How Stable Is the Predictive Power of the Yield Curve? Evidence from Germany and the United States

Arturo Estrella*, Anthony P. Rodrigues* and Sebastian Schich**

September 2000

Abstract

Empirical research over the last decade has uncovered predictive relationships between the slope of the yield curve and subsequent real activity and inflation. Some of these relationships are highly significant, but their theoretical motivations suggest that they may not be stable over time. We use recent econometric techniques for break testing to examine whether the empirical relationships are in fact stable. We consider continuous models, which predict either economic growth or inflation, and binary models, which predict either recessions or inflationary pressure. In each case, we draw on evidence from Germany and the United States. Models that predict real activity are more stable than those that predict inflation and binary models are more stable than continuous models. The model that predicts recessions is stable over our full sample period in both Germany and the United States.

*Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045, USA, e-mail arturo.estrella@ny.frb.org and tony.rodrigues@ny.frb.org.
**OECD, 2, rue André Pascal, 75775 Paris Cedex 16, France, e-mail sebastian.schich@oecd.org.

Views expressed are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of New York, the Federal Reserve System, or the OECD. We thank Muriel Duluc, Marianna Giokas, Sabina Goldina and Elizabeth Reynolds for excellent research support.

1. Introduction

Empirical research over the last decade has uncovered predictive relationships between the slope of the yield curve and subsequent inflation and real economic activity. Mishkin (1990a, 1990b, 1991) finds that the yield curve can predict inflation. He derives his framework from the Fisher equation that expresses a nominal interest rate in terms of a real rate and expected inflation. These results are confirmed and extended, for example, by Jorion and Mishkin (1991), Schich (1999a), Estrella and Mishkin (1997), and Kozicki (1997).

Harvey (1988), Laurent (1988, 1989), Chen (1991), and Estrella and Hardouvelis (1991) find that the yield curve slope helps predict real activity. These authors use a term structure spread to predict either subsequent real output growth or future recessions. Although the motivation for these relationships is more complex than in the case of inflation, the reported results have tended to be statistically stronger. These results have been confirmed and extended by other researchers, for example, Bomhoff (1994), Davis and Henry (1994), Gamber (1996), Davis and Fagan (1997), Estrella and Mishkin (1997, 1998), Kozicki (1997), Bernard and Gerlach (1998), and Filardo (1999).

While some of the predictive relationships are quite close, their theoretical motivations suggest that they may not be stable over time. In the case of inflation, the predictive power depends on the stochastic process for real interest rates and on its relationship to the process for inflation. There is no strong theoretical argument to expect the variances or covariances of these processes to be stable. In the case of real activity, it has been suggested that the predictive power may depend on factors such as a monetary policy reaction function or the relative importance of various shocks, which may change over time. Thus, it seems advisable to verify the stability of even the most significant empirical relationships.

To test for stability, we use recent econometric techniques for break testing. The techniques are straightforward and can be applied to generalized method of moments estimators. Such techniques are derived in Andrews and Fair (1988) and Ghysels and Hall (1993). In addition, some methods do not require that a specific "known" break point be specified. Among these are Andrews (1993) and Ghysels, Guay and Hall (1997).

For both inflation and real activity, we consider two types of models. In one type, the dependent variable is continuous, specifically economic growth (e.g., industrial production or

GDP) or the inflation rate over some specific horizon. Alternatively, we consider binary models, in which the dependent variable is a dichotomous index representing recessions or inflationary pressure. Earlier work on forecasting real activity (e.g., Estrella and Hardouvelis (1991)) suggests that binary indices may be easier to forecast than their continuous counterparts. For that reason, and because it has an intuitive interpretation, we introduce and use a binary model of inflationary pressure, which is defined in section 2.1.

We estimate each type of model with data for both Germany and the United States, where earlier research has found the relationships to be very significant. While we find the yield curve to be generally informative in all models, the stability results differ over the two countries depending on the approach chosen. Real activity models applied to German data are stable in both the economic growth and recession formulations. In contrast some United States real growth models appear to be unstable while the recession models are generally stable. As regards inflation, the evidence for instability is very limited in the case of Germany, but somewhat stronger in the case of the United States, possibly reflecting changes in monetary policy regime.

Overall, the results suggest that models that predict real activity are more stable than those that predict inflation. Further, binary models are more stable than continuous models. Recession prediction models are the most consistently stable in both Germany and the United States.

2. The yield curve and future inflation rates

2.1 Theoretical background

Two different approaches to analyzing the information in the term structure regarding future inflation are considered. One is the well-known Mishkin (1990a, 1990b, and 1991) approach, defining the information content of the yield curve as the ability of its slope to predict future changes in inflation rates. This approach is based on the Fisher decomposition, which states that the *m*-period nominal interest rate, $i_t^{(m)}$, can be divided into two components: the *m*-period *ex ante* real interest rate, denoted $E_t r_t^{(m)}$, and the expected inflation rate over the next *m* periods, denoted $E_t \pi_t^{(m)}$:

$$i_t^{(m)} = E_t r_t^{(m)} + E_t \pi_t^{(m)}. \tag{1}$$

If expectations are rational, the expected inflation rate can be written as the realized inflation $\pi_t^{(m)}$ plus an error term $\varepsilon_{t+m}^{(m)}$ that is orthogonal to information at time *t*:

$$\pi_t^{(m)} = E_t \pi_t^{(m)} + \varepsilon_{t+m}^{(m)}. \tag{2}$$

Substituting for $E_t \pi_t^{(m)}$ from equation (1), one obtains:

$$\pi_t^{(m)} = i_t^{(m)} - E_t r_t^{(m)} + \varepsilon_{t+m}^{(m)}. \tag{3}$$

Hence, the difference between (i) inflation over the next *m* years and (ii) inflation over the next *n* years (with *m* > *n*) can be written, in estimable form, as follows:

$$\pi_t^{(m)} - \pi_t^{(n)} = a_1^{(m,n)} + b_1^{(m,n)} \left(i_t^{(m)} - i_t^{(n)} \right) + \eta_{1,\, t+n}^{(m,n)}, \tag{4}$$

where $a_1^{(m,n)} = -\left(E_t r_t^{(m)} - E_t r_t^{(n)} \right)$ is the slope of *ex ante* real rates and $\eta_{1,\, t+n}^{(m,n)} = \varepsilon_{t+n}^{(m)} - \varepsilon_{t+m}^{(n)}$ is an error term. Assuming that the former is constant over time and that the latter has standard properties, the literature analyzes the information content of the term structure by testing if $b_1^{(m,n)} = 0$. If this hypothesis is rejected, the term spread contains significant information concerning inflation. The higher the coefficient of determination R^2, the more “informative” is the term structure.

Our second approach tests whether the term spread helps predict the *direction* of future inflation changes, that is, whether inflation will increase or decrease. For this purpose, the following binary variable $\Delta\ \overline{\pi}_t^{(n,m)}$ is defined:

$$\Delta\ \overline{\pi}_t^{(n,m)} = \begin{cases} 1 & for\ \ \pi_t^{(m)} - \pi_t^{(n)} > 0 \\ 0 & for\ \ \pi_t^{(m)} - \pi_t^{(n)} \leq 0. \end{cases} \tag{5}$$

The variable is 1 in month *t* if the average (annual) inflation is higher over the longer period of the next *m* years than in the shorter period of the next *n* years and 0 otherwise. It can be interpreted as an indicator of inflationary pressures; *i.e.*, if the indicator is equal to one, average inflation will rise over the period from *n* to *m* years. Using a binary variable to focus on the question of whether future inflation is increasing or not has some advantages. First, this allows one to address a question that frequently confronts monetary policymakers. Specifically, forecasts of the discrete indicator are estimates of the probability that inflation will increase over some specific interval in the future. Second, a binary approach may be more appropriate than the standard linear least squares method if the information variable, in this case the term structure spread, helps predict the direction of inflation change but not its magnitude.

To estimate the relationship between the yield spread and this variable, we use a binary response model, where the principle of maturity matching of interest rate spreads and inflation changes is followed. The estimated equation is:

$$P\left(\Delta\overline{\pi}_t^{(m,n)} = 1\right) = F\left(a_2^{(m,n)} + b_2^{(m,n)}\left(i_t^{(m)} - i_t^{(n)}\right)\right), \tag{6}$$

where F is the normal cumulative distribution function and the parameters are estimated using maximum likelihood. To gauge the performance of this indicator, we use the pseudo-R^2 statistic developed by Estrella (1998).

2.2 Previous empirical results

Various empirical studies have investigated the information content of the yield curve regarding inflation using the first methodological approach described above. The second methodological approach has not been considered in any previous paper on the subject. While most studies have focused on the United States, more recently a growing number of studies have investigated data for other countries, including Germany. Most of the studies on the United States and Germany agree that the yield curve is informative, but the precise results differ depending on the maturity segment and sample period chosen.

