Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve*

Glenn D. Rudebusch†   John C. Williams‡

April 2007

Abstract

We show that professional forecasters have essentially no ability to predict future recessions a few quarters ahead. This is particularly puzzling because, for at least the past two decades, researchers have provided much evidence that the yield curve, specifically the spread between long- and short-term interest rates, does contain useful information at that forecast horizon for predicting aggregate economic activity and, especially, for signalling future recessions. We document this puzzle and suggest that forecasters have generally placed too little weight on yield curve information when projecting declines in the aggregate economy.

*Vuong Nguyen and Stephanie Wang provided excellent research assistance. The views expressed in this paper are our own and not necessarily those of others at the Federal Reserve Bank of San Francisco.
†Federal Reserve Bank of San Francisco; www.frbsf.org/economists/grudebusch; Glenn.Rudebusch@sf.frb.org.
‡Federal Reserve Bank of San Francisco; John.C.Williams@sf.frb.org.

First on Monday and then again on Thursday, [former Fed Chairman] Greenspan upset stock markets merely by uttering the word "recession" and saying that one might but probably would not occur by the end of this year. . . . Mr. Greenspan's use of the R-word helped push down the Dow Jones industrial average by about 200 points on Thursday morning, . . .

— *The New York Times*, March 2, 2007

1. Introduction

Recessions remain scary times—for workers who suffer job losses, for investors who endure asset price declines, for entrepreneurs who risk bankruptcy. Recessions are periods of greater dislocation and anxiety, higher unemployment and suicide rates, and lower output and profits. Over time, recessions have become less frequent and less severe; however, non-recessionary episodes have also become more stable, so in relative terms, as the market sensitivity in the epigraph suggests, recessions appear to many to be as perilous as before.[1] Therefore, any ability to predict recessions remains highly profitable to investors and very useful to policymakers and other economic agents. Accordingly, there remains a keen and widespread interest in predicting recessions, and our paper examines what economic forecasters know about the likely occurrence of a recession and, most importantly, when do they know it.

Our analysis focuses on two divergent strands in the recession prediction literature. First, it is common wisdom that economists are not very good at forecasting recessions. For example, Zarnowitz and Braun (1993) showed that economic forecasters made their largest prediction errors during recessions, and Diebold and Rudebusch (1989, 1991a, b) provide a pessimistic assessment of the ability of the well-known index of leading indicators to actually provide useful signals of future recessions. In this paper, we provide new evidence on this issue by examining the information content of economic forecasts provided by participants in the Survey of Professional Forecasters (SPF). We find that these forecasters have little ability to predict recessions—especially at a forecast horizon of a few quarters ahead.

A second strand of the recession prediction literature uses financial data, notably, the slope of the yield curve, to predict recessions. In distinct contrast to the widely acknowledged weak performance of professional forecasters, the received wisdom from the yield curve studies is that the spread between long- and short-term interest rates is a fairly good predictor of recessions.

[1] We disagree with those who consider the key attributes of recessions to be increased consumption of leisure and greater home production.

This point was made by a variety of authors in the late 1980s and has been reinforced by a large subsequent literature.[2]

This paper documents that the puzzling conflict between these two literatures has not disappeared over the past two decades. In particular, even after the predictive power of the yield curve had been well publicized, it appears that SPF participants did not incorporate all of the available yield curve information into their forecasts of economic activity. Indeed, we find that a simple model for predicting recessions that uses only realtime yield curve information would have produced better forecasts of recessions than the professional forecasters provided.

The paper proceeds as follows. In the next section, we provide a simple definition of recessions in terms of real GDP growth. In Section 3, we describe a variety of alternative realtime probability forecasts for these GDP-based recessions. In Section 4, we assess the accuracy of these forecasts, and we conclude with some speculation about possible resolutions to the puzzle of the enduring relative power of the yield curve for predicting recessions.

2. Defining Recessions

As a first step, it is necessary to define the object of interest. The National Bureau of Economic Research (NBER), which has been dating recessions for almost 80 years, provides the most widely accepted definition of a recession (NBER 2003):

> A recession is a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales. A recession begins just after the economy reaches a peak of activity and ends as the economy reaches its trough. Between trough and peak, the economy is in an expansion.

Thus, in determining the dates of business cycle peaks and troughs, the NBER does not rely on any single macroeconomic indicator, but instead examines a large collection of monthly variables. These individual series have idiosyncratic movements but also display substantial co-movement and correlation, and the NBER selects the overall business cycle peak and trough months to best capture the general consensus among the various series of the high and low points in economic activity.[3] Based on this methodology, the NBER publishes a historical chronology of the monthly

[2] Early researchers that described the predictive power of the yield spread for real activity include Harvey (1989), Stock and Watson (1989), and Estrella and Hardouvelis (1991). More recent contributions include Dotsey (1998), Estrella and Mishkin (1998), Estrella (2005), Chauvet and Potter (2005), Ang, Piazzesi, Wei (2006), and Wright (2006).

[3] For detailed discussion of the NBER methodology, see Diebold and Rudebusch (1992). For a modern interpretation, see Diebold and Rudebusch (1996).

dates of past business cycle peaks and troughs, which delineate recessions and expansions.

The NBER business cycle dating methodology requires substantial judgment in its application, and the resulting chronology is not without measurement error. Diebold and Rudebusch (1992) describe some of the uncertainty about the precise monthly dating of the general turns in business activity; however, even if one accepts the NBER dates of cyclical peaks and troughs as stated, there remains some ambiguity about when a recession starts and ends. Specifically, the peak and trough months could be classified as the first month of a recession and the first month of an expansion, respectively, or as the last month of an expansion and the last month of a recession, respectively.[4] The latter convention, in which the recession starts in the month following the peak month and ends on the date of the subsequent trough month, is the more common one (e.g., Diebold and Rudebusch 1989). A further complication arises from the translation of these monthly dates into designations of peaks and troughs at a quarterly frequency, which we employ in our analysis because the SPF is quarterly. Such a translation can be done in a variety of ways, and we follow the most common convention (e.g., Estrella and Trubin, 2006) and assume that a recession starts in the quarter that follows just after the quarter containing the peak month and that it ends in the quarter containing the trough month.[5] Our resulting chronology of quarterly NBER recessions and expansions is show in the middle column of table 1, with a one for a recession quarter and a zero otherwise. This binary variable will be denoted as $RNBER_t$.

For our analysis, we employ a rule that links changes in real GDP to recessions. Real GDP is obviously crucial to determining recessions. As noted by the NBER (2003): "The NBER considers real GDP to be the single measure that comes closest to capturing what it means by 'aggregate economic activity.' The [NBER] therefore places considerable weight on real GDP and other output measures." Although the NBER does not have a fixed rule that defines recessions in terms of GDP growth, a rule of thumb often discussed is that two consecutive declines in real GDP constitute a recession. We denote this definition as the R2 rule. Specifically, if real GDP falls in quarter t, then it is an R2 recession quarter if either (or both) of the quarters t-1 and t+1 post negative growth as well.

[4] The NBER's view is that the *day* when the economy turns around is contained within the peak or trough month, so those months are mongrel mixes of both recession and expansion phases.

[5] We obtained similar results using an alternative convention in which a recession quarter was defined to contain two or more recession months or using the convention in Wright (2006), in which both the peak quarter and trough quarter were counted as recession quarters.

