The Predictive Content of the Interest Rate Term Spread for Future Economic Growth

Michael Dotsey

Predicting economic activity is important for numerous reasons. It is important for business firms because it aids in deciding how much capacity will be needed to meet future demand. It is important for various government agencies when forecasting budgetary surpluses or deficits. And it is important for the Federal Reserve (the Fed) in deciding the stance of current monetary policy. One set of variables that are potentially useful in forecasting economic activity are financial variables.

Financial market participants are forward-looking, and as a result the prices of various securities embody expectations of future economic activity. This pricing behavior implies that data from financial markets may reasonably be expected to help forecast the growth rate of the economy. Using financial variables to aid in economic projections, therefore, is fairly commonplace. In particular, the yield curve spread between long- and short-term interest rates has received a lot of recent attention. Although not the first to consider the implications that the spread has for predicting economic activity, Stock and Watson (1989) provided much of the impetus for further research by finding that the spread was an important component of their newly constructed index of leading economic indicators. Estrella and Hardouvelis (1991) also thoroughly document the significant relationship between interest rate spreads and future output growth.

Unfortunately, one of the spread's major predictive failures occurred immediately after the publication of these influential articles. Namely, the spread failed to predict the 1990–91 recession. In light of that occurrence, a number

I wish to thank Mark Watson for several useful discussions and Jeff Walker for valuable research assistance. I also wish to thank Arturo Estrella for providing some of his computer programs. Tom Humphrey, Yash Mehra, Pierre Sarte, and Roy Webb provided a number of useful suggestions. The views expressed herein are the author's and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

of papers reinvestigated the spread's predictive content. Among these are the works of Estrella and Mishkin (1997, 1998), Haubrich and Dombrosky (1996), Plosser and Rouwenhorst (1994), and Dueker (1997). These studies mainly concluded that the spread still contains significant information for predicting economic activity.

This article reinforces the view that the spread is generally a useful variable in predicting future growth in real GDP but also indicates that it has become less useful in recent years. In particular, the recent accuracy of the spread's prediction of GDP growth, both in-sample and out-of-sample, is less precise than over earlier sample periods. In fact, adding the spread to a VAR containing lagged output growth and short-term interest rates increases the root mean squared error of the out-of-sample forecast errors over the period 1985 to 1997.

After briefly reviewing relevant literature, I informally characterize the joint behavior of output growth and the spread. From this characterization it is clear that there is a relationship between the two variables, although that relationship is far from perfect. I then attempt to expand on the existing literature by analyzing the predictive content of the spread along a number of new dimensions. In particular, I examine whether there are nonlinearities in the relationship and whether the predictive content of the spread is closely associated with the stance of monetary policy. Further, the results here indicate the important differences between evaluations based on in-sample versus out-of-sample predictive power. Presumably, it is the latter that is most relevant for judging the ability to forecast.

1. RELATED LITERATURE

There is a wide and growing literature that examines the term structure of interest rates' predictive content for economic activity. The review given here is selective and focuses on articles that significantly influenced the statistical tests carried out later in this article.[1] One of the most influential studies is that of Stock and Watson (1989), which systematically attempts to construct a new index of leading economic indicators. Their approach is to examine combinations of 55 various macroeconomic variables and select the combination that best predicts future economic activity. To make their search manageable, they limit their index to seven variables—as does the current National Bureau of Economic Research (NBER) list of leading indicators. One of the variables that is an important component of their leading economic indicator is the spread between the ten-year and one-year U.S. Treasury bond. Because their search

[1] Other papers that look at the predictive content of the spread for real economic activity include Laurent (1988, 1989), Harvey (1988), Frankel and Lown (1994), Bonser-Neal and Morley (1997), and Kozicki (1997).

for a leading indicator series is fairly exhaustive, the finding that the yield spread is an important element of their indicator lent impetus to exploring the predictive content of this variable in isolation.

One article that supports using the spread alone in predicting economic growth is by Estrella and Hardouvelis (1991). Examining data over the period 1955 to 1988, they document that the spread between the yield on the ten-year Treasury bond and the three-month Treasury bill is a useful predictor of both cumulative economic growth up to four years in the future and marginal economic growth rates up to seven quarters in the future. They also find that the spread contains information for future economic growth not already embodied in the current level of real interest rates, in current economic growth, in the current growth rate of the index of leading economic indicators, or in the inflation rate. Further, they find the spread useful in forecasting the probability of a recession. An important implication of this article is its rule of thumb applicability. By concentrating largely on the spread's predictive content, the article's forecasting message is easy to apply and doesn't require sophisticated econometric tools or the application of large economic data sets.

Immediately after these two articles were written, the economy provided another test of the predictive power of the spread. In this case, although the spread narrowed and predicted somewhat weaker economic activity, it failed to predict the 1990–91 recession. As a result, other researchers revisited the issue. For example, Estrella and Mishkin (1997) examine the period 1973 to 1994 and find that the basic results of Estrella and Hardouvelis (1991) continue to hold in the United States as well as in a number of European countries. Haubrich and Dombrosky (1996) also find that over the period 1961:1 to 1995:3, the yield spread is a relatively accurate predictor of four-quarter economic growth but that its predictive content has changed over time. For example, they find that the yield spread was not a very good predictor of economic activity over the period 1985 to 1995.

Plosser and Rouwenhorst (1994) examine the predictive content of the spread between various maturities of long-term bonds and the three-month bill rates for a variety of countries over the period August 1973 to December 1988. A novel feature of their paper is the use of discount equivalent yields and the fact that they match the maturity structure of the interest rate spread with the forecast horizon being studied. They find that the term spread has significant in-sample predictive content for future cumulative changes in industrial production of up to five years but that this predictability is largely due to the spread's ability to predict activity at horizons of up to two years. Also, by looking at the effects of the term spread on forward rates, they are able to show that information in the longer end of the term structure is useful in predicting future economic activity.

Other papers have concentrated on another feature of the Estrella and Hardouvelis (1991) paper, namely, the ability of the term spread to signal the probability of a recession. Estrella and Mishkin (1998), for example, using data over the period 1959:1 to 1995:1, show that the spread between the yield on the ten-year and three-month Treasury securities is the best out-of-sample predictor of the probability of a recession occurring in the next four quarters. For shorter horizons, they find that adding movements in various stock price indexes improves forecast accuracy. Dueker (1997) also finds that the yield spread is a relatively good in-sample predictor of recessions. He adds a lagged-state-of-the-economy variable and finds that it helps his model predict the severity and duration of big recessions; but as in other studies, he finds that milder recessions are harder to predict.

