1.054	¡	0.984	1.017	0.990	0.993	1.000	¡
	8-qtr Horizon						12-qtr Horizon					
M1	0.776*	0.973	0.936	0.934	0.953	0.977	0.722	0.802	0.781	0.798	0.818	0.842
OLS1	0.851	0.970	1.068	1.099	1.143	1.190	0.742	0.841	0.869	0.893	0.927	0.955
M2	0.777	0.963	0.924	0.919	0.935	0.956	0.715*	0.780	0.762	0.777	0.795	0.817
OLS2	0.880	0.983	1.086	1.122	1.170	1.225	0.762	0.859	0.894	0.924	0.966	1.002
M3	¡	0.843	0.840	0.844	0.857	0.871	¡	0.723	0.729	0.743	0.755	0.768
OLS3	¡	0.908	1.079	1.148	1.212	1.281	¡	0.797	0.918	0.991	1.048	1.116
VAR	0.849	0.877	0.843	0.849	0.865	¡	0.745	0.779	0.729	0.732	0.741	¡

NOTE: Table entries are RMSE ratios relative to an AR(1). The out-of-sample period is 1990:Q1-2001:Q4 and we start the estimation in 1964:Q1. We mark the lowest RMSE ratios for each horizon with asterisks.

The …rst striking result in Table 7 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 advocate using short rates as predictors. Holding the regression speci…cation constant, the model RMSE ratios for the short rate (term spread = 1 column) are always lower than the RMSE ratios in the other term spread columns. This result is also repeated for the OLS RMSE ratios, with an exception for OLS1 and OLS2 at the 1-quarter horizon. At the other longer horizons, both the OLS1 and OLS2 speci…cations with the short rate beat the forecasts using any term spread. While the OLS in-sample results (Table 5) advocate the use of the term spread to forecast GDP, the out-of-sample results con…rm the predictions from the yield-curve model that stress the use of the short rate.

Second, the out-of-sample forecasts implied by the yield-curve model are generally better than the corresponding OLS forecasts. In particular, for the trivariate speci…cations, M3 always beats its OLS counterpart, OLS3, at every horizon. In the other regression speci…cations using less information, such as the basic spread regression (equation (1)), the results are more mixed. For the …rst and second regression speci…cations there is no clear winner. Nevertheless, the advantages of the model can be seen when all three factors are used to predict GDP since the M3 RMSE ratios are among the lowest numbers in the table. This result is not unexpected, since the yield-curve model is able to incorporate information from the whole curve more e¢ciently than OLS, where the regressors are exogenous. This demonstrates that adding no-arbitrage restrictions can improve forecasts of GDP.

Third, incorporating lagged GDP and the short rate together with the term spread produces superior forecasts than just using the term spread. The theoretical R^2 results were slightly higher for including all three factors and this is con…rmed by looking at the RMSE ratios for the trivariate speci…cation implied by the model (M3). When including the spread, M3 outperforms both the M1 and M2 speci…cations that omit lagged GDP and the short rate, and the short rate only, respectively. However, the better forecasts obtained by using short rates, lagged GDP and term spreads are not shared by the OLS regressions. For example, at the 12-quarter horizon, using all three variables in OLS3 produces worse forecasts than using only the spread (OLS1) for spread maturities greater than 8 quarters. This is consistent with the in-sample OLS evidence in Table 5, where the short rate coe¢cient is not signi…cant when placed together with the spread. It is only by imposing no-arbitrage restrictions that we obtain superior forecasts by using short rates and term spreads together, and the bene…t, at long horizons, comes mainly from the inclusion of the short rate.

Finally, unconstrained long-horizon VAR forecasts are beaten by M3 forecasts (except for using the 12 and 16-quarter spreads at the 12-quarter forecasting horizon). The VAR forecasts are obtained by using a trivariate speci…cation of the short rate, term spread and GDP growth. A di¤erent VAR is estimated for each di¤erent term spread and we compute the long-horizon forecast implied by the unconstrained VAR. For the case of the 1-quarter maturity, we …t a bivariate VAR. While the forecasts from the yield-curve model can also be interpreted as using a VAR to infer long-horizon coe¢cients (equation (7)), our model additionally imposes restrictions from the absence of arbitrage. The e¢ciency gains from the no-arbitrage restrictions do lead to better out-of-sample performance. However, the main advantage of the model comes from using the VAR factor dynamics (equation (3)).

The superior model forecasting results relative to OLS come mainly from the fact that the model-implied regressor coe¢cients place more weight on the short rate and less weight on the spread than OLS (see Table 5), especially at long horizons. In particular, the model allows the short rate to have greater impact (with a negative sign) on future GDP and about a third of the weight as an unconstrained OLS regression for the term spread coe¢cient. At a 3-year horizon, the model places almost zero weight on the term spread. 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 Figure 1 and Table 1).