Two versions of a recession chronology using the R2 rule are shown in Table 1. The lefthand side uses realtime GDP data, which are defined as the "first final" data—the estimates released about three months after the end of the quarter. The righthand side uses the current vintage (February 2007) of GDP data. We include the current vintage data for comparison purposes, but our focus is almost exclusively on the first final realtime data. Although we obtained similar results using current vintage data, the realtime GDP data more closely corresponds to what forecasters were trying to predict, that is, before any definitional revisions or other changes are incorporated. In any case, using either vintage of data, the R2 rule appears a bit too stringent to provide a good match to the NBER business-cycle dating methodology. For example, using the R2 rule on realtime data produces only 10 recession quarters that match one of the 21 NBER recession quarters, with 11 quarters of missed recession signals. Perhaps most distressingly, the R2 rule using realtime GDP data completely misses the 1980 and 2001 NBER recessions. The R2 rule also produces 2 false calls of recession quarters relative to the NBER definition (in 1969 and 1991). Overall then, the R2 rule gives a total of 13 quarters of incorrect recession signals.

As alternatives to the R2 rule, we considered a variety of other GDP-based recession dating schemes. A straightforward rule that works quite well is the R1 rule, which simply defines any single quarter of negative real GDP growth as a recession quarter. The associated binary variable of recession quarters is denoted as $R1_t$. As shown in Table 1, the R1 rule on realtime data produces 14 recession quarters that match one of the 21 NBER recession quarters, with only 7 missed calls of recession. The R1 rule also produces 5 false calls of recession (including one by a whisker in 1978:Q1). Therefore, there are 12 total quarters of incorrect signals, which is one fewer than the R2 rule produced, and the mix of false and missed call is also more balanced with the R1 rule.[6] Furthermore, the R1 rule has advantages over the R2 rule in terms of simplicity. Although the R1 rule is the focus of our analysis below, we obtained qualitatively similar results using other GDP-based rules for defining recessions, including the R2 rule.

3. Recession Probability Forecasts

In this section, we describe various alternative realtime probability forecasts of R1 recession quarters. These are based on information from the SPF or from the yield curve. In the next section, we will provide a formal examination of their accuracy.

[6] Using the current vintage data, the R1 and R2 rules both produce 11 incorrect signals, but the R2 rule still produces a more lopsided mix of mistakes.

3.1. SPF reported probability forecasts

In every quarter since the end of 1968, the SPF has asked its participants to provide estimates of the probability of negative real GDP growth in that quarter as well as probabilities of negative growth in each of the next four quarters.[7] The wording of the specific survey question is very clear and has changed little over time, and in 2007 it went as follows:

> Indicate the probability you would attach to a decline in real GDP (chain-weighted basis, seasonally adjusted) in the next five quarters. Write in a figure that may range from 0 to 100 in each of the cells (100 means a decline in the given quarter is certain, i.e. 100 percent, 0 means there is no chance at all, i.e. 0 percent).

The median probability forecast for quarter t that was reported in response to this question asked in the survey in quarter $t-h$ is denoted $P^{SPF}_{t|t-h}$.

Importantly, these are direct realtime probability forecasts of what we have termed R1 recessions that require no ex post adjustments or filtering. The solid lines in five panels of Figure 1 plot these probabilities for the current quarter ($h=0$) and for the future four quarters. (Note that in each panel $P^{SPF}_{t|t-h}$ is plotted in quarter t, regardless of the value of the forecast horizon h.) The current-quarter negative growth probabilities ($P^{SPF}_{t|t}$) plotted in the upper left panel of Figure 1 correspond quite closely to the R1 recession quarters, which are shown as the shaded bars in the figure. However, as the forecast horizon increases, the R1 recession predictive ability of the professional forecasters drops off very quickly. Even at a horizon of two quarters, and certainly at three and four quarters ahead, the probability forecasts appear to have little relationship with historical recessions.

3.2. SPF implied probability forecasts

Our second sequence of recession probability forecasts is derived from the real output forecasts reported in the SPF. We use the median real GNP/GDP forecasts for quarter t from the SPF in quarter $t-h$ and the historical error variances associated with forecasts at various horizons to compute implied recession probabilities $P^{GDP}_{t|t}$. We assume that the forecast errors are normally distributed with the variance equal to the SPF forecast error variance over the full sample. We apply the same forecast error variance to all forecasts; therefore, these are not true real time recession probabilities because they are computed as if forecasters knew the full-sample

[7] Croushore (1993) and Croushore and Stark (2001) describe the properties of the SPF, and Campbell (2004) and Lahiri and Wang (2006) specifically analyze SPF probability forecasts. Note that occasionally in 1968 and 1969, probability forecasts were only provided at horizons up to three quarters in length.

distribution of forecast errors. Still, these probabilities are based on the realtime GDP point forecasts. The resulting forecasts are shown by the dashed lines in Figure 1.

It is instructive to examine the differences between the reported R1 recession probabilities in the SPF and the probabilities implied by the SPF GDP projection at each of the five horizons, as shown in Figure 2. First, the differences are generally small and fairly uniform over time, which suggests that the SPF participants provide GDP point forecasts that are fairly consistent with their reported negative growth probability forecasts. This helps validate the probability forecasts as serious predictions. Furthermore, it suggests that there has been little change in the perceived forecast error volatility over time. Recall the implied probabilities assume a constant forecast error variance, which is set to its average over the full sample. This appears to be a pretty good estimate. Notably, the differences between the reported and implied probabilities show no clear trend and are not significantly larger or smaller at the beginning or end of the sample. This lack of trend in the conditional volatility is a little surprising but not inconsistent with the well-known "Great Moderation" in the unconditional volatility of real GDP growth over this period.[8] Finally, the time series of errors show some interesting episodes. In particular, in 1979 and 1980, which was a period of a relatively high reported and implied probabilities of negative growth, the SPF participants still underestimated that likelihood relative to the GDP forecast. Similarly, during much of the 1990s, the implied probabilities were lower than might be expected based on the GDP forecasts and historical forecast error distributions.

3.3. Naive probability forecasts

We also construct simple "naive" realtime recession probability forecasts to use as a benchmark for the other forecasts. To construct the naive recession forecast, we first assume a forecast of output growth equal to real average GDP growth rate over the past 10 years. Given that output forecast, we then compute the probability of negative growth in the current quarter, assuming a mean-zero, normally-distributed forecast error. We set the variance of the forecast error equal to the sample variance of SPF forecast errors. Note that, by construction, at any given point in time, the naive probability of a negative quarter is the same for the current quarter, the next quarter, etc., that is, formally, $P_{t|t}^{NAI} = P_{t+h|t}^{NAI}$, for all h. The real-time predictions of this naive recession forecasting model are shown by the dotted lines in Figure 1.

[8] See Campbell (2004) and Tulip (2005) for further discussion.

3.4. Yield-curve probability forecasts

We now consider forecasts of an R1 recession that are based on the yield spread. Following Estrella and Hardouvelis (1991), we define the yield spread, S_t, to be the difference between the yield on a 10-year U.S. Treasury note, i_t^L, and the yield on a 3-month Treasury bill, i_t^S: $S_t \equiv i_t^L - i_t^S$. We construct this yield spread using quarterly averages of the constant-maturity yields for each Treasury security. Our basic yield-curve recession prediction model is a probit of the form:

$$\Pr[R1_t = 1 | I_{t-h}] = \mathrm{N}[\alpha + \beta S_{t-h-1}],$$

where the variable $R1_t$ equals one if realtime real GDP growth is negative in quarter t and zero otherwise and N[•] denotes the cumulative normal distribution. The forecast horizon is varied so that $h = 0$, 1, 2, 3, 4. Note that the yield curve information available for a forecast made at time $t - h$ includes the average spread in quarter $t - h - 1$. This is consistent with the timing of the information set of the SPF forecasters. Given that yield curve data are not revised and available immediately, forecasters would have knowledge of the spread in the prior quarter when forming their forecasts in the first month of the quarter.

