GDP Growth Predictions through the Yield Spread. Time-Variation and Structural Breaks*

Pierangelo De Pace†

First Draft: May 2009 - This Version: February 2011‡

Abstract

We use TVP models and real-time data to describe the evolution of the leading properties of the yield spread for output growth in five European economies and in the US over the last decades and until the third quarter of 2010. We evaluate the predictive performance of benchmark term-structure models and identify structural breaks in the marginal processes of term spreads and government bond yields to shed light on the dynamic characteristics of the yield curve. Econometric analysis shows that: (i) the predictive content of the term spread is not always significant over time and across countries; (ii) the spread significantly contributes to the forecast performance of simple growth regressions in Europe, but not in the US in recent years; (iii) the variance of the random shocks to the term spreads tends to fall in all countries. This decline is accompanied by vanishing leading properties from the mid-1990s. Such properties reappear after 2008.

JEL Classification: *C22, C32, C53; E37; E43, E47.*

Keywords: *Real-Time Data, Term Spread, TVP Models, Structural Breaks.*

---

*I thank Jon Faust, Damiano Sandri, and Jonathan Wright for precious suggestions and comments, Kyongwook Choi and Massimo Guidolin for useful discussions, and Andrew T. Levin and Jeremy Piger for sharing their codes. I completed part of this work at the Division of International Finance of the Board of Governors of the Federal Reserve System, whose hospitality and support I gratefully acknowledge.

†Pomona College, Department of Economics; Carnegie Building, Room 205 - 425, N College Avenue - Claremont, CA 91711, USA. E-mail: pierangelo.depace@pomona.edu. Telephone #: +1 909 621 8744.

‡Full results and a Companion Tecnhical Appendix are available on request or downloadable from `http://sites.google.com/site/pierangelodepace/`.

1 Introduction

The slope of the term structure of interest rates is often cited as a useful leading economic indicator.[1] Conventional wisdom maintains that a negative slope is able to forecast business cycle downturns and recessions a few quarters ahead in the US and in other OECD countries.

The theoretical economic literature has proposed several explanations for the predictive power of the term spread – i.e., the difference between a long-term nominal interest rate and a short-term nominal rate.[2] Its practical relevance for policy decisions is controversial, though, and even recent empirical literature calls into question its usefulness for forecasting. Estrella, Rodrigues, and Schich (2003) use econometric techniques for break testing to study the stability properties of the relationship between the slope of the yield curve and subsequent real activity. They consider continuous models – to predict economic growth – and binary models – to predict recessions – for Germany and the US and document that the marginal predictive content of the spread for US output growth recently disappeared. Similar evidence is found in Dotsey (1998) for the United States. Giacomini and Rossi (2006) use new tests for forecast breakdown and a variety of in-sample and out-of-sample evaluation procedures to show the presence of structural breaks in the relationship between the slope of the yield curve and the US real output growth. They find forecast breakdowns during the Burns-Miller and Volcker monetary policy regimes and argue that the yield curve was a more reliable leading indicator during the early part of the Greenspan era.[3]

Some of these authors point out that the features of the relationship between the spread and economic activity may change following major economic shocks. In their attempt to explain the breakdowns, other researchers stress the role of globalization and the main central banks, which successfully achieved remarkable degrees of price stability, fostered sustained growth, and induced weaker and less-frequent shifts in the term spread for prolonged periods of time. Kucko and Chinn (2009) re-examine the evidence in the United States and some

[1] For example, see Stock and Watson (1989) and (1992).

[2] The term spread is also known as the yield spread, or the interest rate spread.

[3] Wright (2006) considers a number of probit models using the yield curve to forecast recessions in the US and argues that not only the level but also the shape of the yield curve should be used to gain useful information about the likely odds of a recession.

European countries. They find that the predictive power of the yield curve deteriorated over the years and claim that there are reasons to believe that European-country models perform better than non-European-country models with recent data. In a survey of the existing literature Wheelock and Wohar (2009) document that the term spread predicts output growth and recessions up to one year in advance, but its usefulness varies across countries and over time. However, while the ability of the spread to forecast economic growth diminished lately, the slope of the term structure has remained a reliable predictor of recessions.

The latest international economic events, including the recent US financial crisis and global recession, may have affected the predictive power of the yield curve and motivate a new cross-country analysis on its leading properties.[4] We adopt a systematic approach to estimate the time-variation in the predictive content of the term spread for future GDP growth, to assess its stability, and to examine its relative forecasting performance in six major OECD countries. The period of analysis is country-specific: generally, it is 1960-2010 for Germany, 1980-2010 for Spain and Italy, 1970-2010 for France, 1978-2010 for the UK, and 1964-2010 for the US. The contribution to the empirical literature is threefold. First, while most empirical studies either assume the relationship between future GDP growth and the interest rate spread to be constant or just focus on its stability properties by testing for structural breaks, we model and estimate its evolution through time-varying-parameter (TVP) models and real-time data, allowing for smooth transitions at each point in time. Second, using a real-time dataset, we study the out-of sample forecast performance of a set of simple, widely used, benchmark GDP growth regressions including the term spread as an explanatory variable and compare it with that of autoregressive models. To shed novel light on the dynamic characteristics of the yield curve, we estimate autoregressive models for long-term interest rates and yield spreads and test for breaks in the model parameters and innovation variance using a battery of state-of-the-art structural stability tests. Third, we

[4] In June of 2004 the Fed started tightening its policy. They raised the Federal Funds Target from 1% to 5.25% at seventeen consecutive meetings. Short rates followed the Target and moved in the same direction. However, long maturity rates fell. In a 2005 testimony at the Congress, Alan Greenspan defined the strange behavior of the spread between long and short rates a *conundrum*. This US-specific phenomenon further motivates the present piece of research.

document the reappearance of the predictive content of the spread following the events of 2007 and 2008 that led to the global economic downturn.

We derive the following results. (i) The term spread is not a reliable predictor of output growth. Its predictive content is significant in the early parts of our country-specific samples, then vanishes in later periods. Its leading properties, weak or non-existent between the mid-1990s and 2008, significantly reappear after then. (ii) The out-of-sample forecast accuracy of GDP growth regressions is time-varying. It generally improves over time until 2008, after which we observe a sharp, synchronized, deterioration in all countries. Benchmark term-structure models exhibit a better forecast accuracy than autoregressive models in Europe, unlike in the US in recent years. (iii) The variance of the random shocks to the term spreads falls in all countries, consistently with the facts of the Great Moderation. This decreasing variability is accompanied by weaker and declining leading properties until 2008.

2 The Term Spread as a Leading Economic Indicator

According to the *preferred habitat theory*, investors with heterogeneous investment horizons require a premium to buy bonds with maturities outside their preferred habitat. If short-term investors are prevalent in the fixed-income markets, long-term rates tend to be bigger than short-term rates and the yield curve *naturally* slopes upwards due to the term premia. Similar implications can be found in the *liquidity premium theory,* according to which there exists a term premium that increases with maturity.[5]

The most common explanation of why the term spread should predict output growth is related to countercyclical monetary policy. If the central bank lowers the policy interest rate, nominal and real long- and short-term rates tend to decline. Long-term rates tend to fall less than short-term rates because the monetary expansion raises long-term inflation expectations and the monetary authority is expected to switch to a contractionary stance in the future to respond to potential increases in inflation. The yield curve gets steeper and, since real

[5] In the *liquidity premium theory* the interest rate on a long-term bond equals an average of short-term interest rates expected to occur over the life of the long-term bond plus a premium that depends on the supply and demand conditions for that bond.

interest rates will remain low for a while, output growth is likely to be above average.[6]

Estrella (2005) formally derives the link between the spread and economic activity in a small dynamic rational expectations model containing a Phillips curve, a dynamic IS curve, the Fisher equation, the expectations hypothesis, and a monetary policy rule.[7] In this framework the positive link between the yield spread and expected future output is not structural but influenced by the monetary policy regime. It is stronger when the monetary policy response to output is small, weaker or nonexistent when the response is large. Changes in the leading properties should then mirror changes in the monetary policy stance.

The consumption capital asset pricing model (CCAPM) implies a positive relationship between the slope of the real yield curve and future real consumption growth. In real business cycle models, based on the same first-order condition as the CCAPM, expected positive productivity shocks increase future output. As agents substitute current for future consumption, future real interest rates go up and the real yield curve gets steeper.[8]

3 Predicting Cumulative GDP Growth Using the Term Spread

As customary in this strand of empirical literature, the focus is on a simple benchmark term-structure model for predicting cumulative GDP growth,[9]

$$g_{t,t+k} = \alpha + \beta s_t + \nu_t, \tag{1}$$

and on its two variants,

$$g_{t,t+k} = \alpha + \beta s_t + \gamma g_{t-k,t} + \nu_t \tag{2}$$

and

$$g_{t,t+k} = \alpha + \beta s_{t-1} + \gamma g_{t-k,t} + \nu_t, \tag{3}$$

[6] According to this story, the predictive content of the term spread is a correlation between endogenous variables, whose (co)movements are affected by monetary policy actions.

[7] The expectations hypothesis of the term structure states that the interest rate on a long-term bond equals an average of the short-term interest rates expected to occur over the life of the long-term bond.

[8] These models have implications for the real interest rates. The role of inflation expectations is then crucial, since the term structure is expressed in nominal terms.

[9] See also Estrella and Hardouvelis (1991) and Estrella and Mishkin (1997).

where $g_{t,t+k} = \frac{400}{k} ln\left(\frac{Y_{t+k}}{Y_t}\right)$ is cumulative growth between time t and $t+k$, Y_t is real GDP, $s_t = i_{10yr,t} - i_{3m,t}$ is the term spread, $i_{10yr,t}$ is an annualized ten-year government bond yield, and $i_{3m,t}$ is an annualized three-month money market or interbank rate.[10] The coefficient associated with s_t, β, and the R^2 of a model incorporate the basic information on the predictive content of the spread for output growth. With a positive β, an inversion of the term structure would predict a real downturn k (or $k+1$) quarters in advance. A high informativeness – i.e., a high R^2 – would empirically corroborate this intuition.

4 The Econometric Methodologies

Previous studies document that instability is a feature of the leading properties of the term spread.[11] Ignoring it may have negative consequences on inference and forecasting. Two are the main approaches to instability and change-point modeling: (i) a predominant strategy, based on the estimation of models with a small number of change-points, usually one or two; (ii) a more infrequent solution, based on the estimation of TVP models, where the parameters change with each new observation as random walks or stationary autoregressive processes.

We first assume that the model coefficients in (1) – (3) are time-varying. A compelling critique of this in-sample estimation approach is that the models are estimated using data that were not available at the time of the observation(s) being fitted. To circumvent this problem, we also propose a real-time analysis. Recursive OLS regressions on subsequent vintages of data describe the features of the *long-run* convergence of the coefficient estimates over the sample. Moving OLS regressions, based on a fixed-length moving window of ten years, capture the *short-run* variation and the stability characteristics of the leading properties. At a second stage, we run a battery of state-of-the art tests for breaks at unknown dates in the marginal processes of government bond yields and term spreads. We use classical and Bayesian tests for one or multiple breaks in the AR parameters and/or in the innovation variance of simple autoregressive models describing the time evolution of these variables.

[10] This expression for cumulative growth is appropriate with quarterly data. k varies between 1 and 4.
[11] For example, Benati and Goodhart (2007).

4.1 TVP Models for Cumulative GDP Growth

The starting point is models (1) – (3), for whose coefficients we assume specific time-varying properties. In the state-space specifications used for estimation, $g_{t,t+k}$ and s_t are the observable variables included in the measurement equations, the coefficients β and γ are the unobservable state variables, assumed to be time-varying and following the transition equations that incorporate the characteristics of their time evolution.[12] Such evolution may be the result of slow changes in the process or some form of nonlinearity in the data. TVP models change their parameters automatically and optimally to reflect the variations in the nature of the time series.[13] The specifications of our TVP models – all reported in Appendix B – are conventional and assume either random walks or stationary $AR(1)$ processes as state equations. We use an efficient algorithm, which allows for the optimal, robust, and unbiased estimation of dynamic regression models as discussed in Young et al. (2007).

4.2 Breakpoint Tests on Interest Rate Dynamics

We estimate univariate $AR(K)$ models for the term spread or government bond yield,

$$s_t = \mu + \Sigma_{i=1}^{K} \alpha_i s_{t-i} + \varepsilon_t, \tag{S1}$$

where ε_t is a serially uncorrelated random error term and μ is the intercept. We select the lag order, K, using the Schwarz Information Criterion (SIC). Then we estimate structural breaks at unknown dates in the model parameters and/or the innovation variance. The classical tests for breaks are based on Hansen (2000) and Qu and Perron (2007). Levin and Piger (2004) are the reference for the Bayesian comparison of alternative breaks models. Using Qu and Perron (2007), we also test for structural breaks in a system including equations (1) and ($S1$).

[12] A TVP model can be interpreted as a model with $T-1$ breaks in a sample of size T. With a small number of structural breaks, the magnitude of the change in the coefficients after a break is not typically restricted. The implicit assumption is that, after the last estimated break, there will be no more. In contrast, in TVP models, there is always a probability equal to one of a break in the next new observation. The size of the break is limited by the assumption that the coefficients evolve according to a specified stochastic process.

[13] The TVP methodology is robust to the uncertainty concerning the specific form of time-variation present in the data and is generally capable of successfully tracking processes subject to structural breaks.

5 Empirical Results

What follows is a description of the main findings. Detailed tables and select figures are commented here and reported at the end of the paper.

5.1 The Data

The sample includes six OECD countries: Germany, Spain, France, Italy (in the Euro area), the UK, and the US. We consider annualized ten-year government bond yields (long-term interest rates) and annualized three-month money market rates or interbank offer rates (short-term rates). The real GDP series are expressed in millions of national currency (volume estimates, OECD reference year). The data on real GDP and the interest rates are taken from the OECD database. The source of real-time data on GDP in Germany, Spain, France, and Italy is the OECD Real-Time Data and Revisions Database. The UK real-time data are downloaded from the Bank of England GDP Real-Time Database. The US real-time series are collected from the Philadelphia Fed's Real-Time Data Set for Macroeconomists (RTDSM). Full details on the samples and, in the case of real-time data, vintages are reported in Appendix A, where we also describe some minor issues in terms of missing observations. Unless noted otherwise, all series are quarterly and seasonally adjusted. The vintages and observations in the real-time dataset are also quarterly.