2. THE SPREAD AND ECONOMIC ACTIVITY

Before beginning a detailed statistical analysis, it is instructive to take a more casual view of the data and to consider why the spread may be a good predictor of economic activity. Figure 1 displays the behavior of (1) the spread between the discount equivalent yield on the ten-year U.S. Treasury bond and the three-month Treasury bill and (2) the four-quarter growth rate of real GDP. The NBER recession dates are shaded in. The first thing to notice is that movements in the spread precede changes in real GDP growth and that these two series are positively correlated. Thus the spread seems to indicate whether future output growth will be strong or weak. Also, prior to a number of business cycle peaks, namely, the 1969:4, 1973:4, 1980:1, and 1981:3 peaks, the spread inverted with the short-term rate exceeding the rate on the long bond. The spread also remained negative over most of these recessions. The spread flattened significantly prior to the 1990:3 peak, but as the recession progressed, the yield curve steepened. Such behavior typically indicates renewed strength in the economy. Consequently, it appears that the spread did not perform quite as well in this episode. Less-than-perfect performance is also observed around the 1957:3 and 1960:2 peaks. Further, one notices that the spread became negative in late 1967, and the economy remained strong.

Figures 2a and 2b highlight the behavior of the spread around business cycle peaks and troughs. Figure 2a reemphasizes the point that prior to most recessions, the yield curve becomes inverted and usually remains inverted for a good part of the recession. Figure 2b indicates that the yield curve, although inverted during most recessions, begins to steepen prior to each business cycle trough. Thus it seems reasonable that economic forecasters would find the yield spread a useful but imperfect guide of future economic activity.

The imprecision associated with the spread can be gauged by looking at Table 1. In this table, I record the number of true and false signals of recessions over the period 1956 to 1996. I look at two definitions of a signal. The first definition labels the signal as true if the yield curve is inverted and a recession occurs either contemporaneously or within one-to-four quarters of the signal.

Figure 1 Real GDP Growth and Spread 1956–1997

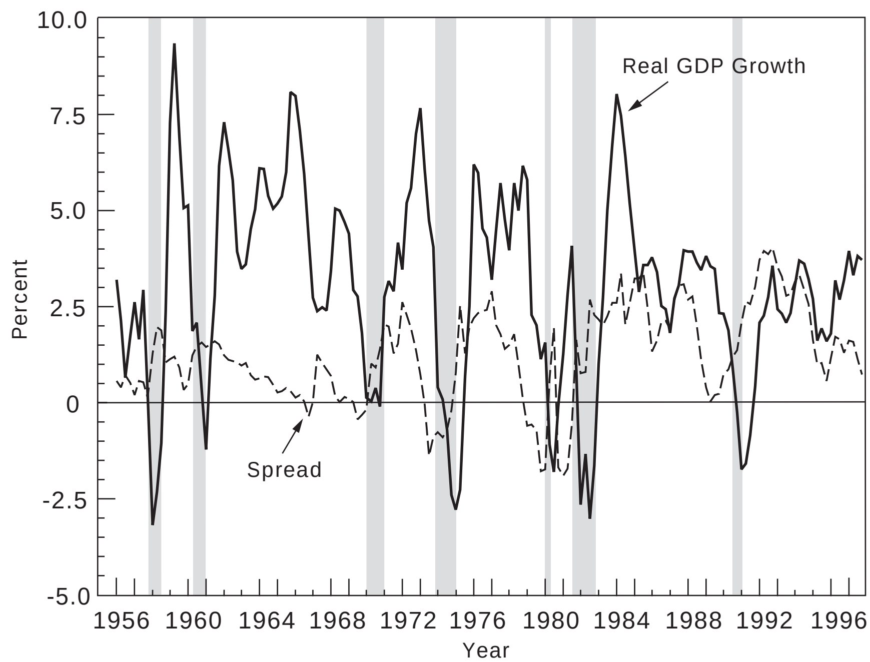


Table 1 Signal Value of the Term Spread

	Spread < 0	**Spread < 0.25**	**Spread < 0 $ff(t) - ff(t-2) > 0.5$**	**Spread < 0.25 $ff(t) - ff(t-2) > 0.5$**
True Signals	18	23	15	17
False Signals	2	13	2	8
Total Signals	20	35	17	25
True Signals within recession	8	6	2	2
Pr (True Signal)	0.833	0.552	0.867	0.652

Notes: The sample is quarterly from 1955:1 to 1995:4. The spread equals the ten-year Treasury bond rate minus the monthly average of the three-month Treasury bill rate. $ff(t)$ is the federal funds rate at time t.

Figure 2a Behavior of Spread around Business Cycle Peaks

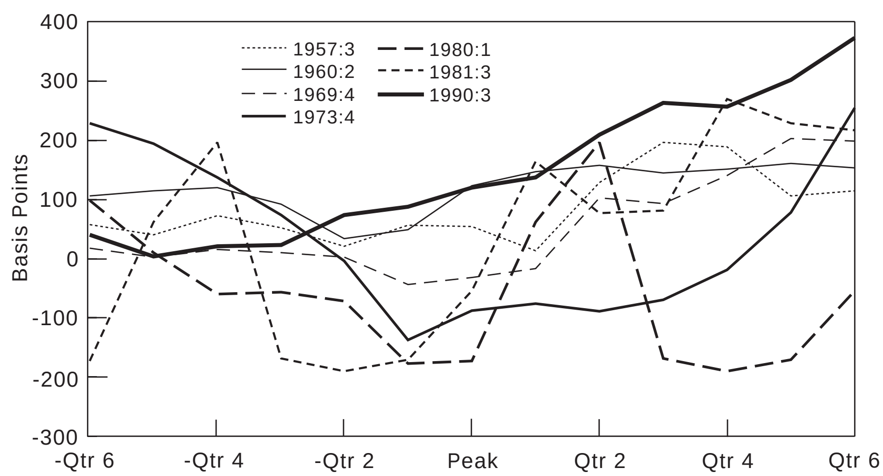
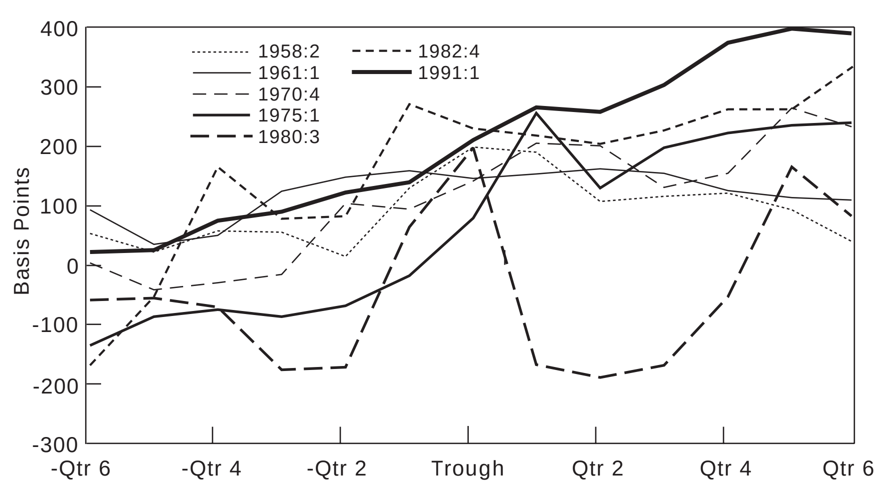