Of course, in real time, a forecaster could only estimate the probit over the sample of available past data. We assume that a forecaster reestimates five probit regressions (one for each forecast horizon) in each quarter $t - h$ using data up to quarter $t - h - 1$ and then uses those probit coefficient estimates to produce five recession probability forecasts of varying horizons. Sequences of realtime probit slope and intercept coefficient estimates are shown in Figure 3.[9] To simplify the figure, coefficient sequences are only shown for the current-quarter, two- and four-quarter ahead forecasts, as the omitted one- and three-quarter ahead sequences are similar. These are expanding sample (or "recursive") estimates based on a sample that always starts in 1955:Q1 and ends in the quarter plotted. Given these realtime coefficient estimates, we define the realtime yield spread R1 recession probability forecast as

$$P_{t|t-h}^{YS} = \mathrm{N}[\hat{\alpha}_{t-h} + \hat{\beta}_{t-h} S_{t-h-1}],$$

where $\hat{\alpha}_{t-h}$ and $\hat{\beta}_{t-h}$ are the sequences of realtime estimates. The resulting yield spread recession probability forecasts—based on realtime yield spread data and realtime probit estimates—are

[9] Consistent with the earlier literature, the estimated coefficients on the spread are typically highly statistically signficant and economically meaningful.

shown as solid lines in Figure 4, with the SPF reported probability forecasts repeated as the dotted line. (Again the timing of this display plots $P^{YS}_{t|t-h}$ in quarter t.)

4. Assessing Probability Forecasts

In assessing the various R1 recession probability forecasts, we first compare the SPF and naive forecasts and then assess the relative accuracy of the forecasts based on the yield spread.

4.1. Comparing SPF and naive probability forecasts

Before providing some statistical comparisons of accuracy, it is useful to examine the differences between the reported R1 recession probabilities in the SPF and the naive forecast probabilities at the five forecast horizons, as shown in Figure 5. Based on these differences, it does appear that the SPF probability forecasts do vary significantly from the naive ones but only at short horizons. Specifically, for forecasting the current quarter, the SPF participants appear to be able to delineate periods of weak or negative real GDP growth. However, the SPF predictive information regarding R1 recessions quickly erodes as the forecast horizon increases. Certainly for three- and four-quarter-ahead forecasts, it does not appear that the SPF forecasts have much if any informational advantage over the simple naive forecasts.

While the Figure 5 is suggestive, it is very useful to conduct a statistical analysis of relative forecast accuracy. To do this, we use two standard forecast accuracy measures: the mean absolute error (MAE) and the root mean squared error (RMSE). These measures are defined for the SPF reported probability forecasts, for example, at a horizon of h as

$$\text{MAE(SPF,}h) = \frac{1}{T}\sum_{t=1}^{T} |P^{SPF}_{t|t-h} - R1_t|$$

and

$$\text{RMSE(SPF,}h) = \sqrt{\frac{1}{T}\sum_{t=1}^{T} (P^{SPF}_{t|t-h} - R1_t)^2}.$$

These two measures evaluate the probability forecasts in terms of accuracy or closeness, on average, of the predicted probabilities to the observed recession realizations, as measured by the zero-one $R1_t$ dummy variable denoting the R1 recession quarters. Analogous measures can be computed for each forecast, and table 2 provides these measures for the each of the four probability forecasts (SPF reported and implied, naive, and yield spread) at each of the five forecast horizons ($h = 0, 1, 2, 3, 4$) and for our full sample (1968-2007) and for a post-1987 sample (1988-2007). Not surprisingly, for the full sample of current-quarter forecasts, the MAE for the SPF reported probabilities (of 0.168) is lower than the MAE of the naive probability forecast (of 0.198). Indeed, the SPF forecast is more accurate (in both MAE and RMSE terms) than the naive one at every forecast horizon and for both samples.

To examine the significance of these differences, we apply the Diebold-Mariano (1995) test of relative forecast accuracy. This test is based on the mean accuracy differential (or, more generally, the loss differential) between two forecasts. For example, the two differential series for the SPF and naive forecasts at a horizon h for the MAE and RMSE accuracy measures are

$$\text{Diff}_t(\text{MAE}, \text{SPF}, \text{NAI}, h) = |P^{SPF}_{t|t-h} - R1_t| - |P^{NAI}_{t|t-h} - R1_t|$$

and

$$\text{Diff}_t(\text{RMSE}, \text{SPF}, \text{NAI}, h) = \sqrt{(P^{SPF}_{t|t-h} - R1_t)^2} - \sqrt{(P^{NAI}_{t|t-h} - R1_t)^2}.$$

The Diebold-Mariano test can be simply based on the t-statistic for the hypothesis of a zero population mean differential taking into account the fact that these differential time series are not necessarily white noise. We compute the test by regressing a differential time series on an intercept and testing the significance of that intercept using standard errors that are corrected for possibly heteroskedastic and autocorrelated residuals. We denote the SPF probability forecasts that are significantly more accurate than the naive forecasts with an asterisk.[10] By this metric, the SPF reported probabilities are more accurate than the naive forecasts at two quarters ahead in both samples and for both accuracy measures. There is also scattered evidence that the SPF forecasts are more accurate at one- and three quarter horizons. Interestingly however, none of the differences between the SPF and naive *current-quarter* forecasts appear to be significant.

The lack of significance of the relative current-quarter forecast performance of the SPF is surprising given Figure 5, which shows much higher SPF probabilities during, for example, the 1974-75 and 1990-91 recessions. Some understanding of this discrepancy can be gleaned from Figure 6, which shows the SPF and naive probability forecast errors for each forecast horizon. These are measured as forecast minus actual, so positive entries represent false signals of recession and negative entries represent missed signals of recession. Unsurprisingly, the current-quarter

[10] Asterisks in the MAE (RMSE) column are based on a MAE (RMSE) loss function. The underlying distribution theory is asymptotic, but Diebold and Mariano (1995) provide evidence that this test is well-sized in small samples.

naive forecast (the dotted line) exhibits large missed signals in each recession quarter. However, although the SPF probability forecasts do better during recessions, they make substantial false signals of recession in the quarters just before and after recessions, notably in 1974-75, 1979-81, and 2001. Because the MAE and RMSE are symmetrical accuracy measures, they weight these false and missed signals equally and so evaluate the SPF and naive forecasts are roughly comparable. Of course, other measures could give difference results. Indeed, there is a large literature and long history of formal evaluations of probability forecasts, particularly in meteorology, with a variety of measures of accuracy. (See, for example, Diebold and Rudebusch 1989 and Lahiri and Wang 2006 for discussion of these measures as well as other forecast attributes, such as calibration or resolution.) However, these alternative measures are also typically symmetric and seem likely to give similar results. The underlying impediment in adopting asymmetric measures is that they require some specificity about why false alarms and missed calls are being treated differently. Such asymmetric weighting requires knowledge of the particular decision-making context in which the forecasts are being used. For example, given the inherent uncertain lags in the transmission of monetary policy, a central banker may not attach much if any cost to a false signal that a recession will occur in the second quarter if one actually does occur in the third quarter. In contrast, a market trader may attach a high cost to a false signal with such a one-quarter miss in timing. Obviously, like so many other researchers before us, we cannot provide general guidance on this issue but only highlight the problem.

4.2. Assessing yield spread probability forecasts

We now consider yield spread predictions and compare the realtime recession probability forecasts based on the yield spread with the SPF reported probabilities. Figure 7 displays the differences between these two forecasts at the five horizons. For the current-quarter and one-quarter-ahead forecasts, it does not appear that the yield spread information is as good as the SPF, as this difference becomes negative during the 1974-75, 1990-91, and 2001 recessions. However, for the three- and four-quarter-ahead forecasts, the yield spread and SPF differences appear clearly to be negative during the non-recession quarters and a bit positive during recessions. This general impression of superior yield spread predictions is confirmed by the evidence in Table 2, where the dagger symbol denotes yield curve forecasts that are significantly more accurate that the SPF reported forecast at the 5 percent level according to the Diebold-Mariano test. At a four-quarter-ahead horizon, the yield spread forecast dominates the SPF forecast according to both the MAE and RMSE measures. Importantly, during the post-1987 sample, which is after the early papers trumpeting the predictive power of the yield curve were already published or widely circulated, the yield spread still contains predictive information that appears not to have been taken into account by the SPF participants.