While there is hardly any information about future inflation in maturities up to one year in the United States and Germany (Mishkin, 1991), going beyond the short-term horizons generally results in significant estimates for the slope of the yield curve for both countries (Jorion and Mishkin, 1991). Keeping the short rate equal to one year and increasing the longer-term rate raises the R^2 in the United States (Mishkin, 1990). A similar result holds for Germany, but going well beyond the medium term segment implies that the R^2 falls again (Schich, 1999a and Gerlach, 1997). Overall, the closest agreement of the empirical results with the theoretical implications is found for the medium term segment up to eight years.

The results also depend on the choice of the sample period. For example, Mishkin (1990) finds that the estimated slope parameter is higher, and for some maturity combinations even significantly higher (at the 10% level), in the pre-October 1979 samples than for later samples in the United States. Also, studying the dynamics of the real interest rate in the United States, Huizinga and Mishkin (1986) identify two breakpoints in October 1979 and October 1982. Other results also point to sample-dependency of results. On the one hand, using a very short sample, Koedijk and Kool (1995) do not find evidence of significant information content about inflation in the term structures in Germany and the United States. On the other hand, Gerlach (1997), Schich (1999a) and Estrella and Mishkin (1997), each using a much larger sample, find that the term structure slope is significantly positively related to future inflation changes in Germany and the United States, respectively, thus confirming the earlier results of Jorion and Mishkin (1991). Thus, while overall the results point to a significant information of the term structures in the United States and Germany, they nevertheless seem to vary over time (Schich, 1999b).

2.3 Stability issues

To see why the information content might vary, we follow Hardouvelis (1988) and Mishkin (1990a, 1991) who show that the estimated slope coefficient $b_1^{(m,n)}$ in (3), assuming rational expectations, can be interpreted as a function of the moments of the inflation change and the real rate spread:

$$b_1^{(m,n)} = \frac{1 + \rho^{(m,n)} \quad q^{(m,n)}}{1 + \left(q^{(m,n)}\right)^2 + 2\rho^{(m,n)} q^{(m,n)}} \tag{7}$$

where

$$\rho^{(m,n)} = corr \left(E_t \left(\pi_t^{(n)} - \pi_t^{(m)}\right), E_t \left(r_t^{(n)} - r_t^{(m)}\right)\right)$$

is the correlation between the expected inflation change and the ex-ante real rate spread and

$$q^{(m,n)} = \frac{\sigma \left(E_t \left(\pi_t^{(n)} - \pi_t^{(m)}\right)\right)}{\sigma \left(E_t \left(r_t^{(n)} - r_t^{(m)}\right)\right)}$$

is the ratio of the standard deviation of the expected inflation change to the standard deviation of the ex-ante real rate spread. This equation demonstrates that changes in the slope parameter estimates could reflect changes in either (i) the correlation between the expected inflation change and the ex-ante real rate spread or (ii) the ratio of the standard deviation of the ex-ante real rate spread to the standard deviation of the expected inflation change or (iii) both.

These factors are likely to vary because of changes in behavior or in exchange rate or monetary policy regimes. For example, Huizinga and Mishkin (1986) suggest that the change in the Federal Reserve operating procedure away from interest-rate smoothing in October 1979 and the de-emphasis of monetary aggregates by the Fed in October 1982 are natural candidates for monetary policy regime shifts. In a similar vein, it could be argued that the adoption of monetary aggregate targeting in Germany in December 1974 represents another natural candidate for a monetary policy regime shift. As a consequence, the relationship between the yield spread and forward inflation rates may not be stable over time. To check for possible breakpoints, the empirical estimates are subjected in section 5 to two types of stability tests: stability tests where the breakpoint is unknown and tests where the breakpoint dates are assumed to be known.

3. The term structure and future real activity

In this section, we consider the power of the term structure to predict changes in real economic activity. There is an extensive literature that documents a positive empirical relationship between the slope of the yield curve and various measures of subsequent real activity, and we will review those results in section 3.2. Before considering the empirical evidence, however, we examine possible reasons for the observed relationship.

3.1 Theory

In the case of inflation, as illustrated in section 2.1, only a pair of simple relationships is needed to establish a possible connection between the slope of the term structure and future changes in inflation. In the case of real activity, it is necessary to extend the model to include some measure of output, as well as relationships between interest rates, output and inflation. Moreover, some of the explications provided in the literature are much less formal than the ones in section 2.1. We review five possible explanations for the relationships: countercyclical monetary policy, a monetary reaction function, the consumption capital asset pricing model (CCAPM), a real business cycle (RBC) model, and a simple dynamic Keynesian IS-LM model.

(a) Countercyclical monetary policy.[1] Consider the policy reaction to weakness in economic activity. First, the central bank increases the money supply. Short-term interest rates, both nominal and real, will tend to decline as a result of the change. Long-term rates, however, will tend to move less than the short rates for two reasons: (1) the monetary easing will tend to raise long-term inflation expectations and (2) the central bank may be expected to move to contractionary policy in the future to revert to a more neutral stance and to respond to future increases in inflation. The net result is a steepening of the yield curve and, since real interest rates will remain low for a period, a pickup in real economic activity. The effects of a monetary tightening in response to unsustainably high real activity are symmetrical.

(b) Monetary policy reaction function. In a small dynamic rational expectation model, Estrella (1997) looks at the relationship between the parameters of a monetary policy reaction function of the Taylor (1993) rule type and the predictive power of the term structure for real activity. The model contains a Phillips curve, a dynamic IS curve, the Fisher equation, the expectations hypothesis and the monetary policy rule. Results of the paper suggest that there is generally a positive relationship between the yield curve slope and expected future growth in real activity. The relationship is strongest when the central bank targets output only and is weakest or nonexistent when the central bank targets inflation only.

(c) CCAPM. The CCAPM, e.g., as in Harvey (1988), implies that there is a positive functional relationship between the slope of the *real* yield curve and future real consumption growth. The relationship follows directly from a standard first-order condition of the form $u'(C_t) = \beta E_t\{u'(C_{t+1})(1+\rho_{t+1})\}$, where C is the level of consumption, u is the utility function, β is a subjective time-discount factor, ρ is the real one-period interest rate, and E is the expectations operator.[2] Since the empirical relationship we examine here involves nominal yields, the applicability of the CCAPM to this issue relies on the assumption that the inflation expectations embedded in nominal rates play only a secondary role (for instance, if expected inflation is the same regardless of maturity).[3]

(d) RBC models.[4] Although the focus is somewhat different, the implications of general equilibrium RBC models for our purposes are analogous to those of the CCAPM.

---

[1] See, e.g., Estrella and Hardouvelis (1991), Dotsey (1998).
[2] For a discussion of how the CCAPM may be derived from the first-order condition, see Romer (1996, section 7.5).
[3] Note that the Fisher-equation approach of section 2, in contrast, assumes that it is the real rate that is constant.
[4] See, e.g., Kydland and Prescott (1988), Chen (1991), Plosser and Rouwenhorst (1994).

Specifically, RBC models rely on the same first order condition stated in the foregoing paragraph.[5] In this case, an interpretation is that expected positive productivity shocks are expected to increase future output, which in turn leads to a higher real interest rate as agents substitute current for future consumption. There is thus a tendency for the *real* yield curve to steepen as a consequence of those shocks. The caveat regarding the role played by inflation expectations, raised in the context of the CCAPM, applies here as well.

(e) Simple dynamic IS-LM models.[6] These Keynesian models provide to some extent a nominal analogue to the CCAPM and RBC results. With sticky prices (constant expected inflation) in the short run, short-term fluctuations are manifest mostly in real output changes. A clear example of these models is provided by Blanchard and Fisher (1989, pp. 532-536) in a model where instantaneous real returns on short-term and long-term bonds differ only by a constant. Consequently, the difference between the long-term (consol) and short-term (instantaneous) yields is given by the sum of a positive constant, which corresponds to "neutral" conditions, minus the capital-gains rate on the consol. Ruling out cases where capital gains (or losses) on the consol grow without bound restricts the economy to a saddle path on which the rates of change of output and the consol rate have the same sign. (When output is rising, interest rates must be increasing to maintain equilibrium in the money market.) Hence, when output is expected to grow, the consol rate is below its long run level and is expected to increase, and the yield curve spread is above the neutral level. Conversely, when output is expected to decline along the saddle path, so is the consol rate, and the yield curve spread is below the neutral level. Note that the yield curve is not necessarily inverted unless the neutral level is zero.

Thus, there are several theoretical lines of reasoning that can help explain a positive relationship between the slope of the term structure and future real activity. Interestingly, models that are usually considered to be incompatible have very similar implications as far as this issue is concerned. One consequence of these similarities is that we have (almost) every reason to expect to find a positive empirical relationship between the yield curve and real activity. Perhaps a more interesting (in the sense of being empirically testable) consequence is that several of the

[5] See Romer (1996, section 4.4) for a discussion.
[6] See, e.g., Dornbusch (1980), Blanchard and Fischer (1989), Estrella and Hardouvelis (1991), Berk (1998).

theories lead us to suspect that the relationship may not be entirely stable over time if circumstances change. We come back to this point in section 3.3.