5.2 Benchmark OLS Estimates

Table 1 shows the OLS estimates of models $(1)-(3)$. Depending on the time horizon over which cumulative growth is computed, the adjusted samples range from 1980.1 (Spain and Italy), 1960.1 (Germany), 1970.1 (France), 1978.1 (UK), or 1964.3 (USA), to 2009.2-2010.2.[14]

What emerges is a mixed picture where conventional wisdom is confirmed only to some extent. In Germany, France, the UK, and the USA the slope coefficients associated with the term spread are significant and positive in all models and at all forecast horizons. The size of

[14] Adjacent growth figures are calculated from overlapping data points, which likely cause problems of serial correlation in the error terms of the models. Newey-West heteroskedasticity and serial correlation robust standard errors are used in the regression analysis.

the estimates is large, generally well above 0.5, and the corresponding levels of informativeness are usually high, with few exceptions at the shorter horizons. No significant predictive content can be found in Spain and Italy. The impression is that the relationship between the spread and economic growth is dissimilar across countries, or at least not consistently significant. However, the standard OLS approach is likely not to capture some important features of the data. More sophisticated techniques would allow us to better describe the stability properties of the model parameters and the time-variation in the relationship under investigation.

5.3 Time-Variation in GDP Growth Regressions

In this section we describe the time-varying properties of β in models (1) – (3). We estimate TVP models, then perform a real-time analysis and assess the out-of-sample forecast performance of the benchmark term-structure equations relative to autoregressive models.

5.3.1 TVP Models

Figures 1a-c show select time-varying estimates of β in model (1).[15] Alternative estimates from the other models, where we allow γ to either vary with time or stay constant, provide similar evidence.[16] $AR(1)$ or random-walk (RW) variation in the yield-spread coefficient is chosen in each case using the R^2 as a criterion for model selection.

In Germany the point estimates of β – generally positive for all values of k – slope downward between 1960 and 2002, then move upwards. The two-standard-error confidence bands cover zero almost always with $k = 3, 4$, except for the period following 2006, when β becomes statistically positive. With $k = 1, 2$ β is statistically positive between 1960 and 1985, then becomes insignificant. The $\widehat{\beta}$s exhibit more variation in the other European countries, but the associated confidence bands usually cover zero. A downward sloping term-spread coefficient is estimated in the US, significantly positive at the beginning of the sample (from 1965), statistically negligible at the end. The statistical disappearance of the US leading properties can be dated in the second half of the 1980s, at the end of the Volcker era and

[15] Systems 1.a-b in Appendix B.

[16] The intercept term, α, is kept constant in all models.

the period of high inflation post oil shocks.[17] The UK point estimates of β sharply increase in 2004/2005 and reach a significantly positive peak around 2008/2009. The US term-spread coefficient picks up a bit around 2009, too, but remains statistically insignificant.[18]

In all cases, the TVP models exhibit a better in-sample performance than their OLS counterparts, as indicated by the bigger coefficients of determination.

5.3.2 Real-Time Analysis

We recursively estimate models (1) $-$ (3) on the vintages of the real-time datasets. First, we run *recursive regressions* by estimating the models over the full samples of each vintage. The window size increases by one quarter at each step, as we switch from a vintage to the one that follows. In this way we capture the long-run evolution of β as GDP revisions are incorporated in the set of data.[19] Then, we run *moving regressions* with a fixed window size on the last forty quarters of each vintage. The attempt is to exclude remote information from the estimates and describe the short-run time-variation incorporated in the coefficients.[20]

Figures 2a-b show the moving regressions estimated on model (1). In France and Italy, the slope coefficient are stable and, most of the times, significantly positive for a few years after 1999. The statistical significance of the β parameters vanishes in these countries in 2004. The coefficients of determination, fairly high in the previous quarters, fall to almost zero simultaneously, remain low until 2007/2008, then rapidly increase and accompany a significantly positive variation in the βs during the global recession period. The German leading properties are non-significant until 2009, then quickly become statistically positive. In Spain they are statistically positive between 2001 and 2004 and then again from 2009. In the UK and, particularly, the US, we observe a steady decline of the predictive content over

[17] TVP models often produce large standard errors and confidence bands. Most likely, we fail to reject the null of statistical non-significance too often – i.e., we have low power. The spread might have had significant predictive content for longer periods in all countries and, in the US, the disappearance of the leading properties could be probably placed at a later date.

[18] Four of the countries in the sample out of six have been in the Euro area and have had a common monetary policy and similar interest rates since 1999. Their currencies were already closely tied from the mid-1990s. However, the TVP point estimates of β do not reveal the existence of similar evolutions in the last 15 years.

[19] The recursive estimates are not reported here but can be found online.

[20] Earlier work in this literature is only based on the most recent vintage of data.

time. The disappearance of the statistical significance of β can be dated in 2002 in the UK and in 1998 in the US. A slow fall of the US informativeness starts around 1984/1985. In both countries the predictive content swiftly picks up in 2008, as indicated by the increasing estimates of β and the corresponding R^2s. Vanishing, weak, or non-existent leading properties are estimated in all countries for most of the ten years between 1998 and 2008. A significant inversion of this trend occurs during the last financial crisis and world recession.

Policy makers are often interested in assessing the difference between the indications they obtain using the available information at the time of their decisions and the indications they would get *ex-post,* if they knew future information and how past data will be revised. This issue is relevant in a forecasting framework. An empirical investigation on the full real-time dataset should be conducted if the goal is to uncover the evolution, subject to error, of a forecasting relationship as new GDP figures get released and old vintages revised. As the last vintage of data is thought to be the series that measures the level of economic activity with least error, the final vintage can be used to verify and, possibly, compare economic relationships, also out of sample. The two approaches thus serve different purposes.

To highlight the discrepancies between the real-time analysis and a standard investigation on the latest data revision, recursive and moving regressions are estimated on the last vintage of each real-time dataset. Such estimates, also reported in Figures 2a-b, are pointwise different from their counterparts based on the full real-time dataset.[21] Using real-time data leads to a concrete risk of misestimating the predictive content of the spread. The difference between the $\widehat{\beta}$s (or R^2s) may occasionally get substantial, determine incorrect analyses, and lead to imprecise policy indications. There is no pattern in the sign of the divergence. However, since the respective confidence bands always overlap, such difference is statistically insignificant.

[21] GDP series are continuously revised, often significantly. Given that we cannot even measure GDP without errors, we cannot expect real-time forecasts of real economic activity to be precise.

Out-of-Sample Forecast Performance of GDP Growth Regressions

We benchmark (1) – (3) in terms of forecast performance against the autoregressive models

$$g_{t,t+k} = \alpha + \gamma g_{t-k,t} + \nu_t \tag{4}$$

and

$$g_{t,t+k} = \alpha + \sum_{j=0}^{3} \gamma_j g_{t-k-j,t-j} + \nu_t. \tag{5}$$

We dynamically and statically forecast the last eight quarters of each vintage in the real-time dataset.[22] The root mean squared forecast error ($RMSFE$) is our metric to compare the abilities of each model to predict growth. Figures 3a-b compare the evolutions of the $RMSFE$s, recursively estimated from the five models.[23]

The forecast errors in France and Italy are stable and approximately of the same size from 1999 to 2008, with a contemporaneous, although mild and temporary, deterioration between 2003 and 2004. The forecast accuracy improves a little in both countries between 2004 and 2008. The errors are more erratic in Germany between 2001 and 2009, but the average magnitude remains similar to that of France and Italy, with peaks in 2002, 2004, and 2008. The Spanish $RMSFE$s peak at the end of 2002 and then decline until 2008. The UK forecasts become less accurate around 1992 (the currency crisis the led the pound sterling out of the European Monetary System), 1997, and 2001. In the US, the accuracy deteriorates in the second half of the 1970s and for a few years at the beginning of the 1980s and 1990s. The first two US deteriorations look similar to those documented by Giacomini and Rossi (2006) for the Burns-Miller and Volcker periods. A smaller deterioration occurs in 2001 under Greenspan. The breakdowns in the US forecast accuracy are coincident with recessions. In all countries, the $RMSFE$s become smaller over time until 2008, after which we observe a

[22] Dynamic forecasting performs a multi-step ahead forecast of the dependent variable. It requires that the data for the exogenous variables be available for every observation in the forecast sample and that the values for any lagged dependent variables be observed at the start of the forecast sample. Static forecasting performs a series of one-step ahead forecasts of the dependent variable. It requires that the data for both the exogenous and any lagged endogenous variables be observed for every observation in the forecast sample.

[23] $k = 4$ in the figures, but we find similar patterns with $k = 1$, 2, and 3 and a forecast sample of one quarter.

sharp, synchronized, worsening of the forecast accuracy of all models.

To statistically compare the term-structure models to models (4) and (5) in terms of out-of-sample forecast performance, we run modified Diebold-Mariano (DM) tests for equality of forecast accuracy on the static forecasts.[24] We would like to test the merits of all the models, but a limitation of this test is that it can be only applied to pairs of non-nested models. Unfortunately, some pairs contain models that are nested.[25] The test statistic of Diebold and Mariano (1995) has a nonstandard distribution under the null hypothesis of equal forecast accuracy if the models are nested, as the models are identical under the null.[26] Thus, not even bootstrap p-values would allow us to make such a comparison.[27]

Tables 2a-b report modified DM statistics for non-nested models with forecast samples of one quarter and eight quarters, respectively, on each vintage.[28] We run the tests on the full samples and on country-specific subsamples.[29] With a forecast sample of one quarter, the term-structure models perform as well as the alternative models, occasionally better, in all European countries. In the US, the autoregressive models perform better over the second subsample, worse in the first subsample, but we never reject the null of equal accuracy in the full sample. With a forecast sample of eight quarters, the term-structure models perform better at forecasting GDP growth than the autoregressive models in all European countries in their specific first subsamples and full samples. In the US, the term-structure models do a better job in the first period, just a marginally better job over the full sample, but are

[24] Harvey, Leybourne, and Newbold (1997) propose a modified DM statistic based on an unbiased estimator of the asymptotic (long-run) variance of $\sqrt{T}\overline{d}$ in the DM statistic, where T is the sample size and $\overline{d}$ is the sample average of the loss differential (in this work the difference of the $RMSFE$s) the test is based on. They show that, with small samples, a Student's t distribution is more appropriate than a standard normal distribution for the computation of the critical values.

[25] Model (1) is nested in (2); model (4) is nested in (2), (3), and (5).

[26] Clark and McCracken (2001).

[27] Faust and Wright (2009).

[28] A complication of the DM test (and its modified version) regards the estimation of the asymptotic variance of $\sqrt{T}\overline{d}$. The standard practice is to estimate this variance by taking a weighted sum of the available sample autocovariances. Optimal k-step ahead forecast errors are at most $(k-1)$-dependent – i.e., autocorrelated up to the $(k-1)$-th order. $(k-1)$-dependence implies that only $(k-1)$ sample autocovariances should be used. Since the forecast horizon of our models is 1, 2, 3, or 4 quarters, we use the sample variance and autocovariances up to the third order. In the event that a negative estimate arises, we treat it as zero and automatically reject the null hypothesis of equal forecast accuracy. See Diebold and Mariano (1995) for details.

[29] The breaks for the determination of the subsamples are estimated in the middle 70% of each full sample using a recursive algorithm that maximizes the absolute average difference of the average $RMSFE$s of the five models over two subsequent subsamples, for each value of k.

outperformed by the autoregressive models in the second. At least in recent years, the US results are consistent with the conclusions in Faust and Wright (2009). Using a new dataset of vintage data consisting of a large number of variables, as observed at the time of each Greenbook forecast since 1979, they show that a univariate $AR(4)$ model better forecasts US GDP growth than alternative specifications including other explanatory variables. Such a finding is not externally valid for the other countries in the sample, for which the term spread provides a significant contribution to the forecast performance.

5.4 Structural Breaks Evidence for Term Spreads and Bond Yields

Tables 3a-b show breaks in the parameters of $(S1)$ based on Hansen (2000).[30] Tables 4a-b and 5a-b describe the outcomes of the Qu and Perron (2007) tests on $(S1)$ and on a system of two equations including (1) and $(S1)$.[31] Using Bayesian methods, Table 6 compares the marginal log-likelihoods of model $(S1)$ in each country with two of its one-break versions.[32]

Despite some heterogeneity in the estimated breaks, most shifts in the government bond yields occur at the following dates: 1970/1971 and 1994.4 (Germany); 1995.1 and 2005 (Spain); 1978.1 and 1994.4 (France), 1996/1997 and end of 2000 (Italy); 1994.2 and 2004/2005 (UK); 1979.3 and 1991.1 (USA). In the case of the interest rate spreads – which is what we should pay more attention to, since it is the variable we employ to predict output growth – the breaks are mainly clustered around: 1969.2 and 1981.3 (Germany); 1994 and 2005 (Spain); 1978 and beginning of 1995 (France), 2000 and 2005 (Italy); 1993/1994 and 2005 (UK); end of 1971, end of 1982, and 1984.4 (USA).[33]

The breaks are similar within each country, independently of whether we assume shifts in the innovation variance, in the coefficients, or both. Consistently with the facts of the Great Moderation, we observe a decline in the volatility of reduced-form random shocks in

[30] The unreported p-values are derived as in Andrews (1993) and Hansen (2000). We also compute heteroskedasticity-robust bootstrap p-values based on $100,000$ bootstrap replications. In this exercise, Andrews (1993)'s asymptotic critical values provide similar inference as the bootstrap.

[31] We test the null of no breaks against the alternatives of one break and, when appropriate, two breaks.

[32] Spreads and bond yields are better fitted by models with a break in the innovation variance than by models with a break in both the innovation variance and the coefficients. See Appendix C for further details.