Our forecasting results are subject to a number of important quali…cations. For example, 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 5), there may be issues of structural stability of these relations, particularly since monetary policy has undergone several di¤erent 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 signi…cant 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^2's 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 7, RMSE ratios for longer term spread maturities should be lower, but they are not.

While we have presented results for beginning the estimation in 1964:Q1, we stress that if we start the estimation in 1952:Q2 and use the same out-of-sample period, the results are even more favorable to the yield-curve model. We report these results in Table 8, which repeats all the major …ndings of Table 7, except, the performance of the yield-curve model is even better. The yield-curve model is a better predictor than OLS for every regression speci…cation and the model always outperforms an unconstrained VAR speci…cation.

Table 8: Out-of-Sample GDP Forecasts Starting 1952:2

	Term Spread Maturity											
	1	4	8	12	16	20	1	4	8	12	16	20
	1-qtr Horizon						4-qtr Horizon					
M1	1.007	1.027	1.034	1.052	1.064	1.073	0.822	0.942	0.920	0.926	0.933	0.942
OLS1	1.109	1.093	1.089	1.113	1.129	1.143	0.933	0.974	1.020	1.059	1.087	1.124
M2	0.946	0.978	0.978	0.987	0.993	0.998	0.810*	0.924	0.899	0.900	0.905	0.911
OLS2	1.009	1.007	1.006	1.020	1.031	1.040	0.909	0.948	0.998	1.034	1.058	1.094
M3	¡	0.942*	0.943	0.951	0.957	0.961	¡	0.824	0.816	0.820	0.824	0.830
OLS3	¡	1.032	1.021	1.034	1.044	1.051	¡	0.908	0.978	1.016	1.043	1.075
VAR	1.009	1.032	1.021	1.034	1.044	¡	0.899	0.908	0.922	0.939	0.951	¡
	8-qtr Horizon						12-qtr Horizon					
M1	0.831*	0.963	0.962	0.968	0.985	1.002	0.809	0.835	0.847	0.868	0.886	0.902
OLS1	0.927	0.984	1.114	1.193	1.258	1.306	0.844	0.853	0.907	0.978	1.041	1.068
M2	0.836	0.970	0.963	0.964	0.978	0.992	0.807	0.835	0.843	0.861	0.876	0.890
OLS2	0.937	0.995	1.128	1.210	1.277	1.326	0.853	0.862	0.921	0.994	1.059	1.088
M3	¡	0.855	0.858	0.862	0.870	0.877	¡	0.795*	0.808	0.819	0.828	0.835
OLS3	¡	0.959	1.127	1.202	1.269	1.309	¡	0.836	0.922	0.996	1.063	1.086
VAR	0.902	0.904	0.924	0.947	0.966	¡	0.833	0.853	0.845	0.862	0.878	¡

NOTE: The estimation starts in 1952:Q2 as opposed to 1964:Q1, as in Table 7.

In summary, while the point estimates of the predictive coe¢cients in the various regression speci…cations are well within OLS con…dence bounds (see Table 5 and Figure 3), the magnitudes of the model-implied coe¢cients are quite di¤erent from the OLS estimates. In out-of-sample forecasting, this di¤erence becomes important, particularly for the weight assigned to the short rate. By imposing no-arbitrage restrictions, our estimates are more e¢cient 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 e¢ciency gains matter.

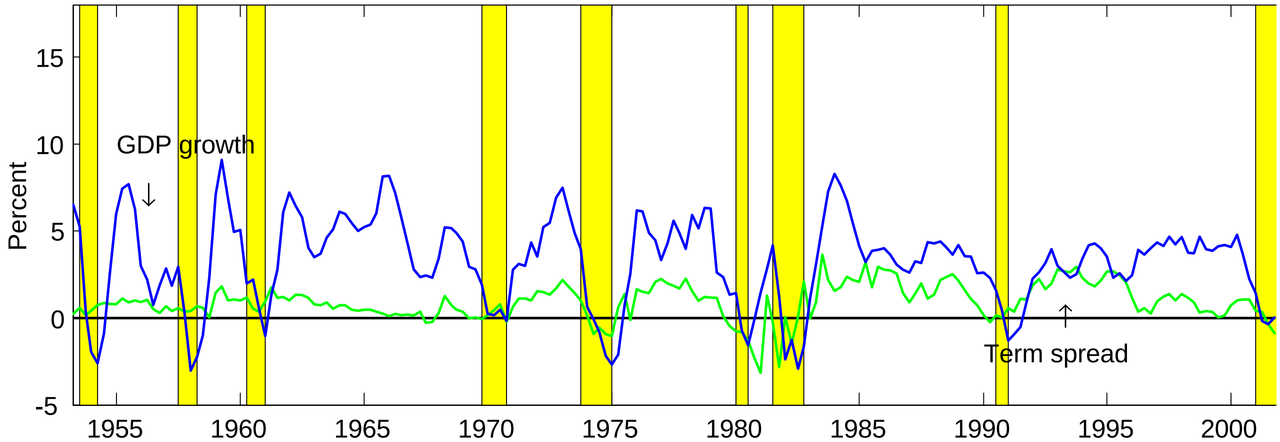
4.5 What is Driving the Short Rate Predictability?