3.2 Previous empirical results

The slope of the term structure has been most often represented in the empirical literature as the spread between long-term (e.g., 5 or 10 years) and short-term (e.g., 3 months or 1 year) government rates. A few papers have varied from this pattern by using, e.g., the federal funds rate as the short rate, or by trying different maturity combinations matched with different predictive horizons. Two types of dependent variables have been tested: continuous variables such as the growth in real GDP or industrial production, and discrete variables such as a recession dummy. To some extent, such dummy variables are analogous to our binary variable in section 2.1. For example, peaks and troughs in Germany, which define recessions, are identified by the OECD solely on the basis of total industrial production. Alternatively, the recession dummies may be defined somewhat independently from any single continuous real activity variable, as is the case with the NBER-dated recession variable, which we use for the United States.

The first papers dealing with the predictability of continuous variables, Laurent (1988, 1989), Stock and Watson (1989), Chen (1991) and Estrella and Hardouvelis (1991) focus on US data. All of these papers find highly significant relationships between the term structure spread and real activity with lead times ranging roughly from 1 to 8 quarters. Estrella and Hardouvelis (1991) look explicitly at the question of the "optimum" horizon and find that the results are most significant between 4 and 6 quarters ahead. Similar analyses, with consistent results, have been presented by Bomhoff (1994), Davis and Henry (1994), and Davis and Fagan (1997). Kozicki (1997) and Dotsey (1998) have recently surveyed this literature.

An interesting variation within the continuous dependent variable models is developed by Plosser and Rouwenhorst (1994), who break down the term structure spread into forward spreads with various non-overlapping horizons, and look at growth rates in GDP or GNP. For the US, they find that the long end of the yield curve (beyond 2 years) has the strongest predictive power and that the predictions are most accurate for shorter horizons.

Plosser and Rouwenhorst (1994) also investigate the predictive power in countries other than the US (UK and Germany in their case). They as well as Davis and Henry (1994) (UK and

Germany), Davis and Fagan (1997) (EU), and Estrella and Mishkin (1997), show that the term structure predicts real activity in several other countries, with particularly strong results in Germany.

The results for predictions of continuous dependent variables using a term structure spread have been consistently strong across time and across countries. However, the reliability of the predictions seems at times to fall short of being uniformly strong. For instance, predictive equations for GDP or GNP growth in the US that are estimated with data from the 1950s to the 1990s seem to exhibit signs of parameter instability.

Models with discrete or binary dependent variables have performed as well or better than the continuous models. For example, Stock and Watson (1989) develop a methodology for extracting indices of coincident and leading indicators from a set of macroeconomic time series. As part of this analysis, they define a recession indicator as a function of their coincident indicator, and calculated the probability of a recession implied by their leading indicators. A yield curve spread (10-year minus 1-year Treasury rates) was initially included in the leading indicators, although they subsequently dropped this variable from their model. Their results for the term structure are good, but not impressive. The main reason seems to be that the horizons examined by Stock and Watson (up to 2 quarters) are shorter than those over which the term structure exhibits its best performance.

Estrella and Hardouvelis (1991) introduce a probit model to predict the probabilities of US recessions, based on NBER dating, one year ahead. This model was quite successful and apparently stable with US data. The analysis has been extended by Bernard and Gerlach (1998), who look at data for 8 countries, and by Estrella and Mishkin (1998), who examine a range of predictive horizons as well as out-of-sample predictive performance, as compared with other financial indicators. Filardo (1999) is a useful recent survey that considers several very distinct modeling approaches.

3.3 Stability issues

There are good reasons to suspect that the relationships may not be stable over time. For instance, the theory suggests that results may be different if the economy is responding to real (productivity) or monetary shocks, or if the central bank is targeting output or inflation.

Empirically, there is some evidence that these concerns may be well founded, particularly in the models with continuous dependent variables.

It seems important, thus, to subject the various models to stability tests, particularly since the tests developed in the econometric literature since the late 1980s are especially suitable for these purposes. We turn to this issue in detail in section 5.

4. Stability tests: econometric issues

In this section, we describe the econometric methodology used to test the stability of the predictive models of inflation and real activity. Our approach follows the large econometric literature deriving asymptotically valid tests for the existence of unknown break points that developed in the past decade. This literature typically uses test statistics from conventional likelihood- or generalized-methods-of-moments criteria for testing known break dates. The innovation in the literature is to derive the asymptotic distributions of the statistics when the break dates are selected to maximize the deviation (as measured by the statistic) from the no-break null hypothesis. In the case where the possible break dates are known a-priori, the asymptotic distributions reduce to the standard chi-square tests familiar from standard asymptotics that are limiting cases of the *F* tests familiar from normal regression theory. The principal statistics we use are the Lagrange multiplier (LM) tests of Andrews (1993) and predictive (PR) tests of Ghysels, Guay and Hall (1997).[7]

Let the model to be tested be $y_t = f(\theta, x_t) + u_t$, where θ is a vector of parameters, and assume that the model is estimated by imposing orthogonality conditions of the form $E(g) = 0$, where g is a vector function of the data and the parameters. Typically $g = \sum_t u_t \cdot z_t$ where z_t is a vector of instrumental variables. As in Hansen (1982) and Newey and West (1987), assume that the estimate θ is obtained by minimizing over θ the quadratic form $g'Wg$, where W is a matrix of weights. Also, let $D = \partial g / \partial \theta$ and let S represent the Newey-West (1987) matrix of weighted orthogonality condition autocovariances. Then a consistent estimator of the variance of θ is

[7] In simulations we performed with stylized models whose autoregressive properties were similar to those of our actual models, the LM and PR tests appear to have reasonable characteristics in sample sizes like ours. The true size of these tests tended to be somewhat less than the nominal size. In contrast, the commonly used Wald test (which

$V = (D'WD)^{-1} D'WSWD(D'WD)^{-1}$. This variance estimator and some of the test statistics are simplified by defining the matrix, $M = (D'WD)^{-1} D'W$, which yields $V = M'SM$.

We also need to define subsample statistics, assuming a breakpoint. Let g_1 *and* g_2 represent the orthogonality conditions computed over the first and second subsample, respectively. Further let S_1 *and* S_2 represent the Newey-West subsample estimators of the matrix of weighted orthogonality condition autocovariances and let D_1 *and* D_2 be the derivatives of the subsample orthogonality conditions with respect to the parameters. Then the Lagrange multiplier, following Andrews and Fair (1988), is

$$\text{LM} = \frac{1}{\pi_1 \pi_2} g_1' M' V^{-1} M g_1$$

where π_i indicates the proportion of the data before (*i*=1) or after (*i*=2) the breakpoint and $\pi_1 + \pi_2 = 1$. The Ghysels-Hall (1990) predictive statistic is

$$\text{PR} = g_2'(S_2 + D_2 V_1 D_2')^{-1} g_2 .$$ [8]

Andrews and Fair (1988) and Ghysels and Hall (1990), respectively, show that under standard regularity conditions with potential break points known a-priori, LM and PR are distributed asymptotically as chi-squared with degrees of freedom equal to the number of parameters that may change across subsamples.[9]

In follow-up articles, Andrews (1993) and Ghysels, Guay and Hall (1997) derive the asymptotic distribution of *sup* LM and *sup* PR, where the *sup* is taken over an interior portion of the full sample that excludes some observations (a fraction π_0 of the total observations) at each end. These test statistics are known to be unbounded in the limit if the potential breakpoints include the endpoints of the sample. They may be used to test for a break when the break point is unknown. We use the sup break test statistics because the estimated break date is also of interest in the cases where we detect breaks. Andrews (1993) shows that the LM and PR test statistics

---

roughly corresponds to a Chow test) had a size much larger than the nominal size. E.g., the true size of a 5% test with 100 observations and 4 parameters was 17%.

[8] This statistic was called the TS statistic in Ghysels and Hall (1990) and renamed PR in Ghysels, Guay and Hall (1997).

[9] In many cases, the formulas above simplify considerably. For instance, if the estimates are obtained by least squares, then $g = X'u$ and $D = W^{-1} = X'X$. If instrumental variables are used, then $g = Z'u$, $D = Z'X$, and $W = (Z'Z)^{-1}$.

converge in distribution to the square of a standardized tied-down Bessel process under fairly general conditions. For a fixed breakpoint, this process has a chi-squared distribution. The distribution of the *sup* of this process and expressions for the cdf of related processes have been derived in the statistics literature. We apply the methodology for computing the exact probability values following results in DeLong (1981).

5. Stability tests of inflation predictions

5.1 Yield curve and future inflation rates

Empirical results for Germany and the United States

We use monthly data on government security interest rates from one to ten years from January 1955 to December 1998 for the United States and from January 1967 to December 1998 for Germany. The U.S. data are from McCulloch and Kwon (1993) and the Federal Reserve Bank of New York, while the German data are obtained from the Bundesbank. They are zero-coupon rates and yields-to-maturity, respectively.[10] While strictly speaking, zero-coupon interest rate would be required according to theory, the choice between zero-coupon interest rates or yields-to-maturity does not seem to matter in practice for purposes such as ours (Schich, 1996). The advantage of using yields-to-maturity for Germany, is that larger samples are available, which is useful when testing for stability of results. Forward inflation rate changes in continuous and binary form are calculated using consumer price series obtained from the U.S. Bureau of Labor Statistics (CPI) and the OECD.

