Is Less More ? Alternative Yield Curve Measures and their Time-varying Predictive Content for Real Activity

Abstract

This paper examines the time-varying predictive power of the yield curve for real activity. Our main focus is the role of time-varying bond risk premia and the short rate in addition to the conventionally used term spread. Another novelty is the analysis of time-varying predictability from an out-of-sample (OOS) perspective. There is strong evidence that the predictive power of the term spread has declined in the US, Germany and the UK. We also document a substantial time-variation of OOS performance. The additional role of bond risk premia and/or the short rate differs across countries, forecast horizons and the time-period.

Keywords: Term structure of interest rates, forecasting, moving block bootstrap, bond risk premia, GDP growth

JEL Classification: E43; E44; E47.

I. Introduction

The slope of the yield curve is one of the most widely followed economic variables. Amongst others, a strong reason for the alertness of economists, market watchers and central bankers is certainly the large empirical literature which has documented the term spread's usefulness for predicting future GDP growth with an inverted yield curve signalling recessions [See e.g. the seminal contributions by Harvey (1989) and Estrella and Hardouvelis (1991) and more recent work by Hamilton and Kim (2002) or Stock and Watson (2003)]. Recently, however, concerns have been raised that the predictive performance of the term spread may be time-variant [e.g. Estrella, Rodrigues, and Schich (2003) and Stock and Watson (2003)]. Other authors have pointed out that additional information contained in the yield curve such as time-varying risk premia or the level of the yield curve may provide additional information for future GDP growth [e.g. Ang, Piazzesi, and Wei (2005) or Wright (2006)]. The purpose of our paper is to study these two issues in greater detail using data from three major capital markets: Germany, United Kingdom and the United States. Thus, we complement the literature which has been primarily focused on the US.

While the in-sample predictive performance of the yield curve is well studied and established, less is known on the time-varying nature of the relationship. A major motivation of this paper is therefore to take a closer look at the time-varying forecasting performance of the yield curve for real output growth.[1] Most of the papers addressing the issue so far focus on an in-sample analysis of time-varying predictive ability, mainly using parameter stability tests [e.g. Estrella et al. (2003) and Giacomini and Rossi (2005).]. However, one may argue that out-of-sample forecast accuracy is more important for market participants from a practical perspective. Hence, our paper distinguishes itself from the remaining literature by its focus on the time-varying out-of-sample (OOS) forecasting performance.

Most papers in the previous literature have been concerned so far with the slope of the yield curve as a predictive variable for future output growth. The main economic explanation for its predictive

[1] There is a strong theoretical reason to believe that the relationship may be subject to variation over time. As noted by Estrella et al. (2003), for instance, the predictive power may depend on underlying factors like the monetary policy reaction function or the relative importance of real and nominal shocks. Both may be subject to variation over time. Thus, it makes sense to investigate the time-varying nature of the forecasting relationship in greater detail.

power is that it serves as an indicator for the effectiveness of monetary policy [See e.g. Estrella et al. (2003)]. If the central bank raises short-term interest rates and market participants expect this policy to be effective for curbing inflation in the longer run, long-term rates (averages of future expected short rates according to the expectations hypothesis) should rise in a smaller proportion. Thus, a restrictive monetary policy tends to flatten the yield curve and at the same time slows down the economy.[2]

However, empirical evidence has accumulated that the expectations hypothesis does not hold empirically for the U.S. [See e.g. Fama and Bliss (1987), Campbell and Shiller (1991) and most notably the recent work by Cochrane and Piazzesi (2005)]. In the light of this evidence, it is now commonly accepted that there are risk premia in the bond-market which evolve in a countercyclical fashion.[3] Hence, in contrast to earlier papers such as Estrella and Hardouvelis (1991) focusing on the term spread only, we assess whether there is also a role for time-varying bond risk premia for predicting future GDP growth. In a seminal contribution, Hamilton and Kim (2002) suggest that the contribution of the term spread can be decomposed into two components: one part due to the expectations hypothesis and the other due to time-varying risk premia. Their empirical approach based on instrumental variable estimation relies on future ex-post realizations of short term interest rates. However, using leads is not suitable for an assessment of out-of-sample predictive performance, which is the main concern in our context. Our paper is most closely related to Wright (2006) who has recently incorporated information about the deviation of the expectations hypothesis (using the return forecasting factor by Cochrane and Piazzesi (2005) as a proxy for time-varying risk premia) into econometric models for predicting the probability of a recession. Contrary to Wright's work, our paper is not concerned about predicting recessions in a binary probit framework, but our dependent variable is defined as real GDP growth over various forecasting horizons.

So far, the empirical evidence whether time-varying risk premia are beneficial for forecasting out-

[2] Estrella (2005) presents a theoretical model of this kind. Another formal model to explain the phenomenon is provided by Eijfinger, Schaling, and Verhagen (2002).

[3] International tests of the validity of the expectations hypothesis are provided by Jorion and Mishkin (1991), Hardouvelis (1994) and Bekaert, Wei, and Xing (2002). In a recent contribution,Tang and Xia (2005) provide a comprehensive international reassessment and find strong evidence against the expectations hypothesis in various markets including Germany, the UK and the US.

put growth remains largely ambiguous. Other important work for the US includes Favero, Kaminska, and Söderström (2005) who use a recursively estimated VAR to separate the expectations hypothesis component from the term premium component. They find strong evidence for a positive effect of risk premia for predicting future GDP growth. Ang et al. (2005) also assess the predictive performance of term premia obtained from a joint modeling approach of GDP growth, long-term rates and the short rate in a VAR framework which explicitly imposes no-arbitrage restrictions. Unlike Hamilton and Kim (2002) and Favero et al. (2005), they find no role for time-varying bond risk premia in predicting future GDP growth. Rudebusch, Sack, and Swanson (2006) study the theoretical implications of time-varying risk premia for future real activity and provide an empirical analysis using a term premium measure obtained from the affine term structure model used by Kim and Wright (2005). According to their analysis, increases in the term premium have a (marginally) significant negative predictive ability for GDP growth four quarters ahead. Contrary to these papers, the measure of time-varying risk premia used in our paper is the Cochrane-Piazzesi return forecasting factor, which is easy to compute for a relatively long sample period. Moreover, our paper addresses explicitly whether accounting for time-varying risk premia is beneficial for a market participant from an out-of-sample forecasting perspective and whether it is useful to consider them in addition to the conventionally used term spread.

Besides allowing for time-varying risk premia, we also investigate the predictive performance of the short term interest rate (level of the yield curve). Ang et al. (2005) report an important role of the short term interest rate (level of the yield curve) for predicting US real GDP growth, in particular during the 1990s. Another recent contribution by Galvao (2006) focuses on the relative importance of the level and the slope of the yield curve. We shed some further light on these issues by using our international data set.

Our methodology to study the time-variation of predictive content of the yield curve differs along several lines from the previous literature. Stock and Watson (2003), among others, report a good predictive ability of the term spread during the period 1971-1985 for the US the UK and Germany, but they find that the good performance vanishes during the period 1985-1999 for the three countries. Stock and Watson draw their conclusions from the performance of linear predictive regression models in certain sub-samples. How to split the sample into several sub-samples, however, remains more or less arbitrary. Other papers try to answer the question of time-varying predictive ability by conducting tests for structural breaks [e.g. Estrella et al. (2003) and Giacomini and Rossi (2005)]. Contrary to these papers, we use a different approach focusing on time-variation of out-of-sample performance. We first illustrate the dynamics of forecasting ability by simple diagnostic plots displaying the evolution of squared forecast errors over time compared to a naive benchmark model. This approach has recently been put forth by Goyal and Welch (2003) and Goyal and Welch (2004) in the field of stock return predictability. To evaluate the time-varying OOS forecast performance from a statistical perspective, we then apply a test procedure for equal predictive ability following the recent work by Clark and West (2006). Applying this procedure recursively allows us to illustrate the evolving path of the test of equal predictive ability (compared to a naive benchmark model) from a statistical point of view. Therefore, our approach is free from the sample-splitting problem mentioned above and can fully characterize the time-varying nature of the predictive power of the yield curve.

Our paper follows the extant prior literature using long-horizon linear regressions in order to examine the predictive power of the yield curve for output growth. Due to overlapping data, the errors of the predictive regression have a MA-structure under the null hypothesis, which must be accounted for when conducting inference. For this purpose, we apply a moving block bootstrap (MBB) methodology whose asymptotic and finite-sample properties have recently been studied by Goncalves and White (2005). It is well known that standard kernel-based estimators of the long-run variance matrix such as the Newey and West (1987) procedure may suffer from finite sample problems [See e.g. Ang and Bekaert (2005), Ang et al. (2005), Goncalves and White (2005)]. Thus, poor finite sample properties may vastly overstate the predictive content of the yield curve for the future development of real activity. Therefore, our approach – which is well suited for finite samples – is likely to provide more accurate inference than the commonly applied methods. Indeed, we find that our bootstrapped standard errors sometimes overturn the conclusions based on inappropriate standard errors.

Our main empirical results can be summarized as follows. We find that – judged from an in-sample perspective – the term spread plays a dominant role for predicting real activity compared to the other measures. The short rate only has a limited predictive content. Results for time-varying risk premia differ across countries and forecasting horizons. From an out-of-sample perspective, the slope of the yield curve again plays a central role. In all three countries, we find no strong evidence that considering the short rate beyond the term spread significantly improves the OOS forecast performance. Similarly, a dominating role of the term spread over our measure of time-varying bond risk premia is found in the case of Germany and the UK. The results for the US, however, indicate an important role of time-varying risk premia (also from an out-of-sample perspective). We document a substantial time-variation of OOS forecasting performance. Despite the generally good in-sample and OOS forecasting ability of the term spread, there is strong evidence that its predictive power has declined over the reported sample period in all three countries.

The remainder of this paper is structured as follows. Section II discusses the methodology employed in the paper with a special focus on the assessment of time-varying out-of-sample prediction. Section III provides a brief overview of our data. Section IV reports our main empirical findings from both the in-sample and out-of-sample analysis for the three capital markets under investigation. Section V concludes.

II. Methodology

In this section we briefly outline the econometric methodology used in the paper. First, we discuss the empirical framework for analyzing the in-sample predictive content of the yield curve and how we deal with several econometric pitfalls. Then we turn to the methodology for judging the time-variation of out-of-sample (OOS) performance.

Following the extant literature [e.g. Estrella and Hardouvelis (1991) or Stock and Watson (2003)], we use a regression-based test to investigate which information contained in the yield curve helps to predict future GDP growth.[4] The predictive regression takes the following form:

[4] Unfortunately, our approach has only limited capacity for delivering structural interpretations as to why the term spread or the bond risk premia may predict future GDP growth. Nevertheless, given the difficulties of deriving testable implications from a more structural model, as noted by Rudebusch et al. (2006), we adhere to the majority of the empirical literature and use a reduced-form approach.

$$y^{(k)}_{t+k} = \beta^{(k)}_0 + \beta^{(k)}_1 X_t + \beta^{(k)}_2 Z_t + \epsilon^{(k)}_{t+k}. \tag{1}$$

$y^{(k)}_{t+k}$ denotes the (log) growth rate of real GDP from t to $t+k$ [(annualized) cumulative real GDP growth] and is defined as $y^{(k)}_{t+k} = (400/k)\ln(Y_{t+k}/Y_t)$, where Y_t is the level of real GDP as of period t. X_t contains the specific information measures of the yield curve we are interested in [i.e. a term spread, the short rate or a measure of time-varying bond risk premia]. Z_t consists of control variables (other potential predictors of GDP growth) considered in some specifications.[5] The examination of the predictive power is based on the OLS regression in Equation (1), testing the statistical significance of the parameter $\hat{\beta}^{(k)}_1$. The economic significance of the predictive power can further be assessed by the adjusted R^2 in addition to the significance of $\hat{\beta}^{(k)}_1$.

Although cumulative GDP growth is commonly used to assess the predictive capacity of the yield spread, some authors also use marginal GDP growth [e.g. Estrella and Hardouvelis (1991), Dotsey (1998), Hamilton and Kim (2002)]. Marginal GDP growth is defined as $y^{(4)}_{t+k} = \frac{1}{4}(y^{(1)}_{t+k} + y^{(1)}_{t+k-1} + y^{(1)}_{t+k-2} + y^{(1)}_{t+k-3})$. Since marginal GDP growth is useful for examining how far into the future the predictive power of the yield curve can reach, we also report estimation results with marginal GDP growth as dependent variable.

A. Overlapping Observations

Despite the apparent simplicity of the predictive linear regression in (1), the approach is plagued by econometric problems due to overlapping observations of the dependent variable. As is well known, the overlap induces serial correlation of the errors of order $(k-1)$ in Equation (1), which must be accounted for when conducting inference.

A remedy for this problem often used in the literature is to use kernel-based HAC standard errors, e.g. according to Hansen and Hodrick (1980) or Newey and West (1987), which are robust against heteroskedasticity and serial correlation. Although these commonly applied standard kernel-based estimators of the long-run covariance matrix deliver consistent estimates, they are known to suffer

[5] For details on the construction of control variables, please refer to the data appendix.

from finite-sample problems [See e.g. Ang and Bekaert (2005), Ang et al. (2005)]. This may vastly overstate the predictability of yield curve measures for future real activity.[6]

In a framework like ours, the usual sample size casts doubt on the validity of the asymptotic argument. We therefore use a moving block bootstrap (MBB) methodology which is particularly suitable in a finite-sample setting with dependent data. The consistency of the MBB standard error estimator has recently been proved by Goncalves and White (2005). Contrary to the (parametric) bootstrap approach put forth by Kilian (1999), the MBB is a (non-parametric) bootstrap which draws blocks of re-sampled observations randomly with replacement from the time series of original observations, where the block length can be fixed or data-driven.[7] In a simulation study, Goncalves and White (2005) show that inference based on MBB standard errors may be considerably more accurate in small samples than inference based on closed-form asymptotic estimates such as Newey-West. As we will discuss in section III, we indeed find that MBB standard errors sometimes overturn the conclusions based on inappropriate standard errors.

B. Out-of-sample Statistics

A main focus of this paper is the analysis of the OOS forecast performance of the yield curve with particular emphasis on its time-varying nature. To generate a series of (pseudo-) OOS forecasts, we estimate the particular model using a recursive scheme. For the first forecast, R in-sample observations are used. Based on the parameters obtained from fitting the model in-sample, a forecast is generated as $\hat{y}^{(k)}_{R+k} = \hat{\beta}^{(k,R)}_0 + \hat{\beta}^{(k,R)}_1 X_R + \hat{\beta}^{(k,R)}_2 Z_R$. The performance of the forecasts from conditional models based on the yield curve are compared to those of a naive (unconditional) benchmark model, which takes the mean of the dependent variable known in R as forecast for the value of real GDP growth from R to $R + k$. The next forecast is then based on the estimation of the model using $R + 1$ observations, and so forth. This procedure provides us with $T - R - k - 1$

[6]Most of the extant literature on the predictive power of the yield curve uses Newey-West standard errors. Recognizing the finite sample limitations of the Newey-West estimator reported by Ang and Bekaert (2005), Ang et al. (2005) use Hodrick (1992) standard errors. To our knowledge, no other paper so far has used the MBB approach to investigate the predictive power of the yield curve for real activity.