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. In Figure 4, we plot four-quarter GDP growth shown with 20-quarter term spreads and 1-quarter short rates lagged four quarters. For example, in 1990, we plot GDP growth from 1989-1990 together with the term spread and short rate at 1989.

Figure 4 shows 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 4 shows that prior to 1970, the term spread is almost ‡at and has little to do with GDP ‡uctuations. 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 4, lagged short rates and GDP are negatively correlated throughout the whole sample period. Table 9 makes this clear.

Table 9 reports unconstrained OLS coe¢cients 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 signi…cantly large negative weights on the short rate. Prior to 1970, we see that the 5-year spread is also insigni…cant (but also has negative point estimates) and the coe¢cients on short rates are large in magnitude, highly statistically signi…cant, and negative. It is only from 1971-1989 that the term spread signi…cantly 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 5). Hence, although the 1990's favor the short rate, this period is by no means special. OLS over the full sample favors the term spread rather than the short rate. In contrast, the forecasting coe¢cients implied by our term structure model always favor the short rate instead of the term spread.



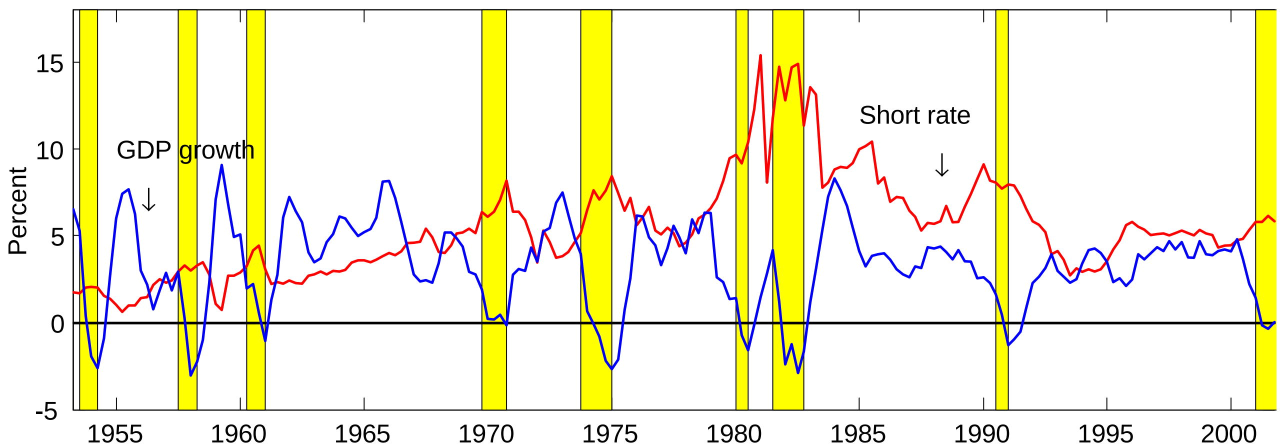


Figure 4: Four-quarter GDP growth, with 20-quarter term spreads and 1-quarter short rates lagged four quarters.

Table 9: Forecasting GDP Growth

	1964:Q1 - 1970:Q4			1971:Q1 - 1989:Q4			1990:Q1 - 2001:Q4		
Horizon 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.

Is it In‡ation 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 in‡ation. 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 in‡ation over the last quarter. Since we use the ex-post real rate, the two sum up to the nominal short rate. To decompose the e¤ect of the nominal rate into the real rate and in‡ation, we …t a 4-variable VAR of in‡ation, the real rate, the 5-year term spread and GDP. We compute long-horizon coe¢cients of each variable for forecasting in‡ation k-quarters out.

Table 10 reports the GDP forecasting coe¢cients where we distinguish between the real rate and in‡ation. We compare the 4-variable in‡ation-real rate system with the last 3 columns, which report the results from the yield-curve model based on the nominal rate. We can see that the forecasting power from the nominal rate is mostly due to in‡ation. While both in‡ation and the real rate have negative coe¢cients, the in‡ation coe¢cients are larger in magnitude than the real rate coe¢cients. Moreover, the real rate coe¢cients are all insigni…cant while the in‡ation coe¢cients are all signi…cant at the 5% level. The magnitude of the in‡ation coe¢cient is also roughly comparable to the coe¢cient on the nominal rate in the yield-curve model. We can conclude that it is the in‡ation content of the nominal rate that has predictive power for future economic growth. These results are consistent with Stock and Watson (1999a), who …nd that the nominal rate is a leading business-cycle indicator, while the ex-post real rate is less cyclical.[8] Bernanke and Blinder (1992) argue that the forecasting power of the nominal rate may be due to monetary policy.