It can be partially illuminated by examining the time series of yield spread and SPF forecast errors shown in Figure 8. At a forecast horizon of 4 quarters, it is clear that one of the failings of the SPF reported probabilities is the sustained 20 percent or so chance of a recession during the long expansions of the past two decades. The yield curve forecasts in contrast registered lower probabilities in the zero to ten percent range. The enduring power of the yield curve remains a puzzle.

5. Conclusion

As witnessed by the public attention to pronouncements of probabilities of recession during the past year, there is a great deal of interest in predicting recessions. Nonetheless, economists have a very spotty track record of predicting future downturns. One possible explanation is that recessions are simply unpredictable. But, this view is contradicted by evidence that the yield curve provides useful information for forecasting future periods of expansion and contraction. In this paper, we show that the yield curve has significant realtime predictive power for distinguishing between expansions and contractions several quarters out relative to the predictions of professional macroeconomic forecasters. This conclusion remains true during the past twenty years, despite the fact that the yield curve model's usefulness has been widely known since the late 1980s.

There are a number of potential reconciliations for this puzzle. First, as noted above, the professional forecasters' loss functions may be quite different from the symmetric ones used in our evaluations, and under their alternative loss functions, the differences that we document may not be significant. Examining such alternative losses in future research would be useful. Second, forecasters may have downweighted the yield curve information because they systematically underestimated the macroeconomic repercussions of changes in the stance of monetary policy as proxied for by shifts in the slope of the yield curve. Indeed, the relationship between output and interest rates is estimated very imprecisely and subject to econometric difficulties that may bias estimates of the interest-sensitivity of output downward. Nonetheless, the longevity of this puzzle makes one question why forecasters have not caught on to this mistake. Finally, it is interesting to note that many times during the past twenty years forecasters have acknowledged the formidable past performance of the yield curve in predicting expansions and recessions but argued that this past performance did not apply in the current situation. That is, signals from the yield curve have often been dismissed because of supposed changes in the economy or special factors influencing interest rates. This paper, however, shows that the relative predictive power of the yield curve does not appear to have diminished much, if at all.

References

Ang, Andrew, Monika Piazzesi, and Min Wei (2006), "What Does the Yield Curve Tell Us about Growth?" *Journal of Econometrics* 131, 359-403.

Campbell, Sean D. (2004), "Macroeconomic Volatility, Predictability and Uncertainty in the Great Moderation: Evidence from the Survey of Professional Forecasters," Finance and Economics Discussion Series, 2004-52, Federal Reserve Board.

Chauvet, Marcelle, and Simon Potter (2005), "Predicting a Recession: Evidence from the Yield Curve in the Presence of Structural Breaks," *Journal of Forecasting* 24, 77-103.

Croushore, Dean (1993), "Introducing: The Survey of Professional Forecasters," Federal Reserve Bank of Philadelphia *Business Review*, November/December, 3–13.

Croushore, Dean, and Tom Stark (2001), "A Real-Time Data Set for Macroeconomists," *Journal of Econometrics*, 105, 111–130.

Diebold, Francis X., and R.S. Mariano (1995), "Comparing Forecast Accuracy," *Journal of Business and Economics Statistics* 13, 253-265.

Diebold, Francis X., and Glenn D. Rudebusch (1989), "Scoring the Leading Indicators," *Journal of Business* 62, 369-391.

Diebold, Francis X., and Glenn D. Rudebusch (1991a), "Forecasting Output with the Composite Leading Index: A Real-Time Analysis," *Journal of the American Statistical Association* 86, 603-610.

Diebold, Francis X., and Glenn D. Rudebusch (1991b), "Turning Point Prediction with the Composite Leading Index: An Ex Ante Analysis," in *Leading Economic Indicators: New Approaches and Forecasting Records*, edited by Lahiri and Moore, Cambridge: Cambridge University Press, 231-256.

Diebold, Francis X., and Glenn D. Rudebusch (1992), "Have Postwar Economic Fluctuations Been Stabilized?," *American Economic Review* 82, 993-1005.

Diebold, Francis X., and Glenn D. Rudebusch (1996), "Measuring Business Cycles: A Modern Perspective," *Review of Economics and Statistics* 78, 67-77

Dotsey, Michael (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, Arturo (2005), "Why Does the Yield Cruve Predict output and Inflation?" *The Economic Journal* 115, 722-744.

Estrella, Arturo, and Gikas A. Hardouvelis (1991), "The Term Structure as a Predictor of Real Economic Actviity," *The Journal of Finance*, 46, 555-576.

Estrella, Arturo, and Frederic S. Mishkin (1998), "Predicting U.S. Recessions: Financial Variables as Leading Indicators," *The Review of Economics and Statistics* 80, 45-61.

Estrella, Arturo, and Mary R. Trubin (2006), "The Yield Curve as a Leading Indicator: Some Practical Issues," Federal Reserve Bank of New York, *Current Issues in Economics and Finance* 12.

Haubrich, Joseph G. (2006), "Does the Yield Curve Signal Recession?" Federal Reserve Bank of Cleveland *Economic Letter*.

Havey, Campbell R. (1989), "Forecasts of Economic Growth from the Bond and Stock Markets," *Financial Analysts Journal* September/October, 38-45.

Lahiri, Kajal, and George J. Wang (2006), "Subjective Probability: Forecasts for Recessions," *Business Economics* 41, 26-36.

NBER (2003), "The NBER's Business-Cycle Dating Procedure," manuscript.

Stock, James H., and Mark W. Watson (1989), "New Indices of Coincident and Leading Indicators," in *NBER Macroeconomic Annual 4*, edited by Oliver Blanchard and Stanley Fischer, 351-394.

Stock, James H., and Mark W. Watson (2003), "How Did Leading Indicator Forecasts Perform During the 2001 Recession?", Federal Bank of Richmond *Economic Quarterly* 89, 71-89.

Tulip, Peter (2005), "Has Output Become More Predictable? Changes in Greenbook Forecast Accuracy," Finance and Economics Discussion Series, 2005-31, Federal Reserve Board.

Wright, Jonathan (2006), "The Yield Curve and Predicting Recessions," Finance and Economics Discussion Series, 2006-07, Federal Reserve Board.

Zarnowitz, Victor and Phillip A. Braun (1993), "Twenty-two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance," in *Business Cycles, Indicators, and Forecasting*, edited by James H. Stock and Mark W. Watson, Chicago: University of Chicago Press.