Figure 2b Behavior of Spread around Business Cycle Troughs



The second definition uses a 25-basis-point cutoff. A signal is labeled false if no recession occurs despite one of the above signals occurring. Looking at the relative frequency of true and false signals will help establish the reliability of the yield curve for predicting recessions. Note that this procedure says nothing about instances when the yield spread failed to flatten or invert prior to a recession. The exercise lets us determine if the yield curve is like the boy who cried wolf or, in other words, if it correctly predicts a weakening in the economy.

I also investigate whether adding an indicator of monetary policy helps refine the signal. In this case a signal is labelled true if the spread inverted or was less than 25 basis points, respectively, and the funds rate was increased by more than 50 basis points in the preceding two quarters. The results in Table 1 confirm the graphical analysis that the spread is a useful but imperfect indicator of declines in economic activity. Looking at column 1, the spread inverts 18 times over the sample period, and on only two occasions does it erroneously signal a recession. Those occasions are in 1966:4 and 1979:1. The latter is labeled false only because it occurred five quarters prior to the onset of a recession. The true signals are clustered around the peaks. There are two true signals prior to and including the 1969:4 peak, three predate the 1973:4 peak, four precede the 1980:1 peak, and four precede the 1981:3 peak. Five of the occurrences are during recessions, which trivially do not signal an impending recession. Therefore, if the yield curve inverts, there is a high probability (83 percent) of an impending recession. The other columns confirm the yield curve's value as a strong signaler of a recession. Generally, most of the false signals occur in the mid- and late 1960s. Also, the character of the signals is not very different when an indicator of monetary policy is used. Consequently, there is not much evidence that the stance of monetary policy contributes to the quality of the signal.

While at first glance it appears that the spread contains information about future economic activity, it is not clear why this is the case. I am unaware of any formal economic model that investigates this issue. The spread contains direct information on a number of economic variables. Because it is a difference in nominal interest rates on bonds of different maturities, it is composed of a real term spread, the expected difference in inflation, and a term premium. Also, only temporary changes in these variables affect the spread. A permanent increase in either inflation or the real rate of interest will have the same effect on both the long- and short-term interest rates.

Often when there is an increase in expected inflation, as depicted by a steepening of the yield curve, the Fed engages in contractionary monetary policy by increasing short-term rates. In many of these episodes the long rate also initially rises, but not by as much as the short rate, and the spread narrows. Subsequently, as inflationary expectations subside, the long rate often falls and the yield curve inverts. The result of the monetary tightening is often a recession.[2] Correspondingly, when economic activity is weak, the Fed often

[2] An excellent documentation of a number of such episodes is provided by Goodfriend (1993).

loosens monetary policy by decreasing the short-term interest rate. This action generally causes the yield curve to steepen, and if an increase in inflationary expectations results from the easing, the yield curve may steepen substantially. Monetary easing often results in an increase in economic growth. Thus the result of easy monetary policy is often a steepening of the yield curve and increased economic activity. If these were the only reasons that movements in the spread were associated with economic activity, then adequately capturing the stance of monetary policy would leave little additional explanatory power for the spread in forecasting economic growth.

There are, however, other reasons why the spread may communicate future economic behavior. For example, Plosser and Rouwenhorst (1994) note that the spread's behavior is consistent with real business cycle theory. In a real business cycle model, relatively high expected future growth would imply rising real interest rates and a steepening of the term structure. The converse would occur if growth was expected to slow. Accordingly, the spread could signal expected changes in the economy that are due to nonmonetary shocks.

3. IN-SAMPLE PREDICTIVE CONTENT

In this section I examine the in-sample predictive content of the term spread between the discount equivalent yield on a ten-year U.S. Treasury bond and a three-month Treasury bill. The sample period begins in 1955:1 and extends to 1997:4. Data on the discount equivalent yield is obtained by splicing McCulloch's (1987) data set with data received from the Federal Reserve Board. The Board's data set begins in June 1961 and is not calculated in exactly the same way as McCulloch's; but except for a few years in the mid-1970s, the two series are indistinguishable. Discount equivalent yields are used for comparability purposes. I also use the ten-year, three-month spread to be comparable to most other studies but generally find that the main results of the analysis are not sensitive to the particular spread used. Results using the two-year, three-month spread and the five-year, three-month spread are similar to those reported below.