The results presented in table 1 and table 2 are obtained from estimates of the linear (3) and the probit inflation equation (5). The equations are estimated using data from 1967:01 to 1998:12 for both countries, so that the results are comparable.[11] Due to the fact that the inflation rates are forward-looking, some observations are lost, and the actual sample ends between

[10] The U.S. rates are estimated using the tax-adjusted, cubic spline methodology in McCulloch (1975) from 1955 to 1991 and the smoothing splines methodology (Fisher, Nychka, Zervos, 1995) thereafter. The German rates are estimated using the linear-logarithmic approach traditionally used by the Bundesbank, and described in Schich (1996).
[11] The stability test results are similar when the larger data sample, starting in 1955:01 is used. They are available on request.

1988:12 (when *m*, the longer maturity in years, is equal to ten) and 1996:12 (when *m* is equal to two). As the overlapping data generate a moving average error term of order (12*m*-1), the standard errors are calculated through the Generalized Methods of Moments (GMM) estimator with Newey-West adjustment. The associated *t*-statistics of the slope parameter estimate in the linear regression and the analogues for the non-linear model are reported in squared brackets. LM and PR give the probability value (*p*-value) of the supremum of a set of possible breakpoints, and its date, of the Lagrange multiplier (LM) test and the predictive (PR) test. The set of possible breakpoints includes all observations, except for 25 percent at the beginning and at the end of the sample.[12] The reduction of the test interval to 50 percent of the total sample is larger than customary, but is necessary to ensure that there is variation in the dependent variable in each of the sub-samples used in the stability test of the probit model. The same reduction factor was applied to the continuous model for comparability of results.

The full-sample results using the continuous variable are consistent with those from previous studies. Specifically, the term structure of interest rates is informative about future inflation in both the United States and Germany, and the exact results depend, among other things, on the choice of the maturity combination. For a given short maturity of *n* equal to one, the R^2 follows a hump-shaped pattern as the long maturity *m* is increased from one. The R^2 first rises, peaks and then falls again. For Germany, this hump-shape is very pronounced, and it can be observed for *n* equal to two as well. For the United States, for any given *n* the R^2 peaks earlier, and when n is equal to two and three, the R^2 decreases (almost) monotonically as *m* is increased. When *n* is increased from one to two and from two to three, the results mostly improve in the case of Germany, but they worsen in the case of the United States. Thus, while the values of the maximum R^2s in the case of the United States and Germany are similar with .38 and .39, respectively, they are obtained in the short to medium-term segment for the United States (4-1 years) and in the medium-term segment for Germany (5-3 years). Summarizing, the German term structure is most informative at horizons beyond two years, while the U.S. term structure is most informative when the short-term horizon is included.

The results using the binary inflation variable are different for the two countries, and also depend on the choice of the maturity combination. On the one hand, the German results suggest

[12] For example, breakpoints are tried from 1974:07 to 1989:06 with *m* equal to two, and from 1973:08 to 1972:05 with *m* equal to ten.

that the term structure is informative about the direction of future inflation changes. The Pseudo-R^2s from the probit regressions are broadly similar to R^2 from the linear regressions. More specifically, while they are generally lower for maturity combinations up to 7-2 years (table 2), they are generally higher for the other maturity combinations, and a maximum Pseudo-R^2 is obtained for the maturity combination 6-3 years at .43. On the other hand, there seems to be more limited information in the U.S. term structure about the direction of future inflation changes. The Pseudo-R^2 obtained in the case of the United States are generally much lower than R^2 from the linear regressions. While the maximum Pseudo-R^2 is obtained for the same maturity combination as in the linear case, it amounts to only .15. For many other maturity combinations the Pseudo-R^2 is very low or even negative. Summarizing, the results regarding the information content of the term structure about the direction of future inflation changes are somewhat mixed.

Stability tests

To see whether the models are stable, two versions of the stability tests are employed. First, the stability tests are applied to both countries assuming that the break points are *unknown*. This has the advantage that the break point dates, if present, are estimated by the statistical procedure. Second, the stability tests are applied to the two countries where we assume that the break points are *known*. This has the advantage that additional information is used, possibly improving the efficiency of the stability tests. Specifically, the assumed breakpoint dates are October 1979 and October 1982 in the case of the United States and December 1974 in the case of Germany.

Using the stability tests with *unknown* breakpoint dates, the results are overall supportive of the hypothesis that the information content of the term structure is stable over time, but the exact results differ between the two countries, and according to the dependent variable (continuous or binary) and test statistic chosen. For example, in the case of the United States, there is no evidence for instability, regardless of the choice of variable and test statistics. While the LM-test statistics are generally higher than those of the PR-test, neither of the two rejects stability for any maturity combination, regardless of the formulation of the dependent variable (continuous or binary). A clear pattern relating the values of the LM-test statistics to the estimated R^2 is not observable. For example, while some of the relatively high LM-statistics are obtained for the very informative maturity combinations between 3-1 and 6-1 years, there does not appear to exist a systematic positive relation between the two statistics, and some of the highest LM-statistics are obtained when the R^2 is very low. But the results are generally more stable if the binary rather than the continuous variable formulation is used, and the relationship between the former and term structure is generally weaker in terms of the R^2 than the relationship between the latter and the term structure. Of course, if the relationship is very weak, the stability of results is also less useful from a monetary policy point of view.

In contrast, when the continuous variable formulation is used for Germany, there is evidence of instability for a few maturity combinations. Specifically, the combinations 8-3, 9-3, 10-3 and 9-4 are significant at the 5 percent level, and 10-2 and 7-3 at the 10 percent level. However, these spreads are cases in which the yield curve is less informative.

The rejections in the case of the continuous variable formulation in the German case do not coincide with rejections in the case of the binary variable formulation. On the contrary, in the case of the binary variable formulation, no evidence of instability in the relationship between the term structure and future inflation changes can be found at the 5 percent level. There are, however, two rejections at the 10 percent level.

Using tests for a single *known* breakpoints, the results also differ between the two countries. In the case of Germany, we do not find evidence of a breakpoint in 1974:10, regardless of the estimation specification.[13] In the case of the United States, there is evidence for a breakpoint in 1979:10 and in 1982:10 for short horizons (table 3). Specifically, in the continuous variable formulation the null hypothesis of stability is rejected for both dates at the 5 percent level of significance for the maturity combinations 2-1 and 3-1. There is also evidence of instability at the 10 percent level for nine other maturity combinations.

There is only one rejection (2-1) at the 5 percent level in the binary variable formulation, and only a few other rejections at the 10 percent level. In general, the maturity combinations where the results suggest stability are generally less informative than those maturity combinations for which evidence of instability is found. Thus, the instability may be a problem for forecasting.[14]

[13] To save space the results are not reported here. They are available on request.

[14] The models examined in the text reflect the standard practice in the literature on inflation forecasts using the yield curve and are derived from the Fisher equation. Thus, they contain no lagged values of inflation even though this history might help forecast future inflation. Models that include lagged inflation may in fact outperform those in the

5.2 Interpretation of results

The results point to a break point in October 1979 in the United States. While the tests for a single breakpoint are consistent with a breakpoint in October 1982, sequential tests for two breakpoints do not provide strong evidence for a *second* breakpoint in October 1982. Focusing on the breakpoint in October 1979, we find that the explanatory power of the yield curve slope about future inflation, measured in terms of the R^2, decreases from .79 before October 1979 to .47 afterwards. And the slope coefficient estimate falls from 3.38 to 1.31. This reflects an increase in the correlation between the expected inflation change and the ex-ante real rate spread. To estimate $\rho^{(m,n)}$ and $q^{(m,n)}$, the procedure outlined in Mishkin (1990) is used, yielding an estimated increase in $\rho^{(m,n)}$ from -.99 to -.82. There was a rise in the ratio of the variability of expected inflation versus ex ante real rate changes $q^{(m,n)}$ (from 1.32 to 1.50), which *ceteris paribus* raises the estimates of the slope parameter $b_1^{(m,n)}$ according to (7). But as long as $q^{(m,n)}$ is greater than one, an increase in $\rho^{(m,n)}$ is associated with a decrease in $b_1^{(m,n)}$. In the event, the rise of $q^{(m,n)}$ was overcompensated by the fall in $\rho^{(m,n)}$, and the estimated $b_1^{(m,n)}$ decreased.

There is more evidence of instability in the case of the continuous variable than in the case of the binary variable. To see why the results depend on the choice of the dependent variable, we suggest the following interpretation. In the continuous variable formulation, an attempt is made to simultaneously estimate the magnitude of upward and downward movements of the inflation rate, and the dates of their occurrence. In contrast, in the binary variable formulation, the focus is just on the latter. This eliminates one source of potential instability; specifically, substantial changes in the estimated slope coefficient $b_1^{(m,n)}$, which are not accompanied by a change in its sign. Indeed, the estimates of $b_1^{(m,n)}$ seemed to have varied substantially, without changing signs.