[7]As recommended by Goncalves and White (2005), we use a data-driven block length, following the procedure by Andrews (1991).

model-based forecasts which we can compare to the corresponding realizations, where T denotes the overall sample size.

One of the measures of forecast evaluation we report is the mean forecast error (ME). A mean forecast error which is significantly different from zero can be interpreted as evidence against the hypothesis of forecast unbiasedness. Another simple descriptive measure of forecast evaluation is Theil's U, which is the ratio of the RMSE of the conditional model to the RMSE of the particular benchmark model. If the conditional forecast is superior to the benchmark (given a quadratic loss), Theil's U should be less than one.

While Theil's U is widely used in the forecasting literature, we draw our main conclusions according to a test recently proposed by Clark and West (2006). This test is designed for comparing a parsimonious null model to a larger model which nests the null model, as is the case in our context. The central idea of Clark and West test is to adjust the mean squared forecast error of the larger unrestricted model for upward bias. The reason is that – under the null hypothesis (additional regressors in the larger model are not necessary for forecasting) – the unrestricted model attempts to estimate parameters that are zero in population, which introduces noise in the forecast. MSFE_{adj} takes these considerations into account

$$\text{MSFE}_{adj} = P^{-1}\sum_{t=R}^{T-k}(y_{t+k}^{(k)} - \hat{y}_{2t,t+k}^{(k)})^2 - P^{-1}\sum_{t=R}^{T-k}(\hat{y}_{1t,t+k}^{(k)} - \hat{y}_{2t,t+k}^{(k)})^2, \tag{2}$$

where the GDP growth forecast (k-quarter ahead) based on the information set at time t is denoted as $\hat{y}_{2t,t+k}^{(k)}$ for the case of the (unrestricted) model of interest and $\hat{y}_{1t,t+k}^{(k)}$ for the case of the benchmark model. P is the number of OOS predictions: $P = T - R - k - 1$. Note that the first term in Equation (2) corresponds to the usual mean squared forecast error of the (unrestricted) model of interest, and the second term is the adjustment term discussed above. A computational convenient way to test equal predictive performance of the two nested models is to regress

$$\hat{f}_{t+k}^{(k)} = (y_{t+k}^{(k)} - \hat{y}_{1t,t+k}^{(k)})^2 - [(y_{t+k}^{(k)} - \hat{y}_{2t,t+k}^{(k)})^2 - (\hat{y}_{1t,t+k}^{(k)} - \hat{y}_{2t,t+k}^{(k)})^2] \tag{3}$$

onto a constant. The t-statistic can be used to assess whether the difference of the MSFE of the restricted benchmark model and the adjusted MSFE of the larger model is statistically different from zero (one-sided test).[8] Again, we use the MBB to obtain the autocorrelation consistent standard error (for $k > 1$).

C. Time-variation of OOS Performance

We investigate the time-variation of OOS performance using diagnostic plots, which are motivated by the recent work of Goyal and Welch (2003) and Goyal and Welch (2004) in the context of stock return predictability.[9] Goyal and Welch suggest to plot the cumulative sum of squared forecast errors from a benchmark model minus the squared errors from the conditional model, that is: Net-SSE(T_1)=$\sum_{t=T_0}^{T_1}[(y_{t+k}^{(k)} - \hat{y}_{1t,t+k}^{(k)})^2 - (y_{t+k}^{(k)} - \hat{y}_{2t,t+k}^{(k)})^2]$, where T_0 is the starting date and T_1 is the end date. When the graph is above the zero horizontal line, it indicates that the model of interest outperforms the benchmark model in terms of squared forecast errors up to period T_1. This graph is a simple but rather informative diagnostic for comparing the relative performance of the competing models.

However, the Net-SSE measure fails to take the upward bias in forecast errors into account (induced by the need to estimate additional parameters in the more general model). Therefore, following the logic by Clark and West (2006), we propose an adjusted Net-SSE to address this concern. Instead of using the cumulative unadjusted squared error of the unrestricted model, we use the cumulative adjusted squared error, $\sum_{t=T_0}^{T_1}(y_{t+k}^k - \hat{y}_{2t,t+k}^k)^2 - \sum_{t=T_0}^{T_1}(\hat{y}_{1t,t+k}^k - \hat{y}_{2t,t+k}^k)^2$, as in Clark and West (2006). The convenient way of plotting the adjusted Net-SSE is to graph the cumulative sum of $\hat{f}_{t+k}^{(k)}$ from Equation (3) over time, which is equivalent to the cumulative sum of the squared errors from the benchmark model minus the squared adjusted forecast errors from the model of interest.

Note that the (adjusted) Net-SSE provides only a descriptive statistic of time-varying OOS performance. Therefore, in order to judge time-varying OOS performance from a statistical perspective we use a recursive application of the Clark-West test. Based on the first 40 forecasts we calcu-

[8]Their simulation shows that the use of MSFE$_{adj}$ with standard normal critical values is as accurate as other competing tests, while the power is as good or better.

[9]In an extensive analysis for the US stock market Goyal and Welch (2004) question the existence of stock return predictability based on their finding of poor OOS performance relative to a naive benchmark.

late the first Clark-West statistic and its standard error based on the MBB.[10] We then update the information set each quarter and recalculate the test statistic each time. The obtained series of the Clark-West test statistic is then plotted together with the corresponding critical value based on the standard normal distribution. This plot illustrates the evolving path of the test of equal forecast performance. Our approach is free from sample splitting problems and can fully characterize the time-varying nature of predictability of the yield curve. In this sense our approach can reveal some important and interesting results that might be neglected if one applies in-sample stability tests only.

III. Empirical Results

A. Data Overview

Our dataset consists of time series of real GDP, yields of zero coupon bonds and three-month interest rates for Germany, the USA and the UK. The sample period ranges from 1972:Q4 to 2006:Q1 for two countries: Germany and the US. Unfortunately, comprehensive yield curve data for the UK are only available for a shorter period from 1980:Q1 to 2006:Q1. The maturities of zero bond yields cover $n = 1, \ldots, 10$ years for all countries. Additional control variables used for multivariate analysis in the case of Germany are taken from the Bundesbank time-series database, Reuters-Ecowin and provided by the ifo institute. Further detailed information on the data, their sources and data transformation is provided in the appendix A.

B. The In-sample Predictive Performance of the Yield Curve

Table I reports the estimation results for the predictive power of the term spread for German real GDP growth at different forecasting horizons ($k = 1, 4, 8, 12$ quarters) and for different bond ma-

[10]Since asymptotic results in Clark and West (2006) depend on the ratio of in-sample observations (R) to OOS forecast observations (P), we use the first 40 forecast observations for the first calculation of the recursive Clark-West statistic. As argued by Clark and West (2006), the behavior of MSFE_{adj} is rationalized, when the ratio P/R is 2 or above. Since in our case R is set to 20 quarters, we use 40 observations for the first calculation of the Clark-West statistic.

turities for the longest yield ($\tau = 1, 3, 5, 7, 10$ years). The shortest yield is always the three-month interest rate. [11]

[Insert Table I About Here]

Overall, we obtain the well known picture from studies with US data. The term spread has a significant (in-sample) predictive power for real activity. The best predictive performance in Germany is obtained for 4 to 8 quarter ahead. It is worth noting that the term spread with the longest time to maturity (in our case 10 years) relative to the three-month interest rate does not necessarily have the best predictive power, lending no support for the common recommendation in the extant literature [e.g. Estrella and Trubin (2006)].

Panel B of Table I reports estimation results with marginal real GDP growth as the dependent variable. This allows us to assess how far into the future the predictive performance of the yield curve reaches. In fact, the marginal predictive power of the term spread vanishes substantially after a predictive horizon of 8 quarters, which is somewhat longer than in the US and the UK (See Tables V and VII). Also note that at a horizon of 8 quarters ahead, the tests based on Newey-West and the MBB deliver different conclusions, indicating that Newey-West standard errors lead to an overstatement of the predictive power of the term spread.[12]

The Role of the Short Rate and Time-Varying Risk Premia

In the following we investigate whether other measures from the yield curve have predictive content for future real activity (taken individually or in combination with each other). Estimation

[11] We report estimates for Germany only since the US relationship is well-studied already. In this table and the remainder of the paper, for each horizon, two types of t-statistics are reported in parentheses below the parameters: t-statistics in $(\cdot)$ are based on Newey-West standard errors whereas those in $\langle\cdot\rangle$ are based on MBB standard errors (100,000 replications). For one-step ahead forecasts ($k = 1$), no HAC standard errors are computed since there is no overlap.

[12] Table A.1 in the appendix indicates that even when controlling for other predictive variables, the term spread still retains much of its predictive content for future output growth, in particular beyond a horizon of 4 quarters. The role of other macro variables is not our major concern in this paper. Therefore, we only report the results of the German case for interested readers.

results for the predictive content of the short rate (Model 2) for cumulative real output growth are reported in Tables II, IV and VI. The significant and negative coefficient is in line with the conventional wisdom that higher interest rates tend to slow down the economy by adding higher costs to investment. As shown by the tables, the predictive content of the short rate is mainly concentrated at shorter horizons. The predictive performance clearly decreases with increasing forecast horizons. It is worth noting, that the predictive power reaches least far into the future in the case of the UK. For marginal GDP growth, the results for Model 2 in Tables III, V and VII show that the predictive power of the short rate is in most cases insignificant beyond a four quarter horizon. Based on this result, we conclude that the short rate's in-sample predictive power is quite limited.

An issue of major interest in this paper is the role of time-varying risk premia for predicting future output growth. As noted above, we use the Cochrane-Piazzesi factor (CP-factor) as our proxy for time-varying bond risk premia, following Wright (2006). A major advantage compared to other measures is that it can be computed back easily for a relatively long time period.

Under the expectations hypothesis, expected bond excess returns should be constant over time. Evidence for predictability of bond excess returns by forward rates suggests, however, that there are time-varying risk premia for holding longer period bonds, which implies that the expectations hypothesis is violated. To see whether it makes sense to investigate the impact of time-varying risk premia, it is useful to study first whether the expectations hypothesis holds in our dataset. We therefore run Fama and Bliss (1987) and Cochrane and Piazzesi (2005) regressions, and the results are summarized in the appendix. As shown in Table A.2 in the appendix, the Cochrane-Piazzesi regressions provide strong evidence against the expectations hypothesis. Note also that the evidence against the expectations hypothesis is less strong in the case of the Fama-Bliss regressions.[13] As shown in Figure A.1, we observe the typical tent-shape pattern in all three countries, which corroborates the recent findings by Cochrane and Piazzesi (2005) using a different dataset.[14]

[Insert Tables II, IV and VI About Here]

[13] Due to limited space only results for the German bond market are reported. Results for the UK and the US are similar and can be delivered by the authors upon request.

[14] Cochrane and Piazzesi (2005) use the Fama-Bliss data set with bond maturities from 1 to 5 years, whereas our dataset covers maturities from 1 to 10 years. In order to avoid problems from multicollinearity we do not use all forwards on the right-hand side.

We now examine the predictive ability of time-varying risk premia for the subsequent real development using the CP-factor as a proxy. As shown in Tables II, IV and VI, we document a positive and significant predictive power for cumulative real GDP growth especially in the case of Germany and the US. In the case of Germany, the CP-factor predicts cumulative real GDP growth up to a horizon of 12 quarters ahead. For the US, the predictive power for real output growth is even stronger with significant coefficients at all reported horizons and adjusted R^2 of about 27% for horizons of 4 and 8 quarter ahead. Not surprisingly, the marginal predictive power decays after a forecast horizon of 6 quarters for Germany and USA. In the case of the UK, the results with marginal GDP growth as dependent variable show that the predictive power is mainly concentrated at a longer horizon (8 and 12 quarters). Putting all results together, we conclude that time-varying risk premia have a certain in-sample predictive ability for real GDP growth, though the predictive power differs across countries in terms of both significance and forecast horizons.

The (significantly) positive coefficient we document is in line with recent findings by Favero et al. (2005) and Wright (2006). Nevertheless, this result contradicts the conventional reasoning that low time-varying risk premia may act as a stimulus to the economy [expressed for instance by chairman Ben Bernanke in his recent speech before the economic club of New York [Bernanke (2006)]. For ourselves, we are somewhat reluctant to draw strong structural conclusions from this result. One way of rationalizing the result may be the following. Bond risk premia (as well as risk premia in the stock market) tend to be higher during a recession when expected business conditions are poor. Eventually, as the economy moves out of the recession and economic conditions improve. Hence, empirically we observe a positive relationship over horizons beyond 4 quarters.

A main question of the paper is whether it is useful to consider several measures of yield curve or whether using less information may actually prove more beneficial from a forecasting perspective. Therefore, we now look at combinations of different yield curve measures, where we assess the predictive power of a particular yield curve measure when controlling for the other ones. Again, the results for cumulative real GDP growth are provided in Tables II, IV and VI, and those for marginal real GDP growth are provided in Tables III, V and VII.

Model specification 4 includes the term spread and the short rate jointly. Including the short rate is motivated by the recent work by Ang et al. (2005) and Galvao (2006) who have found an important role of the short rate for predicting real activity, in particular for recent sample periods. As shown by the tables, the term spread largely maintains its predictive power in most cases, whereas the short rate tends to lose its limited forecasting ability in all three countries. We also examine the joint role of the term spread and time-varying bond risk premia (model 5). Again, for all three countries the term spread mostly maintains its informative role for subsequent real GDP growth. The CP-factor continues to play a role beyond the spread only in the case of the US. Model 6 examines the joint predictive power of the short rate and the CP-factor. As becomes clear from the tables, the CP-factor usually predicts better than the short rate, although the t-statistics decline relative to the case where it is considered solely. From our investigation of the joint role of various yield curve measures, we conclude that the term spread maintains a dominant predictive power relative to the other two variables. A notable exception is the role of the CP-factor in the US. Although the CP-factor and the short rate have predictive content of their own, the predictive power is limited when controlling for the term spread.