Simple Regressions

First, let's examine regressions that analyze the predictive content of the spread and various transformations of the spread for future GDP growth. I explore how sensitive the results are over different sample periods. The main finding is that the spread has predictive content for future output growth but that the regression coefficients change somewhat over different sample periods. I analyze the predictive content both cumulatively, up to two years, and marginally. Specifically, the regressions for cumulative growth are of the form

$$(400/k)\ln(y_{t+k}/y_t) = \alpha_0 + \alpha_1 s_t + e_t, \tag{1}$$

Table 2 Cumulative and Marginal Predictions of GDP Growth
(t-statistics in parentheses)

		Cumulative		Marginal	
Sample Period	***k***	α_1	$\overline{R^2}$	α_1	$\overline{R^2}$
1955:1–1997:4	2	0.96 (3.70)	0.134		
	4	0.88 (3.86)	0.182	0.80 (3.10)	0.092
	6	0.70 (3.25)	0.170	0.38 (1.68)	0.016
	8	0.53 (2.74)	0.127	−0.01 (−0.04)	−0.006
1955:1–1973:4	2	2.03 (3.24)	0.151		
	4	1.98 (4.19)	0.249	1.95 (3.40)	0.133
	6	1.41 (3.80)	0.190	0.22 (0.36)	−0.013
	8	1.29 (3.50)	0.186	0.43 (0.67)	−0.009
1973:1–1989:4	2	1.43 (6.37)	0.418		
	4	1.27 (7.04)	0.516	1.10 (3.16)	0.235
	6	1.07 (5.99)	0.528	0.71 (2.30)	0.092
	8	0.88 (6.50)	0.457	0.17 (0.54)	−0.010
1973:1–1997:4	2	1.06 (5.36)	0.274		
	4	0.95 (5.40)	0.348	0.85 (3.03)	0.172
	6	0.83 (4.98)	0.376	0.60 (2.39)	0.084
	8	0.67 (4.94)	0.322	0.15 (0.60)	−0.004
1985:1–1997:4	2	0.47 (1.82)	0.079		
	4	0.50 (1.66)	0.139	0.51 (2.53)	0.097
	6	0.55 (1.77)	0.243	0.57 (2.57)	0.125
	8	0.56 (1.94)	0.327	0.43 (1.40)	0.064

where y is quarterly real GDP and s is the spread. Values for k are 2, 4, 6, and 8. The regressions for marginal predictability are of the form

$$(400/2)\ln(y_{t+k}/y_{t+k-2}) = \alpha_0 + \alpha_1 s_t + e_t \tag{2}$$

and analyze whether the spread helps predict two-quarter output growth k periods in the future.

The first set of regression results are shown in Table 2. With the exception of the 1985 to 1997 sample period, the spread is significant at the 5 percent level in predicting cumulative output growth up to two years into the future.[3] In the latter period it is only significant at the 10 percent level. One notices, however, that the coefficients on the spread vary over different sample periods as does the informativeness of the spread as measured by the regression's adjusted R^2.

[3] All standard errors have been adjusted using the methodology suggested in Newey and West (1987). I also look at sample periods that conform to high and low inflation environments, namely, 1955:1 to 1972:4, 1973:1 to 1983:4, and 1984:1 to 1997:4, without any significant change in the nature of the results.

For example, the spread is an exceptionally good predictor of output growth over the 1973 to 1989 period.

The marginal predictive power of the spread is documented in the last two columns of Table 2. Consistent with the results in Estrella and Hardouvelis (1991) and Plosser and Rouwenhorst (1994), the spread generally has predictive content for economic growth only up to six quarters. That is, it is helpful at predicting two-quarter growth rates two quarters in the future and four quarters in the future. The spread is not informative about two-quarter growth rates at more distant horizons. Consequently, its ability to predict cumulative growth two years into the future is solely due to its strong association with near-term growth. As in the cumulative regressions, the flavor of the results would not be changed by using a spread that is composed of two-year or five-year long bond rates.

Alternative Specifications

For several reasons, one might expect that the predictive content of the spread could be improved by analyzing some alternative specifications. First, many of the episodes in which the spread inverts are also associated with contractionary monetary policy. It may be that combining an increase in the funds rate with a narrowing of the spread indicates tight monetary policy, and it is only these episodes in which the spread has predictive content. Thus the spread's signal value could be enhanced by adding an interactive term that incorporates tight monetary policy. Second, as mentioned, the spread contains a term premium that may add noise to any signal that the spread provides about the expected course of real interest rates. If this is so, then extreme values of the spread may have more predictive content than the spread itself. Also, if only large and unexpected changes in monetary policy significantly affect real economic activity, then it may be that only large movements in the spread are associated with changes in economic growth. By decomposing the spread into three components—unusually high values, normal values, and unusually low values—and by testing to determine if these different ranges imply a different relationship between the spread and economic growth, one could uncover nonlinearities in this relationship.

Specifically, the regression used for analyzing the combined effect of a monetary tightening and the spread is given by

$$(400/k)\ln(y_{t+k}/y_t) = \alpha_0 + \alpha_1 d_t s_t + \alpha_2 s_t + e_t, \qquad (3)$$

where d_t is a dummy variable that takes on the value of 1 if the funds rate is raised by 50 basis points or more over the preceding two quarters.[4] To investigate the presence of nonlinearities, I run the following regression:

[4] Using cutoffs of 75 basis points or 100 basis points produces similar results.

$$(400/k)\ln(y_{t+k}/y_t) = \alpha_0 + \alpha_1 hs_t + \alpha_2 ms_t + \alpha_3 ls_t + e_t, \qquad (4)$$

where hs_t takes on the value of the spread when the spread exceeds its average value by more than 0.425 standard deviations and is zero elsewhere. Similarly, the variable ls_t equals the spread when the spread is below its mean by more than 0.425 standard deviations. Otherwise it takes on the value zero. The variable ms_t equals the spread when each of the previous variables is zero and is zero elsewhere. The value 0.425 is chosen so that each variable equals the spread approximately one-third of the time. Also, the sum of the three variables is the spread itself. By dividing the spread into high, low, and intermediate values, one can check if output growth is more responsive to extreme values of the spread.

The results of these two investigations are depicted in Table 3. The sample periods are representative of the general results. The top half of the table shows that including tight monetary policy into the regressions does not significantly affect the forecasting ability of the spread. When the interactive term $d_t s_t$ is entered by itself, the adjusted R^2 is lower than in the comparable regressions using the spread by itself. Also, when both variables are entered simultaneously, only the spread retains its statistical significance.

The bottom half of the table shows the results of the analysis regarding nonlinearities. One can make a case for nonlinearity in the relationship between future output growth and the spread. Output growth responds more strongly to low values of the spread. This result may be due to the short, sharp nature of recessions, which tend to be associated with inversions in the yield curve. Both high values and intermediate values of the spread are significant over the entire sample, but high values are more likely to be significant in each subsample. Indeed, intermediate values do not have a statistically significant effect on output growth over the periods 1973:4 to 1989:4 and 1985:1 to 1997:4. For the entire sample period one can reject the equality of the coefficients. Equality, however, cannot be rejected over any of the subsamples.[5] The case for nonlinearities is, therefore, not overwhelming.