Table 1

Dating Recessions From Real GDP Growth

	Realtime Data			NBER	Final Data		
Date	Real GDP Growth	R1 Recessions	R2 Recessions	Recession Quarters	Real GDP Growth	R1 Recessions	R2 Recessions
1968:Q4	3.5	0	0	0	1.7	0	0
1969:Q1	2.6	0	0	0	6.5	0	0
1969:Q2	2.0	0	0	0	1.2	0	0
1969:Q3	2.2	0	0	0	2.5	0	0
1969:Q4	-0.4	1*	1*	0	-1.9	1*	1*
1970:Q1	-2.9	1	1	1	-0.7	1	1
1970:Q2	0.6	0*	0*	1	0.7	0*	0*
1970:Q3	1.4	0*	0*	1	3.6	0*	0*
1970:Q4	-3.9	1	0*	1	-4.2	1	0*
1971:Q1	7.9	0	0	0	11.6	0	0
1971:Q2	4.9	0	0	0	2.3	0	0
1971:Q3	2.7	0	0	0	3.2	0	0
1971:Q4	5.9	0	0	0	1.2	0	0
1972:Q1	6.5	0	0	0	7.3	0	0
1972:Q2	9.4	0	0	0	9.8	0	0
1972:Q3	6.4	0	0	0	3.9	0	0
1972:Q4	8.0	0	0	0	6.7	0	0
1973:Q1	8.6	0	0	0	10.6	0	0
1973:Q2	2.4	0	0	0	4.7	0	0
1973:Q3	3.4	0	0	0	-2.1	1*	0
1973:Q4	1.6	0	0	0	3.9	0	0
1974:Q1	-7.0	1	1	1	-3.4	1	0
1974:Q2	-1.6	1	1	1	1.2	0*	0*
1974:Q3	-1.9	1	1	1	-3.8	1	1
1974:Q4	-9.0	1	1	1	-1.6	1	1
1975:Q1	-11.4	1	1	1	-4.7	1	1
1975:Q2	1.9	0	0	0	3.0	0	0
1975:Q3	11.9	0	0	0	6.9	0	0
1975:Q4	5.0	0	0	0	5.4	0	0
1976:Q1	9.0	0	0	0	9.3	0	0
1976:Q2	4.5	0	0	0	3.0	0	0
1976:Q3	3.9	0	0	0	1.9	0	0
1976:Q4	2.6	0	0	0	2.9	0	0
1977:Q1	7.5	0	0	0	4.9	0	0
1977:Q2	6.2	0	0	0	8.1	0	0
1977:Q3	5.1	0	0	0	7.4	0	0
1977:Q4	3.9	0	0	0	0.0	0	0
1978:Q1	-0.1	1*	0	0	1.3	0	0
1978:Q2	8.7	0	0	0	16.7	0	0
1978:Q3	2.6	0	0	0	4.0	0	0

Note: Asterisks denote R1 or R2 recession quarters that differ from NBER recession dating.

Table 1 (Continued)

Date	Real Time Data: Real GDP Growth	Real Time Data: R1 Recessions	Real Time Data: R2 Recessions	NBER Recession Quarters	Final Data: Real GDP Growth	Final Data: R1 Recessions	Final Data: R2 Recessions
1978:Q4	6.9	0	0	0	5.3	0	0
1979:Q1	1.1	0	0	0	0.8	0	0
1979:Q2	-2.3	1*	0	0	0.1	0	0
1979:Q3	3.1	0	0	0	2.9	0	0
1979:Q4	2.0	0	0	0	1.2	0	0
1980:Q1	1.2	0	0	0	1.3	0	0
1980:Q2	-9.6	1	0*	1	-7.8	1	1
1980:Q3	2.4	0*	0*	1	-0.4	1	1
1980:Q4	3.8	0	0	0	7.6	0	0
1981:Q1	1.0	0	0	0	8.4	0	0
1981:Q2	-1.6	1*	0	0	-3.1	1*	0
1981:Q3	1.4	0	0	0	4.9	0	0
1981:Q4	-4.5	1	1	1	-4.9	1	1
1982:Q1	-5.1	1	1	1	-6.4	1	1
1982:Q2	2.1	0*	0*	1	2.1	0*	0*
1982:Q3	0.7	0*	0*	1	-1.5	1	0*
1982:Q4	-1.1	1	0*	1	0.4	0	0*
1983:Q1	2.6	0	0	0	5.0	0	0
1983:Q2	9.7	0	0	0	9.3	0	0
1983:Q3	7.6	0	0	0	8.1	0	0
1983:Q4	5.0	0	0	0	8.4	0	0
1984:Q1	10.1	0	0	0	8.1	0	0
1984:Q2	7.1	0	0	0	7.1	0	0
1984:Q3	1.6	0	0	0	3.9	0	0
1984:Q4	4.3	0	0	0	3.3	0	0
1985:Q1	0.2	0	0	0	3.7	0	0
1985:Q2	1.9	0	0	0	3.4	0	0
1985:Q3	3.0	0	0	0	6.4	0	0
1985:Q4	0.8	0	0	0	3.1	0	0
1986:Q1	3.8	0	0	0	3.9	0	0
1986:Q2	0.6	0	0	0	1.6	0	0
1986:Q3	2.8	0	0	0	3.9	0	0
1986:Q4	1.1	0	0	0	2.0	0	0
1987:Q1	4.4	0	0	0	2.7	0	0
1987:Q2	2.5	0	0	0	4.5	0	0
1987:Q3	4.4	0	0	0	3.7	0	0
1987:Q4	4.8	0	0	0	7.2	0	0
1988:Q1	3.4	0	0	0	2.0	0	0
1988:Q2	3.0	0	0	0	5.2	0	0
1988:Q3	2.5	0	0	0	2.6	0	0

Note: Asterisks denote R1 or R2 recession quarters that differ from NBER recession dating.

Table 1 (Continued)

	Real Time Data			NBER	Final Data		
Date	Real GDP Growth	R1 Recessions	R2 Recessions	Recession Quarters	Real GDP Growth	R1 Recessions	R2 Recessions
1988:Q4	2.4	0	0	0	5.4	0	0
1989:Q1	3.8	0	0	0	4.1	0	0
1989:Q2	2.5	0	0	0	2.4	0	0
1989:Q3	3.0	0	0	0	2.9	0	0
1989:Q4	1.1	0	0	0	1.0	0	0
1990:Q1	1.7	0	0	0	4.7	0	0
1990:Q2	0.4	0	0	0	1.0	0	0
1990:Q3	1.4	0	0	0	0.0	0	0
1990:Q4	-1.6	1	1	1	-3.0	1	1
1991:Q1	-2.8	1	1	1	-2.0	1	1
1991:Q2	-0.5	1*	1*	0	2.6	0	0
1991:Q3	1.8	0	0	0	2.0	0	0
1991:Q4	0.4	0	0	0	1.9	0	0
1992:Q1	2.9	0	0	0	4.2	0	0
1992:Q2	1.5	0	0	0	3.9	0	0
1992:Q3	3.4	0	0	0	4.0	0	0
1992:Q4	4.7	0	0	0	4.5	0	0
1993:Q1	0.7	0	0	0	0.5	0	0
1993:Q2	1.9	0	0	0	2.0	0	0
1993:Q3	2.9	0	0	0	2.1	0	0
1993:Q4	7.0	0	0	0	5.5	0	0
1994:Q1	3.3	0	0	0	4.1	0	0
1994:Q2	4.1	0	0	0	5.3	0	0
1994:Q3	4.0	0	0	0	2.3	0	0
1994:Q4	5.1	0	0	0	4.8	0	0
1995:Q1	2.7	0	0	0	1.1	0	0
1995:Q2	1.3	0	0	0	0.7	0	0
1995:Q3	3.3	0	0	0	3.3	0	0
1995:Q4	0.5	0	0	0	3.0	0	0
1996:Q1	2.0	0	0	0	2.8	0	0
1996:Q2	4.7	0	0	0	6.7	0	0
1996:Q3	2.1	0	0	0	3.4	0	0
1996:Q4	3.8	0	0	0	4.8	0	0
1997:Q1	4.9	0	0	0	3.1	0	0
1997:Q2	3.3	0	0	0	6.2	0	0
1997:Q3	3.1	0	0	0	5.1	0	0
1997:Q4	3.7	0	0	0	3.0	0	0
1998:Q1	5.6	0	0	0	4.5	0	0
1998:Q2	1.8	0	0	0	2.7	0	0
1998:Q3	3.7	0	0	0	4.7	0	0

Note: Asterisks denote R1 or R2 recession quarters that differ from NBER recession dating.