[33] Unreported breakpoint Chow tests, used for further validation, signal that most of these shifts (taken as exogenous) are significant at conventional levels.

the term-spread marginal processes. The ratios between innovation variances and variances of the spreads exhibit a similar evolution. This drop is accompanied by a generally weaker link between the term spread and the real growth rate from the mid-1990s to 2008.[34]

6 Conclusions

In this paper we estimate the time-variation in the predictive content of the term spread for future GDP growth in six major OECD countries and study the forecasting properties of a set of simple benchmark GDP growth regressions that include the term spread as an explanatory variable. To shed light on the dynamic characteristics of the yield curve, we estimate autoregressive models for long-term interest rates and yield spreads and test for breaks in the model parameters and innovation variance with a battery of state-of-the-art structural stability tests. Our investigation is based on an extensive use of TVP models, real-time datasets, classical and Bayesian tests for structural breaks at unknown dates.

We argue that the spread is not a reliable predictor of output growth. To some extent, its predictive content is statistically and economically significant in the early parts of our country-specific samples, especially in the US and UK. It vanishes in later periods, but reappears in all countries after 2008, during the global downturn. Such leading properties are characterized by time-variation and instability. The real-time analysis shows that the out-of-sample forecast accuracy of simple benchmark GDP growth models is markedly time-varying, but improves over time until 2008, then deteriorates with the financial crisis and world recession. The benchmark term-structure models exhibit a better forecast accuracy than the alternative atheoretical autoregressive models in Europe, but the term spread does not significantly contribute to forecasting growth in the US in recent years. Finally, the structural breaks evidence indicates that the variance of the random shocks to the spreads is declining, consistently with the facts of the Great Moderation. This decreased variability is accompanied by weaker leading properties for approximately ten years until 2008.

[34] This pattern of decreased volatility of the random shocks is less clear, or at least not as pronounced, in the case of government bond yields.

7 References

[1] Benati, Luca and Charles Goodhart (2007). "Investigating Time-Variation in the Marginal Predictive Power of the Yield Spread," *ECB Working Paper Series* No. 802.

[2] Bryson, A.E. and Y.C. Ho (1969). "Applied Optimal Control, Optimization, Estimation and Control," Waltham: Blaisdell Publishing Company.

[3] Chib, Siddhartha (1998). "Estimation and Comparison of Multiple Change-Point Models," *Journal of Econometrics*, Vol. 86; 221-241.

[4] Clark, Todd E. and Michael W. McCracken (2001). "Tests of Equal Forecast Accuracy and Encompassing for Nested Models," *Journal of Econometrics*, Vol. 105; 85-110.

[5] Diebold, Francis X. and Roberto S. Mariano (1995). "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, Vol. 13; 253-265.

[6] Dotsey, Michael (1998). "The Predictive Content of the Interest Rate Term Spread for Future Economic Growth," Federal Reserve Bank of Richmond *Economic Quarterly*.

[7] Estrella, Arturo (2005). "Why Does the Yield Curve Predict Output and Inflation?," *Economic Journal*, Royal Economic Society, Vol. 115 No. 505; 722-744.

[8] Estrella, Arturo and Gikas A. Hardouvelis (1991). "The Term Structure as a Predictor of Real Economic Activity," *The Journal of Finance*, Vol. 46 No. 2; 555-576.

[9] Estrella, Arturo and Frederic Stanley Mishkin (1997). "The Predictive Power of the Term Structure of Interest Rates in Europe and the United States: Implications for the European Central Bank," *European Economic Review*, Vol. 41; 1375-1401.

[10] 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," *The Review of Economics and Statistics*, Vol. 85 No. 3; 629-644.

[11] Faust, Jon and Jonathan H. Wright (2009). "Comparing Greenbook and Reduced Form Forecasts Using a Large Real Time Dataset," *Journal of Business and Economic Statistics*, Vol. 27 No. 4; 468-479.

[12] Giacomini, Raffaella and Barbara Rossi (2006). "How Stable is the Forecasting Performance of the Yield Curve for Output Growth?," *Oxford Bulletin of Economics and Statistics*, Department of Economics, University of Oxford, Vol. 68 Issue 1; 783-795.

[13] Hansen, Bruce E. (2000). "Testing for Structural Change in Conditional Models," *Journal of Econometrics*, Vol. 97; 93-115.

[14] Harvey, David; Stephen Leybourne, and Paul Newbold (1997). "Testing the Equality of Prediction Mean Squared Errors," *International Journal of Forecasting*, Vol. 13; 281-291.

[15] Jeffreys, H. (1961). *Theory of Probability*, Third Edition, Oxford University Press.

[16] Kass, Robert E. and Adrian E. Raftery (1995). "Bayes Factor," *Journal of the American Statistical Association*, Vol. 90 No. 430; 773-795.

[17] Kim, Chang-Jin and Charles R. Nelson (1999). *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, Cambridge and London, MIT Press.

[18] Kucko, Kavan and Menzie D. Chinn (2009). "The Predictive Power of the Yield Curve across Countries and Time," manuscript.

[19] Levin, Andrew T. and Jeremy Piger (2003). "Is Inflation Persistence Intrinsic in Industrial Economies?," *Computing in Economics and Finance 2003* 298, Society for Computational Economics.

[20] Qu, Zhongjun and Pierre Perron (2007). "Estimating and Testing Structural Changes in Multivariate Regressions," *Econometrica*, Vol. 75 No. 2; 459-502.

[21] Stock, James H. and Mark W. Watson (1989). "New Indexes of Coincident and Leading Economic Indicators," *NBER Macroeconomics Annual*, Vol. 4; 351-394.

[22] ————— (1992). "A Procedure for Predicting Recessions With Leading Indicators: Econometric Issues and Recent Experience," *NBER Working Papers* 4014.

[23] Wheelock, David C. and Mark E. Wohar (2009). "Can the Term Spread Predict Output Growth and Recessions? A Survey of the Literature," Federal Reserve Bank of St. Louis *Review*, Vol. 91 No. 5, Part 1; 419-440.

[24] Wright, Jonathan (2006). "The Yield Curve and Predicting Recessions," *Finance and Economics Discussion Series*, Divisions of Research and Statistics and Monetary Affairs, Federal Reserve Board.

[25] Young, P.C.; C.J. Taylor, W. Tych, and D.J. Pedregal. (2007). The Captain Toolbox. Centre for Research on Environmental Systems and Statistics, Lancaster University, UK. Internet: `http://www.es.lancs.ac.uk/cres/captain`.

8 Technical Appendix

We provide details on the dataset and on the estimation of TVP and Bayesian models. Further information is given in the Companion Tecnhical Appendix.

Appendix A. Data Description

The first table describes the samples for each variable in each country. The second table provides information on the GDP real-time dataset. Data are quarterly, as well as vintages and observations in the real-time dataset. Interest rates are never revised.

	Short-Term Interest Rate	Long-Term Interest Rate	Term Spread	Real GDP
Germany	1960.1-2010.3	1956.3-2010.3	1960.1-2010.3	1960.1-2010.3
Spain	1977.1-2010.3	1980.1-2010.3	1980.1-2010.3	1960.1-2010.3
France	1970.1-2010.3	1960.1-2010.3	1970.1-2010.3	1949.1-2010.2
Italy	1978.4-2010.3	1980.1-2010.3	1980.1-2010.3	1960.1-2010.3
UK	1978.1-2010.3	1960.1-2010.3	1978.1-2010.3	1955.1-2010.3
USA	1964.3-2010.3	1953.2-2010.3	1964.3-2010.3	1947.1-2010.3

Real-Time Data for Real GDP

	Vintages	Observations
Germany	1999.1-2010.3	1991.1-2010.2
Spain	1999.1-2010.3	1980.1-2010.2
France	1999.1-2010.3	1960.1-2010.2
Italy	1999.1-2010.3	1970.1-2010.2
UK	1990.1-2010.2	1970.1-2010.1
USA	1965.4-2010.3	1947.1-2010.2

Note the following. In the case of Spain: vintage 2005.2 starts in 2000.1 and the vintages from 2005.3 to 2010.3 start in 1995.1. In the case of France: the vintages from 1999.4 to 2009.2 start in 1978.1. In the case of Italy: the vintages from 2000.1 to 2001.1 start in 1982.1; the vintages from 2003.3 to 2004.3 start in 1980.1; and the vintages from 2006.2 to 2010.3 start in 1981.1. In the case of the USA: the vintages from 1992.1 to 1992.4 and from 1999.4 to 2000.1 start in 1959.1; the vintages from 1996.1 to 1997.1 start in 1959.3. All these missing observations might cause some minor imperfections in the recursive and moving OLS estimates, which are usually solved by adjusting the samples or dropping some vintages.

Appendix B. TVP Models

The table on the next page summarizes the TVP models we estimate to document the time-varying properties of the predictive content of the term spread under the assumptions that:

- ν_t, υ_t, and ξ_t are normally distributed, with zero mean and constant variances;
- $Cov(\nu_t, \upsilon_t) = Cov(\nu_t, \xi_t) = 0$;
- the initial stochastic states, β_0 and γ_0, are independent of ν_t, υ_t, and ξ_t for every t;
- the variances of ν_t, υ_t, and ξ_t, the covariance between υ_t and ξ_t, the system parameters α, δ, and ϑ are estimated through maximum likelihood prior to the application of the recursive algorithm that provides estimates of the states;
- the initial conditions for the states and their covariance matrix are unknown.

Results are obtained using the CAPTAIN Toolbox for MATLAB, which implements an efficient algorithm that allows for the optimal, robust, and unbiased estimation of dynamic regression models.[35] This formulation of the estimation problem allows the recursive algorithms, which estimate the state vector of time-varying parameters from measured data, to provide an optimal solution based on the minimization of the associated mean squared errors. State variables are estimated sequentially by the Kalman Filter whilst working through the data in temporal order. When all the time series data are available for analysis, this filtering operation is accompanied by optimal recursive smoothing. The estimates obtained from the forward pass filtering algorithm are updated sequentially whilst working through the data in reverse temporal order using a backwards-recursive Fixed Interval Smoothing (FIS) algorithm.[36] The noise-to-variance ratio – that is, the ratio between the variance/covariance matrix of υ_t and ξ_t and the variance of the error term in the measurement equation, ν_t – is

[35] See Young et al. (2007) for detailed information.

[36] Bryson and Ho (1969).

estimated by maximum likelihood based on the prediction error decomposition.

TVP Models

System 1.a	System 2.a	System 3.a
$g_{t,t+k} = \alpha + \beta_t s_t + \nu_t$	$g_{t,t+k} = \alpha + \beta_t s_t + \gamma_t g_{t-k,t} + \nu_t$	$g_{t,t+k} = \alpha + \beta_t s_{t-1} + \gamma_t g_{t-k,t} + \nu_t$
$\beta_t = \beta_{t-1} + \upsilon_t$	$\beta_t = \beta_{t-1} + \upsilon_t$	$\beta_t = \beta_{t-1} + \upsilon_t$
	$\gamma_t = \gamma$	$\gamma_t = \gamma$
System 1.b	System 2.b	System 3.b
$g_{t,t+k} = \alpha + \beta_t s_t + \nu_t$	$g_{t,t+k} = \alpha + \beta_t s_t + \gamma_t g_{t-k,t} + \nu_t$	$g_{t,t+k} = \alpha + \beta_t s_{t-1} + \gamma_t g_{t-k,t} + \nu_t$
$\beta_t = \delta\beta_{t-1} + \upsilon_t$	$\beta_t = \delta\beta_{t-1} + \upsilon_t$	$\beta_t = \delta\beta_{t-1} + \upsilon_t$
	$\gamma_t = \gamma$	$\gamma_t = \gamma$
	System 2.c	System 3.c
	$g_{t,t+k} = \alpha + \beta_t s_t + \gamma_t g_{t-k,t} + \nu_t$	$g_{t,t+k} = \alpha + \beta_t s_{t-1} + \gamma_t g_{t-k,t} + \nu_t$
	$\beta_t = \beta_{t-1} + \upsilon_t$	$\beta_t = \beta_{t-1} + \upsilon_t$
	$\gamma_t = \gamma_{t-1} + \xi_t$	$\gamma_t = \gamma_{t-1} + \xi_t$
	System 2.d	System 3.d
	$g_{t,t+k} = \alpha + \beta_t s_t + \gamma_t g_{t-k,t} + \nu_t$	$g_{t,t+k} = \alpha + \beta_t s_{t-1} + \gamma_t g_{t-k,t} + \nu_t$
	$\beta_t = \delta\beta_{t-1} + \upsilon_t$	$\beta_t = \delta\beta_{t-1} + \upsilon_t$
	$\gamma_t = \vartheta\gamma_{t-1} + \xi_t$	$\gamma_t = \vartheta\gamma_{t-1} + \xi_t$

Note: k=1, 2, 3, 4.

In this work we only report the smoothed estimates of β in either System 1.a or 1.b. All the other estimates can be found online.