A Closer Look at the Information Content of the Term Structure

In this section I explore the additional information contained in the spread. Previous works, for example, Estrella and Hardouvelis (1991), Plosser and Rouwenhurst (1994), and Estrella and Mishkin (1997), have investigated this issue to some extent. Basically, these papers have simultaneously included

[5] The relevant statistic for the test of equality among the coefficients is distributed Chi-squared with 2 degrees of freedom. The test statistic is significant at the 5 percent level for the whole sample; the levels are 0.017 for $k = 2$ and 0.011 for $k = 4$. For the period 1955:1 to 1973:4, the significance levels are 0.843 and 0.962. For the sample 1973:1 to 1989:4, they are 0.780 and 0.775. And for the sample 1985:1 to 1997:2, they are 0.117 and 0.690.

Table 3 Predictive Content of Term Spread Using Alternative Specifications

Monetary Tightening

(t-statistics in parentheses)

Sample Period	***k***	α_1		α_2		$\overline{R^2}$
1955:1–1971:4	2	1.04	(2.86)			0.069
	2	0.50	(1.42)	0.81	(3.82)	0.143
	4	0.66	(2.05)			0.043
	4	0.09	(0.33)	0.85	(3.48)	0.178
1973:1–1989:4	2	1.46	(3.56)			0.208
	2	0.38	(0.92)	1.27	(5.23)	0.417
	4	1.14	(3.53)			0.205
	4	0.13	(0.50)	1.20	(5.32)	0.503

Nonlinearities

Sample Period	***k***	α_1		α_2		α_3		$\overline{R^2}$
1955:1–1977:4	2	0.73	(3.41)	1.33	(2.56)	2.44	(3.64)	0.181
	4	0.70	(2.97)	1.27	(2.75)	2.07	(3.69)	0.236
1955:1–1973:4	2	2.04	(3.34)	2.17	(3.35)	2.93	(2.14)	0.131
	4	2.02	(4.73)	2.00	(3.93)	2.41	(1.81)	0.229
1973:1–1989:4	2	1.31	(4.82)	1.77	(1.44)	1.76	(2.21)	0.409
	4	1.13	(4.00)	1.19	(1.47)	1.61	(3.12)	0.504
1985:1–1997:4	2	0.70	(2.00)	0.75	(1.20)	4.87	(1.99)	0.096
	4	0.62	(1.56)	0.78	(1.02)	2.45	(0.97)	0.113

another leading indicator or an index of indicators, a contemporaneous short-term interest rate or monetary aggregate, or the current growth rate of output. None of them have added a number of lags of other economic variables as is typically done in the VAR literature. Here I add two lags of output growth and four lags of the short-term nominal interest rate and test if the spread retains any significant predictive ability. The tests are performed with respect to cumulative output growth two and four quarters into the future. Thus a typical regression is given by

$$(400/k)\ln(y_{t+k}/y_t) = a_0 + \sum_{j=0}^{1}(400/k)b_j \ln(y_{t-jk}/y_{t-(j+1)k}) \qquad (5)$$

$$+ \sum_{j=0}^{3} c_j r_{t-j} + ds_t + e_t,$$

Table 4 Additional Information in Spread
(t-statistics in parentheses)

Sample Period	*k*	Σ*b*		Σ*c*		*d*		$\overline{R^2}$
1955:1 – 1971:4	2	0.20	(1.62)	−0.29	(−2.47)	0.48	(1.83)	0.23
	4	0.05	(0.31)	−0.30	(−2.71)	0.32	(1.22)	0.31
1955:1 – 1973:4	2	0.15	(0.80)	−0.29	(−0.99)	1.71	(1.91)	0.18
	4	0.24	(0.82)	−0.34	(−1.10)	1.74	(1.97)	0.27
1973:1 – 1997:4	2	0.25	(1.73)	−0.11	(−0.70)	0.60	(2.04)	0.27
	4	0.05	(0.32)	−0.14	(−0.99)	0.22	(1.11)	0.28
1985:1 – 1997:4	2	0.33	(2.20)	−0.50	(−1.91)	−0.28	(−0.91)	0.17
	4	0.13	(0.30)	−0.48	(−1.52)	−0.07	(−0.14)	0.14

where r is the interest rate on the three-month Treasury bill.[6] The results of this experiment are reported in Table 4. Over the entire sample period the spread is significant at the 10 percent level when predicting growth six months ahead but not statistically significant when predicting growth four quarters ahead. The spread is helpful in predicting two-quarter- and four-quarter-ahead growth rates over the 1955:1 to 1973:4 period and in predicting six-month growth over the 1973:1 to 1997:4 period. However, for this latter period the coefficient on the spread is insignificant when predicting four-quarter-ahead growth. This outcome is somewhat surprising given the results in Tables 2 and 3. Consistent with the results in Haubrich and Dombrosky (1996), the spread does not appear to be statistically significant over the most recent sample period of 1985:1 to 1997:4.

Hence, the results of this exercise indicate that the information content of the spread is reduced once other variables such as past output growth and past levels of short-term interest rates are taken into account. One must be a little guarded about the last statement. Estrella and Mishkin (1997), among others, stress that in-sample and out-of-sample predictive content are two very different things. Their work indicates that although parsimonious specifications may not perform as well in-sample, they often provide more accurate out-of-sample forecasts. In the next section, therefore, I investigate the out-of-sample predictive properties of the various models considered above.

[6] A distributive lag of past spreads was statistically insignificant. Also, longer lag lengths on past output growth were generally insignificant as well.