Table 1 (Continued)

Date	Real Time Data: Real GDP Growth	Real Time Data: R1 Recessions	Real Time Data: R2 Recessions	NBER Recession Quarters	Final Data: Real GDP Growth	Final Data: R1 Recessions	Final Data: R2 Recessions
1998:Q4	6.0	0	0	0	6.2	0	0
1999:Q1	4.3	0	0	0	3.4	0	0
1999:Q2	1.9	0	0	0	3.4	0	0
1999:Q3	5.7	0	0	0	4.8	0	0
1999:Q4	7.3	0	0	0	7.3	0	0
2000:Q1	4.8	0	0	0	1.0	0	0
2000:Q2	5.7	0	0	0	6.4	0	0
2000:Q3	2.2	0	0	0	-0.5	1*	0
2000:Q4	1.0	0	0	0	2.1	0	0
2001:Q1	1.3	0	0	0	-0.4	1*	0
2001:Q2	0.3	0*	0*	1	1.2	0*	0*
2001:Q3	-1.3	1	0*	1	-1.4	1	0*
2001:Q4	1.7	0*	0*	1	1.6	0*	0*
2002:Q1	5.0	0	0	0	2.7	0	0
2002:Q2	1.3	0	0	0	2.2	0	0
2002:Q3	4.0	0	0	0	2.4	0	0
2002:Q4	1.4	0	0	0	0.2	0	0
2003:Q1	1.4	0	0	0	1.2	0	0
2003:Q2	3.3	0	0	0	3.5	0	0
2003:Q3	8.2	0	0	0	7.5	0	0
2003:Q4	4.1	0	0	0	2.7	0	0
2004:Q1	4.5	0	0	0	3.9	0	0
2004:Q2	3.3	0	0	0	4.0	0	0
2004:Q3	4.0	0	0	0	3.1	0	0
2004:Q4	3.9	0	0	0	2.6	0	0
2005:Q1	3.8	0	0	0	3.4	0	0
2005:Q2	3.3	0	0	0	3.3	0	0
2005:Q3	4.1	0	0	0	4.2	0	0
2005:Q4	1.7	0	0	0	1.8	0	0

Note: Asterisks denote R1 or R2 recession quarters that differ from NBER recession dating.

Table 2

Evaluation of Accuracy of Realtime Probability Forecasts

Probability forecast	Full sample MAE	Full sample RMSE	Post-1987 sample MAE	Post-1987 sample RMSE
	Current-quarter R1 recession prediction			
SPF reported forecast	0.168	0.261	0.129	0.193
SPF implied forecast	0.183	0.277	0.167	0.214
Naive forecast	0.198	0.335	0.146	0.229
Yield spread forecast	0.205	0.313	0.163	0.250
	One-quarter-ahead R1 recession prediction			
SPF reported forecast	0.212	0.291	0.154*	0.196
SPF implied forecast	0.229	0.312	0.198	0.229
Naive forecast	0.238	0.337	0.194	0.248
Yield spread forecast	0.186†	0.296	0.150*	0.238
	Two-quarter-ahead R1 recession prediction			
SPF reported forecast	0.234*	0.313*	0.173*	0.214*
SPF implied forecast	0.259	0.338	0.231	0.261
Naive forecast	0.263	0.342	0.222	0.263
Yield spread forecast	0.195*†	0.303	0.151*	0.222*
	Three-quarter-ahead R1 recession prediction			
SPF reported forecast	0.248*	0.330	0.194*	0.244*
SPF implied forecast	0.258	0.341	0.236	0.272
Naive forecast	0.270	0.344	0.229	0.268
Yield spread forecast	0.201*†	0.307*	0.144*	0.218*†
	Four-quarter-ahead R1 recession prediction			
SPF reported forecast	0.255	0.338	0.205	0.259
SPF implied forecast	0.255	0338	0.237	0.277
Naive forecast	0.271	0.345	0.230	0.269
Yield spread forecast	0.206*†	0.311*†	0.147*†	0.220*†

Note: The asterisk denotes SPF or yield curve forecasts that are significantly more accurate than the Naive forecast at the 5 percent level. The dagger denotes yield curve forecasts that are significantly more accurate than the SPF reported forecast at the 5 percent level.