Appendix C. Bayesian Comparison of Breaks Model

We make use of simple Bayesian methods to compare the likelihoods of alternative models with breaks or no breaks. We estimate the model

$$s_t = \mu + \rho s_{t-1} + \Sigma_{i=1}^{K-1} \phi_i \Delta s_{t-i} + \varepsilon_t, \tag{S1.a}$$

which is equivalent to equation $(S1)$, where $\rho = \Sigma_{i=1}^{K} \alpha_i$ is a persistence parameter and the ϕ_is are transformations of the AR coefficients, α_i. The error term is normally distributed with zero mean and variance σ_t^2. In a model without breaks, σ_t^2 is thought to be constant – i.e., $\sigma_t^2 = \sigma^2\ \forall t$. Alternatively, we model the variance of ε_t by allowing for the presence of a one-time structural shift, so that $\sigma_t^2 = \sigma_0^2 (1 - D_t) + \sigma_1^2 D_t$, where D_t is a dummy variable that controls for the shift. We compute the marginal likelihoods of the models as in Chib (1998) and assume that D_t is a discrete latent variable with Markov-transition probabilities $Prob\,(D_{t+1} = 0 | D_t = 0) = q$ and $Prob\,(D_{t+1} = 1 | D_t = 1) = 1$, with $q \in (0, 1)$. The implication is that there is a constant positive probability, $(1 - q)$, for a break to occur in any period, if it has not occurred yet. Once the break has occurred at a specific date t_0, then $D_t = 1$, $\forall t \geq t_0$ (*absorbing state*).[37] We estimate the breakpoint date with the posterior mean of the posterior distribution of q. For the estimation of the model without breaks, we assume that $\mu | \sigma^2 \sim N\left(0, 3\sigma^2\right)$, $\rho | \sigma^2 \sim N\left(1, 3\sigma^2\right)$, $\phi_i | \sigma^2 \sim N\left(0, 3\sigma^2\right)\ \forall i$, and $\sigma^2 \sim InvGamma\,(1, 2)$. In the model with one break in the innovation variance, $\mu \sim N\,(0, 3)$, $\rho \sim N\,(1, 3)$, $\phi_i \sim N\,(0, 3)$ $\forall i$, $\sigma_{0,1}^2 \sim InvGamma\,(1, 2)$, and $q \sim Beta\,(8, 0.05)$. We impose that μ, ρ, and the ϕ_is are statistically independent of each other.

The relatively informative priors are a compromise between the need of *letting the data speak* and the necessity of incorporating the *a-priori* information coming from an informal inspection of the data.[38] The distributional structure imposed to the model without breaks

[37] For the technical details about how to estimate Markov-Switching models in a Bayesian setting through Gibbs sampling, see Chapter 9 in Kim and Nelson (1999).

[38] First partial autocorrelations of ten-year government bond yields are usually close to one; they are smaller for interest rate spread series. Higher-order partial autocorrelations are generally close to zero. The standard Beta distribution ensures that the domain of the probability measure q is over the interval $[0, 1]$. The chosen parameters imply that much of the mass of the distribution is spread around values close to one. This

assigns priors for μ, ρ, and ϕ_i that are elicited conditional on σ^2. This makes the linear model fit the *Normal-Gamma* framework and the computation of many relevant quantities analytically feasible.[39] For each model we choose the lag order, $1 \leq K \leq 4$, that maximizes the marginal likelihood. The equations are estimated through the Gibbs sampler, a Markov Chain Monte Carlo ($MCMC$) technique that computes marginal posterior distributions for the parameters through the likelihood function of the model and by means of complex numerical methods that simulate draws from the joint posterior.

Following a similar approach, we estimate models where the parameters are allowed to break at the same date as the error variance,

$$s_t = \mu_0 + D_t\mu_1 + (\rho_0 + D_t\rho_1)\, s_{t-1} + \Sigma_{i=1}^{K-1} (\phi_{0i} + D_t\phi_{1i})\, \Delta s_{t-i} + \varepsilon_t, \tag{S1.b}$$

with $\mu_{0,1} \sim N(0,3)$, $\rho_{0,1} \sim N(1,3)$, $\phi_{0,1;i} \sim N(0,3)\ \forall i$, $Var(\varepsilon_t) = \sigma_t^2 = \sigma_0^2 (1 - D_t) + \sigma_1^2 D_t$, $\sigma_{0,1}^2 \sim InvGamma(1,2)$, and $q \sim Beta(8, 0.05)$.

All the variables are assumed to be independent of each other.[40]

specification gives more prior probability to late breakpoint dates in the sample. Different calibrations for the prior of q do not alter much the estimated changepoints.

[39] The Normal-Gamma framework is a particular case of a two-level hierarchical Bayesian model, in which a conjugate prior distribution is specified at the first stage and a non-informative or weakly informative prior is generally assumed at the second stage.

[40] The likelihood of a model with respect to another can be assessed by comparing the corresponding Bayes factors and following the *rules of thumb* in Jeffreys (1961) and Kass and Raftery (1995).

9 Tables and Figures

		Model 1				Model 2				Model 3			
		k=1	*k=2*	*k=3*	*k=4*	*k=1*	*k=2*	*k=3*	*k=4*	*k=1*	*k=2*	*k=3*	*k=4*
Germany	beta	**0.682*****	**0.681*****	**0.649*****	**0.659*****	**0.652*****	**0.655*****	**0.619*****	**0.650*****	**0.657*****	**0.634*****	**0.597*****	**0.602*****
		(0.210)	(0.193)	(0.173)	(0.158)	(0.200)	(0.188)	(0.162)	(0.146)	(0.184)	(0.174)	(0.154)	(0.145)
	R Squared	0.055	0.101	0.136	0.180	0.057	0.102	0.173	0.236	0.056	0.091	0.159	0.206
	Sample	1960.1-2010.2	1960.1-2010.1	1960.1-2009.4	1960.1-2009.3	1960.2-2010.2	1960.3-2010.1	1960.4-2009.4	1961.1-2009.3	1960.2-2010.2	1960.3-2010.1	1960.4-2009.4	1961.1-2009.3
Spain	beta	0.066	0.086	0.142	0.177	0.072	0.121	0.167	0.160	0.069	0.128	0.170	0.178
		(0.268)	(0.246)	(0.225)	(0.214)	(0.239)	(0.142)	(0.153)	(0.162)	(0.204)	(0.106)	(0.120)	(0.130)
	R Squared	0.001	0.004	0.011	0.018	0.016	0.368	0.339	0.288	0.016	0.371	0.338	0.289
	Sample	1980.1-2010.2	1980.1-2010.1	1980.1-2009.4	1980.1-2009.3	1980.1-2010.2	1980.1-2010.1	1980.1-2009.4	1980.1-2009.3	1980.2-2010.2	1980.2-2010.1	1980.2-2009.4	1980.2-2009.3
France	beta	**0.515****	**0.603*****	**0.647*****	**0.662*****	**0.387*****	**0.527*****	**0.645*****	**0.684*****	**0.464*****	**0.508*****	**0.604*****	**0.604*****
		(0.204)	(0.188)	(0.175)	(0.164)	(0.146)	(0.173)	(0.188)	(0.189)	(0.147)	(0.176)	(0.190)	(0.187)
	R Squared	0.082	0.149	0.195	0.227	0.312	0.448	0.392	0.363	0.319	0.426	0.354	0.300
	Sample	1970.1-2010.1	1970.1-2009.4	1970.1-2009.3	1970.1-2009.2	1970.1-2010.1	1970.1-2009.4	1970.1-2009.3	1970.1-2009.2	1970.2-2010.1	1970.2-2009.4	1970.2-2009.3	1970.2-2009.2
Italy	beta	-0.085	-0.066	-0.059	-0.075	-0.025	0.070	0.113	0.074	-0.015	0.008	0.015	-0.045
		(0.183)	(0.185)	(0.185)	(0.188)	(0.148)	(0.184)	(0.209)	(0.225)	(0.128)	(0.143)	(0.171)	(0.182)
	R Squared	0.002	0.002	0.001	0.003	0.217	0.193	0.125	0.069	0.223	0.192	0.120	0.070
	Sample	1980.1-2010.2	1980.1-2010.1	1980.1-2009.4	1980.1-2009.3	1980.1-2010.2	1980.1-2010.1	1980.1-2009.4	1980.1-2009.3	1980.2-2010.2	1980.2-2010.1	1980.2-2009.4	1980.2-2009.3
UK	beta	**0.550*****	**0.584*****	**0.600*****	**0.603*****	**0.484****	**0.460*****	**0.516*****	**0.553*****	**0.529****	**0.412*****	**0.466*****	**0.507*****
		(0.187)	(0.176)	(0.168)	(0.159)	(0.196)	(0.118)	(0.136)	(0.143)	(0.210)	(0.122)	(0.143)	(0.150)
	R Squared	0.079	0.141	0.169	0.185	0.109	0.376	0.343	0.299	0.116	0.351	0.311	0.263
	Sample	1978.1-2010.2	1978.1-2010.1	1978.1-2009.4	1978.1-2009.3	1978.1-2010.2	1978.1-2010.1	1978.1-2009.4	1978.1-2009.3	1978.2-2010.2	1978.2-2010.1	1978.2-2009.4	1978.2-2009.3
USA	beta	**0.680*****	**0.732*****	**0.708*****	**0.683*****	**0.571*****	**0.653*****	**0.679*****	**0.675*****	**0.603*****	**0.562*****	**0.600*****	**0.616*****
		(0.164)	(0.158)	(0.143)	(0.132)	(0.143)	(0.137)	(0.131)	(0.128)	(0.186)	(0.150)	(0.150)	(0.148)
	R Squared	0.100	0.173	0.197	0.216	0.178	0.265	0.256	0.244	0.182	0.220	0.204	0.198
	Sample	1964.3-2010.2	1964.3-2010.1	1964.3-2009.4	1964.3-2009.3	1964.3-2010.2	1964.3-2010.1	1964.3-2009.4	1964.3-2009.3	1964.4-2010.2	1964.4-2010.1	1964.4-2009.4	1964.4-2009.3

Model 1: *one exogenous regressor (interest rate spread)* - growth(t,t+k)=alpha+beta*Spread(t)+v(t);
Model 2: one pre-determined regressor and one exogenous regressor (interest rate spread) - growth(t,t+k)=alpha+beta*growth(t-k,t)+gamma*Spread(t)+v(t) ;
Model 3: *one pre-determined regressor and one exogenous regressor (lagged interest rate spread)* - growth(t,t+k)=alpha+beta*growth(t-k,t)+gamma*Spread(t-1)+v(t);
Note 1: figures in bold are statistically significant at least at the 10% level. One asterisk indicates statistical significance at the 10% level, two asterisks statistical significance at the 5% level, three asterisks statistical significance at the 1% level.
Note 2: growth(t,t+k)=(400/k)*(log(Y(t+k))-log(Y(t))). Data are quarterly.

Table 1. Growth Regressions, Benchmark OLS Estimates

Compared Models	k=1 1st Sample	k=1 2nd Sample	k=1 Full Sample	k=2 1st Sample	k=2 2nd Sample	k=2 Full Sample	k=3 1st Sample	k=3 2nd Sample	k=3 Full Sample	k=4 1st Sample	k=4 2nd Sample	k=4 Full Sample
Modified Diebold-Mariano Test for Equality of Forecast Accuracy (Type of Forecast on each Vintage: One-Step Ahead; Forecast Sample: One Observation)												
						Germany						
M1-M3	0.000	-0.152	-0.771	-0.812	-0.219	-1.200	-0.756	-0.227	-1.159	-0.727	-0.263	-1.159
M1-M4	1.315	-0.921	-0.693	1.077	-0.411	-0.313	0.425	-0.786	-0.649	-0.362	-0.313	-1.037
M1-M5	-0.479	-0.618	**-2.101**	-0.261	-0.075	-0.464	-0.100	-0.288	-0.671	-0.044	-0.271	-0.655
M2-M3	1.314	-0.192	-0.121	1.151	-0.226	-0.193	1.298	-0.359	-0.298	1.656	-0.380	-0.382
M2-M5	-0.015	-0.058	-0.304	0.565	0.110	0.820	0.524	-0.175	-0.177	0.860	-0.317	-0.274
M3-M5	-0.458	-0.023	-0.248	0.276	0.149	0.769	0.179	-0.057	-0.056	0.589	-0.342	-0.210
						Spain						
M1-M3	0.286	1.243	**2.427**	-0.835	0.847	1.331	-0.041	0.573	1.092	0.066	0.370	0.897
M1-M4	-0.591	0.377	1.042	0.670	0.253	0.873	0.481	0.386	1.027	0.765	0.420	1.131
M1-M5	-1.298	0.263	0.313	-0.672	0.322	0.763	-0.045	1.005	1.137	0.526	0.732	1.118
M2-M3	-0.497	-0.152	-0.701	**-2.497**	-0.435	**-1.725**	-1.414	**NAN**	**-1.972**	-1.253	-0.736	-1.550
M2-M5	-1.570	-0.982	**-2.643**	-0.424	-1.024	-1.524	-0.572	-0.843	-1.310	**NAP**	-0.302	-0.838
M3-M5	-1.494	-0.896	**-2.509**	0.048	-0.924	-1.379	0.010	-0.826	-1.168	1.487	-0.227	-0.521
						France						
M1-M3	0.460	0.224	0.904	0.864	0.306	1.316	0.679	**17.167**	1.324	0.929	0.214	1.272
M1-M4	-0.059	0.079	0.206	0.196	0.044	0.261	0.369	-0.536	-0.238	0.422	**NAN**	-0.018
M1-M5	0.784	0.092	0.761	-0.394	0.058	-0.184	-0.261	-1.609	-0.824	0.204	**NAN**	0.081
M2-M3	0.230	0.008	0.204	-0.449	-0.273	-0.912	-0.244	-0.701	-1.039	0.070	-0.490	-0.801
M2-M5	0.371	-0.109	-0.100	-1.130	-0.221	-1.324	-0.909	-0.987	-1.696	-0.425	**NAN**	-1.315
M3-M5	0.316	-0.101	-0.131	-1.104	-0.161	-1.111	-0.793	-0.711	-1.548	-0.576	**NAN**	-1.454
						Italy						
M1-M3	-0.074	0.149	0.470	0.033	0.041	0.151	0.792	-0.032	0.159	-0.380	-0.245	-0.867
M1-M4	0.140	-0.138	-0.049	-0.474	-0.078	-0.548	-1.138	-0.046	-0.929	-0.784	-0.084	-0.802
M1-M5	-0.935	-0.737	**-1.864**	-0.605	-0.052	-0.663	-1.475	-0.158	-1.590	-0.788	-0.092	-0.818
M2-M3	1.152	-0.014	1.121	0.360	**NAN**	0.381	0.581	-0.470	-0.007	0.446	-0.464	0.129
M2-M5	-0.397	-0.260	-0.977	-0.369	-0.051	-0.361	-1.381	-0.051	-1.125	-0.601	1.153	-0.379
M3-M5	-0.592	-0.263	-1.089	-0.403	-0.051	-0.387	-1.252	-0.007	-0.977	-0.568	1.270	-0.316
						UK						
M1-M3	-0.896	0.858	-0.135	**1.812**	0.956	**2.215**	0.849	0.518	1.216	0.198	0.438	0.609
M1-M4	-0.271	0.870	0.303	-0.310	0.323	-0.013	-0.402	0.222	-0.203	-0.677	0.235	-0.491
M1-M5	-0.164	0.766	0.460	-0.239	0.289	0.030	-0.453	0.220	-0.257	-0.551	0.224	-0.375
M2-M3	1.001	0.935	1.487	0.623	-0.888	-0.012	-0.140	-0.642	-0.491	-1.055	-0.569	-1.360
M2-M5	0.834	0.595	1.161	-0.910	-1.056	-1.603	-0.906	-0.703	-1.233	-0.756	-0.621	-0.957
M3-M5	0.380	0.224	0.501	-1.048	-0.942	-1.653	-0.856	-0.637	-1.116	-0.603	-0.605	-0.756
						USA						
M1-M3	0.569	0.734	0.942	-0.983	**1.999**	0.219	-0.827	**1.956**	-0.258	-0.493	1.266	0.003
M1-M4	-1.072	**3.073**	0.839	-1.596	1.654	-0.220	-1.518	1.387	-0.226	**-1.767**	1.214	-0.130
M1-M5	-1.619	**2.852**	0.344	**-2.860**	**1.781**	-0.588	**-3.129**	1.609	-0.394	**-2.932**	**1.659**	-0.046
M2-M3	0.359	-1.617	-0.263	-1.234	0.530	-0.880	-1.106	0.618	-1.017	-0.832	0.217	-0.870
M2-M5	**-1.982**	**1.940**	-0.472	**-2.912**	1.326	-0.988	**-3.589**	1.280	-0.737	**-3.063**	1.413	-0.438
M3-M5	**-1.960**	**2.278**	-0.265	**-2.143**	1.223	-0.746	-1.558	1.195	-0.282	-1.402	1.480	-0.054