Table 10: Forecasting GDP

Horizon	VAR with In‡ation and Real Rate				Yield-Curve Model		
k-qtrs	In‡ation	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: Sample Period 1964:Q1 to 2001:Q4.

[8] GDP growth also has predictive power for in‡ation, as Stock and Watson (1999b) and others …nd.

4.6 Forecasting Recessions

Up to now we have considered only forecasting GDP growth. We now shift our attention to forecasting recessions, which are dummy variables. We consider two definitions of recessions. In Table 11, the 'NBER recessions' refer to a dummy variable which is 1 if the NBER declared a recession for that quarter and 0 otherwise. 'Actual' refers to a dummy variable which is 1 if GDP growth is negative in that quarter and 0 otherwise. The correlation between these two variables is 69%, as indicated in the second column of Table 11. To assess how well various models can forecast these recessions, we compute the correlation between the model forecast and these dummy variables. We concentrate on two sets of right-hand side variables. The first set, M1 and OLS1 use only the spread, and is the set-up of Estrella and Hardouvelis (1991). The second set, M3 and OLS3, is the full trivariate specification (the short rate, spread and lagged GDP growth) implied by the yield-curve model.

Table 11: Recession forecasts

	NBER recessions					Actual Negative GDP Growth			
Spread	Actual	M1	OLS1	M3	OLS3	M1	OLS1	M3	OLS3
1	0.689	0.378	0.378	–	–	0.266	0.266	–	–
4	0.689	0.218	0.160	0.586	0.623	0.165	0.146	0.360	0.373
8	0.689	0.233	0.226	0.616	0.622	0.191	0.190	0.380	0.383
12	0.689	0.254	0.251	0.620	0.624	0.187	0.190	0.378	0.380
16	0.689	0.251	0.250	0.622	0.624	0.191	0.191	0.380	0.382
20	0.689	0.275	0.274	0.626	0.626	0.192	0.192	0.380	0.380

NOTE: Table entries are correlations between in-sample forecast and recession dummies computed over the sample 1964:Q1 to 2001:Q4.

At each point in time, we compute the probability of negative GDP growth implied by the model and OLS for the various regressor specifications. That is, using the model-implied or OLS coefficients, we can compute the probability that GDP growth next quarter will be negative, since this probability is a function of the conditional mean and volatility. We compare different term spreads as regressors. In Table 10, 'spread' refers to the maturity of the spread used in the specification, ranging from 4 to 20 quarters. The row corresponding to '1' refers to the 1-quarter yield itself.

The exercise we conduct in Table 11 is in-sample. It is very hard to run an out-of-sample exercise for forecasting recessions, rather than GDP growth as in the previous section, because the frequency of recessions is very low. This is especially true over the 1990's, which experienced mostly one large economic expansion, that would comprise the majority part of any out-of-sample period. Even for the in-sample exercise, we note that there are only 7 recessions during the 1964:Q1-2001:Q4 period, so we must interpret this exercise with caution.

The yield-curve model does slightly better at forecasting NBER recessions than the unrestricted OLS regression, when only the term spread is used. When we compare columns M1 and OLS1 in the NBER recession panel, we see that this is true for all spread maturities. When we include all right-hand side variables in the M3 and OLS3 columns, the yield-curve model does worse when the 4-quarter spread is used and is comparable to OLS for the other spread maturities. When we forecast actual quarters of negative GDP growth in the "Actual Negative GDP Growth" panel, our model and OLS have roughly the same performance. Overall, our model does slightly better when we forecast NBER recessions than actual negative periods of GDP growth. The reason may be that NBER recessions represent more dramatic economic downturns than quarters of negative GDP growth and thus are more likely to show up in the yield curve.

5 Conclusion

Imposing restrictions from theory usually only helps in extracting information from data when the theory is right. The absence of arbitrage in bond markets is an assumption which is extremely reasonable. We present a ‡exible arbitrage-free model of yields and GDP growth that can be easily estimated and gives us a number of advantages to forecasting future economic growth. First, the yield-curve model guides us in choosing the right spread maturity in forecasting GDP growth. We …nd that the maximal maturity di¤erence is the best measure of slope in this context. Second, the nominal short rate dominates the slope of the yield curve in and out of sample in forecasting GDP growth. This …nding is robust to the maturity of the yields used to compute the slope. Third, lagged GDP growth is informative about the future economic activity, and should not be omitted from the predictive regression speci…cations, especially for short forecasting horizons. Finally, imposing no-arbitrage restrictions allows us to predict GDP out-of-sample better than OLS. This …nding is independent of the forecasting horizon and the combination of right-hand side variables used. While our model does not take a stance on the equilibrium structure of the economy, we can certainly improve upon unrestricted forecasts by imposing no-arbitrage restrictions.