---

text, at least in-sample, but the qualitative results regarding stability are generally unchanged. For instance, if we add two lagged values of annual inflation to the models of tables 1 and 2, these lags are significant in some cases. With the exception of some very short maturities, the significance of the yield curve spread is not much affected. Moreover, the number of rejections of stability in the equations with these lags is roughly the same as without the lags, as in the tables. An exception is the U.S. data where we find one rejection of stability with lagged inflation where we found none when using only the current term structure slope.

6. Stability tests of equations for predicting real activity

Earlier research has established that equations in which a yield curve spread is used to predict some measure of real activity perform quite well in both Germany and the United States. However, as discussed in section 3, there are theoretical and empirical reasons to question whether these equations are stable over time. In this section, we examine data for Germany and the United States and we focus on (1) linear equations in which the dependent variable is growth in industrial production over 1, 2 or 3 years ahead, and (2) probit equations in which the dependent variable is a binary variable with the same 1, 2 or 3 year horizon that has a value of 1 during recessions and 0 otherwise. We look at the fit of the equations, to confirm that our specifications produce good results comparable to those in the earlier literature, and we apply the LM and PR tests for an unknown break point.

Define the following cumulative growth rates:

$$y_t^{(c,k)} = \frac{1200}{k} \log\left(I_{t+k}/I_t\right),$$

where I_t is the index of industrial production at time *t* and *k* is the forecast horizon in months, and define analogous marginal growth rates as

$$y_t^{(m,k)} = \frac{1200}{k} \log\left(I_{t+k}/I_{t+k-12}\right).$$

Note that, by definition, marginal and cumulative rates are the same for a one-year horizon, i.e., when *k*=12. Our linear equations for each country are of the form:

$$y_t^{(j,k)} = b_0 + b_1 \left(i_t^{(q)} - i_t^{(n)}\right) + \varepsilon_t, \tag{8}$$

where *j*=*c* or *m*. That is, we use the spread between nominal yields on *q* and *n* period bonds to predict the annualized growth rate in industrial production over the subsequent *k* months, or some portion thereof.

The first and third panels of table 4 contain results for the marginal and cumulative forms of this equation for 1, 2, and 3 year horizons (*k* = 12, 24, 36), long rates of 2, 5, and 10 years, and a short rate of 1 year. For each combination, we report three values: the R^2 of the equation, the *p* value of a *sup LM* test and the analogous *p* value of a *sup* PR test. These results cover a sufficiently large set of values so as to give an accurate picture of the dependence of the results on each of the parameters.

First, these results confirm that the performance of the equations is good for both countries, with an R^2 of at least .23 for Germany and .32 for the US with a 1-year horizon. Marginal results for the second year are not as good, with the R^2 falling to about .10 in both cases, and there is hardly any marginal predictive power for the third year. The fit of the cumulative equations gets progressively better as the horizon lengthens, although the gains beyond the 1-year results tend to be small.

In the inflation equations of the previous section, the combination of two Fisher equations implied that interest rate maturities and inflation horizons were exactly matched. In the case of real activity, theory is less helpful in pinning down the appropriate maturity combination for a given horizon, so we experiment with a number of maturity combinations for each predictive horizon. In earlier empirical work, Plosser and Rouwenhorst (1994) find that there is marginal information in interest rates of maturities between 2 and 5 years for predicting changes in industrial production in Germany and the US.

Our results in table 4 suggest that the information in the term structure is captured effectively by any of the maturity combinations reported. For predicting industrial production, there is a slight drop-off in the R^2 when the longer maturity is 5 or 10 years and the predictive horizon is one year, but even in these cases the results for different maturity combinations are fairly comparable. The results for Germany are particularly similar across maturity combinations. Plosser and Rouwenhorst (1996) and Estrella and Mishkin (1997) have shown evidence that the elements of the term structure in Germany are highly correlated, so that a dominant first principal component accounts for most of the explanatory power.

We now turn to the tests of stability for the industrial production equations. In Germany, there seems to be little evidence of instability; the lowest p value for any of the tested equations is .35.

For the US, in contrast, there is stronger evidence – albeit at the 10 percent level – of a break around September 1983 with a one-year horizon. There are at least two possible explanations for this result. One relates to the time series properties of the industrial production series. In the US, there is evidence of a change in trend productivity growth around the end of 1973. The signals from the stability tests indicate a much later date, but the trend change may have some influence on these results.

Perhaps a more compelling connection is to the change in monetary policy regime associated with the advent of Fed Chairman Volcker in late 1979. Some fundamental changes took place in October of that year, but arguably further major changes in the approach to policy were introduced as late as 1982. By 1983, the move to a monetary policy regime more concerned with inflation than earlier regimes had essentially been completed. These dates coincide much more closely with the date identified by the stability tests.

The evidence of instability for the US appears only with a one-year horizon, for which the marginal results are most significant. For horizons of 2 and 3 years, there is no evidence against the stability of the industrial production equations.

The results for the probit models used to predict recessions are just as strong as those are for industrial production, but there is less evidence of instability. The probit model is defined as

$$P\left(R_{t+k}^{(j)}=1\right)=\Phi\left(a_0+a_1\left(i_t^{(q)}-i_t^{(n)}\right)\right) \tag{9}$$

where R is an indicator that has value 1 if the observation is a recession and 0 otherwise. The superscript $j=c$ or m, as before, makes a distinction between a cumulative recession observation (a recession occurred on any month between t and $t+k$) and a marginal recession observation ($t+k$ is a recession month). In table 4, we present evidence for the same horizons and maturity combinations as we have for the industrial production equations.

For Germany, the pseudo R^2s are comparable to those of the industrial production equations. For a one-year horizon, the fit is a bit better when longer-term rates are used in the term structure spread, either with marginal or cumulative recession indicators. The 2-year marginal results, in contrast, are somewhat worse for the recession indicator, whereas the 3-year horizon results are equally weak for the two dependent variables.

For the US, there is a substantial difference in the recession results with a one-year horizon between the marginal and cumulative cases. The marginal R^2s are somewhat lower than those with industrial production, whereas the cumulative R^2s are somewhat better. The 2-year horizon results for the recession equation are of course not as good for the marginal equations, but are in fact much better that the 1-year for the cumulative recession equation.

One feature of the recession equation results is that there is no evidence of instability at any horizon, for any maturity combination, for either country. This represents a clear improvement over the industrial production equations, in which there are important cases in which the equations may not be completely stable over time.[15]

7. Conclusions

In this paper, we have examined the intertemporal stability of models that use the yield curve spread to predict inflation or real activity. While earlier work has shown that such models fit German and United States data well at some horizons, their usefulness in forecasting also depends on whether the models are stable. Since we have summarized the evidence for each type of model in the earlier sections, we focus here on a comparison of the inflation and real activity models and on general lessons we may draw from the overall exercise.

7.1 Comparison of results for predictions of inflation and real activity

Earlier work has shown that models that use the yield curve to predict inflation tend to have poorer performance than those that predict real activity. This is particularly true of short-term inflation predictions, which Mishkin (1990b) found were not very accurate. We have seen here that, in addition, models that predict inflation tend to exhibit more instability than those for real activity.

A similar pattern with regard to stability emerges when we compare continuous and binary models that predict either inflation or real activity. The binary models, in which the possible values of the dependent variable are clearly very limited, tend to perform better than their continuous counterparts. When combined, these two patterns imply that the most stable models are those that predict recessions, whereas the least stable are those that predict the change in inflation. Although the exact results are certainly not identical in Germany and the United States, the overall pattern is very similar in both countries.

---

[15] As in the case of the inflation equations, we experimented with models that add lags of the dependent variable to the equation (see the previous footnote). In contrast with the inflation case, lags of the real activity variables are generally insignificant and do not affect the significance of the yield curve spread. The stability results are therefore essentially unchanged. The one exception is the cumulative recession models, in which the dependent variable (e.g., an indicator of at least one recession month in the next twelve) is highly persistent. These models, however, are not realistic models for forecasting recessions, since they assume that forecasters know whether the previous or very recent months were recessions. However, the beginning of a recession is usually only obvious some time after the recession is underway. Further, recession dating from NBER (U.S.) or OECD (Germany) is typically only available with lags of six months to one year or more.

7.2 General conclusions

One general conclusion that emerges from our results is that models that use the yield curve spread to predict recessions may be employed with a certain level of confidence in their continued reliability. In most other cases, we have identified particular specifications of the models that exhibit some degree of instability. In fact, even in the case of the recession prediction model, we can suggest reasons why the model may not be "structural" in the sense that changes, for instance in monetary policy regime, may lead to changes in the predictive power of the model.