Notes: a positive (negative) value indicates that the second (first) model has a better forecasting power. In the comparison of non-nested models: figures in bold indicate statistical significance at least at the 10% level. **NAN**: the modified Diebold-Mariano statistic is not available, but the test rejects the null of equal forecast accuracy and favors the first model. **NAP**: the modified Diebold-Mariano statistic is not available, but the test rejects the null of equal forecast accuracy and favors the second model. **Loss Function**: Root Mean Squared Forecast Error. **First Sample**: 2001.1-2008.2 (Germany), 2001.1-2008.1 (Spain), 1999.4-2008.1 (France), 1999.4-2008.1 (Italy), 1990.1-2006.1 (UK), 1974.3-1983.1 (USA). **Second Sample**: 2008.3-2010.3 (Germany), 2008.2-2010.3 (Spain), 2008.2-2010.3 (France), 2008.2-2010.3 (Italy), 2006.2-2010.2 (UK), 1983.2-2010.3 (USA). The test is run on real-time data.

Table 2a. Diebold-Mariano Tests - Models Comparison (1)

Compared Models	Modified Diebold-Mariano Test for Equality of Forecast Accuracy (Type of Forecast on each Vintage: One-Step Ahead; Forecast Sample: Eight Observations)											
	k=1			k=2			k=3			k=4		
	1st Sample	*2nd Sample*	*Full Sample*	*1st Sample*	*2nd Sample*	*Full Sample*	*1st Sample*	*2nd Sample*	*Full Sample*	*1st Sample*	*2nd Sample*	*Full Sample*
						Germany						
M1-M3	-0.492	0.441	-0.098	**-2.506**	-0.401	**-2.989**	**-1.820**	-0.054	**-1.828**	**-1.914**	-0.248	**-1.996**
M1-M4	0.936	-0.667	-0.263	0.030	-0.433	-0.638	-0.866	-0.486	-1.450	-1.403	-0.429	**-1.980**
M1-M5	**-3.182**	-1.072	**-4.734**	-1.148	-0.410	**-1.764**	-0.802	-0.232	-1.113	-0.746	-0.012	-0.770
M2-M3	**2.480**	-0.399	1.193	1.228	-0.245	0.221	0.532	-0.203	-0.054	0.845	-0.177	0.122
M2-M5	**-3.899**	-1.044	**-5.416**	0.651	-0.343	0.124	0.118	-0.183	-0.217	0.312	-0.013	0.273
M3-M5	**-4.213**	-1.213	**-6.107**	0.395	-0.418	0.072	0.025	-0.175	-0.199	0.167	0.056	0.243
						Spain						
M1-M3	0.272	1.245	**2.424**	0.913	0.703	1.639	0.557	0.474	1.255	0.251	0.413	0.920
M1-M4	-0.787	1.557	1.032	0.350	1.424	1.542	0.440	0.733	1.313	0.344	0.555	1.094
M1-M5	**-2.398**	**2.304**	-0.740	-0.366	1.126	0.869	-0.108	0.629	0.925	0.067	0.492	0.935
M2-M3	-1.358	-0.768	**-2.587**	-1.297	-1.318	**-2.211**	-1.274	-1.311	**-2.211**	-1.687	-0.721	**-2.461**
M2-M5	**-3.985**	-1.127	**-3.237**	**-2.284**	-0.609	**-2.001**	**-1.872**	-0.306	**-2.140**	-1.182	0.603	0.327
M3-M5	**-3.339**	-1.049	**-3.088**	**-2.199**	-0.554	**-1.913**	-1.661	-0.120	-1.469	-0.486	0.626	0.707
						France						
M1-M3	0.162	0.900	**2.119**	**2.627**	0.456	**2.319**	1.679	0.240	1.637	0.471	0.119	0.591
M1-M4	-0.635	-1.521	**-2.155**	-0.001	-1.139	-0.916	0.301	-0.892	-0.290	0.313	-0.963	-0.147
M1-M5	-1.398	-1.223	**-2.593**	-0.999	-1.120	-1.814	-0.933	-0.901	-1.585	-0.557	-0.649	-1.082
M2-M3	-1.640	1.412	-0.774	**-2.275**	-0.251	**-2.489**	**-1.705**	-0.343	**-2.059**	-0.642	-0.256	-1.000
M2-M5	-1.260	-1.028	**-2.777**	**-2.232**	-0.636	**-2.792**	**-1.774**	-0.854	**-2.615**	-1.347	**-5.677**	**-2.122**
M3-M5	-1.080	-1.088	**-2.699**	**-1.999**	-0.661	**-2.629**	-1.560	-0.987	**-2.415**	-1.239	-1.430	**-1.926**
						Italy						
M1-M3	1.097	1.110	**2.634**	0.052	0.805	1.407	0.328	0.507	1.100	-0.135	-0.098	-0.168
M1-M4	**-2.290**	-0.433	**-2.867**	-1.430	0.573	-1.257	-1.195	0.640	-0.949	-1.010	0.341	-0.653
M1-M5	**-5.099**	-1.229	**-6.478**	**-2.338**	-0.895	**-2.639**	**-1.900**	0.273	**-1.761**	-1.295	0.302	-0.979
M2-M3	**2.114**	-0.993	1.446	1.141	0.184	1.249	0.743	-0.313	0.536	1.084	-0.136	0.953
M2-M5	**-3.978**	-1.145	**-4.177**	-1.640	-0.800	**-2.526**	-1.598	-0.518	**-2.178**	-1.104	0.434	-0.813
M3-M5	**-3.835**	-1.147	**-4.328**	-1.620	-0.819	**-2.551**	-1.457	-0.552	**-1.945**	-1.113	0.337	-0.857
						UK						
M1-M3	**-3.428**	**-2.531**	**-4.341**	**2.071**	0.933	**2.496**	0.741	0.314	0.922	-0.224	-0.365	-0.385
M1-M4	**-2.339**	-0.072	**-2.371**	-1.308	-0.441	-1.446	-0.905	-0.544	-1.059	-0.888	-0.504	-1.022
M1-M5	**-2.972**	-1.127	**-3.312**	-1.320	-0.834	-1.570	-0.909	-0.763	-1.099	-0.744	-0.799	-0.898
M2-M3	0.670	1.259	1.295	-0.363	-0.471	-0.456	-1.261	-1.149	-1.458	-1.554	-1.037	-1.750
M2-M5	-1.103	0.822	-0.693	**-1.995**	-1.160	**-2.669**	-1.289	-0.632	-1.567	-0.826	-0.305	-0.919
M3-M5	-1.268	0.332	-1.172	**-1.911**	-1.128	**-2.568**	-1.157	-0.577	-1.420	-0.619	-0.201	-0.683
						USA						
M1-M3	**2.236**	-0.525	**1.929**	**-2.340**	**3.030**	-0.804	-1.449	1.347	-1.016	-0.863	1.080	-0.291
M1-M4	**-3.612**	**4.291**	-0.323	**-2.797**	**2.634**	-0.665	**-2.282**	**1.860**	-0.516	**-2.446**	1.587	-0.261
M1-M5	**-8.293**	**3.872**	-1.485	**-8.335**	**2.685**	-1.379	**-6.399**	**2.072**	-1.091	**-3.726**	**2.119**	-0.394
M2-M3	**1.945**	**-6.244**	-0.522	**-2.701**	-0.061	**-2.436**	-1.581	-0.409	**-1.681**	**-1.786**	0.193	-1.447
M2-M5	**-8.770**	**2.053**	**-3.050**	**-9.139**	**1.815**	**-1.978**	**-7.335**	**1.676**	-1.377	**-4.673**	**1.870**	-0.756
M3-M5	**-8.786**	**3.523**	**-2.142**	**-6.525**	**1.848**	-1.329	**-3.093**	**1.899**	-0.724	**-2.333**	**2.027**	-0.318

Notes: a positive (negative) value indicates that the second (first) model has a better forecasting power. In the comparison of non-nested models: figures in bold indicate statistical significance at least at the 10% level. **NAN**: the modified Diebold-Mariano statistic is not available, but the test rejects the null of equal forecast accuracy and favors the first model. **NAP**: the modified Diebold-Mariano statistic is not available, but the test rejects the null of equal forecast accuracy and favors the second model. **Loss Function**: Root Mean Squared Forecast Error. **First Sample**: 2001.1-2008.2 (Germany), 2001.1-2008.1 (Spain), 1999.4-2008.1 (France), 1999.4-2008.1 (Italy), 1990.1-2006.1 (UK), 1974.3-1984.3 (USA). **Second Sample**: 2008.3-2010.3 (Germany), 2008.2-2010.3 (Spain), 2008.2-2010.3 (France), 2008.2-2010.3 (Italy), 2006.2-2010.2 (UK), 1984.4-2010.3 (USA). The test is run on real-time data.

Table 2b. Diebold-Mariano Tests - Models Comparison (2)

Country	Variable	K	Break	Innovation Variance 1st Subsample	Innovation Variance 2nd Subsample	Average Term Spread 1st Subsample	Average Term Spread 2nd Subsample
Germany	10yr Government Bond Yield	2	**1970.4**$^{(E,A)}$	0.062 (0.013)	0.127 (0.014)	1.729	1.064
	Yield Spread	2	**1975.2**$^{(M,E,A)}$	1.037 (0.230)	0.213 (0.036)	1.378	1.133
Spain	10yr Government Bond Yield	3	**1995.1**$^{(M,E,A)}$	0.614 (0.107)	0.094 (0.022)	-0.525	1.301
	Yield Spread	3	**1987.3**$^{(M,E,A)}$	1.821 (0.469)	0.257 (0.059)	-0.603	0.732
France	10yr Government Bond Yield	2	**1979.1**$^{(M,E,A)}$	0.050 (0.012)	0.203 (0.033)	1.219	0.989
	Yield Spread	2	**1996.1**$^{(M,E,A)}$	0.673 (0.104)	0.138 (0.046)	0.873	1.345
Italy	10yr Government Bond Yield	2	**1996.4**$^{(M,E,A)}$	0.499 (0.117)	0.076 (0.015)	0.016	1.286
	Yield Spread	2	**1994.3**$^{(M,E,A)}$	0.662 (0.164)	0.175 (0.058)	-0.194	1.301
UK	10yr Government Bond Yield	3	**1987.3**$^{(M,E,A)}$	0.503 (0.082)	0.137 (0.024)	0.118	-0.002
	Yield Spread	2	**1994.2**$^{(M,E,A)}$	0.716 (0.156)	0.160 (0.046)	-0.322	0.394
USA	10yr Government Bond Yield	2	**1979.3**$^{(M,E,A)}$	0.068 (0.009)	0.316 (0.055)	0.089	1.112
	Yield Spread	2	**1982.3**$^{(M,E,A)}$	1.070 (0.347)	0.215 (0.028)	-0.150	1.378

Notes: Hansen (2000)'s fixed regressor grid-bootstrap procedure tests the null of no breaks vs the alternative of one break. The test is applied to the marginal process of the indicated variable, as in equation (S1). When the test rejects the null of no breaks on the basis of computed bootstrap p-values (size of the tests is 10%), the estimated break is reported in bold. Innovation variances are estimated over the two subsamples; estimated standard errors are given in parentheses.

$^{(M)}$ Significant using Max F-Statistic. $^{(E)}$ Significant using Exp F-Statistic. $^{(A)}$ Significant using Ave F-Statistic.