Figure 1
SPF and Naive R1 Recession Probabilities

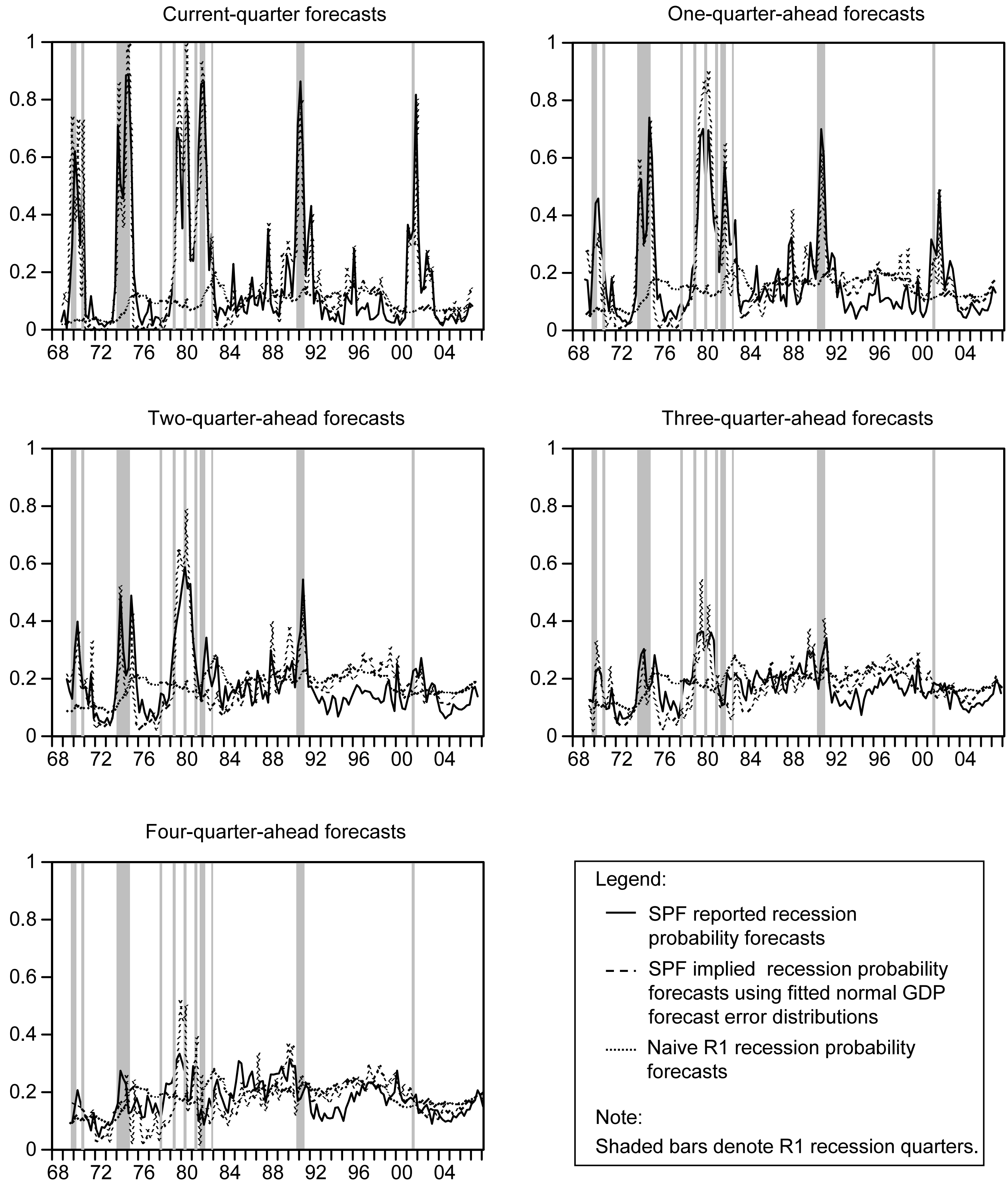


Figure 2
Differences Between SPF Reported and Implied Probabilities

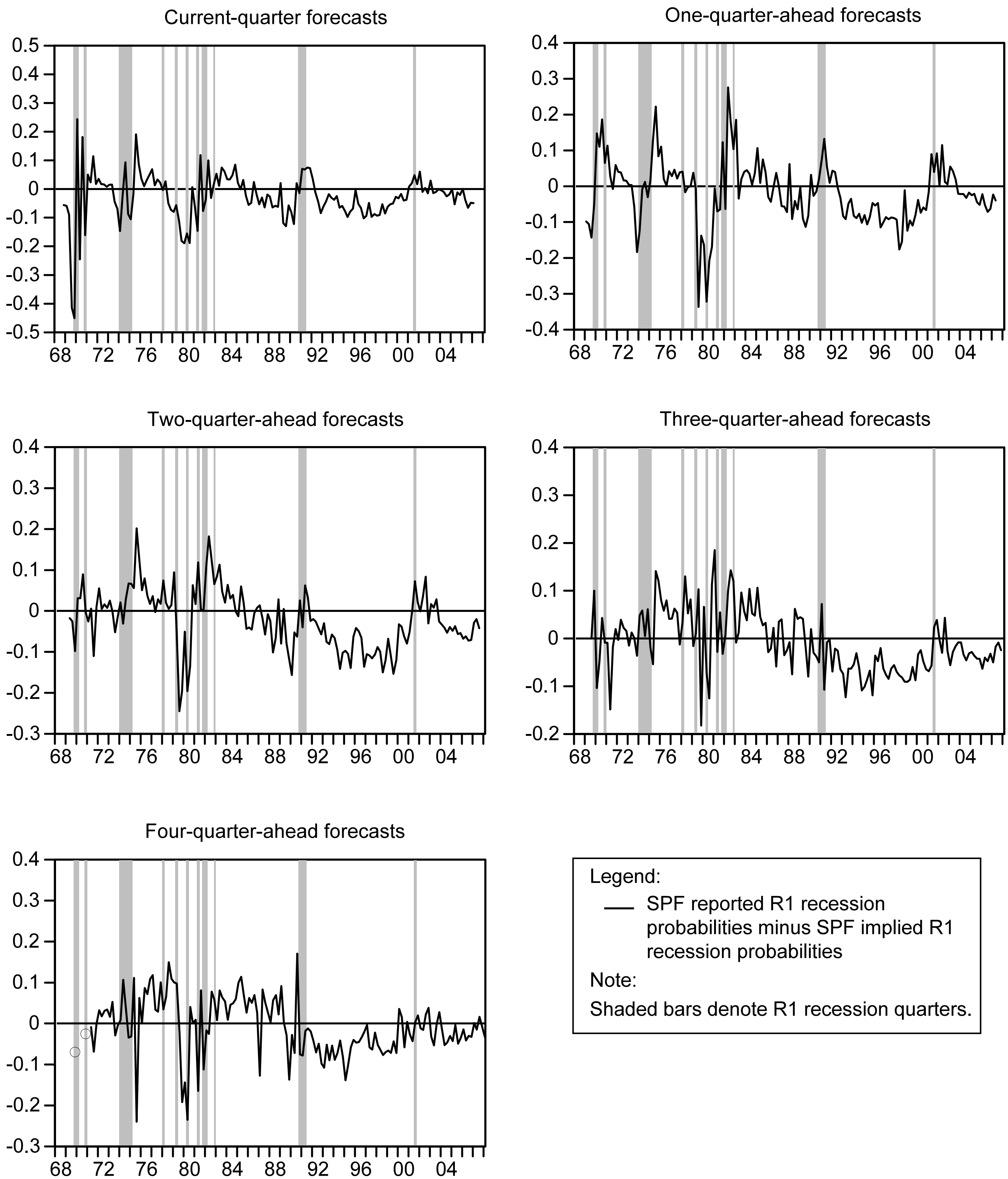
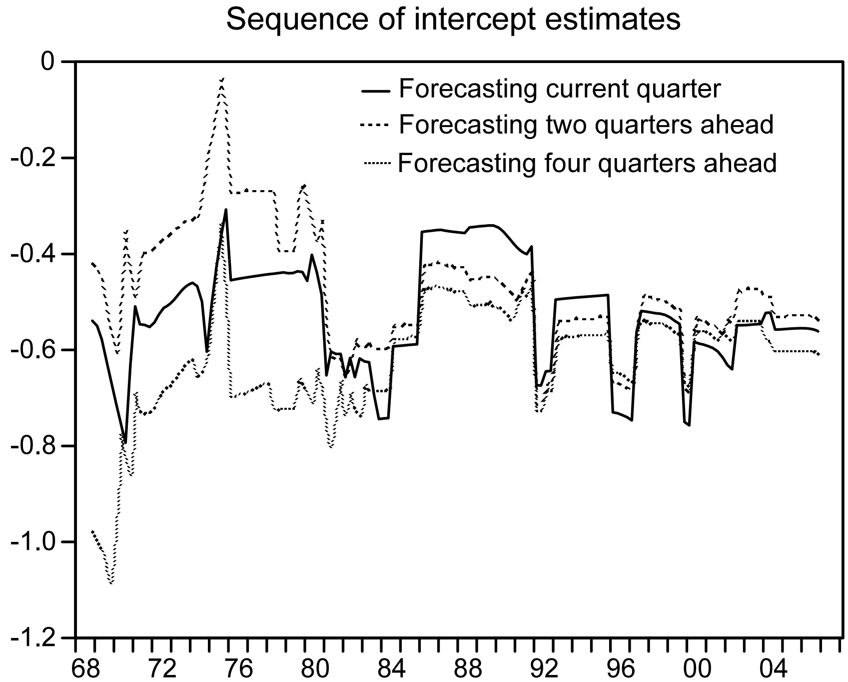
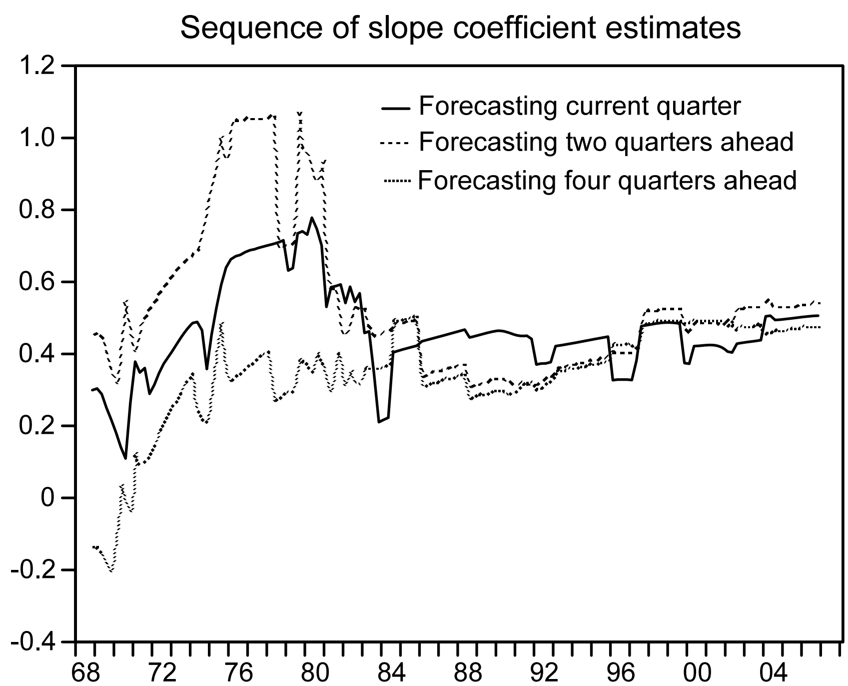


Figure 3
Coefficient Estimates from Realtime Probit Regressions





Note: Probit regressions have R1 recession probability as the left-hand-side variable. Expanding sample realtime estimation begins with 10 years of data starting in 1955Q1. Forecasts for each quarter uses vintage data estimated through a year prior.