Table 5 Root Mean Squared Errors
Forecasts 1970:1 to 1997:4

Specification	RMSE	DM		RMSE	DM	
Start Date 1955:1						
	1970:1–1997:4			1985:1–1997:4		
Equation 5 without the spread	2.171			1.802		
Equation 5	2.170			2.274	2.17	(0.03)
Start Date Advances						
Equation 5 without the spread	2.171			1.848		
Equation 5	2.081	0.56	(0.58)	2.215	2.11	(0.04)
Equation 1	1.950	1.02	(0.31)	2.437	2.87	(0.00)
Equation 3	1.990	0.85	(0.40)	2.455	2.37	(0.02)
Equation 4	1.926	1.12	(0.26)	2.199	2.38	(0.02)

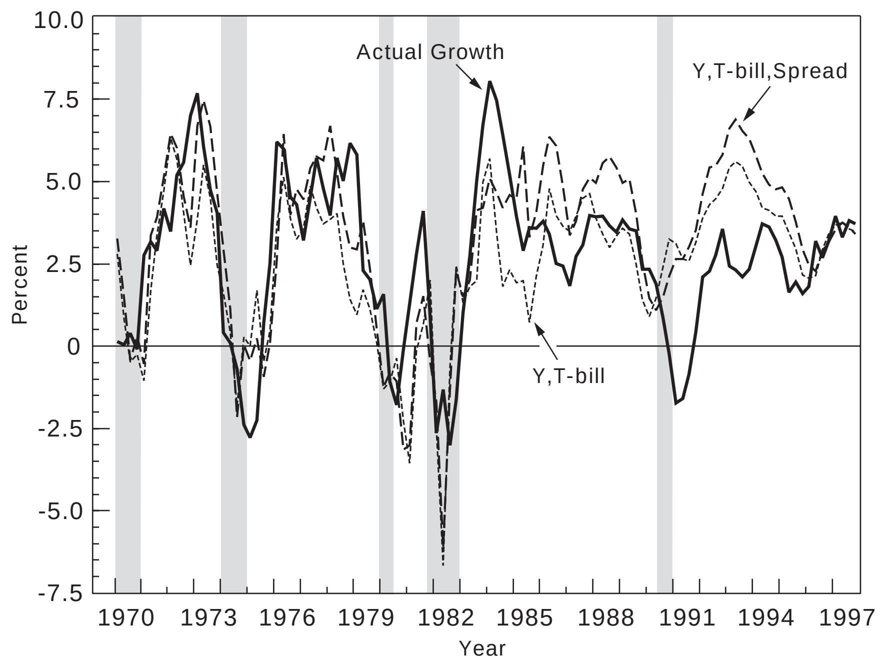
4. OUT-OF-SAMPLE FORECASTS

I now look at the out-of-sample forecast accuracy of one-year-ahead output growth for the variety of specifications considered in the previous section. The forecasts and the actual data are presented in Figures 3 through 5, and the root mean squared errors (RMSE) of the forecasts are given in Table 5. Forecasts are made over the period 1970:1 to 1997:4. The comparative predictive accuracy of the forecasts is analyzed using the methodology of Diebold and Mariano (1995) on differences of the squared forecast errors. The value of their test statistic and its significance level is reported in the columns labeled DM. This comparison is made for the entire forecasting period and for the more recent period of 1985:1 to 1997:4.

In Figure 3, the start date for the regressions is kept fixed, and the end date is continually advanced. Hence, the forecast for output growth over the period 1969:1 to 1970:1 uses data available up to 1968:4. I first examine forecast accuracy using the specification in equation (5), with and without the spread. As one sees from the two forecasts and the reported RMSE's, adding the spread does not significantly improve the out-of-sample forecasts. The root mean squared error declines almost imperceptibly from 2.171 to 2.170.

The in-sample regressions examined in the previous section, however, indicate that the coefficient on the spread varies over different sample periods. This behavior implies that a better forecasting procedure might be to roll the starting date of the regression forward as well to allow the estimated coefficients to change more rapidly. The results of this experiment are depicted in Figure 4. Here there is some improvement, with the RMSE declining from 2.171 to

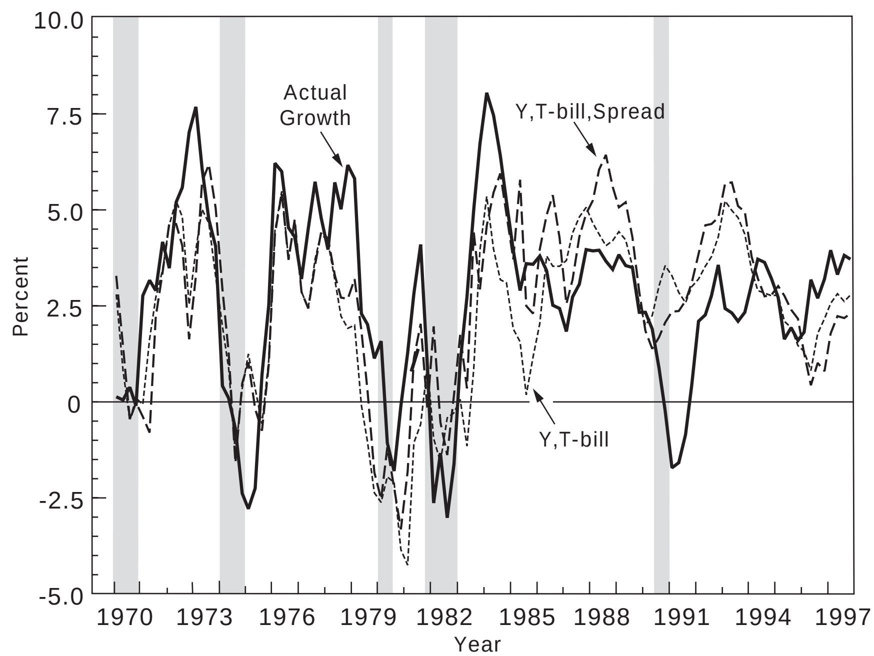
Figure 3 Actual and Out-of-Sample Predictions of Output Growth



2.081. The forecast including the spread does not overpredict the depth of the 1980 recession by quite as much as the specification without the spread, and it does not predict a sharp decline in output in 1985. The specification with the spread also indicates a slightly weaker economy in 1990 and 1991, but neither specification comes close to predicting a recession. On net, including the spread produces only a small gain in forecasting accuracy, and this gain is not statistically significant.

Surprisingly, over the entire forecasting period, only the nonlinear specification produces better out-of-sample forecasts than the spread by itself, and the improvement is minor (an RMSE of 1.926 as opposed to 1.950). Although the spread by itself produces a 10 percent increase in forecasting accuracy, as compared with a model that uses lagged values of output growth and lagged values of short-term interest rates (see Figure 5a), this increase in forecasting accuracy is statistically insignificant using the DM test statistic. Much of this gain is due to the improved forecasts in the early 1980s. Including a dummy variable that indicates tight monetary policy, as in equation (3), does not improve out-of-sample forecasting performance. Consequently, even though a parsimonious

Figure 4 Actual and Out-of-Sample Predictions of Output Growth (Rolling Regression)



specification that uses only the spread produces superior forecasts, the forecasts are not statistically significantly better.