Thus, the main lesson is that all of these models must be used with caution, and that it is advisable to use the methods employed here to test the stability of a particular model if it is to be used for forecasting. Since we cannot rule out instability by theoretical arguments, it becomes an empirical issue. Fortunately, we have at our disposal various tests that can help us decide how much trust to place on a given model.

References

Andrews, D.W.K., and R. C. Fair (1988), "Inference in Non-linear Econometric Models with Structural Change," *Review of Economic Studies*, 55:615-640.

Andrews, D.W.K. (1993), "Tests for Parameter Instability and Structural Change with Unknown Change Point", *Econometrica*, Vol. 61, No. 4, 821-856, July.

Bai, J., and P. Perron (1998), "Estimating and Testing Linear Models with Multiple Structural Changes", *Econometrica*, Vol. 66, No. 1, 47-78, January.

Berk, J.M. (1998), "The Information Content of the Yield Curve for Monetary Policy: A Survey," *De Economist* 146:303-320.

Bernard, H., and S. Gerlach (1998), "Does the Term Structure Predict Recessions? The International Evidence," CEPR Discussion Paper No. 1892.

Blanchard, O.J., and S. Fisher (1989), *Lectures on Macroeconomics*, Cambridge, Mass.: MIT Press.

Bomhoff, E.J. (1994), *Financial Forecasting for Business and Economics*, Academic Press.

Chen, N. (1991), "Financial Investment Opportunities and the Macroeconomy," *Journal of Finance* 46:529-554.

Davis, E.P., and G. Fagan (1997), "Are Financial Spreads Useful Indicators of Future Inflation and Output Growth in E.U. Countries?" *Journal of Applied Econometrics* 12:701-714.

Davis, E.P., and S.G.B. Henry (1994), "The Use of Financial Spreads as Indicator Variables: Evidence for the United Kingdom and Germany," *I.M.F. Staff Papers* 41:517-525.

DeLong, D.M. (1981), "Crossing Probabilities for a Square Root Boundary by a Bessel Process," *Communications in Statistics—Theory and Method*, A10(21):2197-2213.

Dornbusch, R. (1980), *Open Economy Macroeconomics*, New York: Basic Books.

Dotsey, M. (1998), "The Predictive Content of The Interest Rate Term Spread for Future Economic Growth," Federal Reserve Bank of Richmond *Economic Quarterly* 84:31-51.

Estrella, A. (1997), "Why Do Interest Rates Predict Macro Outcomes? A Unified Theory of Inflation, Output, Interest, And Policy," Federal Reserve Bank of New York Research Paper No. 9717.

Estrella, A. (1998), "A New Measure of Fit for Equations with Dichotomous Dependent Variables," *Journal of Business And Economic Statistics* 16:198-205.

Estrella, A., and A. P. Rodrigues (1998), “Consistent Covariance Matrix Estimation in Probit Models with Autocorrelated Errors,” Federal Reserve Bank of New York Staff Report No. 39.

Estrella, A., and F. S. Mishkin (1997), “The Predictive Power of the Term Structure of Interest Rates in Europe and The United States: Implications for The European Central Bank,” *European Economic Review* 41:1375-1401.

Estrella, A., and F. S. Mishkin (1998), “Predicting U.S. Recessions: Financial Variables as Leading Indicators,” *Review of Economics and Statistics* 80:45-61.

Estrella, A., and G. Hardouvelis (1991), “The Term Structure as a Predictor of Real Economic Activity,” *Journal of Finance* 46:555-576.

Filardo, A. J. (1999), “How Reliable are Recession Prediction Models?” Federal Reserve Bank of Kansas City *Economic Review*, 83:35-55.

Fisher, M., D. Nychka and D. Zervos (1995), “Fitting The Term Structure of Interest Rates with Smoothing Splines”, Federal Reserve Board, Washington, D.C., *FEDS Working Paper*, 95-1.

Gamber, E. N. (1996), “The Policy Content of the Yield Curve Slope”, *Review of Financial Economics,* Vol. 5, No. 2, Pp. 163-179.

Gerlach, S. (1997), “The Information Content of the Term Structure: Evidence for Germany,” *Empirical Economics*, 22(2): 161-179.

Ghysels, E., and A. Hall (1990), “A Test for Structural Stability of Euler Conditions Parameters Estimated via the Generalized Method of Moments Estimator: Small Sample Properties”, *International Economic Review*, Vol. 31, No. 2, May.

Ghysels, E., A. Guay and A. Hall (1997), “Predictive Tests for Structural Change with Unknown Breakpoint,” *Journal of Econometrics* 82:209-233.

Harvey, C. (1988), “The Real Term Structure and Consumption Growth,” *Journal of Financial Economics* 22.

Huizinga, J., and F. S. Mishkin (1986), “Monetary Policy Regime Shifts and the Unusual Behavior of Real Interest Rates”, *Carnegie-Rochester Conference Series on Public Policy* 24, North Holland, Pp. 231-274.

Jorion, P., and F. Mishkin (1991), “A Multicountry Comparison of Term-Structure Forecasts at Long Horizons”, *Jounal of Financial Economics*, 29, 59-80, North-Holland.

Kozicki, S. (1997), “Predicting Real Growth and Inflation with The Yield Spread,” Federal Reserve Bank of Kansas City *Economic Review* 82:39-57.

Kydland, F. E., and E. C. Prescott (1988), “The Workweek of Capital and its Cyclical Implications,” *Journal of Monetary Economics* 21:343-360.

Laurent, R. D. (1988), “An Interest Rate-Based Indicator of Monetary Policy,” Federal Reserve Bank of Chicago *Economic Perspectives* 12:3-14.

Laurent, R. D. (1989), “Testing the Spread,” Federal Reserve Bank of Chicago *Economic Perspectives* 13.

McCulloch, J.H., 1975. “The Tax-Adjusted Yield Curve,” *Journal of Finance*, June, pp. 811-30.

McCulloch, J. H. and H-C. Kwon (1993), “U.S. Term Structure Data, 1947-1991,” Ohio State Working Paper #93-6, March 1993.

Mishkin, F. S. (1988), “The Information in the Term Structure: Some Further Results”, *Journal of Applied Econometrics*, Vol. 3, 307-314.

Mishkin, F. S. (1990), “The Information in The Longer Maturity Term Structure about Future Inflation”, The *Quarterly Journal Of Economics*, Vol. CV, Cambridge.

Mishkin, F. S. (1990), “What Does the Term Structure Tell Us about Future Inflation?” *Journal of Monetary Economics*, 25, 77-95, North-Holland.

Mishkin, F. S. (1991), “A Multi-Country Study of the Information in the Shorter Maturity Term Structure about Future Inflation”, *Journal of International Money and Finance*, 10, 2-22.

Plosser, C. I., and K. G. Rouwenhorst (1994), “International Term Structures and Real Economic Growth,” *Journal of Monetary Economics* 33: 133-155.

Romer, D. (1996), *Advanced Macroeconomics*, New York: McGraw-Hill.

Schich, S. (1996), “Alternative Specifications of the German Term Structure and its Information Content Regarding Inflation,” Discussion Paper 8/96, Deutsche Bundesbank, October.

Schich, S. (1999a), “The Information Content of the German Term Structure Regarding Inflation,” *Applied Financial Economics*, 9:385-395.

Schich, S. (1999b), “What the Yield Curves Say about Inflation: Does It Change Over Time? OECD Economics Department Working Paper 227, December 1999.

Sédillot, F. (1999), “La pente des taux contient-elle de l’information sur l’activité économique future?”, *Notes d’études et de recherche*, No. 67, Banque de France, June.

Stock, J.H., and M.W. Watson (1989), “New Indexes of Coincident and Leading Indicators,” In O.J. Blanchard And S. Fischer, Eds., *NBER Macroeconomics Annual*, Cambridge MA: MIT Press.

Taylor, J. B. (1993), “Discretion Versus Policy Rules in Practice,” *Carnegie-Rochester Conference Series On Public Policy* 39:195-214.