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 (3) 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:

$$\min_{\{\lambda_0;\lambda_1\}} \sum_{t=1}^{T} \sum_{n=1}^{N} \left(\hat{y}_t^{(n)} - y_t^{(n)} \right)^2 \tag{10}$$

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 (6). 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 σ in the second stage.

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 (10). The standard errors we compute adjust for the two-stage estimation process. This is done as follows.

Let $\theta_1 = \{\mu; \Phi; \Sigma\}$ and $\theta_2 = \{\lambda_0; \lambda_1; \sigma\}$ 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 Gaussian-VAR(1) process to the observed state variable series and estimate θ_1 by solving the system of equations:

$$\sqrt{T} g_1\left(\bar{\theta}_1\right) = 0; \tag{11}$$

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:

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

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

$$G_{2;2}' \sqrt{T} g_2\left(\bar{\theta}_2; \bar{\theta}_1\right) = 0; \tag{12}$$

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's approximation:

$$\begin{aligned} \sqrt{T} g_1\left(\bar{\theta}_1\right) &= \sqrt{T} g_1(\theta_1) + G_{1;1}\sqrt{T}\left(\bar{\theta}_1 - \theta_1\right); \\ \sqrt{T} g_2\left(\bar{\theta}_2; \bar{\theta}_1\right) &= \sqrt{T} g_2(\theta_2; \theta_1) + G_{2;1}\sqrt{T}\left(\bar{\theta}_1 - \theta_1\right) + G_{2;2}\sqrt{T}\left(\bar{\theta}_2 - \theta_2\right); \end{aligned}$$

which can then be substituted back into the moment conditions (11) and (12) to yield:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -\sqrt{T} g_1(\theta_1) \\ -G_{2;2}'\sqrt{T} g_2(\theta_2; \theta_1) \end{bmatrix} - \begin{bmatrix} G_{1;1} & 0 \\ G_{2;2}' G_{2;1} & G_{2;2}' G_{2;2} \end{bmatrix} \begin{bmatrix} \sqrt{T}\left(\bar{\theta}_1 - \theta_1\right) \\ \sqrt{T}\left(\bar{\theta}_2 - \theta_2\right) \end{bmatrix} \tag{13}$$

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 (13) as:

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

The asymptotic variances of the parameter estimates are thus:

$$\begin{aligned} \operatorname{var}\left[\sqrt{T}\left(\bar{\theta}_1 - \theta_1\right)\right] &= G_{1;1}^{-1}\,\Omega_{11}\,G_{1;1}^{\prime -1} \\ \operatorname{var}\left[\sqrt{T}\left(\bar{\theta}_2 - \theta_2\right)\right] &= \left[-\left(G_{2;2}' G_{2;2}\right)^{-1}\left(G_{2;2}' G_{2;1}\right) G_{1;1}^{-1} \quad \left(G_{2;2}' G_{2;2}\right)^{-1} G_{2;2}'\right]\Omega \\ &\quad \left[-\left(G_{2;2}' G_{2;2}\right)^{-1}\left(G_{2;2}' G_{2;1}\right) G_{1;1}^{-1} \quad \left(G_{2;2}' G_{2;2}\right)^{-1} G_{2;2}'\right]' \end{aligned}$$

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

Computation of Regression Coefficients

In this section we detail the computation of the regression coefficients $\beta_{k;i}^{(n)}$ and the R^2 statistic for the most general regression (9), which regresses the k-quarter GDP growth

$g_{t \to t+k}$ onto the short rate, n-qtr term spread and the current-quarter GDP growth. For 2 yield-curve factors and a GDP factor ordered last, $g_{t\to t+k} = \frac{4}{k} e_3^T \sum_{i=1}^{k} X_{t+i}$.

Under the factor dynamics specified in Section 3, k-quarter GDP growth, $g_{t\to t+k}$, can be decomposed into the part that is in the time-t information set and future shocks:

$$
\begin{aligned}
g_{t\to t+k} &= \frac{4}{k} e_3^T \sum_{i=1}^{k} X_{t+i} \\
&= \frac{4}{k} e_3^T \sum_{i=1}^{k} \left(\sum_{j=0}^{i-1} \Phi^j \mu + \Phi^i X_t + \sum_{j=1}^{i} \Phi^{i-j} \Sigma \varepsilon_{t+j} \right) \\
&= c + 4 e_3^T \Phi \tilde{\Phi}_k X_t + \frac{4}{k} e_3^T \sum_{i=1}^{k} \sum_{j=1}^{i} \Phi^{i-j} \Sigma \varepsilon_{t+j}
\end{aligned} \qquad (14)
$$