[Insert Tables III, V and VII About Here]

C. Evaluation of OOS Forecast Performance

In the following, we discuss the results of OOS forecast evaluation across different model specifications. The (pseudo-) OOS forecasts are generated according to a recursive scheme as outlined in section II. Tables VIII, IX and X summarize the results of forecast evaluation for Germany, USA and UK respectively. The test framework proposed by Clark and West (2006) tests the null hypothesis of equal predictive performance of a larger conditional model and a restricted benchmark model against the alternative of superior forecast performance of the unrestricted model. We therefore consider a one-sided testing framework with critical values +1.282 (for 10% level) or +1.645 (for 5% level). Models which include a single measure of the yield curve (M1, M2, M3) are always compared to a naive model which takes the prevailing mean of the dependent variable as forecast for the next period (denoted as n.M. in the table), whereas those models which jointly consider different measures of the yield curve (M4, M5, M6) are also compared to the two nested models which only include a corresponding subset of the variables.[15]

M1 refers to the model which includes the slope of the yield curve as predictive variable. The model tends to outperform the naive model for several – though not for all – forecasting horizons. The Clark-West statistics indicate a significantly better OOS performance (at the 10% level) of the conditional model based on the spread for very short (1 quarter) and very long (12 quarters) horizons in Germany. The evidence for superior OOS performance is found to be strongest in the US (4, 8, 12 quarters ahead) and weakest for the UK (1 quarter). In some cases, Theil's U is slightly above 1 while the Clark-West statistics still reject the null hypothesis of equal OOS performance. Since Theil's U is only a descriptive statistic, we rely our inference upon the Clark-West statistics. It appears that the term spread has a better OOS forecast performance in the US compared to the other two countries. In the case of Germany (which covers the same sample period as the US), this may well be due to the fact that several structural events have affected the German economy but not the US economy during the sample period (German reunification, introduction of the Euro). The rather weak results for the UK could be ascribed to the relatively short time period for evaluating OOS performance in the UK. The OOS evaluation in the UK is limited to a period, where the predictive performance of the slope of the yield curve appears to have declined at an international level as we will discuss in greater detail below.

[Insert Tables VIII, IX and X About Here]

We now turn to the OOS performance of the model based on the short rate (M2). Our evidence for Germany suggests that the short rate fails to improve upon the random walk model for all reported horizons. These results differ markedly from those for the other two countries. For the US the OOS performance is superior to the one of the benchmark at a 4 quarter horizon (Clark-West statistic is marginally significant at 10% level). In the case of the UK, we find a rather good OOS performance of the short rate, concentrated especially at shorter horizons (1 and 4 quarter ahead).

[15] In order to mimic the real-time experience of a typical investor, only information truly available as of period *t* is used for the recursive forecast scheme. Therefore, we use a recursive estimate of time-varying risk premia in the OOS forecast experiment.

The OOS forecast performance of our measure of time-varying risk premia (M3) also differs substantially across the three countries. In Germany, the performance relative to the naive model remains rather poor. By contrast, we find a fairly good OOS performance for the US market. The results for the UK are somewhat mixed. Considering the CP-factor as a predictive variable improves relative to the benchmark model only at a horizon of 12 quarters, but the Clark-West statistic is highly significant.

In the light of these results it seems natural to ask whether it is necessary at all to account for yield curve measures other than the term spread from an OOS perspective: Is less information actually more when it comes to OOS forecast performance ? Hence, we compare the OOS predictive ability of larger models combining different yield curve measures with those of smaller nested models which are either the naive model or a single yield curve measure only. The superscript at the corresponding Clark-West statistic denotes which benchmark is used for the comparison.

First, we discuss the results using the unconditional naive model as the benchmark. For all three countries, the bivariate models (M4 and M5) nesting the term spread model outperform the naive model only when M1 (the model includes only the term spread) has superior predictive ability relative to the benchmark. This observation clearly indicates the central role of term spread for driving most of the predictive power of the yield curve for real GDP growth. Moreover, it is worth noting that while the model which includes the term spread only outperforms the benchmark in terms of OOS accuracy, the model which combines term spread and short rate generally does not.

Inspection of Tables VIII, IX and X reveals that combining different information measures may often introduce more noise into the forecast such that the OSS forecasting ability declines. Thus, to formally assess whether bivariate models have a superior performance than models including a single predictor, we rely on the Clark-West test using the nested univariate model as the benchmark.

The result from these tests indicate that once the slope of the yield curve is combined with the level (model M4), the model performs never better, but often worse than the model which includes the slope only (M1). When the model with the short rate (M2) is considered as the benchmark, however, M4 performs never worse but sometimes better. This finding holds true for all three countries. We thus conclude that – from an OOS perspective – the dominating information for future real activity appears to be in the slope of the yield curve and not in the level. In this regard our conclusion differs markedly from Ang et al. (2005) and Galvao (2006).

Similarly, we document a dominating role of the term spread over the CP-factor in the case of Germany and the UK. This statement does not hold for the US, however. According to our results for the US, time-varying risk premia play a distinctive role: The model which considers the spread along with the CP-factor (M5) usually performs no worse but often better than the model with the spread only (M1), while it performs no better but sometimes worse than the CP-factor alone (M5).

D. Time-Variation of OOS Forecast Performance

In the following, we discuss our results on the time-variation of OOS forecast performance. Our aim in this context is to provide a complete picture of how the predictive relation evolves over time in Germany, USA and UK. As described in section II, we use a simple but informative graphical approach for illustration (adjusted Net-SSE plots and recursively calculated Clark-West statistics).

First, we discuss briefly how the OLS coefficients estimated using the recursive OOS scheme evolve over time. Figure 1 plots the coefficient on the term spread obtained by estimating M1 in a recursive fashion.[16] When considering a forecast horizon of 1 quarter for Germany, the relationship appears to be rather stable. For a horizon of 4 quarters, there seems to be a downward shift in the estimated parameter in the vicinity of 1991 which implies a drop of the predictive power of the term spread, which may be ascribed to the German reunification. In the US (k=4) a clear decline of the coefficient can be observed over the whole sample period. Nevertheless, the coefficient still remains highly significant at the end of the sample period. The picture for the UK is quite different in that we observe a sudden increase of the recursively estimated coefficient in the early 1990s from around zero to a significantly positive coefficient.

We draw our main conclusions on the time-varying OOS performance of the different models based on the adjusted Net-SSE plots provided in Figure 3. For the ease of comparison, we also show Goyal-Welch type Net-SSE plots (Figure 2), where no adjustment is made to the forecast errors of the larger model (See the discussion in section II). Both plots show the OOS performance

[16] The 95% confidence bands are estimated using MBB standard errors with 10,000 replications.

of the term spread model over time relative to the naive benchmark, with forecasting horizons k=1, 4 for all three countries. The statistical significance of variation in the OOS performance is assessed by recursively estimated Clark-West statistics shown in Figure 4.

The adjusted Net-SSE plots for Germany indicate a superior OOS performance of the model with the term spread relative to the naive benchmark. However, the relative OOS performance clearly varies over time as illustrated by Figure 2. Between 1990 and 1994, one can observe a “U" shape in both sub-figure (a) and (b), which shows that the OOS forecast performance initially fell and then bounced back. We attribute this finding to the effects of German reunification which is a typical example of a real shock to the economy. The unadjusted Net-SSE plots show a more volatile OOS accuracy over time, which is due to the noise induced by the need for estimating additional parameters.

From a statistical perspective, the plot of the recursive Clark-West statistic shows that for a forecast horizon of 1 quarter the OOS performance indeed has undergone changes over time.[17] In the period around 1993-1996, we find evidence for superiority of OOS performance relative to the naive benchmark while afterwards equal OOS performance cannot be rejected. For a 4 quarter horizon – though there is some time-variation of OOS performance – it is less pronounced and not significant.

The adjusted Net-SSE plots for the US show a superior OOS performance of M1 over almost the whole period under scrutiny and for both forecast horizons. The unadjusted Net-SSE, however indicates a very poor relative performance for the one-quarter forecast horizon. For a forecast horizon of 4 quarters, a substantial decline in the relative OOS performance can be observed: from about 1993 onwards the conditional model based on the spread performed almost uniformly worse than the benchmark model in terms of squared forecast errors. The recursive Clark-West plot confirms these results for a forecast horizon of 4 quarters. The relative OOS superiority declines steadily, but from 2000 onwards it has declined less rapidly for the rest of the sample. In the case of one-step ahead forecasts, though there is also decline, equal OOS forecast performance cannot be rejected for the whole period.

[17] Note that the rejection region plotted along with the Clark-West test statistic is based on the 5% significance level.

In the UK there is also some evidence of time-varying OOS performance. Unfortunately, the sample period of the UK yield curve data is relatively short, so the time-variation might not be captured adequately by the figures. The adjusted Net-SSE shows a relatively poor OOS performance at the beginning of the period (especially for k=4) which has somewhat recovered afterwards. According to the unadjusted Net-SSE plot, the OOS performance is rather good at the one-step horizon but worse than the benchmark for 4 quarters ahead. Indeed, the plots of recursive Clark-West statistics indicate a superior OOS performance compared to the benchmark at a one-step horizon, while one cannot reject equal OOS forecast performance for a horizon of 4 quarters.

IV. Conclusion

This paper addresses the question whether different measures from the yield curve (such as the term spread, the short rate and time-varying bond risk premia) have forecasting power for subsequent real GDP growth (separately and beyond each other) and whether the forecasting relationships vary over time. Our findings provide new evidence for Germany, the US and the UK, especially with regard to the time-variation of out-of-sample forecast performance.

From our in-sample analysis we find a dominant role of the slope of the yield curve compared to the short rate and our measure of time-varying bond risk premia. Time-varying risk premia and the short rate have a predictive content of their own, but only to a limited extent beyond the term spread. The notable exception is the case of the US where there is evidence that time-varying risk premia also play an important role. Moreover, our findings show a dominant out-of-sample forecast performance of the term spread relative to the short rate for all three countries. There is some evidence in the case of the US that accounting for time-varying risk premia is beneficial from an out-of-sample forecast perspective. We further document substantial time-variation of out-of-sample forecast performance with strong evidence that the predictive power of the yield curve has declined over the reported sample period.

REFERENCES

Andrews, Donald W K, 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* 59, 817–58.

Ang, Andrew, and Geert Bekaert, 2005, Stock return predictability: Is it there?, forthcoming Review of Financial Studies.

Ang, Andrew, Monika Piazzesi, and Min Wei, 2005, What does the yield curve tell us about GDP growth, *Journal of Econometrics* .

Bekaert, Geert, Min Wei, and Yuhang Xing, 2002, Uncovered interest rate parity and the term structure, Working paper, Columbia University.

Bernanke, Ben S., 2006, Reflections on the yield curve and monetary policy, Remarks by Chairman Ben S. Bernanke, The Federal Reserve Board.

Campbell, John Y., and Robert J. Shiller, 1991, Yield spreads and interest rates: A bird's eye view, *Review of Economic Studies* 58, 495–514.

Clark, Todd E., and Kenneth D. West, 2006, Approximately normal tests for equal predictive accuracy in nested models, *Journal of Econometrics* .

Cochrane, John H., and Monika Piazzesi, 2005, Bond risk premia, *American Economic Review* 95, 138–160.

Dotsey, Michael, 1998, The predictive content of the interest rate term spread for future economic growth, *Economic Quarterly* 84, 31–51.

Eijfinger, S., E. Schaling, and W. Verhagen, 2002, The term structure of interest rates and inflation forecast targeting, CEPR discussion paper No. 2375.

Estrella, Arturo, 2005, Why does the yield curve predict output and inflation, *The Economic Journal* 115, 722–744.

Estrella, Arturo, and Gikas Hardouvelis, 1991, The term spread as a predictor of real economic activity, *Journal of Finance* 46, 555–576.

Estrella, Arturo, Anthony P. Rodrigues, and Sebastian Schich, 2003, How stable is the predictive power of the yield curve? evidence from Germany and the United States, *Review of Economics and Statistics* 85, 629–644.

Estrella, Arturo, and Mary R. Trubin, 2006, The yield curve as a lending indicator: Some practical issues, *Current Issues in Economics and Finance* 12, 1–7.

Fama, Eugene F., and Robert R. Bliss, 1987, The information in long-maturity forward rates, *The American Economic Review* 77, 680–692.

Favero, Carlo A., Iryna Kaminska, and Ulf Söderström, 2005, The predictive power of the yield spread: Further evidence and a structural interpretation, Working paper, Bocconi University.

Galvao, Ana Beatriz, 2006, The changing effect of the yield curve on GDP growth, Bank of Portugal, Working paper.

Giacomini, Raffaela, and Barbara Rossi, 2005, How stable is the forecasting performance yield curve for output growth?, forthcoming Oxford Bulletin of Economics and Statistics.

Goncalves, S., and Halbert White, 2005, Bootstrap standard error estimates for linear regressions, *Journal of the American Statistical Association* 100, 970–979.

Goyal, Amit, and Ivo Welch, 2003, Predicting the equity premium with dividend ratios, *Management Science* 49, 639–654.

Goyal, Amit, and Ivo Welch, 2004, A comprehensive look at the empirical performance of equity premium prediction, forthcoming Review of Financial Studies.

Gürkayanak, Refet, Brian Sack, and Jonathan Wright, 2006, The U.S. treasury yield curve: 1961 to the present, Federal Reserve Board, Fianance and Econonomics Discussion Series.

Hamilton, James D., and Dong Heon Kim, 2002, A reexamination of the predictability of economic activity using the yield spread, *Journal of Money, Credit and Banking* 34, 340–360.

Hansen, Lars Peter, and Robert J. Hodrick, 1980, Forward exchange rates as optimal predictors of future spot rates: An econometric analysis, *The Journal of Political Economy* 88, 829–853.

Hardouvelis, Gikas, 1994, The term structure spread and future changes in long and short rates in the G7 countries: Is there a puzzle?, *Journal of Monetary Economics* 33, 255–284.

Harvey, Campbell R., 1989, Forecasts of economic growth from the bond and stock markets, *Financial Analysts Journal* 49, 38–45.

Hodrick, Robert J., 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357–386.

Jorion, Philippe, and Frederic Mishkin, 1991, A multicountry comparison of term-structure forecasts at long horizons, *Journal of Financial Economics* 29, 59–80.

Kilian, Lutz, 1999, Exchange rate fundamentals: what do we learn from long-horizon regressions, *Journal of Applied Econometrics* 14, 491–510.

Kim, Don H., and Jonathan Wright, 2005, An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates, Discussion paper, Federal Reserve Board.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica: Journal of the Econometric Society* 55, 703–708.

Rudebusch, Glenn D., Brian P. Sack, and Eric T. Swanson, 2006, Macroeconomic implications of changes in the term premium, Working paper, Federal Reserve Bank of San Francisco.

Stock, James H., and Mark W. Watson, 2003, Forecasting output and inflation: The role of asset prices, *Journal of Economic Literature* 151, 788–829.

Tang, Huarong, and Yihong Xia, 2005, An international examination of affine term structure models and the expectations hypothesis, Working paper, Wharton School.

Wright, Jonathan H., 2006, The yield curve and predicting recessions, Working paper, Federal Reserve Board.