Table 1. The yield curve as a predictor of future inflation (unknown change point, continuous dependent variable, 25 percent of sample dropped at each end)

Spreads (*m-n* years)	Sample	Interval	**Germany** Regression parameters $b_1^{(m,n)}$ [t stat]	R^2	Stability tests (p values, dates) LM	PR	**United States** Regression parameters $b_1^{(m,n)}$ [t stat]	R^2	Stability tests (p values, dates) LM	PR
2 – 1	1967:01 - 1996:12	74:07 - 89:06	.23 [.96]	.02	.84 (80:02)	.94 (80:02)	1.40 [4.36]	.24	.23 (78:12)	.49 (79:01)
3 - 1	1967:01 - 1995:12	74:04 - 88:09	.50 [2.06]	.08	.94 (80:02)	.97 (76:09)	1.66 [4.36]	.34	.17 (78:12)	.49 (79:02)
4 - 1	1967:01 - 1994:12	74:01 - 87:12	.75 [3.66]	.16	.83 (79:12)	.91 (76:09)	1.68 [3.67]	.38	.19 (78:06)	.58 (78:12)
5 - 1	1967:01 - 1993:12	73:10 - 87:03	.97 [4.98]	.25	.67 (79:12)	.85 (76:07)	1.55 [3.13]	.36	.26 (78:01)	.70 (76:12)
6 - 1	1967:01 - 1992:12	73:07 - 86:06	1.07 [5.31]	.29	.61 (79:11)	.87 (76:07)	1.38 [2.75]	.32	.37 (77:09)	.78 (75:12)
7 - 1	1967:01 - 1991:12	73:04 - 85:09	1.07 [4.68]	.28	.60 (79:03)	.90 (73:04)	1.50 [3.45]	.33	.41 (80:04)	.84 (73:04)
8 - 1	1967:01 - 1990:12	73:01 - 84:12	.96 [3.87]	.24	.63 (79:03)	.96 (73:01)	1.53 [3.98]	.33	.45 (80:04)	.89 (73:01)
9 - 1	1967:01 - 1989:12	72:10 - 84:03	.88 [3.73]	.23	.53 (79:03)	.98 (72:10)	1.48 [3.99]	.30	.48 (80:04)	.92 (72:10)
10 - 1	1967:01 - 1988:12	72:07 - 83:06	.83 [3.78]	.22	.31 (79:03)	.99 (72:07)	1.35 [3.35]	.25	.50 (80:04)	.94 (72:07)
3 - 2	1967:01 - 1995:12	74:04 - 88:09	.88 [2.77]	.13	.98 (76:06)	.98 (76:06)	1.91 [3.44]	.28	.28 (78:01)	.64 (79:02)
4 - 2	1967:01 - 1994:12	74:01 - 87:12	1.27 [5.25]	.237	.76 (76:06)	.93 (76:06)	1.79 [2.72]	.27	.24 (77:11)	.61 (77:11)
5 - 2	1967:01 - 1993:12	73:10 - 87:03	1.55 [7.34]	.332	.63 (76:06)	.89 (76:06)	1.47 [2.18]	.20	.27 (77:06)	.70 (77:08)
6 - 2	1967:01 - 1992:12	73:07 - 86:06	1.64 [7.79]	.362	.46 (79:03)	.91 (73:07)	1.19 [1.81]	.15	.31 (77:01)	.79 (77:03)
7 - 2	1967:01 - 1991:12	73:04 - 85:09	1.62 [7.06]	.340	.27 (79:03)	.91 (73:04)	1.25 [2.01]	.14	.35 (76:09)	.85 (76:08)
8 - 2	1967:01 - 1990:12	73:01 - 84:12	1.39 [5.90]	.276	.22 (79:03)	.96 (73:01)	1.25 [2.34]	.14	.39 (80:03)	.90 (73:01)
9 - 2	1967:01 - 1989:12	72:10 - 84:03	1.17 [5.36]	.229	.17 (79:03)	.98 (79:08)	1.13 [2.30]	.11	.45 (76:06)	.92 (72:10)
10 - 2	1967:01 - 1988:12	72:07 - 83:06	.97 [4.88]	.180	**.09** (79:04)	.99 (79:09)	.90 [1.67]	.07	.43 (72:07)	.93 (72:07)
4 - 3	1967:01 - 1994:12	74:01 - 87:12	1.74 [6.07]	.279	.72 (75:07)	.94 (75:07)	1.35 [1.84]	.11	.35 (77:02)	.70 (77:11)

Table 1 (continued). The yield curve as a predictor of future inflation (unknown change point, continuous dependent variable, 25 percent of sample dropped at each end)

Spreads (*m-n* years)	Sample	Interval	Germany				United States			
			Regression parameters		Stability tests (p values, dates)		Regression parameters		Stability tests (p values, dates)	
			$b_1^{(m,n)}$ [t stat]	R^2	LM	PR	$b_1^{(m,n)}$ [t stat]	R^2	LM	PR
5 - 3	1967:01 - 1993:12	73:10 - 87:03	2.05 [8.14]	.39	.60 (79:03)	.96 (73:10)	.76 [.97]	.04	.31 (76:12)	.88 (77:03)
6 - 3	1967:01 - 1992:12	73:07 - 86:06	2.04 [8.17]	.38	.34 (79:03)	.92 (73:07)	.37 [.45]	.01	.36 (76:09)	.88 (74:10)
7 - 3	1967:01 - 1991:12	73:04 - 85:09	1.94 [7.39]	.33	**.07** (79:03)	.95 (75:08)	.25 [.28]	.00	.45 (76:05)	.92 (76:07)
8 - 3	1967:01 - 1990:12	73:01 - 84:12	1.58 [5.73]	.24	**.04** (79:03)	.97 (79:10)	.19 [.24]	.00	.37 (78:05)	.91 (76:02)
9 - 3	1967:01 - 1989:12	72:10 - 84:03	1.18 [5.01]	.17	**.01** (79:03)	.98 (79:10)	.07 [.10]	.00	.22 (78:05)	.92 (72:10)
10 - 3	1967:01 - 1988:12	72:07 - 83:06	.82 [4.22]	.10	**.01** (79:03)	.99 (79:10)	-.22 [-.31]	.00	.17 (78:06)	.93 (72:07)
5 - 4	1967:01 - 1993:12	73:10 - 87:03	2.34 [7.55]	.38	.86 (79:04)	1.00 (87:03)	-.10 [-.11]	.00	.50 (74:12)	.72 (74:12)
6 - 4	1967:01 - 1992:12	73:07 - 86:06	2.14 [6.74]	.33	.47 (79:03)	.98 (79:08)	-.45 [-.46]	.01	.46 (78:05)	.83 (74:10)
7 - 4	1967:01 - 1991:12	73:04 - 85:09	1.86 [5.52]	.24	.12 (79:03)	.97 (79:10)	-.57 [-.57]	.02	.43 (78:05)	.95 (75:11)
8 - 4	1967:01 - 1990:12	73:01 - 84:12	1.34 [3.66]	.14	.14 (79:03)	.97 (73:03)	-.60 [-.73]	.02	.33 (78:04)	.92 (75:03)
9 - 4	1967:01 - 1989:12	72:10 - 84:03	.81 [2.57]	.06	**.03** (79:03)	.95 (72:12)	-.72 [-1.13]	.03	.19 (78:05)	.93 (72:10)
10 - 4	1967:01 - 1988:12	72:07 - 83:06	.33 [1.36]	.01	.13 (79:04)	.95 (72:07)	-1.05 [-1.59]	.06	.18 (78:05)	.94 (72:07)

Explanation: Regression parameters are based on full sample estimates. LM and PR are the Lagrange multiplier of Andrews (1993) and the predictive test of Ghysels, Guay and Hall (1997) for one break point, respectively. The first row shows the *p* value of the supremum of the test and the second row the implied breakpoint date.

Table 2. The yield curve as a predictor of future inflationary pressure (unknown change point, binary dependent variable, 25 percent of sample dropped at each end)

Spreads (*m*-*n* years)	Sample	Germany				United States			
		Regression parameters		Stability tests		Regression parameters		Stability tests	
		t-statistics	Pseudo-R^2	LM	PR	t-statistics	Pseudo-R^2	LM	PR
2 - 1	1967:01 - 1996:12	.68	.01	.85 (80:02)	.86 (74:07)	2.55	.09	.28 (82:12)	.40 (79:01)
3 - 1	1967:01 - 1995:12	1.20	.03	.79 (79:12)	.94 (76:09)	2.36	.13	.34 (78:08)	.54 (78:08)
4 - 1	1967:01 - 1994:12	3.06	.13	.80 (76:06)	.90 (76:01)	2.33	.15	.25 (78:03)	.54 (80:05)
5 - 1	1967:01 - 1993:12	4.22	.18	.76 (75:08)	.90 (75:05)	1.62	.10	.31 (78:01)	.75 (78:01)
6 - 1	1967:01 - 1992:12	4.77	.18	.57 (75:05)	.82 (73:07)	.99	.05	.37 (86:06)	.81 (73:07)
7 - 1	1967:01 - 1991:12	4.56	.18	.52 (84:09)	.86 (73:04)	1.57	.09	.48 (77:08)	.82 (73:04)
8 - 1	1967:01 - 1990:12	4.43	.20	.50 (75:05)	.93 (73:02)	1.92	.11	.50 (73:01)	.85 (73:01)
9 - 1	1967:01 - 1989:12	4.20	.19	.48 (79:04)	.94 (73:02)	1.73	.10	.47 (77:04)	.90 (73:12)
10 - 1	1967:01 - 1988:12	1.98	.11	.37 (79:03)	.96 (73:02)	1.32	.06	.51 (76:10)	.91 (73:08)
3 - 2	1967:01 - 1995:12	2.17	.12	.50 (76:05)	.72 (76:05)	2.77	.20	.28 (88:02)	.66 (78:01)
4 - 2	1967:01 - 1994:12	3.84	.23	.40 (75:09)	.81 (75:09)	2.34	.13	.33 (87:12)	.66 (77:11)
5 - 2	1967:01 - 1993:12	4.67	.31	.39 (75:05)	.81 (75:05)	1.11	.05	.32 (87:03)	.76 (77:08)
6 - 2	1967:01 - 1992:12	5.29	.34	.43 (73:07)	.83 (73:07)	.48	.01	.31 (86:06)	.81 (77:04)
7 - 2	1967:01 - 1991:12	5.15	.34	.37 (73:04)	.84 (73:04)	.95	.03	.33 (85:09)	.87 (77:02)
8 - 2	1967:01 - 1990:12	4.84	.29	.31 (79:03)	.92 (73:02)	1.56	.06	.44 (76:10)	.87 (73:01)
9 - 2	1967:01 - 1989:12	4.58	.27	.26 (79:03)	.97 (72:10)	1.34	.03	.36 (76:08)	.89 (72:10)
10 - 2	1967:01 - 1988:12	3.14	.22	.30 (79:04)	.96 (72:09)	.66	.01	.20 (76:04)	.91 (73:08)