where e_i is a column vector that picks the i-th VAR variable,

$$\tilde{\Phi}_i = \frac{1}{k} \sum_{j=0}^{i-1} \Phi^j = \frac{1}{k} (I - \Phi)^{-1} \left[I - \Phi^i \right];$$

and

$$c = 4 e_3^T \sum_{i=1}^{k} \tilde{\Phi}_i \mu = 4 e_3^T (I - \Phi)^{-1} \left(I - \Phi \tilde{\Phi}_k \right) \mu$$

is a constant. Hence the time-t expectation of the regressor is:

$$E_t [g_{t\to t+k}] = c + 4 e_3^T \Phi \tilde{\Phi}_k X_t.$$

The regressors in regression (9) are also linear combinations of the state variables and can be represented as $4(A + BX_t)$ 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 (6).

The regression coefficients $\beta_k^{(n)} = [\beta_{k;1}^{(n)}, \beta_{k;2}^{(n)}, \beta_{k;3}^{(n)}]^\top$ implied by the model can then be computed as:

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

where Σ_X is the unconditional covariance matrix of the factors X_t with $\mathrm{vec}(\Sigma_X) = (I - \Phi \otimes \Phi)^{-1} \mathrm{vec}(\Sigma)$.

Computation of R^2

By switching the order of summation, the last term in (14) can be rewritten as:

$$
\begin{aligned}
\frac{4}{k}e_3^T \sum_{j=1}^{k}\sum_{i=j}^{k} \Phi^{i-j}\Sigma\varepsilon_{t+j} &= \frac{4}{k}e_3^T \sum_{j=1}^{k}\sum_{i=0}^{k-j} \Phi^{i}\Sigma\varepsilon_{t+j} \\
&= \frac{4}{k}e_3^T \sum_{j=1}^{k}\sum_{i=0}^{j-1} \Phi^{i}\Sigma\varepsilon_{t+k-j+1} \\
&= 4e_3^T \sum_{j=1}^{k} \Psi_j \Sigma\varepsilon_{t+k-j+1} \qquad (15)
\end{aligned}
$$

Combining (14) and (15) we can compute the unconditional variance of the k-quarter GDP growth

$$
var\left(g_{t\to t+k}\right) = 4^2 e_3^T \Phi\Psi_k(\Sigma_X\Sigma_X^\top)\Psi_k^T\Phi^T e_3 + 4^2 e_3^T \sum_{j=1}^{k} \Psi_j\Sigma\Sigma^T\Psi_j^T e_3
$$

and the R^2 statistic

$$
R^2 = \frac{1}{4^2}\frac{\left(\beta_k^{(n)}\right)^T A(\Sigma_X\Sigma_X^\top)A^T\beta_k^{(n)}}{e_3^T\Phi\Psi_k(\Sigma_X\Sigma_X^\top)\Psi_k^T\Phi^T e_3 + e_3^T\sum_{j=0}^{k-1}\Psi_j(\Sigma\Sigma^T)\Psi_j^T e_3}:
$$

We compute the standard errors for $\beta_k^{(n)}$ and 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.[9] 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 2-step estimation of the latent yield-curve model, the results from the two approaches are almost identical. The extracted latent factors are almost exact transformations of

[9] 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.

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 5 graphs the implied coefficients from a predictive GDP regression with short rate, term spread and lagged GDP regressors, as given by equation (9). 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 5 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 5 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.