Appendix

A. Data Appendix

Panel A: Germany		
Variable	**Data Source**	**Details on Data Construction**
Zero-Bond Yields	Bundesbank	Zero-Bond yields are taken from the Bundesbank time-series database. The yield curve construction by the Bundesbank follows the Svenson Method. Calculations are based on market prices of Federal Bonds with different times to maturity. Monthly data are transformed into quarterly data. Yields are annualized and expressed in continuous compounding.
Real GDP Growth	Reuters-Ecowin	A seasonally adjusted time series of real GDP is used. The outlier in the growth rate of real GDP due to the German reunification (1991:Q1) is adjusted by interpolation as in Stock and Watson (2003): the corresponding observation was replaced by the median of the three previous and the three following observations.
Three-Month Interest Rate	Bundesbank	Frankfurt Interbank Money Market Rates for ninety days' deposits. Monthly data are transformed into quarterly data. We use annualized yields expressed in continuous compounding.
Money-Market Rate	Bundesbank	Frankfurt Money Market Rates for daily deposits. Monthly data are transformed into quarterly data. Yields are annualized and expressed in continuous compounding.
Stock Market Portfolio	Reuters-Ecowin	We use the MSCI Germany Gross Total Return Index as a representative stock index. Monthly returns based on the MSCI Index are transformed into quarterly data by cumulating monthly log returns.
Business Climate Indicator	Ifo-Institute	Until 1991:Q2 the business climate indicator applies to Western Germany, whereas thereafter data apply to the reunited Federal Republic of Germany. We use the growth rate of the index based on quarterly data.
Inflation Rate	Reuters-Ecowin	The inflation rate is calculated based on the Consumer Price Index (CPI).
Oil Price	Reuters-Ecowin	Oil price (Brent), Changes of the oil price are used.
Panel B: USA		
Zero-Bond Yields	Refet Gürkaynak's Homepage	Zero-Bond yields calculated by Gürkayanak, Sack, and Wright (2006). Daily data are quarterlized. We use annualized yields expressed in continuous compounding.
Real GDP Growth	Reuters-Ecowin	Seasonally adjusted time series of real GDP growth.
Three-Month Interest Rate	Econstats	Three-month Treasury Bills in the Secondary Market. Monthly data are transformed into quarterly data. We use annualized yields expressed in continuous compounding.
Panel C: UK		
Zero-Bond Yields	Econstats	Zero-Bond Yields calculated by the Bank of England. Monthly data are transformed into quarterly data. Yields are annualized and expressed in continuous compounding.
Real GDP Growth	Reuters-Ecowin	Seasonally adjusted time series of real GDP growth.
Three Month Interest Rate	Econstats	Three-month Treasury Bills in the Secondary Market. Monthly data are transformed into quarterly data. We use annualized yields expressed in continuous compounding.

B. Additional Results

[Insert Table A.1 About Here]

[Insert Table A.2 About Here]

[Insert Figure A.1 About Here]

C. Tables and Figures

Table I
Predictive Regressions for Real GDP Growth, Spreads of Different Maturities, Germany

Estimation is based on the following regression specification

$$y_{t+k}^{(k)} = \alpha_n^{(k)} + \beta_n^{(k)} X_t + \epsilon_{t+k;n}^{(k)},$$

where X_t contains the term spread $(i_t^{(n)} - i_t^{(3m)})$ as explanatory variable. Different maturity combinations for the longest maturity n are used. The dependent variable is defined as cumulative real GDP growth (annualized) in Panel A, whereas in Panel B it is defined as marginal real GDP growth. The forecasting horizon is denoted by k. The sample period is 1972:Q4-2006:Q1. $\bar{R}^2$ denotes the adjusted R^2. We report t-statistics based on two methods for calculating standard errors: $(\cdot)$ contains t-statistics based on Newey-West standard errors with $k-1$ lags, $\langle\cdot\rangle$ is based on the bootstrapped standard errors (100,000 replications).

Horizon	Term-Spread (maturity n)									
	n=1		n=3		n=5		n=7		n=10	
	Panel A: Cumulative GDP Growth									
	$\beta_k^{(1)}$	$\bar{R}^2$	$\beta_k^{(3)}$	$\bar{R}^2$	$\beta_k^{(5)}$	$\bar{R}^2$	$\beta_k^{(7)}$	$\bar{R}^2$	$\beta_k^{(10)}$	$\bar{R}^2$
k=1	1.816	0.065	1.005	0.069	0.750	0.058	0.629	0.050	0.529	0.042
	(2.30)		(2.90)		(2.86)		(2.76)		(2.61)	
k=4	1.446	0.181	0.848	0.218	0.682	0.216	0.596	0.206	0.520	0.191
	(5.31)		(4.95)		(4.64)		(4.30)		(3.92)	
	⟨4.40⟩		⟨4.30⟩		⟨3.95⟩		⟨3.60⟩		⟨3.30⟩	
k=8	0.980	0.157	0.583	0.194	0.484	0.206	0.432	0.206	0.387	0.201
	(3.80)		(4.45)		(4.45)		(4.25)		(4.00)	
	⟨3.16⟩		⟨4.02⟩		⟨3.95⟩		⟨3.79⟩		⟨3.59⟩	
k=12	0.643	0.094	0.440	0.157	0.390	0.191	0.357	0.200	0.322	0.198
	(2.51)		(3.91)		(4.27)		(4.25)		(4.11)	
	⟨2.37⟩		⟨3.48⟩		⟨3.64⟩		⟨3.61⟩		⟨3.49⟩	
	Panel B: Marginal GDP Growth									
	$\beta_k^{(1)}$	$\bar{R}^2$	$\beta_k^{(3)}$	$\bar{R}^2$	$\beta_k^{(5)}$	$\bar{R}^2$	$\beta_k^{(7)}$	$\bar{R}^2$	$\beta_k^{(10)}$	$\bar{R}^2$
k=4	1.446	0.181	0.848	0.218	0.682	0.216	0.596	0.206	0.520	0.191
	(5.31)		(4.95)		(4.64)		(4.30)		(3.92)	
	⟨4.41⟩		⟨4.31⟩		⟨3.96⟩		⟨3.62⟩		⟨3.29⟩	
k=8	0.508	0.019	0.315	0.028	0.283	0.037	0.266	0.042	0.251	0.046
	(1.50)		(1.77)		(2.02)		(2.12)		(2.19)	
	⟨1.22⟩		⟨1.55⟩		⟨1.76⟩		⟨1.85⟩		⟨1.87⟩	
k=12	-0.047	-0.008	0.135	-0.002	0.181	0.011	0.180	0.016	0.162	0.015
	(-0.11)		(0.62)		(1.03)		(1.14)		(1.12)	
	⟨-0.09⟩		⟨0.52⟩		⟨0.85⟩		⟨0.94⟩		⟨0.92⟩	

Table II
Predictive Performance of Different Yield Curve Measures for Cumulative Real GDP Growth, Germany

Estimation is based on the following regression

$$y_{t+k}^{(k)} = \beta_0^{(k)} + \beta_1^{(k)} X_t + \epsilon_{t+k}^{(k)},$$

where X_t contains different information measures from the yield curve, with different combinations of yield curve measures from Model 1 through 6. The dependent variable $y_{t+k}^{(k)}$ is defined as cumulative real GDP growth (annualized). The forecasting horizon is denoted by k. All models are estimated using a constant although its estimate is not reported. The sample period is 1972:Q4-2006:Q1. We report t-statistics based on two methods for calculating standard errors: $(\cdot)$ contains t-statistics based on Newey-West standard errors with $k-1$ lags, $\langle\cdot\rangle$ is based on the standard errors from the moving block bootstrap with 100,000 replications.

Horizon		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
k=1	Term Spread	0.563			0.916	0.600	
		(2.50)			(2.20)	(2.30)	
	Short Rate		-0.201		0.268		-0.165
			(-1.70)		(1.24)		(-1.24)
	CP-Factor			0.124		-0.035	0.070
				(1.07)		(-0.28)	(0.55)
	$\bar{R}^2$	0.044	0.009	0.002	0.045	0.036	0.003
k=4	Term Spread	0.444			0.596	0.359	
		(2.96)			(2.58)	(2.23)	
		⟨2.44⟩			⟨1.95⟩	⟨1.92⟩	
	Short Rate		-0.204		0.114		-0.136
			(-1.85)		(0.74)		(-1.21)
			⟨-1.33⟩		⟨0.53⟩		⟨-0.87⟩
	CP-Factor			0.177		0.081	0.133
				(2.55)		(1.28)	(2.01)
				⟨2.19⟩		⟨1.13⟩	⟨1.59⟩
	$\bar{R}^2$	0.146	0.072	0.088	0.146	0.154	0.110
k=8	Term Spread	0.366			0.529	0.288	
		(3.31)			(2.25)	(2.25)	
		⟨2.92⟩			⟨1.78⟩	⟨1.94⟩	
	Short Rate		-0.169		0.125		-0.107
			(-1.83)		(0.77)		(-1.03)
			⟨-1.48⟩		⟨0.58⟩		⟨-0.87⟩
	CP-Factor			0.151		0.074	0.117
				(3.03)		(1.41)	(1.94)
				⟨2.63⟩		⟨1.28⟩	⟨1.55⟩
	$\bar{R}^2$	0.165	0.078	0.108	0.171	0.178	0.129
k=12	Term Spread	0.333			0.573	0.287	
		(3.84)			(2.69)	(2.43)	
		⟨3.22⟩			⟨2.10⟩	⟨2.19⟩	
	Short Rate		-0.142		0.188		-0.091
			(-1.85)		(1.28)		(-0.92)
			⟨-1.71⟩		⟨0.97⟩		⟨-0.83⟩
	CP-Factor			0.119		0.043	0.091
				(3.61)		(0.85)	(1.52)
				⟨2.74⟩		⟨0.80⟩	⟨1.32⟩
	$\bar{R}^2$	0.191	0.072	0.094	0.222	0.193	0.113

Table III
Predictive Performance of Different Yield Curve Measures for Marginal Real GDP Growth, Germany

Estimation is based on the following regression: $y^{(4)}_{t+k} = \beta^{(k)}_0 + \beta^{(k)}_1 X_t + \epsilon^{(k)}_{t+k}$, where X_t contains different information measures from the yield curve, with different combinations of yield curve measures from Model 1 through 6. The dependent variable $y^{(4)}_{t+k}$ is defined as marginal real GDP growth (annualized). The forecasting horizon is denoted by k. All models are estimated using a constant although its estimate is not reported. The sample period is 1972:Q4-2006:Q1. We report t-statistics based on two methods for calculating standard errors: $(\cdot)$ contains t-statistics based on Newey-West standard errors with $k-1$ lags, $\langle\cdot\rangle$ is based on the standard errors from the moving block bootstrap with 100,000 replications.

Horizon		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
k=6	Term Spread	0.301			0.339	0.206	
		(2.09)			(1.24)	(1.17)	
		⟨1.87⟩			⟨1.05⟩	⟨1.06⟩	
	Short Rate		-0.156		0.029		-0.097
			(-1.49)		(0.16)		(-0.85)
			⟨-1.24⟩		⟨0.13⟩		⟨-0.74⟩
	CP-Factor			0.145		0.090	0.114
				(2.57)		(1.33)	(1.83)
				⟨2.36⟩		⟨1.22⟩	⟨1.60⟩
	$\bar{R}^2$	0.065	0.039	0.059	0.058	0.076	0.066
k=8	Term Spread	0.278			0.487	0.205	
		(2.11)			(2.03)	(1.26)	
		⟨1.80⟩			⟨1.80⟩	⟨1.11⟩	
	Short Rate		-0.111		0.160		-0.053
			(-1.09)		(0.93)		(-0.45)
			⟨-0.97⟩		⟨0.81⟩		⟨-0.41⟩
	CP-Factor			0.124		0.069	0.108
				(2.05)		(0.91)	(1.43)
				⟨1.76⟩		⟨0.82⟩	⟨1.34⟩
	$\bar{R}^2$	0.054	0.014	0.041	0.060	0.057	0.037
k=10	Term Spread	0.265			0.626	0.223	
		(1.73)			(2.72)	(1.11)	
		⟨1.41⟩			⟨2.44⟩	⟨1.01⟩	
	Short Rate		-0.073		0.281		-0.022
			(-0.69)		(1.73)		(-0.17)
			⟨-0.57⟩		⟨1.52⟩		⟨-0.16⟩
	CP-Factor			0.099		0.039	0.092
				(1.45)		(0.40)	(1.07)
				⟨1.34⟩		⟨0.42⟩	⟨1.07⟩
	$\bar{R}^2$	0.048	0.001	0.023	0.080	0.043	0.015
k=12	Term Spread	0.234			0.769	0.247	
		(1.43)			(3.84)	(1.11)	
		⟨1.24⟩			⟨3.33⟩	⟨1.08⟩	
	Short Rate		-0.024		0.420		0.008
			(-0.21)		(2.70)		(0.06)
			⟨-0.20⟩		⟨2.45⟩		⟨0.06⟩
	CP-Factor			0.054		-0.012	0.056
				(0.71)		(-0.11)	(0.59)
				⟨0.71⟩		⟨-0.11⟩	⟨0.61⟩
	$\bar{R}^2$	0.036	-0.008	0.000	0.112	0.027	-0.009

Table IV
Predictive Performance of Different Yield Curve Measures for Cumulative Real GDP Growth, USA

Estimation is based on the following regression

$$y_{t+k}^{(k)} = \beta_0^{(k)} + \beta_1^{(k)} X_t + \epsilon_{t+k}^{(k)},$$

where X_t contains different information measures from the yield curve, with different combinations of yield curve measures from Model 1 through 6. The dependent variable $y_{t+k}^{(k)}$ is defined as cumulative real GDP growth (annualized). The forecasting horizon is denoted by k. All models are estimated using a constant although its estimate is not reported. The sample period is 1972:Q4-2006:Q1. We report t-statistics based on two methods for calculating standard errors: $(\cdot)$ contains t-statistics based on Newey-West standard errors with $k-1$ lags, $\langle\cdot\rangle$ is based on the standard errors from the moving block bootstrap with 100,000 replications.