Table 3a. Breakpoint Tests (Hansen, 2000). One Break in Innovation Variance

Country	Variable	K	Break	Innovation Variance		Innovation Variance - Variable Variance Ratio		Average Term Spread	
				1st Subsample	2nd Subsample	1st Subsample	2nd Subsample	1st Subsample	2nd Subsample
Germany	10yr Government Bond Yield	2	**1994.4**$^{(M,E,A)}$	0.119	0.074	0.080	0.061	1.141	1.357
	Yield Spread	2	**1969.2**$^{(M,E,A)}$	0.481	0.382	0.883	0.145	2.115	0.999
Spain	10yr Government Bond Yield	3	1987.1	0.705	0.242	0.123	0.016	-0.345	0.624
	Yield Spread	3	**1993.4**$^{(M,E,A)}$	1.014	0.194	0.371	0.187	-0.765	1.366
France	10yr Government Bond Yield	2	**1981.3**$^{(M,E)}$	0.119	0.150	0.015	0.012	1.215	0.970
	Yield Spread	2	1981.3	0.620	0.409	0.237	0.260	1.215	0.970
Italy	10yr Government Bond Yield	2	1995.2	0.503	0.086	0.047	0.024	-0.054	1.232
	Yield Spread	2	**1992.4**$^{(M,E,A)}$	0.624	0.190	0.423	0.168	-0.389	1.297
UK	10yr Government Bond Yield	3	**1974.4**$^{(M,E)}$	0.142	0.383	0.025	0.031	---	---
	Yield Spread	2	1988.3	0.783	0.260	0.370	0.083	0.117	-0.008
USA	10yr Government Bond Yield	2	**1981.3**$^{(M,E)}$	0.148	0.235	0.022	0.031	-0.186	1.347
	Yield Spread	2	**1980.3**$^{(A)}$	0.601	0.494	0.234	0.224	-0.014	1.202

Notes: Hansen (2000)'s fixed regressor grid-bootstrap procedure tests the null of no breaks vs the alternative of one break. The test is applied to the marginal process of the indicated variable, as in equation (S1). When the test rejects the null of no breaks on the basis of computed bootstrap p-values (size of the tests is 10%), the estimated break is reported in bold. Innovation variances are estimated over the two subsamples.

$^{(M)}$ Significant using Max F-Statistic. $^{(E)}$ Significant using Exp F-Statistic. $^{(A)}$ Significant using Ave F-Statistic.

Table 3b. Breakpoint Tests (Hansen, 2000). One Break in Innovation Variance and Model Coefficients

Country	Variable	K	Break	Innovation Variance 1st Subsample	Innovation Variance 2nd Subsample	Innovation Variance - Variable Variance Ratio 1st Subsample	Innovation Variance - Variable Variance Ratio 2nd Subsample	Average Term Spread 1st Subsample	Average Term Spread 2nd Subsample
Germany	10yr Government Bond Yield	2	**1994.4**	0.117	0.070	0.079	0.060	1.141	1.357
	Yield Spread	2	**1981.3**	0.879	0.134	0.249	0.089	1.382	1.078
Spain	10yr Government Bond Yield	3	**1995.1**	0.581	0.066	0.102	0.019	-0.525	1.301
	Yield Spread	3	**1994.2**	0.984	0.154	0.383	0.157	-0.710	1.382
France	10yr Government Bond Yield	2	**1973.2**	0.025	0.187	0.017	0.014	1.326	1.014
	Yield Spread	2	**1996.1**	0.672	0.122	0.287	0.146	0.873	1.345
Italy	10yr Government Bond Yield	2	**1997.2**	0.481	0.057	0.043	0.129	0.024	1.323
	Yield Spread	2	**1993.3**	0.616	0.160	0.408	0.140	-0.302	1.300
UK	10yr Government Bond Yield	3	**1994.2**	0.450	0.070	0.056	0.035	-0.322	0.394
	Yield Spread	2	**1993.4**	0.708	0.144	0.202	0.073	-0.405	0.452
USA	10yr Government Bond Yield	2	**1979.3**	0.066	0.315	0.018	0.036	0.089	1.112
	Yield Spread	2	**1982.4**	1.005	0.195	0.376	0.119	-0.126	1.375

Notes: the likelihood ratio test is for the null of no breaks vs the alternative of one break in correspondence of the estimated date. The test is applied to the marginal process of the indicated variable, as in equation (S1). When the test rejects the null of no breaks (size of the test is 10%), the estimated break is reported in bold. Innovation variances are estimated over the two subsamples.

Table 4a. Breakpoint Tests (Qu-Perron, 2007). One Break in Innovation Variance and Model Coefficients

Country	Variable	K	Breaks		Innovation Variance			Innovation Variance - Variable Variance Ratio			Average Term Spread		
					1st Subsample	2nd Subsample	3rd Subsample	1st Subsample	2nd Subsample	3rd Subsample	1st Subsample	2nd Subsample	3rd Subsample
Germany	10yr Government Bond Yield	2	**1970.4**	**1994.4**	0.054	0.149	0.070	0.079	0.099	0.060	1.729	0.872	1.357
	Yield Spread	2	**1969.2**	**1981.3**	0.440	0.928	0.134	0.892	0.178	0.089	2.115	0.813	1.078
Spain	10yr Government Bond Yield	3	**1995.1**	**2002.2**	0.581	0.079	0.041	0.102	0.017	0.295	-0.525	1.204	1.387
	Yield Spread	3	**1987.2**	**1993.4**	1.494	0.204	0.182	0.368	0.289	0.183	-0.517	-1.051	1.366
France	10yr Government Bond Yield	2	**1973.2**	**1994.4**	0.025	0.263	0.064	0.017	0.039	0.051	1.326	0.781	1.333
	Yield Spread	2	**1989.2**	**1996.1**	0.526	0.881	0.122	0.303	0.400	0.148	1.289	-0.330	1.345
Italy	10yr Government Bond Yield	2	**1997.2**	**2002.2**	0.481	0.061	0.037	0.043	0.188	0.254	0.024	0.908	1.574
	Yield Spread	2	**1989.1**	**1994.2**	0.465	0.602	0.154	0.428	0.602	0.132	-0.758	0.663	1.322
UK	10yr Government Bond Yield	3	**1972.1**	**1990.2**	0.081	0.652	0.102	0.059	0.186	0.024	---	-0.375	0.285
	Yield Spread	2	**1992.3**	**1997.2**	0.658	0.177	0.149	0.189	0.379	0.087	-0.558	1.720	0.087
USA	10yr Government Bond Yield	2	**1979.3**	**1988.2**	0.066	0.712	0.136	0.018	0.149	0.050	0.089	0.748	1.256
	Yield Spread	2	**1971.3**	**1982.4**	0.144	1.513	0.195	0.284	0.368	0.119	-0.145	-0.113	1.375

Notes: the likelihood ratio test is for the null of no breaks vs the alternative of one break and the null of no breaks vs the alternative of two breaks in correspondence of the estimated dates. The tests are applied to the marginal process of the indicated variable, as in equation (S1). When the tests reject the null of no breaks (size of the tests is 10%), the estimated breaks are reported in bold. Innovation variances are estimated over the three subsamples.

Table 4b. Breakpoint Tests (Qu-Perron, 2007). Two Breaks in Innovation Variance and Model Coefficients

Dependent Variables (Model 1 - Marginal Process)		K	Breaks		Innovation Variance (Bond Yield or Spread) 1st Subsample	2nd Subsample	3rd Subsample	Beta 1st Subsample	2nd Subsample	3rd Subsample	Average Term Spread 1st Subsample	2nd Subsample	3rd Subsample
								Germany					
growth(t,t+1)	10yr Government Bond Yield	2	**1971.1**	**1994.4**	0.058	0.146	0.070	0.324	0.623	1.041	1.697	0.878	1.357
growth(t,t+2)			**1971.1**	**1994.4**	0.058	0.146	0.068	0.272	0.609	1.075	1.697	0.878	1.357
growth(t,t+3)			**1974.3**	**1989.1**	0.086	0.157	0.085	0.483	0.828	0.336	1.323	1.529	0.913
growth(t,t+4)			**1974.3**	**1988.4**	0.085	0.157	0.089	0.629	0.781	0.268	1.323	1.545	0.909
growth(t,t+1)	Yield Spread	2	**1969.2**	**1981.3**	0.441	0.928	0.135	3.981	0.475	0.669	2.115	0.813	1.078
growth(t,t+2)			**1969.2**	**1981.3**	0.441	0.928	0.136	2.663	0.461	0.599	2.115	0.813	1.078
growth(t,t+3)			**1974.3**	**1985.2**	1.011	0.378	0.138	0.761	0.811	0.424	1.323	1.416	1.053
growth(t,t+4)			**1974.3**	**1988.4**	1.014	0.328	0.134	0.881	0.773	0.261	1.323	1.545	0.909
								Spain					
growth(t,t+1)	10yr Government Bond Yield	3	**1995.1**	**2005.4**	0.582	0.075	0.035	0.037	-0.451	0.018	-0.525	1.344	1.206
growth(t,t+2)			**1995.1**	**2005.3**	0.581	0.075	0.037	-0.036	-0.371	0.434	-0.525	1.351	1.197
growth(t,t+3)			**1995.2**	**2005.2**	0.576	0.078	0.045	-0.018	-0.214	0.310	-0.481	1.337	1.190
growth(t,t+4)			**1995.1**	**2005.2**	0.581	0.076	0.057	-0.005	-0.320	0.467	-0.525	1.358	1.190
growth(t,t+1)	Yield Spread	3	**1994.2**	**2005.4**	0.984	0.095	0.296	-0.270	-0.247	0.451	-0.710	1.455	1.206
growth(t,t+2)			**1994.2**	**2005.3**	0.984	0.097	0.290	-0.139	-0.325	0.737	-0.710	1.465	1.197
growth(t,t+3)			**1994.1**	**2005.2**	0.922	0.123	0.292	-0.100	-0.434	1.157	-0.756	1.483	1.190
growth(t,t+4)			**1994.2**	**2005.2**	0.984	0.099	0.292	-0.064	-0.515	0.467	-0.710	1.474	1.190
								France					
growth(t,t+1)	10yr Government Bond Yield	2	**1976.4**	**1994.4**	0.056	0.297	0.062	0.681	0.405	0.968	0.904	0.839	1.333
growth(t,t+2)			**1978.1**	**1994.4**	0.071	0.311	0.063	0.877	0.382	1.124	1.023	0.775	1.333
growth(t,t+3)			**1978.1**	**1994.4**	0.066	0.312	0.063	0.879	0.388	1.225	1.023	0.775	1.333
growth(t,t+4)			**1978.1**	**1994.4**	0.067	0.312	0.064	0.875	0.388	1.316	1.023	0.775	1.333
growth(t,t+1)	Yield Spread	2	**1978.3**	**1995.2**	0.558	0.674	0.126	1.032	0.460	1.149	1.120	0.718	1.350
growth(t,t+2)			**1978.2**	**1995.2**	0.578	0.671	0.127	1.175	0.455	1.220	1.067	0.751	1.350
growth(t,t+3)			**1978.1**	**1995.2**	0.516	0.685	0.129	1.181	0.454	1.262	1.023	0.776	1.350
growth(t,t+4)			**1978.1**	**1995.2**	0.510	0.685	0.131	1.116	0.395	1.351	1.023	0.776	1.350

Notes: this table summarizes breaks estimated through Qu-Perron (2007), applied on a system of two equations (Model 1 + Marginal Process (S1) for either 10yr Government Bond Yields or Yield Spreads, as indicated). Sup likelihood ratio tests of no breaks against the alternative of two breaks reject the null of no breaks in all cases (size of the tests is 10%). **Samples**. Germany: from 1960.1 (10yr gov bond yield), from 1960.3 (yield spread); Spain: from 1980.4 (10yr gov bond yield), from 1980.4 (yield spread). France: from 1970.1 (10yr gov bond yield), from 1970.3 (yield spread).

Table 5a. Two Breaks in Innovation Variances and Models Coefficients, Systems of Equations (1)

Dependent Variables (Model 1 - Marginal Process)	K	Breaks		Innovation Variance (Bond Yield or Spread) 1st Subsample	2nd Subsample	3rd Subsample	Beta 1st Subsample	2nd Subsample	3rd Subsample	Average Term Spread 1st Subsample	2nd Subsample	3rd Subsample
							Italy					
growth(t,t+1) 10yr Government	2	**2000.4**	**2005.3**	0.423	0.052	0.033	0.016	1.121	0.847	0.147	1.647	1.408
growth(t,t+2) Bond Yield		**2000.3**	**2005.3**	0.428	0.049	0.037	-0.009	0.676	0.976	0.144	1.587	1.408
growth(t,t+3)		**2000.3**	**2005.2**	0.428	0.052	0.049	-0.019	0.632	0.632	0.144	1.605	1.401
growth(t,t+4)		**1995.2**	**2000.3**	0.480	0.214	0.049	-0.186	2.230	1.507	-0.054	0.727	1.498
growth(t,t+1) Yield Spread	2	**2000.4**	**2005.3**	0.502	0.034	0.221	-0.071	1.184	1.452	0.147	1.647	1.408
growth(t,t+2)		**2000.3**	**2005.3**	0.508	0.038	0.230	0.001	0.757	1.533	0.144	1.587	1.408
growth(t,t+3)		**2000.2**	**2005.2**	0.514	0.038	0.225	-0.049	0.405	1.694	0.135	1.567	1.401
growth(t,t+4)		**2000.2**	**2005.1**	0.514	0.040	0.217	-0.139	0.275	1.897	0.135	1.575	1.401
							UK					
growth(t,t+1) 10yr Government	3	**1994.2**	**2005.2**	0.468	0.066	0.066	0.754	-0.100	0.734	-0.322	0.358	0.468
growth(t,t+2) Bond Yield		**1994.2**	**2005.2**	0.470	0.067	0.065	0.742	-0.173	0.859	-0.322	0.358	0.468
growth(t,t+3)		**1994.2**	**2005.1**	0.474	0.066	0.063	0.711	-0.095	1.043	-0.322	0.376	0.429
growth(t,t+4)		**1994.2**	**2004.4**	0.473	0.067	0.064	0.671	-0.070	1.348	-0.322	0.389	0.401
growth(t,t+1) Yield Spread	2	**1994.2**	**2005.3**	0.708	0.109	0.228	0.861	-0.033	0.901	-0.322	0.344	0.505
growth(t,t+2)		**1994.2**	**2005.2**	0.708	0.111	0.223	0.861	-0.084	1.078	-0.322	0.358	0.468
growth(t,t+3)		**1994.2**	**2005.1**	0.708	0.112	0.220	0.865	-0.082	1.255	-0.322	0.376	0.429
growth(t,t+4)		**1994.2**	**2004.4**	0.708	0.115	0.230	0.759	-0.067	1.731	-0.322	0.389	0.401
							USA					
growth(t,t+1) 10yr Government	2	**1979.3**	**1987.3**	0.085	0.748	0.139	1.506	1.298	0.215	0.089	0.591	1.268
growth(t,t+2) Bond Yield		**1977.3**	**1984.2**	0.092	0.679	0.157	1.259	1.519	0.281	0.145	-0.160	1.333
growth(t,t+3)		**1991.1**	**2000.1**	0.304	0.138	0.099	1.123	-0.378	0.587	0.311	1.427	1.399
growth(t,t+4)		**1991.1**	**1999.3**	0.304	0.146	0.111	1.053	-0.420	0.596	0.311	1.495	1.347
growth(t,t+1) Yield Spread	2	**1973.2**	**1984.4**	0.215	1.494	0.185	1.395	1.888	0.262	0.063	0.106	1.322
growth(t,t+2)		**1971.4**	**1984.4**	0.153	1.380	0.186	1.791	1.608	0.329	-0.109	0.200	1.322
growth(t,t+3)		**1971.4**	**1984.4**	0.153	1.380	0.185	1.746	1.449	0.368	-0.109	0.200	1.322
growth(t,t+4)		**1991.1**	**1999.3**	0.774	0.103	0.272	1.130	-0.512	0.685	0.311	1.495	1.347