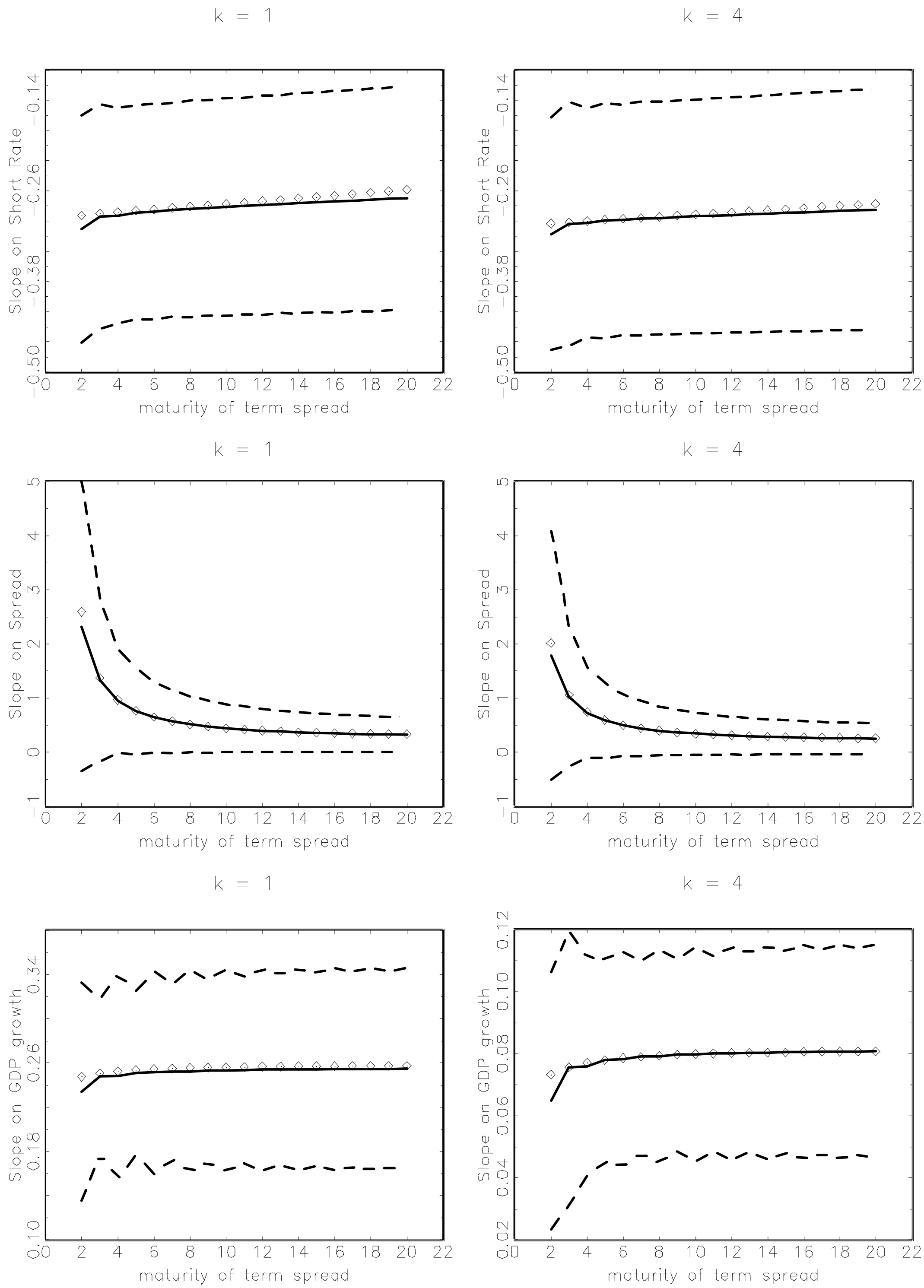


Figure 5: Latent and Observable Yield Curve Model Predictive Coe¢cients

References

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

[2] Ang, Andrew and Geert Bekaert, 2002. "Regime Switches in Interest Rates." Journal of Business and Economic Statistics, 20, 2, 163-189.

[3] Ang, Andrew and Monika Piazzesi, 2002. "A No-Arbitrage Vectorautoregression of Term Structure Dynamics with Macroeconomic and Latent Variables." Forthcoming Journal of Monetary Economics.

[4] Bernanke, Ben S. and Alan S. Blinder, 1992. "The Federal Funds Rate and the Channels of Monetary Transmission." American Economic Review, 82, 4, 901-921.

[5] Bernard, Henri and Stefan Gerlach, 1998. "Does the Term Structure Predict Recessions? The International Evidence." International Journal of Finance and Economics, 3, 195-215.

[6] Berardi, Andrea and Walter Torous, 2002. "Does the Term Structure Forecast Consumption Growth?" Working paper, UCLA.

[7] Campbell, John Y. and Robert J. Shiller, 1988. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." Review of Financial Studies, 1, 195-228.

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

[9] Chen, Nai-Fu, 1991. "Financial Investment Opportunities and the Macroeconomy." Journal of Finance, 46, 529-54.

[10] Chen, Ren-Raw, and Louis 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.

[11] Christiano, Laurence J., Martin Eichenbaum and Charles Evans, 1999. "Monetary Policy Shocks: What Have we Learned and to What End?" in Woodford, Mark and John B. Taylor, eds., Handbook of Macroeconomics, North Holland, Amsterdam.

[12] Clarida, Richard H., Jordi Gali and Mark Gertler, 2000. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." Quarterly Journal of Economics, 65, 147-80.

[13] Cochrane, John Y., 1994. "Permanent and Transitory Components of GNP and Stock Prices." Quarterly Journal of Economics, 109, 241-66.

[14] Cochrane, John Y. and Monika Piazzesi, 2002. "Bond Risk Premia." Working paper, UCLA.

[15] Dai, Qiang and Kenneth Singleton, 2000, "Specification Analysis of Affine Term Structure Models." Journal of Finance, 55, 1943-78.

[16] Dai, Qiang and Kenneth Singleton, 2002. "Expectation Puzzles, Time-varying Risk Premia, and Affine Models of the Term Structure." Journal of Financial Economics, 63, 415-41.