Horizon		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
k=1	Term Spread	0.605			0.451	0.166	
		(2.46)			(1.83)	(0.68)	
	Short Rate		-0.234		-0.125		-0.111
			(-1.82)		(-0.91)		(-1.06)
	CP-Factor			0.244		0.218	0.219
				(3.07)		(2.59)	(3.12)
	$\bar{R}^2$	0.055	0.040	0.124	0.057	0.120	0.126
k=4	Term Spread	0.730			0.648	0.398	
		(3.69)			(2.59)	(1.82)	
		⟨2.93⟩			⟨2.09⟩	⟨1.68⟩	
	Short Rate		-0.228		-0.064		-0.116
			(-2.01)		(-0.50)		(-1.30)
			⟨-1.54⟩		⟨-0.42⟩		⟨-0.99⟩
	CP-Factor			0.225		0.162	0.200
				(4.23)		(2.54)	(3.43)
				⟨3.95⟩		⟨2.43⟩	⟨2.63⟩
	$\bar{R}^2$	0.224	0.106	0.276	0.223	0.317	0.296
k=8	Term Spread	0.655			0.672	0.455	
		(4.16)			(2.87)	(2.88)	
		⟨3.44⟩			⟨2.40⟩	⟨2.52⟩	
	Short Rate		-0.158		0.014		-0.071
			(-1.46)		(0.13)		(-0.73)
			⟨-1.23⟩		⟨0.10⟩		⟨-0.56⟩
	CP-Factor			0.169		0.096	0.154
				(4.49)		(2.98)	(4.01)
				⟨4.07⟩		⟨2.69⟩	⟨2.98⟩
	$\bar{R}^2$	0.313	0.082	0.270	0.308	0.368	0.279
k=12	Term Spread	0.496			0.555	0.425	
		(3.83)			(3.05)	(2.58)	
		⟨3.34⟩			⟨2.48⟩	⟨2.46⟩	
	Short Rate		-0.089		0.052		-0.034
			(-0.92)		(0.58)		(-0.34)
			⟨-0.83⟩		⟨0.41⟩		⟨-0.28⟩
	CP-Factor			0.100		0.033	0.094
				(4.51)		(1.07)	(3.59)
				⟨3.43⟩		⟨0.97⟩	⟨2.17⟩
	$\bar{R}^2$	0.294	0.034	0.158	0.298	0.300	0.155

Table V
Predictive Performance of Different Yield Curve Measures for Marginal Real GDP Growth, USA

Estimation is based on the following regression: $y^{(4)}_{t+k} = \beta^{(k)}_0 + \beta^{(k)}_1 X_t + \epsilon^{(k)}_{t+k}$, where X_t contains different information measures from the yield curve, with different combinations of yield curve measures from Model 1 through 6. The dependent variable $y^{(4)}_{t+k}$ is defined as marginal real GDP growth (annualized). The forecasting horizon is denoted by k. All models are estimated using a constant although its estimate is not reported. The sample period is 1972:Q4-2006:Q1. We report t-statistics based on two methods for calculating standard errors: $(\cdot)$ contains t-statistics based on Newey-West standard errors with $k-1$ lags, $\langle\cdot\rangle$ is based on the standard errors from the moving block bootstrap with 100,000 replications.

Horizon		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
k=6	Term Spread	0.696			0.729	0.483	
		(3.53)			(2.51)	(2.14)	
		⟨3.17⟩			⟨2.22⟩	⟨1.98⟩	
	Short Rate		-0.163		0.026		-0.070
			(-1.32)		(0.19)		(-0.67)
			⟨-1.19⟩		⟨0.16⟩		⟨-0.55⟩
	CP-Factor			0.181		0.104	0.166
				(3.45)		(1.76)	(3.48)
				⟨3.28⟩		⟨1.64⟩	⟨2.83⟩
	$\bar{R}^2$	0.203	0.048	0.175	0.197	0.237	0.177
k=8	Term Spread	0.563			0.690	0.498	
		(4.03)			(2.82)	(3.06)	
		⟨3.43⟩			⟨2.32⟩	⟨2.64⟩	
	Short Rate		-0.074		0.103		-0.012
			(-0.50)		(0.63)		(-0.09)
			⟨-0.46⟩		⟨0.53⟩		⟨-0.07⟩
	CP-Factor			0.111		0.031	0.108
				(1.60)		(0.40)	(2.09)
				⟨1.53⟩		⟨0.39⟩	⟨1.80⟩
	$\bar{R}^2$	0.129	0.003	0.061	0.137	0.125	0.053
k=10	Term Spread	0.339			0.534	0.472	
		(1.51)			(2.53)	(2.36)	
		⟨1.34⟩			⟨2.09⟩	⟨2.23⟩	
	Short Rate		0.030		0.166		0.042
			(0.18)		(1.04)		(0.28)
			⟨0.17⟩		⟨0.84⟩		⟨0.23⟩
	CP-Factor			0.012		-0.063	0.021
				(0.17)		(-0.79)	(0.41)
				⟨0.16⟩		⟨-0.76⟩	⟨0.31⟩
	$\bar{R}^2$	0.040	-0.007	-0.008	0.070	0.046	-0.014
k=12	Term Spread	0.130			0.358	0.374	
		(0.45)			(1.28)	(0.93)	
		⟨0.39⟩			⟨1.06⟩	⟨0.94⟩	
	Short Rate		0.111		0.202		0.088
			(0.78)		(1.43)		(0.65)
			⟨0.65⟩		⟨1.02⟩		⟨0.48⟩
	CP-Factor			-0.056		-0.115	-0.039
				(-0.66)		(-0.90)	(-0.54)
				⟨-0.61⟩		⟨-0.97⟩	⟨-0.37⟩
	$\bar{R}^2$	-0.002	0.014	0.009	0.045	0.041	0.013

Table VI
Predictive Performance of Different Yield Curve Measures for Cumulative Real GDP Growth, UK

Estimation is based on the following regression specification

$$y_{t+k}^{(k)} = \beta_0^{(k)} + \beta_1^{(k)} X_t + \epsilon_{t+k}^{(k)},$$

where X_t contains different information measures from the yield curve, with different combinations of yield curve measures from Model 1 through 6. The dependent variable $y_{t+k}^{(k)}$ is defined as cumulative real GDP growth (annualized). The forecasting horizon is denoted by k. The sample period is 1980:Q1-2006:Q1. We report t-statistics based on two methods for calculating standard errors: $(\cdot)$ contains t-statistics based on Newey-West standard errors with $k-1$ lags, $\langle\cdot\rangle$ is based on the standard errors from the moving block bootstrap with 100,000 replications.

Horizon		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
k=1	Term Spread	0.449			0.398	0.438	
		(3.10)			(2.39)	(3.00)	
	Short Rate		-0.178		-0.040		-0.100
			(-2.16)		(-0.43)		(-1.03)
	CP-Factor			0.123		0.016	0.110
				(1.14)		(0.17)	(1.11)
	$\bar{R}^2$	0.105	0.056	0.013	0.097	0.095	0.014
k=4	Term Spread	0.472			0.419	0.436	
		(2.79)			(2.60)	(2.96)	
		⟨1.83⟩			⟨1.71⟩	⟨2.05⟩	
	Short Rate		-0.193		-0.041		-0.114
			(-1.67)		(-0.39)		(-0.97)
			⟨-1.02⟩		⟨-0.25⟩		⟨-0.56⟩
	CP-Factor			0.159		0.052	0.145
				(1.31)		(0.58)	(1.40)
				⟨1.03⟩		⟨0.46⟩	⟨1.16⟩
	$\bar{R}^2$	0.245	0.140	0.070	0.240	0.244	0.089
k=8	Term Spread	0.447			0.430	0.424	
		(2.20)			(2.30)	(2.24)	
		⟨1.63⟩			⟨1.60⟩	⟨1.70⟩	
	Short Rate		-0.176		-0.014		-0.088
			(-1.22)		(-0.12)		(-0.61)
			⟨-0.90⟩		⟨-0.09⟩		⟨-0.40⟩
	CP-Factor			0.138		0.034	0.128
				(1.18)		(0.54)	(1.33)
				⟨1.10⟩		⟨0.48⟩	⟨1.34⟩
	$\bar{R}^2$	0.293	0.147	0.069	0.285	0.288	0.079
k=12	Term Spread	0.384			0.390	0.337	
		(2.26)			(2.51)	(2.05)	
		⟨1.83⟩			⟨1.77⟩	⟨1.74⟩	
	Short Rate		-0.149		0.005		-0.048
			(-1.07)		(0.04)		(-0.34)
			⟨-0.88⟩		⟨0.03⟩		⟨-0.25⟩
	CP-Factor			0.150		0.069	0.146
				(1.87)		(2.25)	(2.21)
				⟨1.69⟩		⟨1.50⟩	⟨2.01⟩
	$\bar{R}^2$	0.280	0.122	0.114	0.271	0.294	0.110

Table VII
Predictive Performance of Different Yield Curve Measures for Marginal Real GDP Growth, UK

Estimation is based on the following regression: $y_{t+k}^{(4)} = \beta_0^{(k)} + \beta_1^{(k)} X_t + \epsilon_{t+k}^{(k)}$, where X_t contains different information measures from the yield curve, with different combinations of yield curve measures from Model 1 through 6. The dependent variable $y_{t+k}^{(4)}$ is defined as marginal real GDP growth (annualized). The forecasting horizon is denoted by k. The sample period is 1980:Q1-2006:Q1. We report t-statistics based on two methods for calculating standard errors: $(\cdot)$ contains t-statistics based on Newey-West standard errors with $k-1$ lags, $\langle\cdot\rangle$ is based on the standard errors from the moving block bootstrap with 100,000 replications.

Horizon		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
k=6	Term Spread	0.464			0.446	0.483	
		(2.12)			(2.05)	(2.27)	
		⟨1.59⟩			⟨1.60⟩	⟨1.84⟩	
	Short Rate		-0.180		-0.014		-0.103
			(-1.26)		(-0.12)		(-0.71)
			⟨-0.93⟩		⟨-0.10⟩		⟨-0.50⟩
	CP-Factor			0.090		-0.028	0.079
				(0.68)		(-0.37)	(0.70)
				⟨0.61⟩		⟨-0.34⟩	⟨0.68⟩
	$\bar{R}^2$	0.239	0.118	0.015	0.230	0.232	0.027
k=8	Term Spread	0.418			0.471	0.407	
		(1.80)			(2.14)	(1.71)	
		⟨1.47⟩			⟨1.64⟩	⟨1.44⟩	
	Short Rate		-0.136		0.042		-0.028
			(-0.86)		(0.36)		(-0.18)
			⟨-0.67⟩		⟨0.27⟩		⟨-0.14⟩
	CP-Factor			0.115		0.015	0.112
				(1.12)		(0.29)	(1.25)
				⟨1.04⟩		⟨0.23⟩	⟨1.14⟩
	$\bar{R}^2$	0.192	0.060	0.031	0.186	0.182	0.021
k=10	Term Spread	0.313			0.411	0.226	
		(1.57)			(2.24)	(1.04)	
		⟨1.32⟩			⟨1.55⟩	⟨0.91⟩	
	Short Rate		-0.079		0.079		0.050
			(-0.52)		(0.63)		(0.35)
			⟨-0.43⟩		⟨0.44⟩		⟨0.28⟩
	CP-Factor			0.183		0.128	0.189
				(2.97)		(2.57)	(3.16)
				⟨2.28⟩		⟨1.76⟩	⟨2.44⟩
	$\bar{R}^2$	0.103	0.011	0.099	0.105	0.138	0.093
k=12	Term Spread	0.228			0.344	0.136	
		(2.02)			(2.05)	(0.98)	
		⟨1.56⟩			⟨1.40⟩	⟨0.77⟩	
	Short Rate		-0.040		0.095		0.085
			(-0.33)		(0.67)		(0.62)
			⟨-0.27⟩		⟨0.49⟩		⟨0.48⟩
	CP-Factor			0.169		0.137	0.177
				(2.63)		(1.84)	(2.68)
				⟨1.94⟩		⟨1.37⟩	⟨2.05⟩
	$\bar{R}^2$	0.049	-0.007	0.083	0.054	0.089	0.085

Table VIII
Out-of-sample Forecast Evaluation, Germany

A series of (pseudo-) out-of-sample forecasts is generated by recursive estimation of the forecasting regression for cumulative real GDP growth:

$$y_{t+k}^{(k)} = \beta_0^{(k)} + \beta_1^{(k)} X_t + \epsilon_{t+k}^{(k)},$$

where X_t contains different measures from the yield curve. Model 1 (M1) only includes the term spread, M2 uses the short rate, M3 includes the Cochrane-Piazzesi measure of time-varying bond risk premia (CP-factor). M4 is based on the term spread and the short rate, M5 contains the spread and the CP-factor, whereas M6 is based on the short rate and the CP-factor. ME is the mean forecast error, Theil's U is the ratio of the RMSE of the model of interest to the RMSE of a naive benchmark model (n.M.). CW is the difference between the MSFE of the benchmark model and the MSFE (adjusted) of the model of interest. The Clark-West t-statistic for testing the hypothesis of equal predictive performance is reported in parentheses $\langle \cdot \rangle$ based on the MBB. $CW^{\S}$ and $CW^{\dagger}$ compare the MSFE from a larger conditional model to a smaller conditional nested model, with the superscript $\langle \cdot \rangle^a$, $\langle \cdot \rangle^b$ indicating the smaller nested model. A CW-statistic which is significant at the 10 % level or below is bold-printed. Sample period: 1972:Q4-2006:Q1. After 20 quarters of initialization, the models are estimated recursively.

Horizon	Statistic	M1 [a]	M2 [b]	M3 [c]	M4	M5	M6	n.M. [n]
k=1	Mean Error	-0.277	-0.569	-0.183	-0.190	-0.302	-0.546	-0.403
	ME *t*-stat.	⟨-0.82⟩	⟨-1.64⟩	⟨-0.53⟩	⟨-0.56⟩	⟨-0.89⟩	⟨-1.57⟩	⟨-1.20⟩
	Theil's U	1.001	1.026	1.013	1.006	1.008	1.035	
	CW	**1.364**	1.103	0.210	**1.345**	**1.340**	0.987	
	CW *t*-stat.	⟨1.46⟩^{n}	⟨1.09⟩^{n}	⟨0.37⟩^{n}	⟨1.42⟩^{n}	⟨1.39⟩^{n}	⟨0.96⟩^{n}	
	$CW^{\S}$				-0.041	-0.127	-0.174	
	$CW^{\S}$ *t*-stat.				⟨-0.22⟩^{a}	⟨-0.70⟩^{a}	⟨-1.16⟩^{b}	
	$CW^{\dagger}$				**0.589**	0.174	-0.503	
	$CW^{\S}$ *t*-stat.				⟨1.82⟩^{b}	⟨0.23⟩^{c}	⟨-0.71⟩^{c}	
k=4	Mean Error	-0.226	-0.513	-0.056	-0.441	-0.174	-0.415	-0.272
	ME *t*-stat.	⟨-0.59⟩	⟨-1.19⟩	⟨-0.14⟩	⟨-1.03⟩	⟨-0.45⟩	⟨-0.96⟩	⟨-0.74⟩
	Theil's U	1.022	1.084	1.028	1.076	1.028	1.074	
	CW	0.827	1.105	0.410	1.035	0.750	1.017	
	CW *t*-stat.	⟨0.99⟩^{n}	⟨1.06⟩^{n}	⟨0.56⟩^{n}	⟨1.00⟩^{n}	⟨0.85⟩^{n}	⟨0.96⟩^{n}	
	$CW^{\S}$				-0.020	0.006	0.117	
	$CW^{\S}$ *t*-stat.				⟨-0.07⟩^{a}	⟨0.10⟩^{a}	⟨1.07⟩^{b}	
	$CW^{\dagger}$				**0.376**	0.039	-0.246	
	$CW^{\S}$ *t*-stat.				⟨1.53⟩^{b}	⟨0.08⟩^{c}	⟨-0.40⟩^{c}	
k=8	Mean Error	-0.119	-0.395	0.010	-0.352	-0.072	-0.282	-0.113
	ME *t*-stat.	⟨-0.29⟩	⟨-0.87⟩	⟨0.02⟩	⟨-0.71⟩	⟨-0.17⟩	⟨-0.56⟩	⟨-0.29⟩
	Theil's U	1.043	1.130	1.092	1.154	1.076	1.170	
	CW	0.356	0.245	-0.162	0.116	0.145	-0.076	
	CW *t*-stat.	⟨0.85⟩^{n}	⟨0.43⟩^{n}	⟨-0.91⟩^{n}	⟨0.25⟩^{n}	⟨0.40⟩^{n}	⟨-0.17⟩^{n}	
	$CW^{\S}$				-0.238	-0.082	-0.053	
	$CW^{\S}$ *t*-stat.				⟨-0.99⟩^{a}	⟨-0.79⟩^{a}	⟨-0.26⟩^{b}	
	$CW^{\dagger}$				0.120	0.118	-0.212	
	$CW^{\S}$ *t*-stat.				⟨0.53⟩^{b}	⟨0.33⟩^{c}	⟨-0.36⟩^{c}	
k=12	Mean Error	-0.082	-0.322	0.001	-0.125	-0.048	-0.224	-0.094
	ME *t*-stat.	⟨-0.19⟩	⟨-0.64⟩	⟨0.00⟩	⟨-0.22⟩	⟨-0.11⟩	⟨-0.38⟩	⟨-0.22⟩
	Theil's U	0.965	1.057	1.028	1.179	1.021	1.201	
	CW	**0.492**	0.254	-0.001	-0.134	**0.463**	-0.199	
	CW *t*-stat.	⟨1.65⟩^{n}	⟨0.57⟩^{n}	⟨-0.01⟩^{n}	⟨-0.42⟩^{n}	⟨1.34⟩^{n}	⟨-0.52⟩^{n}	
	$CW^{\S}$				**-0.593**	**-0.143**	**-0.413**	
	$CW^{\S}$ *t*-stat.				⟨-2.54⟩^{a}	⟨-2.98⟩^{a}	⟨-1.61⟩^{b}	
	$CW^{\dagger}$				-0.293	0.059	-0.511	
	$CW^{\S}$ *t*-stat.				⟨-0.68⟩^{b}	⟨0.17⟩^{c}	⟨-1.17⟩^{c}	