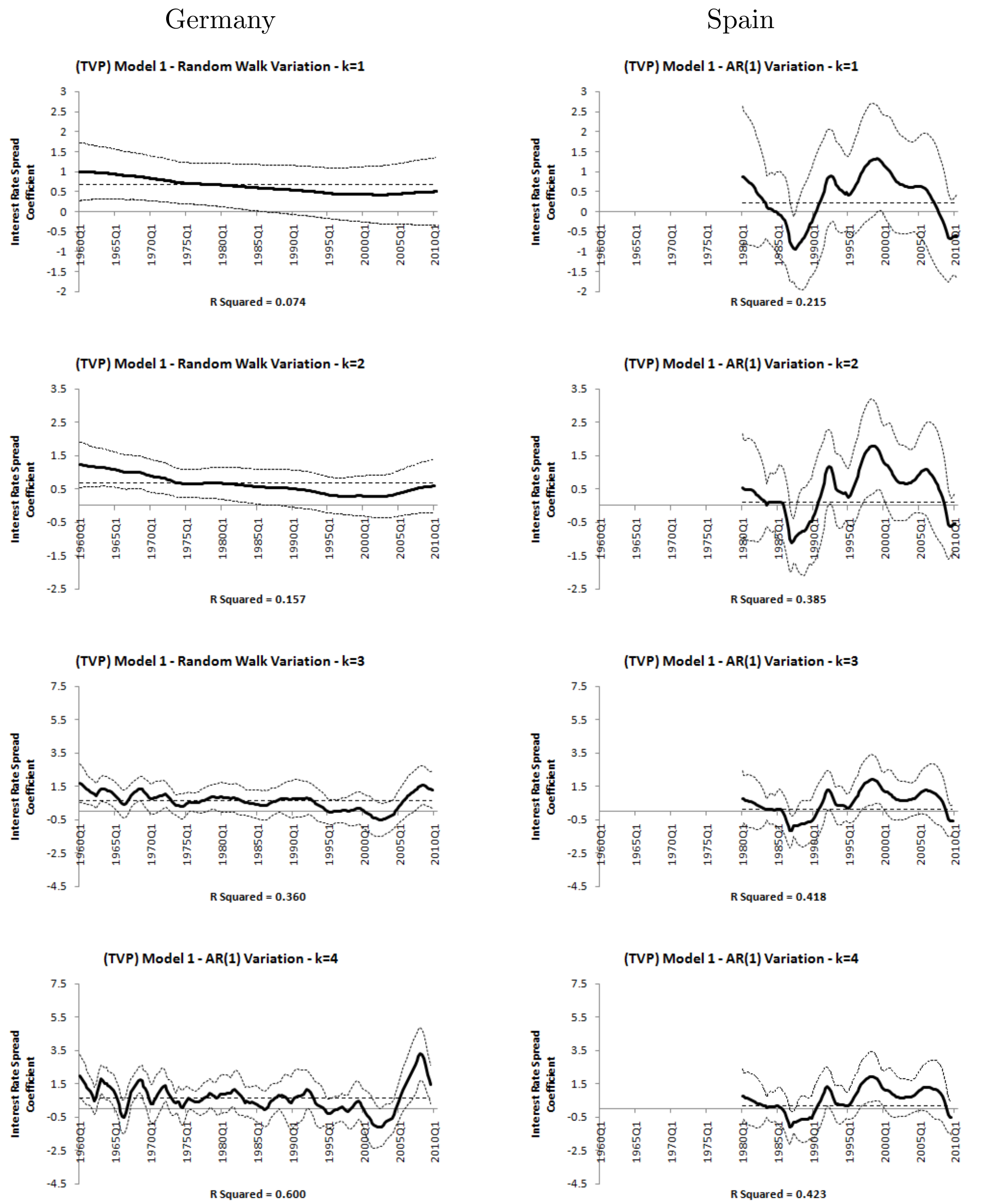
Notes: this table summarizes breaks estimated through Qu-Perron (2007), applied on a system of two equations (Model 1 + Marginal Process (S1) for either 10yr Government Bond Yields or Yield Spreads, as indicated). Sup likelihood ratio tests of no breaks against the alternative of two breaks reject the null of no breaks in all cases (size of the tests is 10%). **Samples**. Italy: from 1980.3 (10yr gov bond yield), from 1980.3 (yield spread); UK: from 1978.1 (10yr gov bond yield), from 1978.3 (yield spread); USA: from 1964.3 (10yr gov bond yield), from 1965.1 (yield spread).

Table 5b. Two Breaks in Innovation Variances and Models Coefficients, Systems of Equations (2)

		Model with No Breaks			Model with One Break in Innovation Variance			Model with One Break in Model Coefficients and Innovation Variance		
Country	**Variable**	**K**	**Break**	**Marginal Log-Likelihood**	**K**	**Break**	**Marginal Log-Likelihood**	**K**	**Break**	**Marginal Log-Likelihood**
Germany	10yr Government Bond Yield	2	---	-84.474	2	1973.3	-88.987	2	1968.3	-92.365
	Yield Spread	2	---	-220.080	2	**1980.2***	-181.611	2	**1981.1**	-191.012
Spain	10yr Government Bond Yield	3	---	-122.232	2	**1994.2***	-100.572	2	**1993.4**	-102.286
	Yield Spread	3	---	-157.003	2	**1987.4***	-123.286	2	**1986.2**	-132.002
France	10yr Government Bond Yield	2	---	-107.395	2	**1976.3***	-102.712	2	**1975.4**	-105.844
	Yield Spread	2	---	-182.012	2	**1995.1***	-165.138	2	**1994.4**	-174.196
Italy	10yr Government Bond Yield	2	---	-113.775	2	**1996.4***	-97.603	2	**1995.4**	-100.630
	Yield Spread	2	---	-128.816	2	**1993.1***	-114.765	2	**1991.3**	-120.520
UK	10yr Government Bond Yield	3	---	-190.918	1	**1993.1**	-173.302	2	**1973.1***	-87.207
	Yield Spread	2	---	-141.885	2	**1994.1***	-125.252	2	**1993.4**	-133.923
USA	10yr Government Bond Yield	2	---	-156.249	2	**1978.2***	-137.715	2	**1978.2**	-146.022
	Yield Spread	2	---	-217.533	2	**1979.3***	-186.054	2	**1981.3**	-194.573

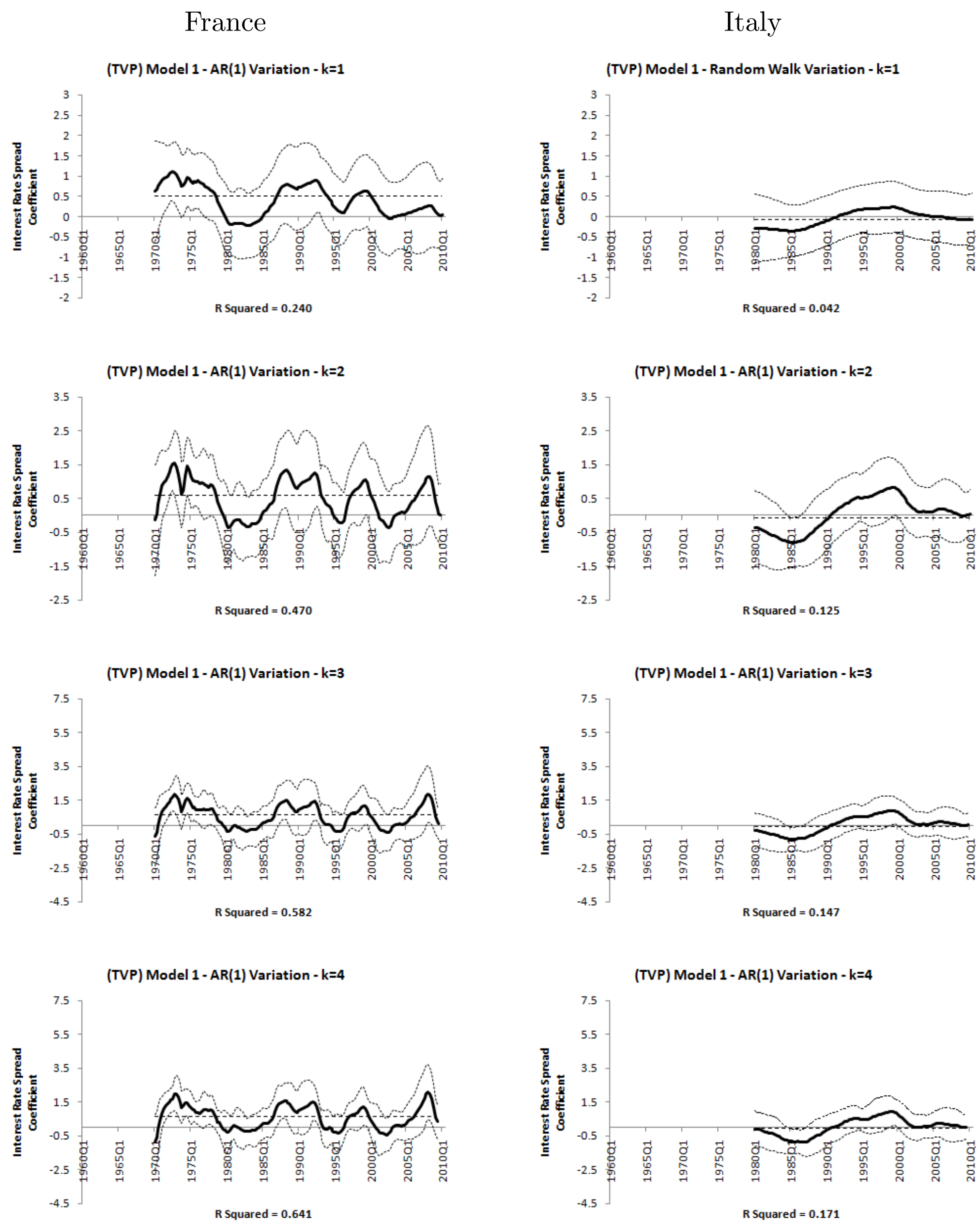
Notes: model (S1) is estimated in a Bayesian fashion as described in the Companion Technical Appendix. Estimated breaks are in bold if the corresponding model is to be preferred to the model without breaks based on marginal log-likelihood comparison. Asterisks denote the *preferred* breaks based on marginal likelihoods.

Table 6. Bayesian Estimates



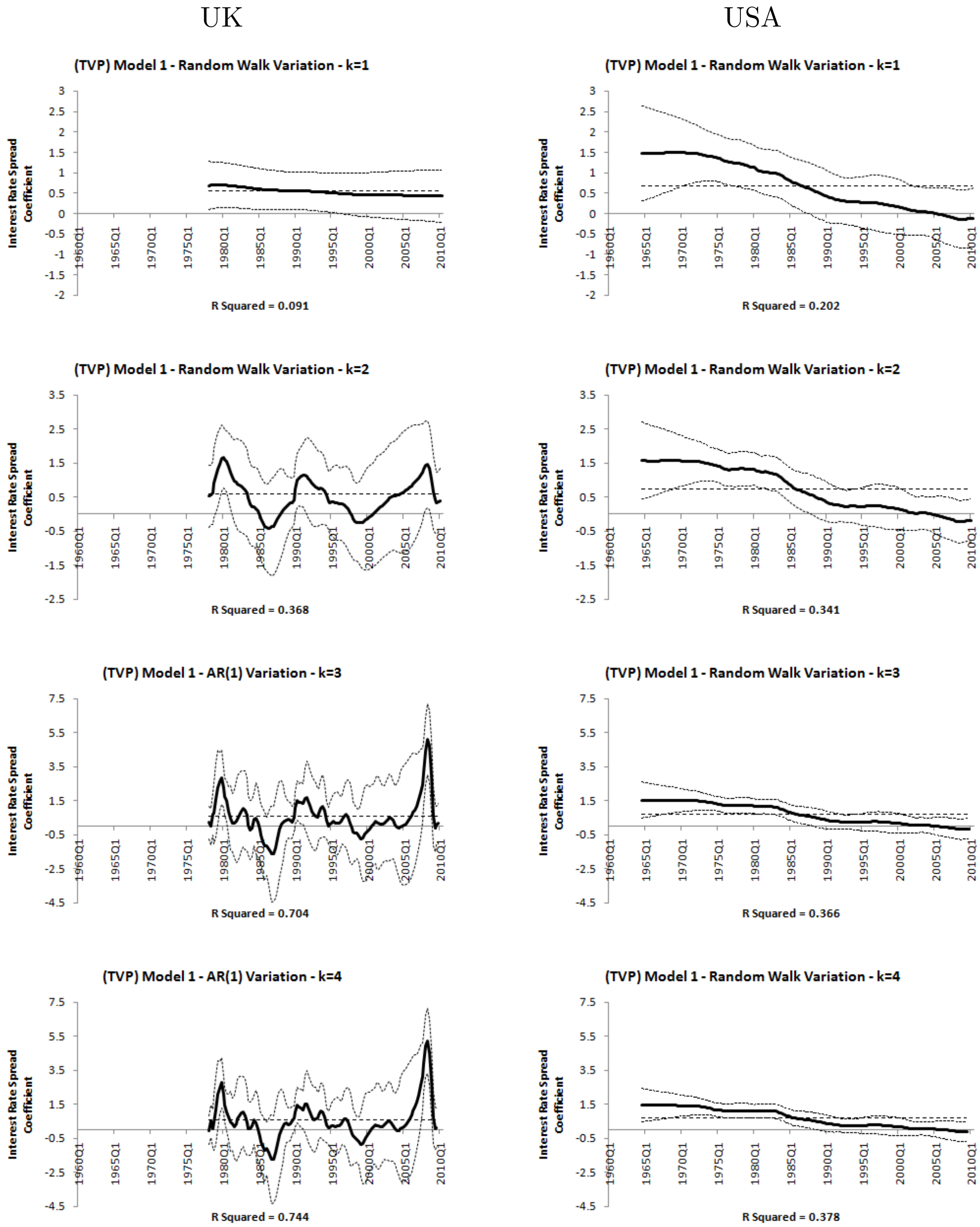
Notes: dotted lines and dashed lines denote two-standard-error confidence bands and OLS point estimates.