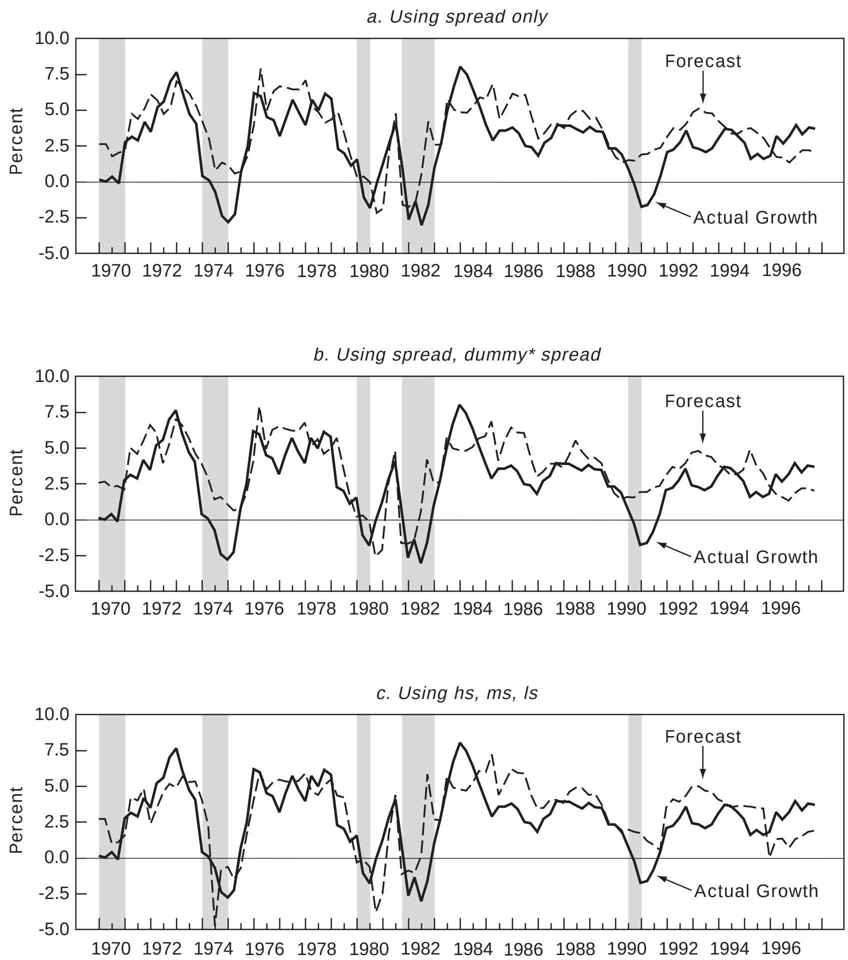
Over the more recent sample period, the results are strikingly different. Here the VAR model without the spread produces the most accurate forecasts, and these forecasts are significantly better.

5. PREDICTING RECESSIONS

In this section, I look at the ability of the spread to predict the onset of a recession using the probit model described in Estrella and Mishkin (1998). Based on the preceding section, the analysis concentrates on out-of-sample predictions but first analyzes some in-sample predictions. The relative ability of the various specifications given in equations (1), (3), (4), and (5) to accurately forecast recessions is indicated by the pseudo R^2.[7] Its values are displayed in Table 6,

[7] The pseudo R^2 is given by $1 - [\log(L_u)/\log(L_c)]^{-(2/n)\log L_c}$, where L_u is the log of the unconstrained likelihood function and L_c is the log of the maximum value of the likelihood function under the constraint that all coefficients except the constant term are zero.

Figure 5 Actual and Out-of-Sample Predictions (Alternative Regression Specifications)



as is the significance of the various coefficients in the probit regressions.[8]

As one sees from the table, the spread by itself predicts the in-sample probability of a recession relatively well. Adding a term that incorporates tight monetary policy does not help forecast recessions, nor does a specification that

[8] The significance levels for individual coefficients are corrected using the procedure in Estrella and Mishkin (1998). I wish to thank Arturo Estrella for sharing his code. The significance levels for joint tests of the coefficients on the lags of GDP growth and the T-bill rate were calculated using likelihood ratio tests that were not corrected for serial correlation.

Table 6 Significance of Variables for Predicting Recessions Using Probit Model

Specification	Variables	Significance Level	Pseudo R^2
1	spread	0.0000	0.277
2	spread	0.0004	0.281
	d*spread	0.45	
3	2 lags of GDP	0.85	0.171
	4 lags of T-bill	0.0000	
4	2 lags of GDP	0.126	0.317
	4 lags of T-bill	0.604	
	spread	0.0000	

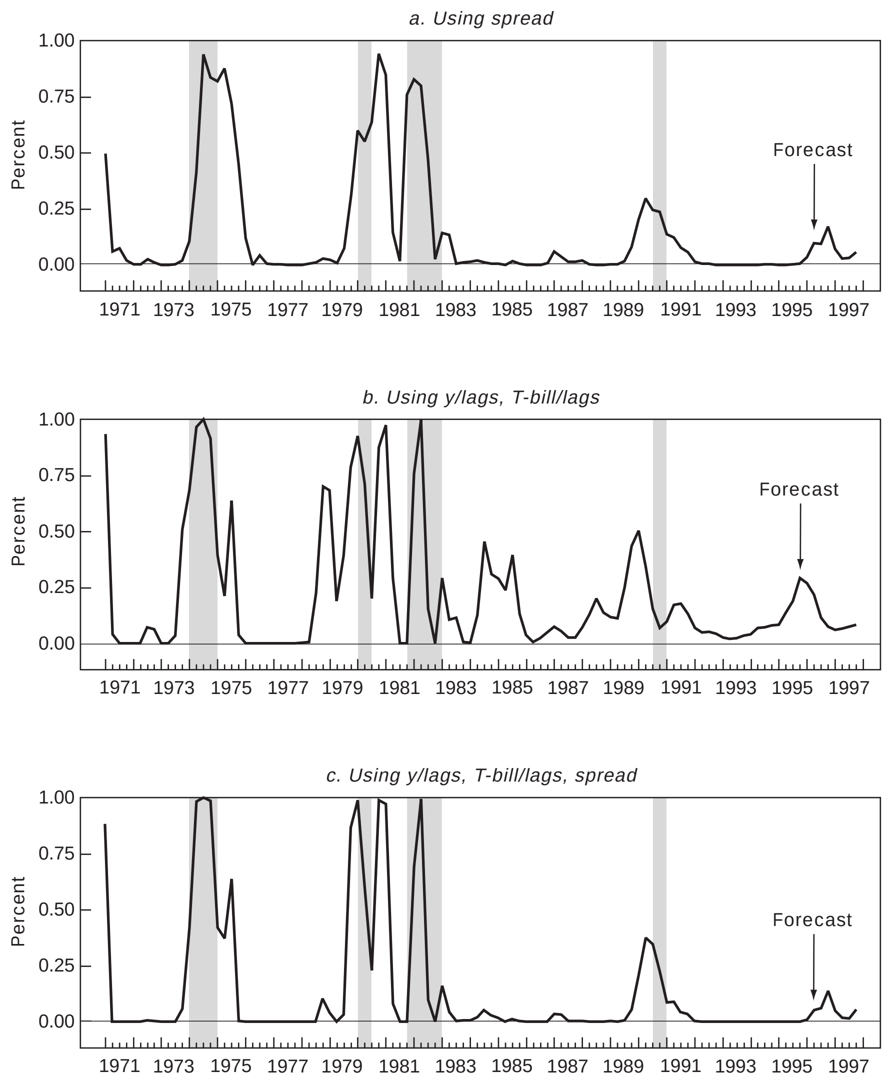
allows for nonlinear effects of the spread (the latter experiment is not reported). Adding the spread to a specification that includes lagged values of GDP growth and lagged values of the Treasury bill rate noticeably improves the in-sample forecasts of a recession.