Table IX
Out-of-sample Forecast Evaluation, USA

A series of (pseudo-) out-of-sample forecasts is generated by recursive estimation of the forecasting regression for cumulative real GDP growth:

$$y_{t+k}^{(k)} = \beta_0^{(k)} + \beta_1^{(k)} X_t + \epsilon_{t+k}^{(k)},$$

where X_t contains different measures from the yield curve. Model 1 (M1) only includes the term spread, M2 uses the short rate, M3 includes the Cochrane-Piazzesi measure of time-varying bond risk premia (CP-factor). M4 is based on the term spread and the short rate, M5 contains the spread and the CP-factor, whereas M6 is based on the short rate and the CP-factor. ME is the mean forecast error, Theil's U is the ratio of the RMSE of the model of interest and the RMSE of a naive benchmark model (n.M.). CW is the difference between the MSFE of the benchmark model and the MSFE (adjusted) of the model of interest. The Clark-West t-statistic for testing the hypothesis of equal predictive performance is reported in parentheses $\langle\cdot\rangle$ based on the MBB. $CW^{\S}$ and $CW^{\dagger}$ compare the MSFE from a larger conditional model to a smaller conditional nested model, with the superscript $\langle\cdot\rangle^a$, $\langle\cdot\rangle^b$ indicating the smaller nested model. A CW-statistic which is significant at the 10 % level or below is bold-printed. Sample period: 1972:Q4-2006:Q1. After 20 quarters of initialization, the models are estimated recursively.

Horizon	Statistic	M1 [a]	M2 [b]	M3 [c]	M4	M5	M6	n.M. [n]
k=1	Mean Error	-0.432	-0.725	-0.183	-0.648	-0.358	-0.358	-0.072
	ME *t*-stat.	⟨-1.36⟩	⟨-2.29⟩	⟨-0.66⟩	⟨-1.98⟩	⟨-1.18⟩	⟨-1.19⟩	⟨-0.23⟩
	Theil's U	1.061	1.082	0.985	1.120	1.057	1.045	
	CW	1.065	1.204	**2.837**	0.285	1.778	**1.603**	
	CW t-stat.	$\langle 1.25\rangle^n$	$\langle 1.13\rangle^n$	$\langle 1.70\rangle^n$	$\langle 0.49\rangle^n$	$\langle 1.26\rangle^n$	$\langle 1.53\rangle^n$	
	$CW^{\S}$				0.298	**0.985**	**1.674**	
	$CW^{\S}$ *t*-stat.				$\langle 0.32\rangle^a$	$\langle 1.57\rangle^a$	$\langle 2.39\rangle^b$	
	$CW^{\dagger}$				0.625	-0.134	0.228	
	$CW^{\S}$ *t*-stat.				$\langle 0.76\rangle^b$	$\langle -0.13\rangle^c$	$\langle 0.28\rangle^c$	
k=4	Mean Error	-0.566	-0.712	-0.129	-0.755	-0.530	-0.274	0.102
	ME *t*-stat.	⟨-1.49⟩	⟨-1.39⟩	⟨-0.38⟩	⟨-1.77⟩	⟨-1.35⟩	⟨-0.56⟩	⟨0.28⟩
	Theil's U	1.015	1.091	0.890	1.144	1.001	1.152	
	CW	**2.458**	**1.232**	**2.810**	**1.113**	**2.846**	0.643	
	CW t-stat.	$\langle 1.84\rangle^n$	$\langle 1.30\rangle^n$	$\langle 1.52\rangle^n$	$\langle 2.11\rangle^n$	$\langle 1.62\rangle^n$	$\langle 0.95\rangle^n$	
	$CW^{\S}$				**-0.582**	**0.444**	0.480	
	$CW^{\S}$ *t*-stat.				$\langle -1.29\rangle^a$	$\langle 1.56\rangle^a$	$\langle 0.56\rangle^b$	
	$CW^{\dagger}$				-0.019	-0.405	-0.933	
	$CW^{\S}$ *t*-stat.				$\langle -0.03\rangle^b$	$\langle -0.73\rangle^c$	$\langle -0.63\rangle^c$	
k=8	Mean Error	-0.419	-0.434	-0.058	-0.277	-0.384	-0.146	0.196
	ME *t*-stat.	⟨-1.16⟩	⟨-0.98⟩	⟨-0.20⟩	⟨-0.66⟩	⟨-1.04⟩	⟨-0.45⟩	⟨0.51⟩
	Theil's U	0.993	1.151	0.892	1.219	0.966	1.059	
	CW	**1.995**	0.037	**1.668**	**0.867**	**2.201**	**0.589**	
	CW t-stat.	$\langle 1.91\rangle^n$	$\langle 0.14\rangle^n$	$\langle 1.64\rangle^n$	$\langle 1.44\rangle^n$	$\langle 1.83\rangle^n$	$\langle 2.32\rangle^n$	
	$CW^{\S}$				**-0.583**	**0.214**	**1.245**	
	$CW^{\S}$ *t*-stat.				$\langle -2.11\rangle^a$	$\langle 2.22\rangle^a$	$\langle 3.10\rangle^b$	
	$CW^{\dagger}$				0.085	-0.159	0.204	
	$CW^{\S}$ *t*-stat.				$\langle 0.11\rangle^b$	$\langle -0.53\rangle^c$	$\langle 0.43\rangle^c$	
k=12	Mean Error	-0.345	-0.253	-0.029	-0.176	-0.329	0.288	0.050
	ME *t*-stat.	⟨-1.14⟩	⟨-0.87⟩	⟨-0.12⟩	⟨-0.46⟩	⟨-1.06⟩	⟨0.82⟩	⟨0.18⟩
	Theil's U	1.065	1.121	0.839	1.319	1.096	1.486	
	CW	**0.890**	-0.063	**0.649**	**0.943**	**0.833**	-0.462	
	CW t-stat.	$\langle 1.96\rangle^n$	$\langle -0.39\rangle^n$	$\langle 1.78\rangle^n$	$\langle 1.39\rangle^n$	$\langle 1.92\rangle^n$	$\langle -0.93\rangle^n$	
	$CW^{\S}$				-0.285	-0.025	-0.294	
	$CW^{\S}$ *t*-stat.				$\langle -1.05\rangle^a$	$\langle -0.29\rangle^a$	$\langle -0.50\rangle^b$	
	$CW^{\dagger}$				-0.166	**-0.433**	-0.821	
	$CW^{\S}$ *t*-stat.				$\langle -0.32\rangle^b$	$\langle -2.44\rangle^c$	$\langle -1.01\rangle^c$	

Table X
Out-of-sample Forecast Evaluation, UK

A series of (pseudo-) out-of-sample forecasts is generated by recursive estimation of the forecasting regression for cumulative real GDP growth:

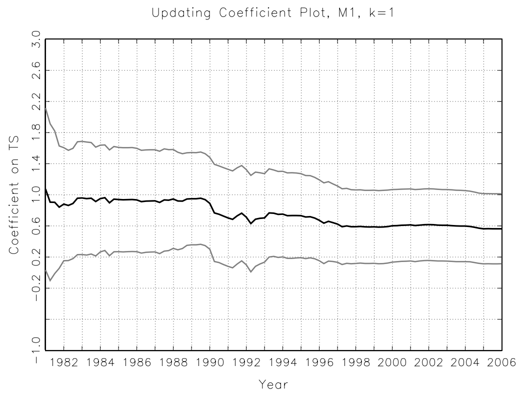
$$y_{t+k}^{(k)} = \beta_0^{(k)} + \beta_1^{(k)} X_t + \epsilon_{t+k}^{(k)},$$

where X_t contains different measures from the yield curve. Model 1 (M1) only includes the term spread, M2 uses the short rate, M3 includes the Cochrane-Piazzesi measure of time-varying bond risk premia (CP-factor). M4 is based on the term spread and the short rate, M5 contains the spread and the CP-factor, whereas M6 is based on the short rate and the CP-factor. ME is the mean forecast error, Theil's U is the ratio of the RMSE of the model of interest and the RMSE of a naive benchmark model (n.M.). CW is the difference between the MSFE of the benchmark model and the MSFE (adjusted) of the model of interest. The Clark-West t-statistic for testing the hypothesis of equal predictive performance is reported in parentheses $\langle\cdot\rangle$ based on the MBB. $CW^{\S}$ and $CW^{\dagger}$ compare the MSFE from a larger conditional model to a smaller conditional nested model, with the superscript $\langle\cdot\rangle^a$, $\langle\cdot\rangle^b$ indicating the smaller nested model. A CW-statistic which is significant at the 10 % level or below is bold-printed. Sample period: 1980:Q1-2006:Q1. After 20 quarters of initialization, the models are estimated recursively.

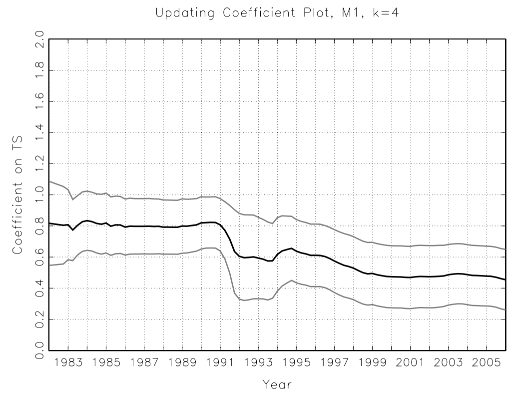
Horizon	Statistic	M1 [a]	M2 [b]	M3 [c]	M4	M5	M6	n.M. [n]
k=1	Mean Error	-0.681	-1.230	-0.538	-1.056	-0.706	-1.228	-0.442
	ME *t*-stat.	⟨-2.56⟩	⟨-6.37⟩	⟨-1.52⟩	⟨-4.88⟩	⟨-2.60⟩	⟨-5.56⟩	⟨-1.24⟩
	Theil's U	0.946	0.981	1.004	0.984	0.961	1.035	
	CW	**1.101**	**1.852**	0.022	**1.515**	**0.981**	0.798	
	CW *t*-stat.	⟨1.91⟩^{n}	⟨1.95⟩^{n}	⟨0.16⟩^{n}	⟨1.46⟩^{n}	⟨1.99⟩^{n}	⟨1.20⟩^{n}	
	$CW^{\S}$				0.085	-0.104	-0.033	
	$CW^{\S}$ *t*-stat.				⟨0.19⟩^{a}	⟨-1.19⟩^{a}	⟨-0.08⟩^{b}	
	$CW^{\dagger}$				**0.340**	0.328	0.143	
	$CW^{\S}$ *t*-stat.				⟨1.36⟩^{b}	⟨0.77⟩^{c}	⟨0.22⟩^{c}	
k=4	Mean Error	-0.912	-1.631	-0.557	-1.343	-0.934	-1.447	-0.429
	ME *t*-stat.	⟨-1.72⟩	⟨-4.79⟩	⟨-1.04⟩	⟨-2.74⟩	⟨-1.75⟩	⟨-3.82⟩	⟨-0.77⟩
	Theil's U	1.075	1.121	1.003	1.147	1.084	1.074	
	CW	0.649	**1.914**	0.075	1.168	0.605	**1.140**	
	CW *t*-stat.	⟨0.93⟩^{n}	⟨1.94⟩^{n}	⟨0.73⟩^{n}	⟨1.22⟩^{n}	⟨0.87⟩^{n}	⟨1.72⟩^{n}	
	$CW^{\S}$				-0.100	**-0.051**	0.743	
	$CW^{\S}$ *t*-stat.				⟨-0.41⟩^{a}	⟨-1.87⟩^{a}	⟨1.04⟩^{b}	
	$CW^{\dagger}$				0.210	-0.510	-0.026	
	$CW^{\S}$ *t*-stat.				⟨0.37⟩^{b}	⟨-0.98⟩^{c}	⟨-0.07⟩^{c}	
k=8	Mean Error	-0.822	-1.707	-0.314	-0.941	-0.839	-1.048	-0.187
	ME *t*-stat.	⟨-1.64⟩	⟨-4.49⟩	⟨-0.71⟩	⟨-1.60⟩	⟨-1.67⟩	⟨-3.21⟩	⟨-0.42⟩
	Theil's U	1.321	1.535	1.012	1.461	1.335	1.090	
	CW	-0.151	0.565	0.023	-0.271	-0.203	0.585	
	CW *t*-stat.	⟨-0.15⟩^{n}	⟨0.44⟩^{n}	⟨0.17⟩^{n}	⟨-0.23⟩^{n}	⟨-0.20⟩^{n}	⟨1.14⟩^{n}	
	$CW^{\S}$				**-0.674**	**-0.078**	**3.582**	
	$CW^{\S}$ *t*-stat.				⟨-5.88⟩^{a}	⟨-1.35⟩^{a}	⟨2.42⟩^{b}	
	$CW^{\dagger}$				0.642	**-1.635**	0.709	
	$CW^{\S}$ *t*-stat.				⟨0.89⟩^{b}	⟨-2.20⟩^{c}	⟨1.02⟩^{c}	
k=12	Mean Error	-0.361	-1.075	0.072	0.325	-0.379	0.099	0.189
	ME *t*-stat.	⟨-1.34⟩	⟨-6.31⟩	⟨0.39⟩	⟨0.99⟩	⟨-1.39⟩	⟨0.33⟩	⟨0.82⟩
	Theil's U	1.416	1.885	0.899	1.860	1.408	1.341	
	CW	0.523	0.852	**0.180**	-0.174	0.566	-0.051	
	CW *t*-stat.	⟨1.12⟩^{n}	⟨1.20⟩^{n}	⟨2.55⟩^{n}	⟨-0.59⟩^{n}	⟨1.17⟩^{n}	⟨-0.26⟩^{n}	
	$CW^{\S}$				0.056	0.020	**3.135**	
	$CW^{\S}$ *t*-stat.				⟨0.12⟩^{a}	⟨0.81⟩^{a}	⟨2.58⟩^{b}	
	$CW^{\dagger}$				0.843	**-0.589**	**1.739**	
	$CW^{\S}$ *t*-stat.				⟨1.28⟩^{b}	⟨-3.06⟩^{c}	⟨1.87⟩^{c}	