Figure 1a. Time-Varying Spread Coefficients (Smoothed Estimates)



Notes: dotted lines and dashed lines denote two-standard-error confidence bands and OLS point estimates.

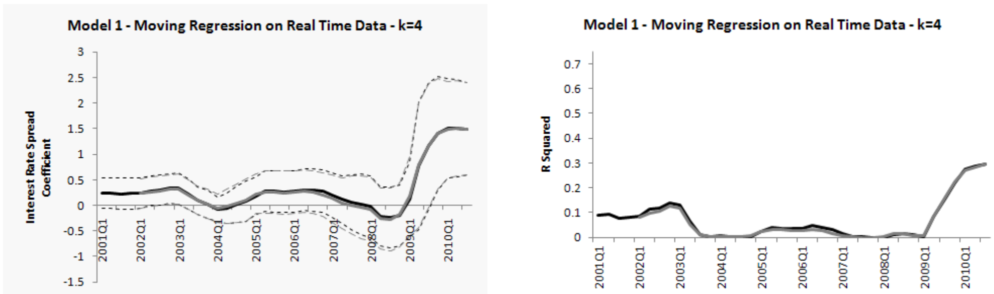
Figure 1b. Time-Varying Spread Coefficients (Smoothed Estimates)



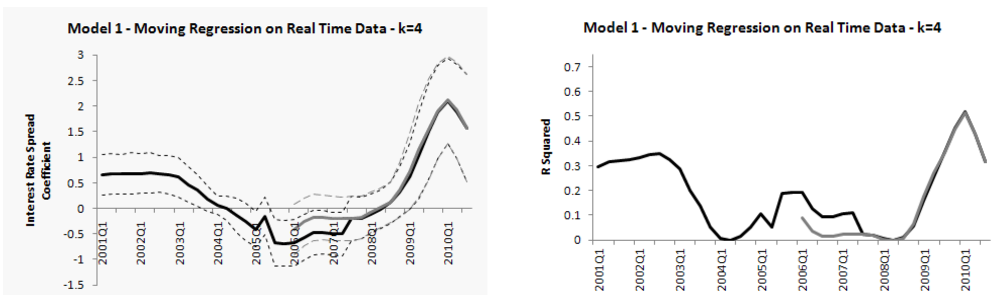
Notes: dotted lines and dashed lines denote two-standard-error confidence bands and OLS point estimates.

Figure 1c. Time-Varying Spread Coefficients (Smoothed Estimates)

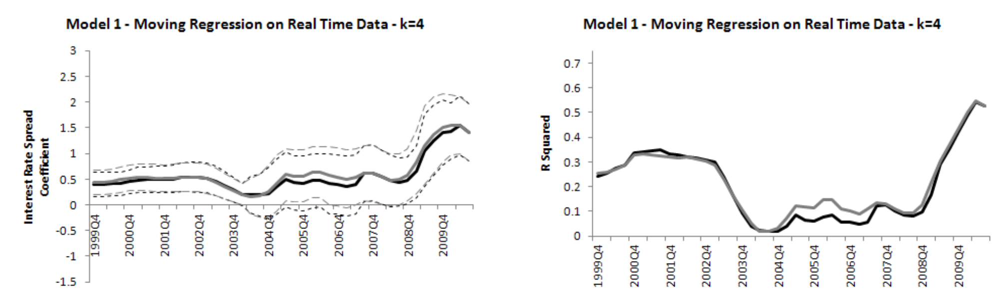
Germany



Spain



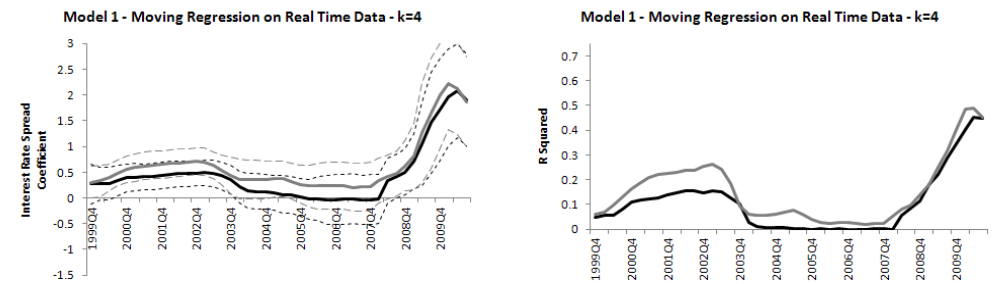
France



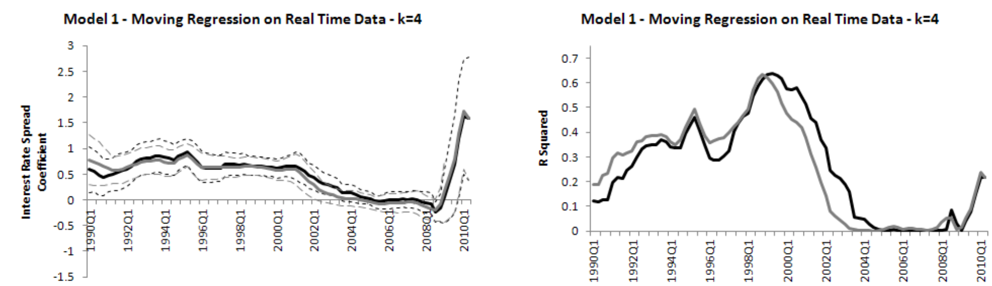
Notes: black solid/dashed lines denote real-time data estimates and two-standard-error confidence bands; gray solid/dashed lines denote estimates on the last vintage of data and corresponding two-standard-error confidence bands.

Figure 2a. Real-Time Estimates of β and Informativeness, Moving Regressions

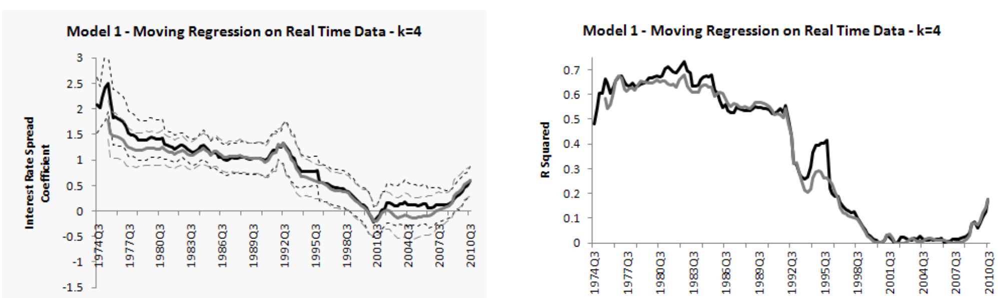
Italy



UK



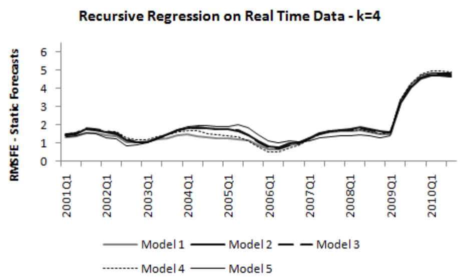
USA

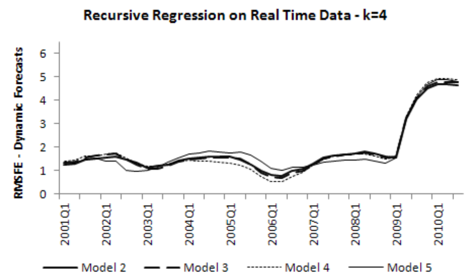


Notes: black solid/dashed lines denote real-time data estimates and two-standard-error confidence bands; gray solid/dashed lines denote estimates on the last vintage of data and corresponding two-standard-error confidence bands.

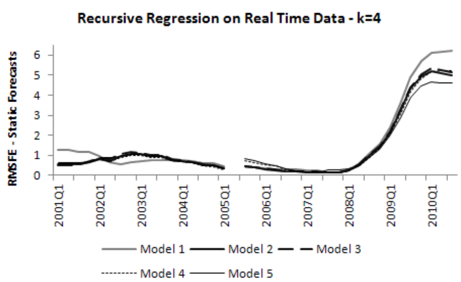
Figure 2b. Real-Time Estimates of β and Informativeness, Moving Regressions

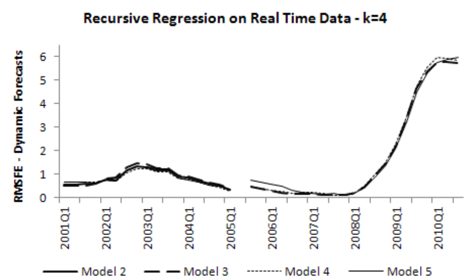
Germany



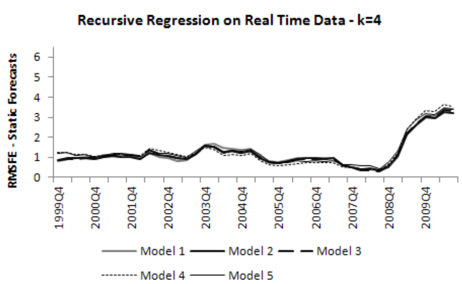


Spain





France



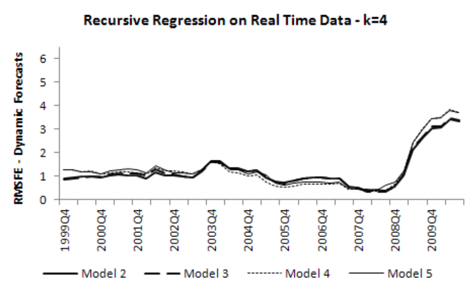
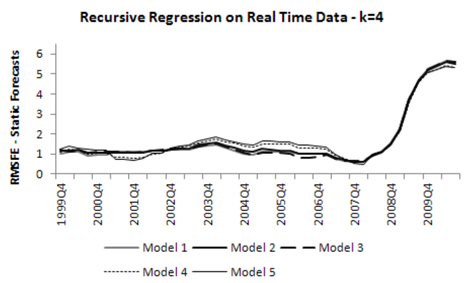
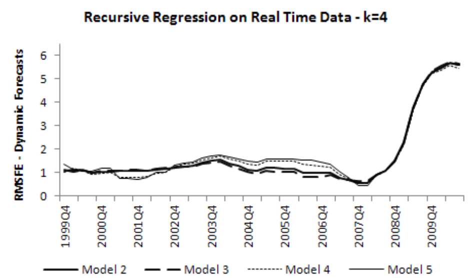


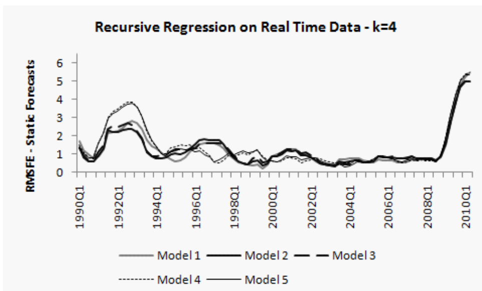
Figure 3a. Real-Time Estimates of $RMSFE$s, Recursive Regressions

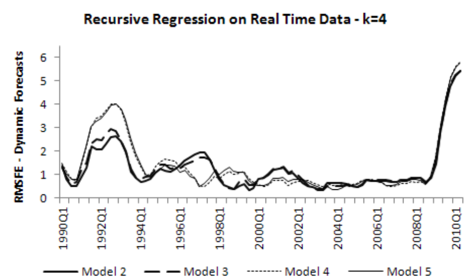
Italy



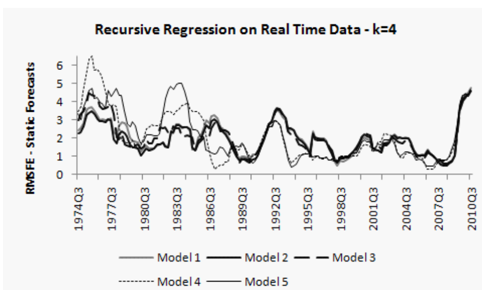


UK





USA



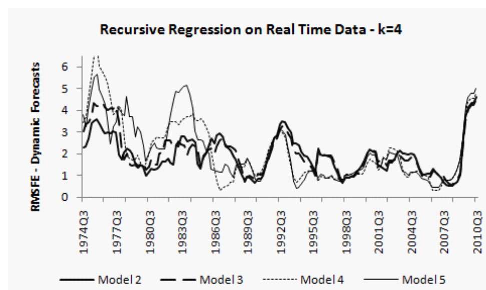


Figure 3b. Real-Time Estimates of $RMSFE$s, Recursive Regressions