[17] Diebold, Francis X. and Roberto S. Mariano, 1995. "Comparing Predictive Accuracy." Journal of Business and Economic Statistics, 13, 3, 253-63.

[18] Dotsey, Michael, 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.

[19] Duffie, Darrell and Rui Kan, 1996. "A Yield-Factor Model of Interest Rates." Mathematical Finance, 6, 379-406.

[20] Estrella, Arturo, and Gikas A. Hardouvelis, 1991. "The Term Structure as Predictor of Real Economic Activity." Journal of Finance, 46, 1991, 555-76.

[21] Estrella, Arturo and Frederic 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.

[22] Estrella, Arturo and Frederic 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.

[23] Estrella, Arturo and Frederic S. Mishkin, 1998. "Predicting U.S. Recessions: Financial Variables as Leading Indicators." Review of Economics and Statistics, 1, 45-61.

[24] Evans, Charles and David 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.

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

[26] Fama, Eugene F. and Robert R. Bliss, 1987. "The Information in Long-Maturity Forward Rates." American Economic Review, 77, 680-692.

[27] Hamilton, James D. and Dong 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.

[28] Harvey, Campbell R., 1986. "Recovering Expectations of Consumption Growth from an Equilibrium Model of the Term Structure of Interest Rates." Ph.D. Dissertation, University of Chicago.

[29] Harvey, Campbell R., 1989. “Forecasts of Economic Growth from the Bond and Stock Markets.” Financial Analysts Journal, 45, 38-45.

[30] Harvey, Campbell R., 1991. “The Term Structure and World Economic Growth.” Journal of Fixed Income, 1, 7-19.

[31] Harvey, Campbell R., 1993. “The Term Structure Forecasts Economic Growth.” Financial Analyst Journal, 49, 6-8.

[32] Hodrick, Robert J., 1992. “Dividend Yields And Expected Stock Returns: Alternative Procedures For In‡uence And Measurement.” Review of Financial Studies, 5, 357-86.

[33] Jorion, Phillipe and Frederic S. Mishkin, 1991. “A Multi-Country Comparison of Term-Structure Forecasts at Long Horizons.” Journal of Financial Economics, 29, 59-80.

[34] Knez, Peter J., Robert Litterman and Jose Scheinkman, 1994. “Explorations Into Factors Explaining Money Market Returns.” Journal of Finance, 49, 5, 1861-882.

[35] Laurent, Robert, 1988. “An Interest Rate-Based Indicator of Monetary Policy.” Economic Perspectives, Federal Reserve Bank of Chicago 12, 3-14.

[36] Meese, Richard A. and Kenneth Rogo¤., 1983. “Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?” Journal of International Economics, 14, 1/2, 3-24.

[37] Mishkin, Frederic S., 1990a. “The Information in the Longer Maturity Term Structure about Future In‡ation.” Quarterly Journal of Economics, 55, 815-28.

[38] Mishkin, Frederic S., 1990b. “A Multi-country study of the Information in the Term Structure about Future In‡ation.” Journal of International Money and Finance, 19, 2-22.

[39] Newey, Whitney K. and Kenneth D. West, 1987. “A Simple, Positive Semi-De…nite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix.” Econometrica, 55, 703-08.

[40] Piazzesi, Monika, 2001. “Macroeconomic Jump E¤ects and the Yield Curve.” Working paper, UCLA.

[41] Plosser, Charles, I. and K. Geert Rouwenhorst, 1994. “International Term Structures and Real Economic Growth.” Journal of Monetary Economics, 33, 133-55.

[42] Stambaugh, Robert F., 1999. “Predictive Regressions.” Journal of Financial Economics, 54, 375-421.

[43] Stock, James H. and Mark W. Watson, 1989. “New Indexes of Coincident and Leading Indicators,” in Olivier Blanchard and Stanley Fischer, Eds., 1989 NBER Macroeconomics Annual, Cambridge MA: MIT Press.

[44] Stock, James H. and Mark W. Watson, 1999a. "Business Cycle Fluctuations in U.S. Macroeconomic Time Series." Handbook of Macroeconomics, edited by J. Taylor and M. Woodford, North Holland, Vol. 1a, 1999, pages 3-64.

[45] Stock, James H. and Mark W. Watson, 1999b. "Forecasting In‡ation." Working paper, Princeton University.

[46] Stock, James H. and Mark W. Watson, 2001. "Forecasting Output and In‡ation: The Role of Asset Prices." NBER Working Paper 8180.

[47] Stock, James H. and Mark W. Watson, 2002. "Has the Business Cycle Changed and Why?" NBER Working Paper 9127.

[48] Taylor, John B., 1993, "Discretion versus Policy Rules in Practice." Carnegie-Rochester Conference Series on Public Policy, 39, 195-214.