Figure 1

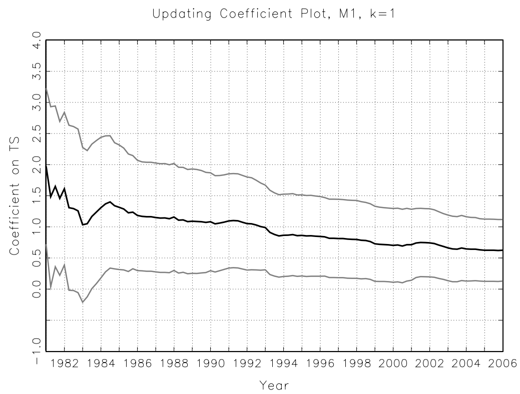
Time-varying Forecast Performance: Recursively Estimated Coefficient on the Term Spread



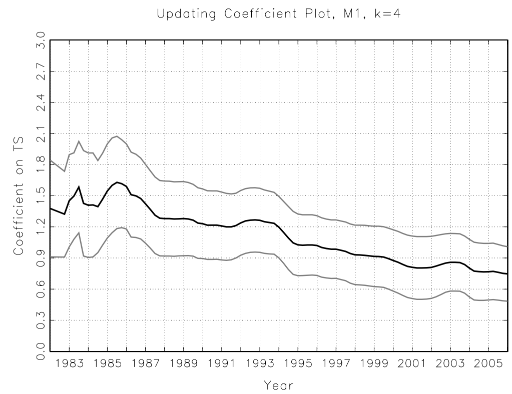
(a) k=1, Germany



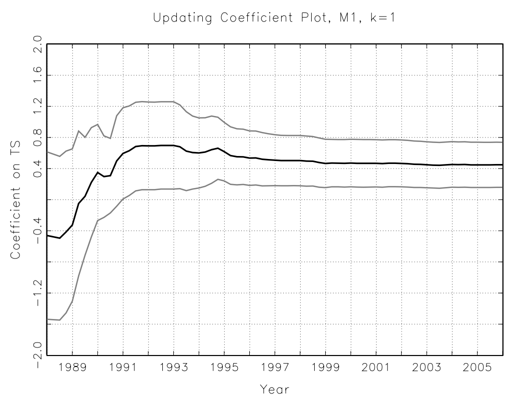
(b) k=4, Germany



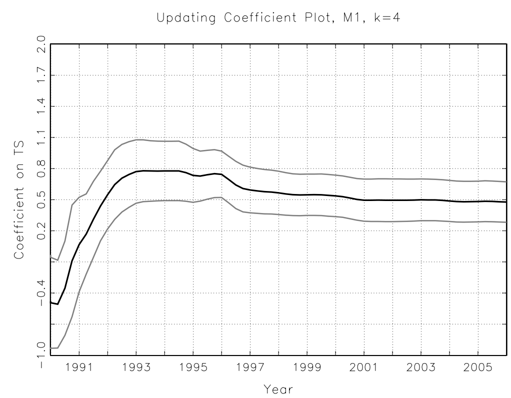
(c) k=1, USA



(d) k=4, USA

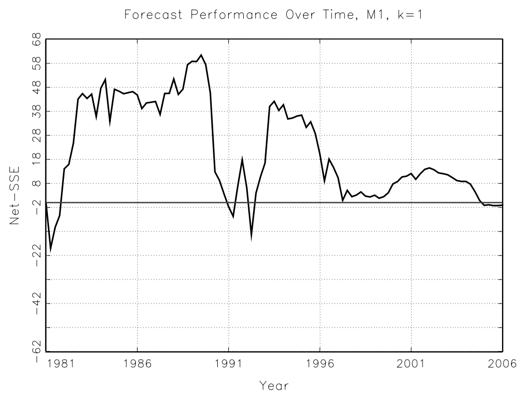


(e) k=1, UK

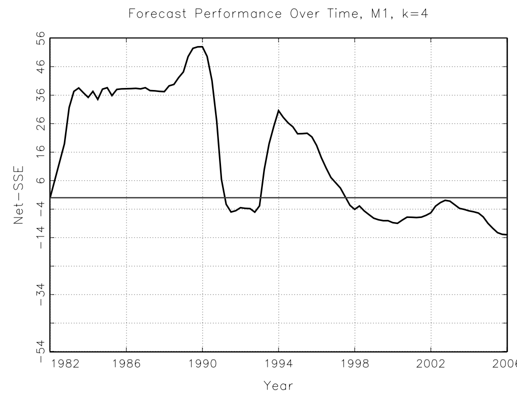


(f) k=4, UK

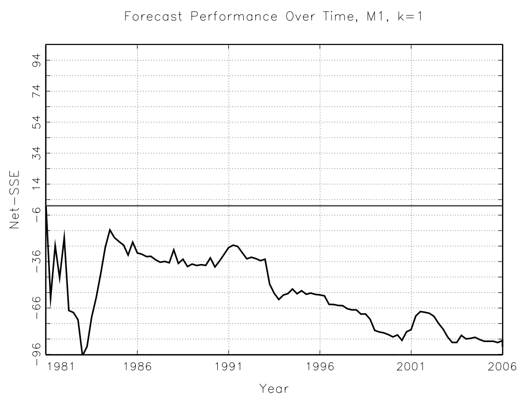
Figure 2
Time-varying Out-of-sample Performance: NET-SSE Plots



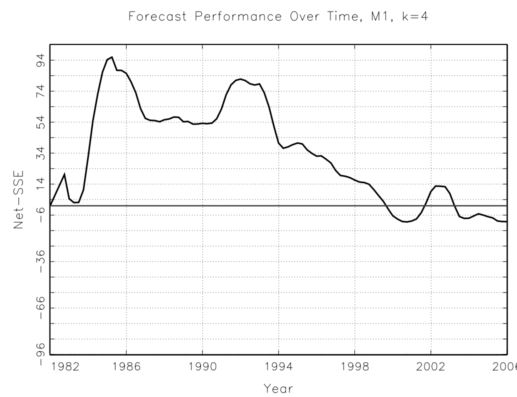
(a) NET-SSE k=1, Germany



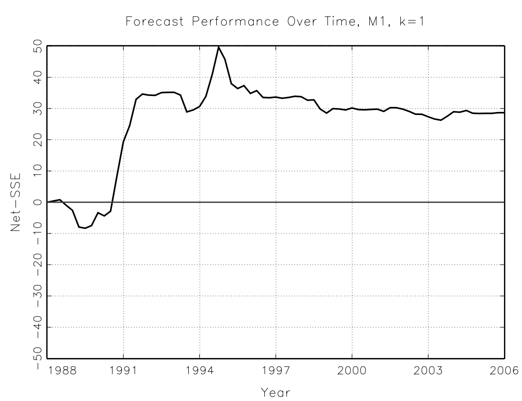
(b) NET-SSE k=4, Germany



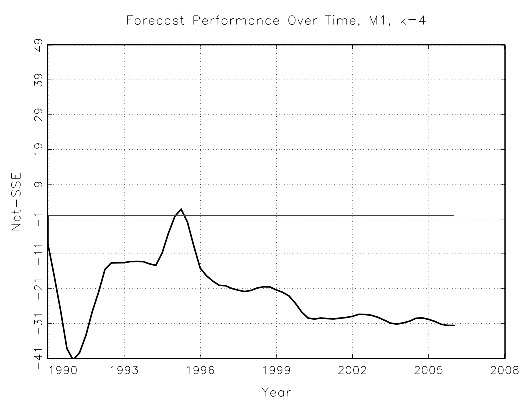
(c) NET-SSE k=1, USA



(d) NET-SSE k=4, USA

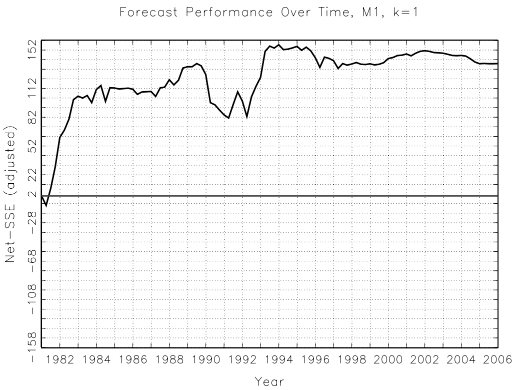


(e) NET-SSE k=1, UK

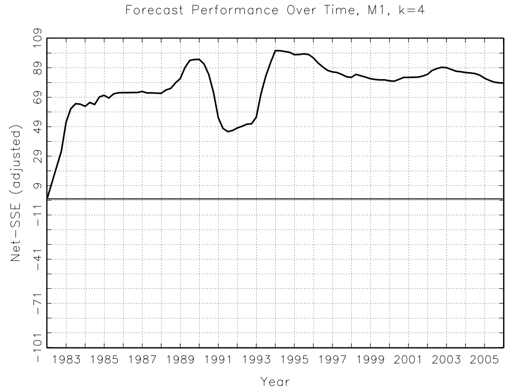


(f) NET-SSE k=4, UK

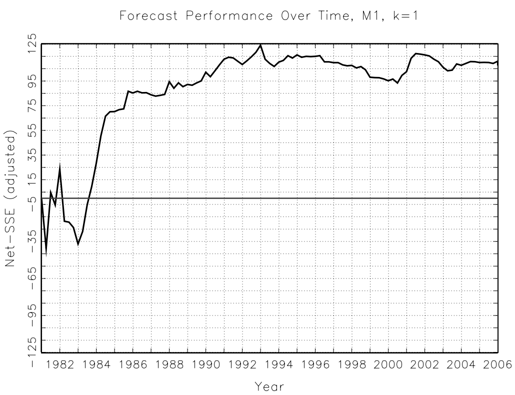
Figure 3
Time-varying Out-of-sample Performance: Adjusted Net-SSE Plots



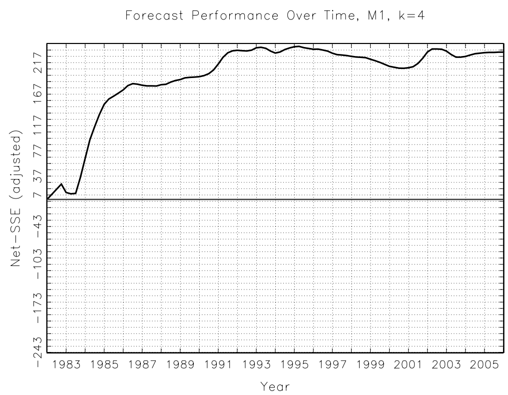
(a) Adjusted NET-SSE k=1, Germany



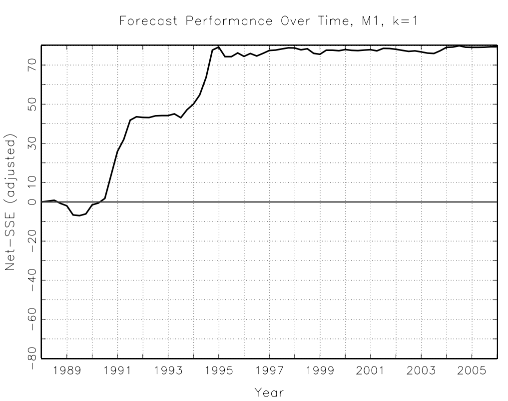
(b) Adjusted NET-SSE k=4, Germany



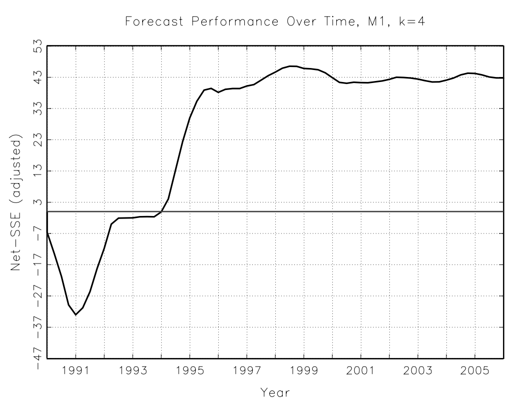
(c) Adjusted NET-SSE k=1, USA



(d) Adjusted NET-SSE k=4, USA

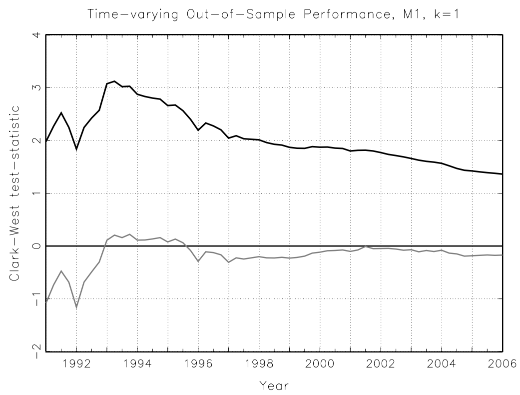


(e) Adjusted NET-SSE k=1, UK

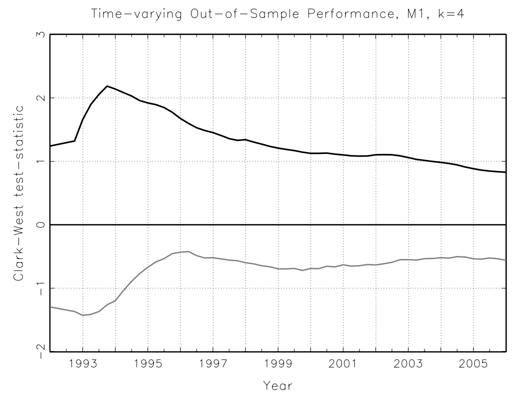


(f) Adjusted NET-SSE k=4, UK

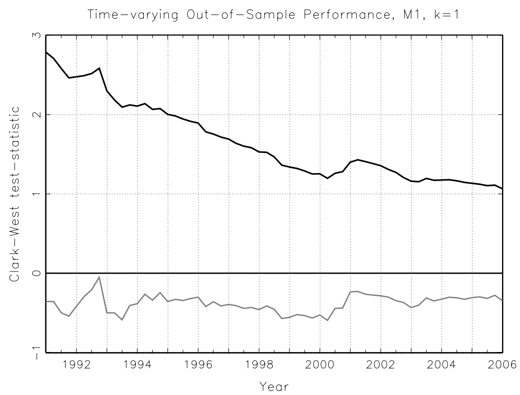
Figure 4
Time-varying Out-of-sample Performance: Clark-West Statistic