The out-of-sample forecasts for specifications 1, 3, and 4 are shown in Figures 6a, 6b, and 6c. Only the pseudo R^2 for the specification using the spread by itself is positive and equals 0.324. The reason the pseudo R^2 is negative for the latter two out-of-sample forecasts is that the measure imposes a significant penalty for predicting a high probability of recession, when in fact no recession occurs. Also, the penalty is nonlinear, rising steeply for big forecast errors. These errors are more frequent in the latter two specifications. In some sense, though, the penalty is overly harsh because it is imposed equally whether the prediction of a recession is off by one quarter or whether the prediction occurs in the middle of an economic boom.

The three figures indicate that using the spread reduces the chance of falsely predicting the onset of a recession. This feature is particularly evident in comparing Figures 6b and 6c, where using the spread significantly reduces the probability of a recession during the mid-1980s. One also notices that while prior to the recessions in the 1970s and 1980s the three specifications forecast a high probability of recession, none of the specifications accurately signaled the 1990–91 recession. This evidence is consistent with that reported in Dueker (1997) and Estrella and Mishkin (1998).

As a final check on the spread's ability to forecast recessions, I compared its performance with that of a naive forecasting model that predicts the economy will be in its current state one quarter into the future. Even though the naive forecast uses more current information, the forecasting ability of the spread is noticeably better than the naive model. The DM statistic, which is based on

Figure 6 Out-of-Sample Predictions of Recessions



squared forecast errors, is 2.30, and the forecasts are, therefore, statistically different at the 2 percent significance level.

6. CONCLUSION

This article has investigated the forecasting properties of the yield spread for economic activity. It mainly concludes that the spread contains useful information—information not contained in past economic activity or past monetary policy. Combined with the work of other authors, most notably Estrella and Hardouvelis (1991), Estrella and Mishkin (1997, 1998), and Plosser and Rouwenhorst (1994), the article adds to the evidence that the spread has been a useful leading indicator of economic activity. That conclusion must be tempered, however, by the observation that over more recent periods the spread has not been nearly as informative as it has been in the past. It is impossible to say whether its reduced predictive content is a function of some permanent change in the economy, or is only transitory, or is simply an outcome of examining a small sample period characterized by relatively little output variability. Given the spread's long history as a useful forecasting tool and the simplicity of its use, it will probably continue to receive significant attention in both the financial press and academic research.

REFERENCES

Bonser-Neal, Catherine, and Timothy R. Morley. "Does the Yield Spread Predict Real Economic Activity? A Multicountry Analysis," Federal Reserve Bank of Kansas City *Economic Review*, vol. 82 (Third Quarter 1997), pp. 37–53.

Diebold, Francis X., and Roberto S. Mariano. "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, vol. 13 (July 1995), pp. 253–63.

Dueker, Michael J. "Strengthening the Case for the Yield Curve as a Predictor of U.S. Recessions," Federal Reserve Bank of St. Louis *Review*, vol. 79 (March/April 1997), pp. 41–51.

Estrella, Arturo, and Gikas A. Hardouvelis. "The Term Structure as a Predictor of Real Economic Activity," *Journal of Finance*, vol. 46 (June 1991), pp. 555–76.

Estrella, Arturo, and Frederic S. Mishkin. "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 (July 1997), pp. 1375–1401.

__________, and __________. "Predicting U.S. Recessions: Financial Variables as Leading Indicators," *Review of Economics and Statistics*, vol. 80 (February 1998), pp. 45–61.

Goodfriend, Marvin. "Interest Rate Policy and the Inflation Scare Problem: 1979–1992," Federal Reserve Bank of Richmond *Economic Quarterly*, vol. 79 (Winter 1993), pp. 1–24.

Haubrich, Joseph G., and Ann M. Dombrosky. "Predicting Real Growth Using the Yield Curve," Federal Reserve Bank of Cleveland *Economic Review*, vol. 32 (First Quarter 1996), pp. 26–34.

Kozicki, Sharon. "Predicting Real Growth and Inflation with the Yield Spread," Federal Reserve Bank of Kansas City *Economic Review,* vol. 82 (Fourth Quarter 1997), pp. 39–57.

Laurent, Robert. "An Interest Rate-Based Indicator of Monetary Policy," Federal Reserve Bank of Chicago *Economic Perspectives,* vol. 12 (January/February 1988), pp. 3–14.

Newey, Whitney, and Kenneth West. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica,* vol. 55 (May 1987), pp. 703–08.

Plosser, Charles I., and K. Geert Rouwenhorst. "International Term Structures and Real Economic Growth," *Journal of Monetary Economics*, vol. 33 (February 1994), pp. 133–56.

Stock, James H., and Mark W. Watson. "New Indexes of Coincident and Leading Economic Indicators," in Olivier J. Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual 1989.* Cambridge, Mass.: MIT Press, 1989, pp. 352–94.