Table 2 (continued). The yield curve as a predictor of future inflationary pressure (unknown change point, binary dependent variable, 25 percent of sample dropped at each end)

Spreads (*m-n* years)	Sample	Germany				United States			
		Regression parameters		Stability tests		Regression parameters		Stability tests	
		t-statistics	Pseudo-R^2	LM	PR	t-statistics	Pseudo-R^2	LM	PR
4 - 3	1967:01 - 1994:12	4.71	.24	.63 (75:05)	.90 (75:05)	1.98	.10	.36 (87:06)	.89 (77:03)
5 - 3	1967:01 - 1993:12	4.13	.41	.53 (83:09)	.90 (73:10)	.63	.02	.35 (86:12)	.90 (77:01)
6 - 3	1967:01 - 1992:12	4.17	.43	.52 (79:04)	.87 (73:07)	-.01	.00	.26 (86:03)	.90 (76:10)
7 - 3	1967:01 - 1991:12	4.36	.41	.43 (79:03)	.91 (73:04)	-.09	.00	.43 (85:09)	.91 (76:05)
8 - 3	1967:01 - 1990:12	4.24	.38	.42 (79:03)	.98 (73:01)	.38	.01	.42 (76:03)	.89 (76:03)
9 - 3	1967:01 - 1989:12	4.08	.35	.59 (79:03)	.99 (79:12)	.17	.00	.29 (76:02)	.87 (76:02)
10 - 3	1967:01 - 1988:12	4.66	.23	.33 (76:06)	.97 (74:10)	.06	.00	.32 (75:12)	.91 (72:07)
5 - 4	1967:01 - 1993:12	4.96	.41	.41 (79:03)	.95 (83:05)	-.01	.00	.52 (86:06)	.73 (74:12)
6 - 4	1967:01 - 1992:12	4.92	.37	.38 (79:03)	.97 (81:10)	-.44	.01	.48 (76:02)	.94 (74:12)
7 - 4	1967:01 - 1991:12	4.90	.26	.83 (79:03)	.94 (74:01)	-.33	.01	.42 (75:11)	.94 (75:11)
8 - 4	1967:01 - 1990:12	3.82	.17	.66 (73:10)	.89 (73:10)	-.06	.00	.43 (75:07)	.88 (75:08)
9 - 4	1967:01 - 1989:12	3.00	.09	**.07** (82:05)	.76 (73:10)	-1.05	.02	.29 (75:09)	.77 (75:09)
10 - 4	1967:01 - 1988:12	2.11	.04	**.08** (82.02)	.75 (73:11)	-1.40	.06	.28 (75:11)	.78 (75:11)

Explanation: Regression parameters are based on full sample estimates. LM and PR as in table 1. The first row shows the *p* value of the supremum of the test and the second row the implied breakpoint date. In the case of the 9-1, 10-1, 10-2 and 9-3 year spread, 30 percent of the sample had to be dropped at each end, and in the case of the 9-4 and 10-4 year spread 37.5 percent and 40 percent, respectively, to avoid that all dependent variables were equal to one.

Table 3. The yield curve as a predictor of future inflation

Tests for a single known break-point for the United States (*p* value)

	Continuous dependent variable				Binary dependent variable			
	LM		PR		LM		PR	
	1979:10	1982:10	1979:10	1982:10	1979:10	1982:10	1979:10	1982:10
2 -1	**.04**	**.03**	**.09**	.17	**.06**	**.05**	**.07**	**.09**
3 - 1	**.04**	**.05**	**.09**	.17	**.08**	**.07**	.11	.12
4 - 1	**.06**	**.08**	.15	.21	**.10**	.11	.17	.19
5 - 1	**.09**	.11	.22	.29	.15	.22	.23	.29
6 - 1	.11	.12	.28	.34	.18	.19	.30	.36
7 - 1	.11	**.09**	.38	.42	.18	.14	.37	.38
8 - 1	.15	.12	.48	.55	.21	.15	.44	.48
9 - 1	.18	.14	.53	.60	.24	.15	.49	.52
10 - 1	.20	.14	.55	.60	.29	.13	.53	.55
3 - 2	**.07**	**.10**	.16	.19	.21	.29	.18	.24
4 - 2	**.10**	.14	.23	.25	.22	.35	.19	.29
5 - 2	.15	.19	.26	.33	.25	.41	.25	.38
6 - 2	.19	.22	.30	.41	.22	.35	.30	.43
7 - 2	.17	.21	.39	.41	.29	.30	.39	.43
8 - 2	.18	.26	.49	.64	.41	.51	.46	.65
9 - 2	.25	.35	.56	.73	.53	.35	.54	.62
10 - 2	.35	.38	.62	.68	.58	.15	.60	.65
4 - 3	.20	.28	.30	.39	.57	.90	.46	.81
5 - 3	.53	.63	.43	.67	.51	.92	.46	.86
6 - 3	.85	.98	.75	.97	.38	.87	.39	.81
7 - 3	.73	.94	.65	.69	.38	.84	.41	.66
8 - 3	.60	.86	.62	.54	.35	.80	.42	.51
9 - 3	.55	.81	.65	.58	.42	.80	.48	.62
10 - 3	.41	.74	.65	.94	.53	.95	.62	.81
5 - 4	.44	.69	.50	.71	.48	.91	.42	.81
6 - 4	.19	.44	.30	.46	.39	.87	.39	.68
7 - 4	.22	.56	.39	.49	.34	.84	.41	.58
8 - 4	.23	.62	.47	.52	.32	.82	.42	.52
9 - 4	.19	.59	.50	.56	.51	.63	.49	.77
10 - 4	.15	.59	.54	.84	.73	.32	.64	.57

Note: Sample dates and full sample regression parameter estimates as in tables 1 and 2. LM and PR as in table 1.

Table 4: Results of stability tests for predictions of real activity

R^2 and p values of LM and PR break tests with $\pi_0 = .25$
Sample: January 1967 to December 1998

	Maturities	1 year marginal			1 year cumulative			2 year marginal			2 year cumulative			3 year marginal			3 year cumulative		
		R^2	LM	PR	R^2	LM	PR	R^2	LM	PR	R^2	LM	PR	R^2	LM	PR	R^2	LM	PR
Industrial	2-1	.24	.57	.70	—†	—	—	.09	.85	.95	.28	.81	.88	.02	.67	.74	.29	.86	.77
production,	5-1	.24	.47	.64	—	—	—	.09	.87	.94	.28	.72	.83	.02	.56	.71	.29	.82	.78
Germany	10-1	.23	.39	.60	—	—	—	.09	.90	.92	.27	.61	.74	.02	.41	.67	.28	.79	.81
Recessions,	2-1	.24	.52	.52	.27	.23	.41	.02	.75	.91	.21	.32	.76	.03	.69	.71	.18	.28	.74
Germany	5-1	.27	.42	.42	.30	.26	.45	.02	.69	.92	.25	.32	.80	.02	.63	.66	.21	.26	.74
	10-1	.30	.32	.31	.33	.33	.54	.03	.59	.92	.29	.36	.83	.02	.55	.61	.25	.25	.73
Industrial	2-1	.35	.11*	.18	—	—	—	.10	.92	.97	.40	.25	.45	.01	.51	.54	.29	.76	.95
production,	5-1	.32	**.09***	.23	—	—	—	.11	.38	.65	.40	.22	.44	.01	.55	.62	.31	.46	.90
US	10-1	.32	**.08***	.24	—	—	—	.09	.26	.66	.37	.22	.51	.01	.60	.67	.27	.55	.93
Recessions,	2-1	.24	.62	.66	.36	.86	.96	.05	.25	.60	.50	.40	.74	.01	.38	.40	.49	.52	.46
US	5-1	.27	.82	.75	.40	.97	.91	.07	.24	.19	.57	.59	.84	.01	.41	.51	.53	.62	.45
	10-1	.29	.90	.79	.43	.98	.84	.06	.25	.66	.58	.52	.89	.01	.42	.60	.53	.72	.46

*Sup of LM statistic attained in September 1983.

Key: † denotes that the marginal and cumulative equations are the same in the case of real growth over one year. LM and PR as in table 1.