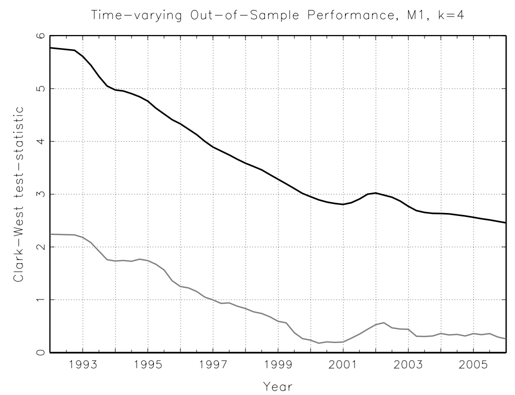
(a) Clark-West, k=1, Germany



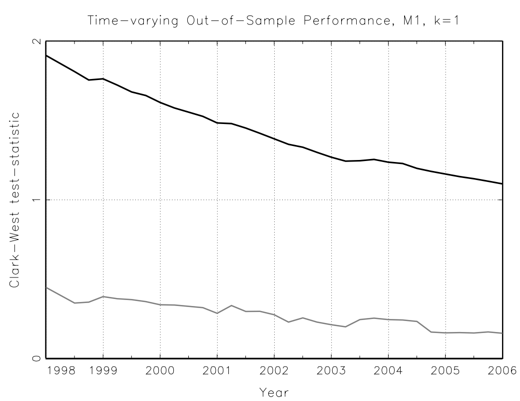
(b) Clark-West, k=4, Germany



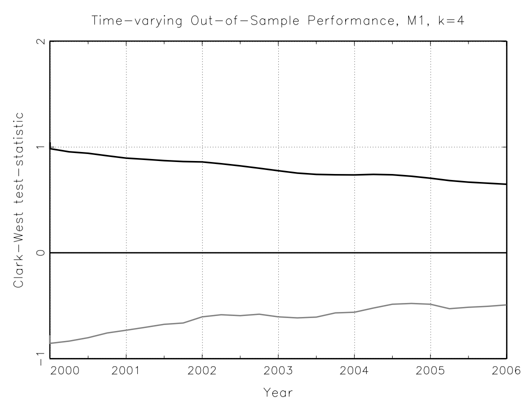
(c) Clark-West, k=1, USA



(d) Clark-West, k=4, USA



(e) Clark-West, k=1, UK



(f) Clark-West, k=4, UK

Table A.1
Predictive Performance of the Spread and other Information Variables for Real GDP Growth, Germany

Estimation is based on the following regression specification

$$y_{t+k}^{(k)} = \beta_0^{(k)} + \beta_1^{(k)} X_t + \theta^{(k)} Z_t + \epsilon_{t+k}^{(k)},$$

where X_t is a vector containing the term spread using maturity 5 (Panel A) and maturity 10 (Panel B). Z_t is a vector of additional control variables: Z_t contains the lagged dependent variable known as of t (coefficient $\theta_{1,k}$), the return on the aggregate stock market (coefficient $\theta_{2,k}$), the growth of the ifo business climate indicator (coefficient $\theta_{3,k}$), the inflation rate (coefficient $\theta_{4,k}$), changes in the oil price (coefficient $\theta_{5,k}$), and the money market rate (coefficient $\theta_{6,k}$). The dependent variable is defined as cumulative real GDP growth (annualized). The forecasting horizon is denoted by k. The sample period is 1972:Q4-2006:Q1. We report t-statistics based on two methods for calculating standard errors: $(\cdot)$ contains t-statistics based on Newey-West standard errors with $k-1$ lags, $\langle\cdot\rangle$ is based on the bootstrapped standard errors.

Horizon	Panel A: n=5								
	$\beta_{0,k}$	$\beta_{1,k}$	$\theta_{1,k}$	$\theta_{2,k}$	$\theta_{3,k}$	$\theta_{4,k}$	$\theta_{5,k}$	$\theta_{6,k}$	$\bar{R}^2$
k=1	1.401	0.598	-0.210	-0.019	0.391	-0.045	0.325	0.092	0.078
	(1.26)	(1.26)	(-2.15)	(-0.53)	(2.64)	(-0.21)	(0.15)	(0.40)	
k=4	1.492	0.470	0.080	0.028	0.109	0.021	-0.775	-0.049	0.233
	(1.79)	(1.99)	(0.66)	(1.26)	(1.54)	(0.30)	(-0.91)	(-0.40)	
	⟨1.53⟩	⟨1.72⟩	⟨0.57⟩	⟨1.17⟩	⟨1.40⟩	⟨0.29⟩	⟨-0.77⟩	⟨-0.34⟩	
k=8	1.005	0.549	0.042	0.025	0.014	-0.023	-0.187	0.074	0.204
	(1.16)	(2.88)	(0.28)	(1.47)	(0.23)	(-0.51)	(-0.34)	(0.69)	
	⟨1.01⟩	⟨2.48⟩	⟨0.25⟩	⟨1.39⟩	⟨0.22⟩	⟨-0.46⟩	⟨-0.26⟩	⟨0.57⟩	
k=12	1.180	0.516	-0.144	0.015	-0.014	-0.024	-0.332	0.125	0.208
	(1.42)	(2.86)	(-1.25)	(0.88)	(-0.27)	(-0.55)	(-0.57)	(1.19)	
	⟨1.10⟩	⟨2.29⟩	⟨-0.91⟩	⟨0.83⟩	⟨-0.25⟩	⟨-0.53⟩	⟨-0.45⟩	⟨0.87⟩	
	Panel B: n=10								
k=1	1.586	0.357	-0.203	-0.021	0.427	-0.027	0.160	0.057	0.068
	(1.14)	(0.87)	(-2.08)	(-0.58)	(2.95)	(-0.13)	(0.07)	(0.22)	
k=4	1.293	0.365	0.102	0.027	0.130	0.028	-0.873	-0.035	0.223
	(1.24)	(1.56)	(0.85)	(1.24)	(1.93)	(0.40)	(-0.98)	(-0.24)	
	⟨1.06⟩	⟨1.37⟩	⟨0.73⟩	⟨1.14⟩	⟨1.76⟩	⟨0.39⟩	⟨-0.84⟩	⟨-0.21⟩	
k=8	0.326	0.539	0.066	0.024	0.032	-0.029	-0.278	0.154	0.218
	(0.33)	(2.87)	(0.44)	(1.46)	(0.52)	(-0.62)	(-0.55)	(1.26)	
	⟨0.28⟩	⟨2.48⟩	⟨0.40⟩	⟨1.36⟩	⟨0.51⟩	⟨-0.57⟩	⟨-0.41⟩	⟨1.04⟩	
k=12	0.387	0.542	-0.117	0.014	-0.002	-0.034	-0.407	0.220	0.243
	(0.41)	(3.00)	(-1.05)	(0.86)	(-0.04)	(-0.75)	(-0.77)	(1.82)	
	⟨0.32⟩	⟨2.46⟩	⟨-0.80⟩	⟨0.78⟩	⟨-0.04⟩	⟨-0.72⟩	⟨-0.61⟩	⟨1.40⟩	

Table A.2
Estimation Results: Cochrane-Piazzesi and Fama-Bliss Regressions, Germany

Panel A reports results from (unrestricted) Cochrane-Piazzesi regressions

$$rx_{t+4}^{(n)} = \beta_0^{(n)} + \beta_1^{(n)} f_t^{(0,1)} + \beta_2^{(n)} f_t^{(2,3)} + \beta_3^{(n)} f_t^{(8,9)} + \epsilon_{t+4}^{(n)},$$

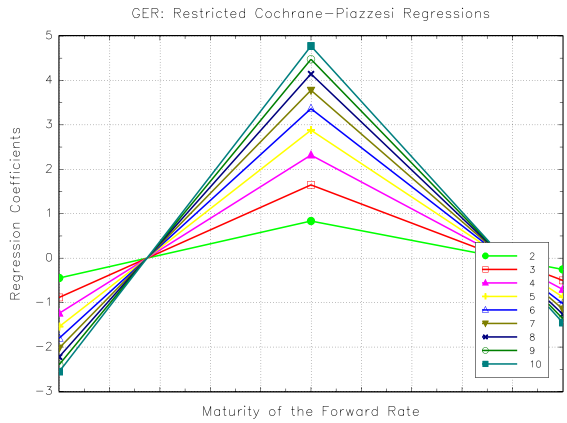
where $rx_{t+4}^{(n)}$ is defined as the holding period log return of a bond maturing in n years over the yield on a one year bond $[rx_{t+4}^{(n)} = p_{t+4}^{(n-1)} - p_t^{(n)} - i_t^{(1)}]$. $p_t^{(\tau)}$ is the log price of a bond maturing in τ periods. $f_t^{(\tau-1,\tau)}$ are forward rates implied by the yield curve: $f_t^{(\tau-1,\tau)} = p_t^{(\tau-1)} - p_t^{(\tau)}$. Panel B reports the results of Fama-Bliss regressions

$$rx_{t+4}^{(n)} = \beta_0^{(n)} + \beta_1^{(n)}(f_t^{(n-1,n)} - f_t^{(0,1)}) + \epsilon_{t+4}^{(n)},$$

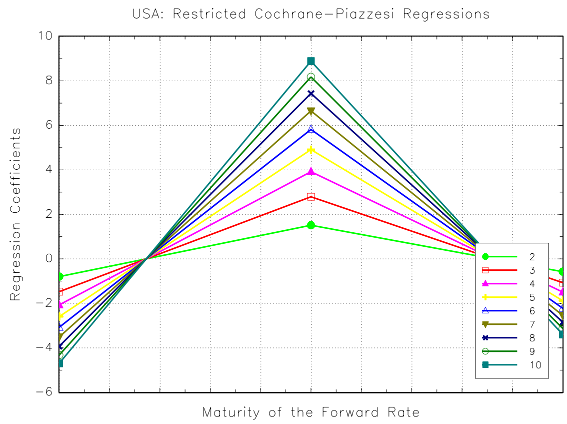
where $rx_{t+4}^{(n)}$ is defined as above. Again $f_t^{(\tau-1,\tau)}$ are forward rates implied by the yield curve. $\bar{R}^2$ denotes the adjusted R^2. The Cochrane/Piazzesi and the Fama/Bliss regressions are run for bond maturities $n = 2, ..., 10$ years. $(\cdot)$ contains t-statistics based on Newey-West standard errors with $k-1$ lags, $\langle\cdot\rangle$ is based on the bootstrapped standard errors (100,000 replications).

	Bond maturity								
	n=2	n=3	n=4	n=5	n=6	n=7	n=8	n=9	n=10
Panel A: Cochrane-Piazzesi regressions									
$\beta_0^{(n)}$	-1.114	-1.704	-1.949	-2.145	-2.513	-3.076	-3.923	-4.993	-6.314
	(-0.84)	(-0.71)	(-0.61)	(-0.55)	(-0.56)	(-0.61)	(-0.69)	(-0.79)	(-0.91)
	⟨-0.67⟩	⟨-0.57⟩	⟨-0.48⟩	⟨-0.44⟩	⟨-0.45⟩	⟨-0.48⟩	⟨-0.57⟩	⟨-0.65⟩	⟨-0.75⟩
$\beta_1^{(n)}$	-0.305	-0.678	-1.075	-1.459	-1.810	-2.113	-2.371	-2.586	-2.750
	(-1.64)	(-1.95)	(-2.21)	(-2.41)	(-2.55)	(-2.64)	(-2.68)	(-2.68)	(-2.64)
	⟨-1.75⟩	⟨-1.96⟩	⟨-2.21⟩	⟨-2.43⟩	⟨-2.60⟩	⟨-2.73⟩	⟨-2.68⟩	⟨-2.70⟩	⟨-2.68⟩
$\beta_2^{(n)}$	0.696	1.554	2.369	3.082	3.653	4.048	4.287	4.373	4.303
	(1.93)	(2.32)	(2.55)	(2.70)	(2.75)	(2.69)	(2.57)	(2.37)	(2.13)
	⟨1.71⟩	⟨2.02⟩	⟨2.27⟩	⟨2.47⟩	⟨2.61⟩	⟨2.63⟩	⟨2.48⟩	⟨2.34⟩	⟨2.13⟩
$\beta_3^{(n)}$	-0.173	-0.524	-0.886	-1.182	-1.358	-1.387	-1.269	-1.012	-0.615
	(-0.42)	(-0.71)	(-0.89)	(-0.99)	(-0.99)	(-0.90)	(-0.74)	(-0.53)	(-0.29)
	⟨-0.37⟩	⟨-0.61⟩	⟨-0.78⟩	⟨-0.87⟩	⟨-0.89⟩	⟨-0.82⟩	⟨-0.68⟩	⟨-0.50⟩	⟨-0.28⟩
R^2	0.102	0.121	0.132	0.139	0.144	0.145	0.145	0.144	0.141
Panel B: Fama-Bliss regressions									
$\beta_1^{(n)}$	0.321	0.566	0.735	0.856	0.935	1.013	1.088	1.154	1.212
	(0.82)	(1.26)	(1.44)	(1.50)	(1.49)	(1.49)	(1.51)	(1.52)	(1.54)
	⟨1.18⟩	⟨1.47⟩	⟨1.50⟩	⟨1.52⟩	⟨1.48⟩	⟨1.46⟩	⟨1.45⟩	⟨1.43⟩	⟨1.43⟩
R^2	0.014	0.036	0.048	0.053	0.053	0.053	0.053	0.052	0.050

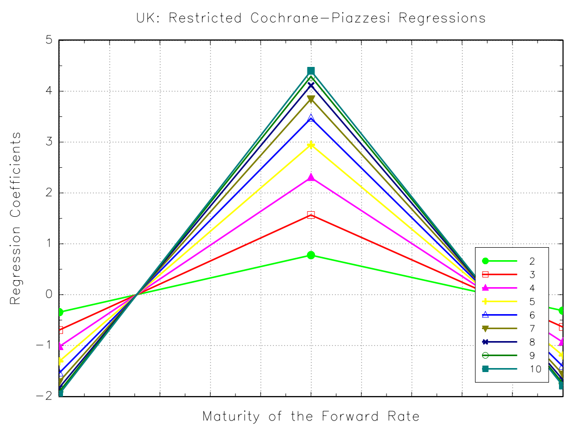
Figure A.1
Restricted Cochrane-Piazzesi Regressions, Tent-shape pattern



(a) Germany



(b) USA



(c) UK

Note: This figure presents regression coefficients of one year bond excess returns from restricted Cochrane-Piazzesi regressions. On the y-axis the coefficients of the restricted Cochrane-Piazzesi regressions are reported. The x-axis gives the different forward rates used as predictive variables. Our selection of forwards follows Tang and Xia (2005): $f_t^{(0,1)}$, $f_t^{(2,3)}$ and $f_t^{(8,9)}$.