Figure 4
SPF and Yield Spread R1 Recession Probabilities

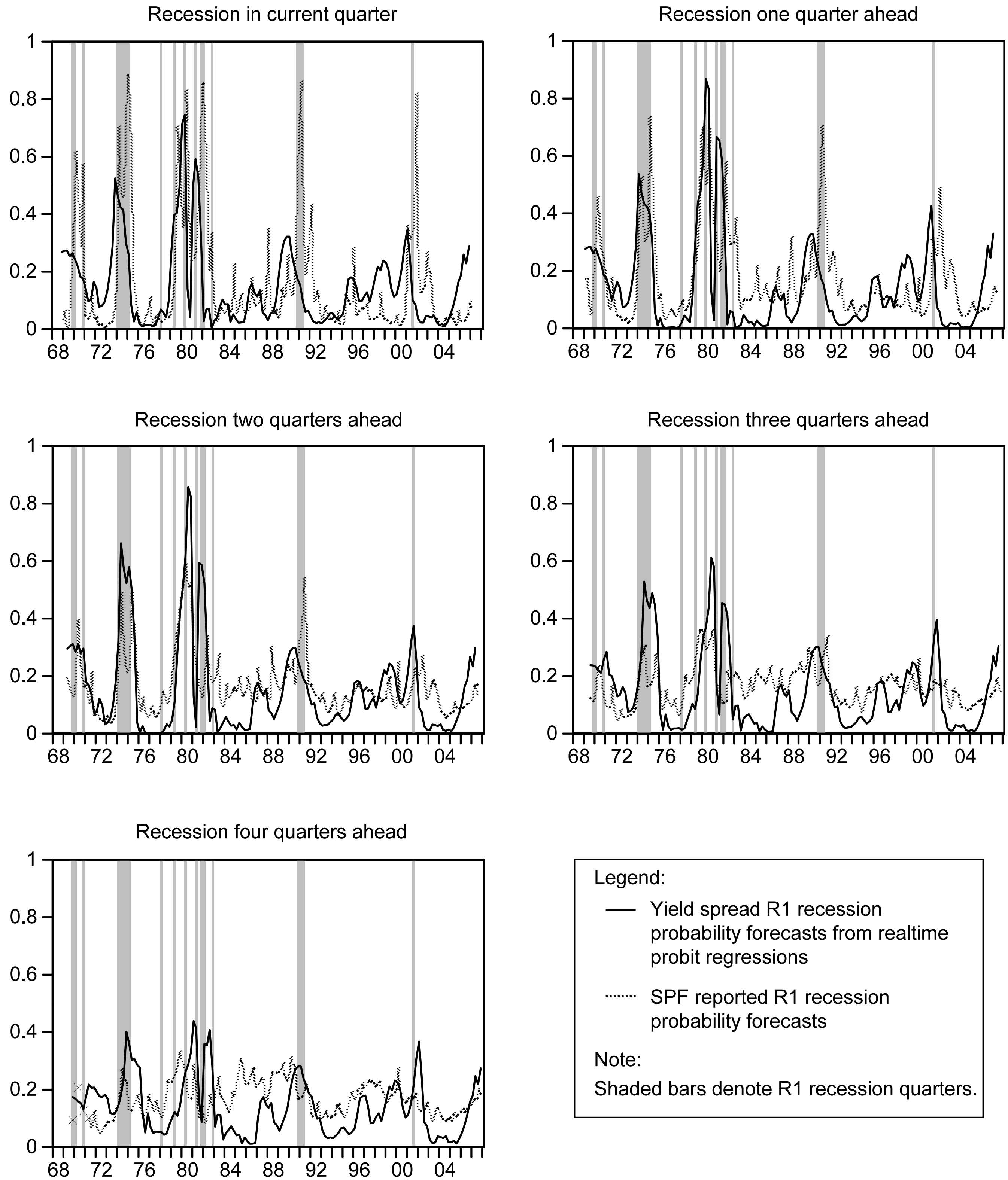
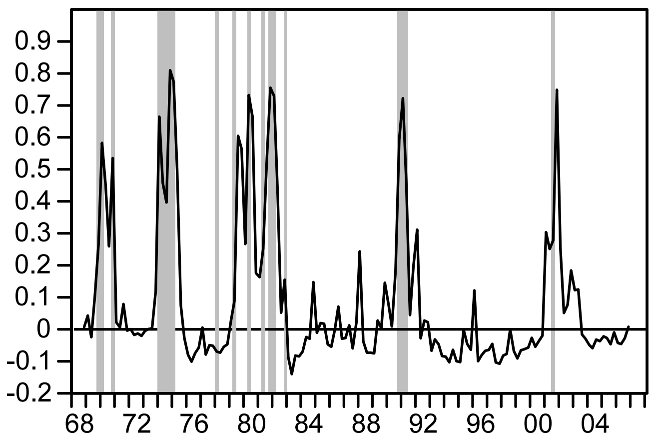
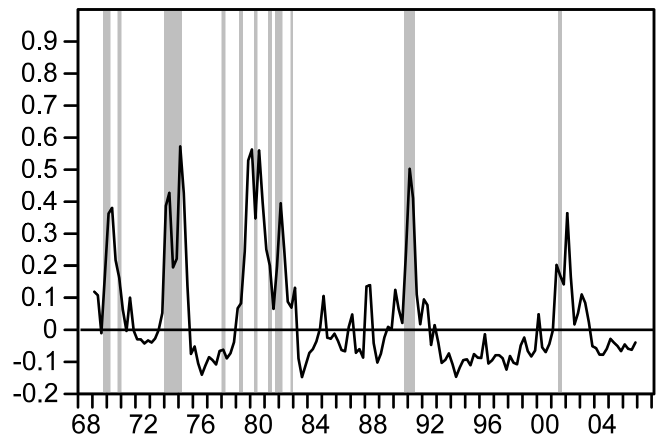


Figure 5
Differences Between SPF and Naive Probabilities

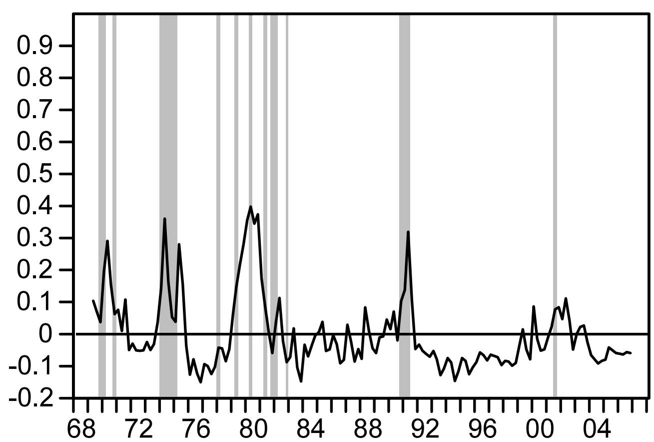
Current-quarter forecasts



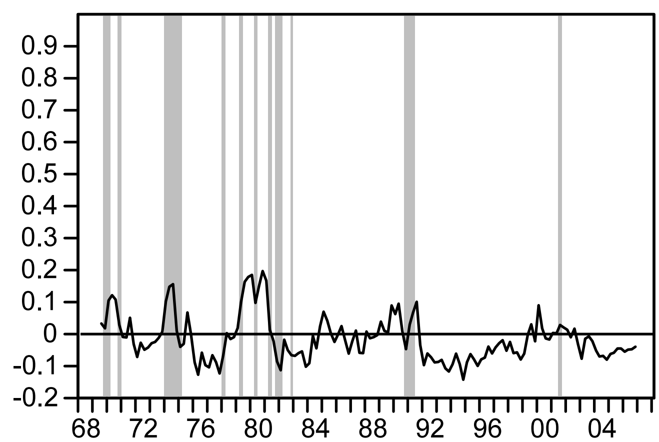
One-quarter-ahead forecasts



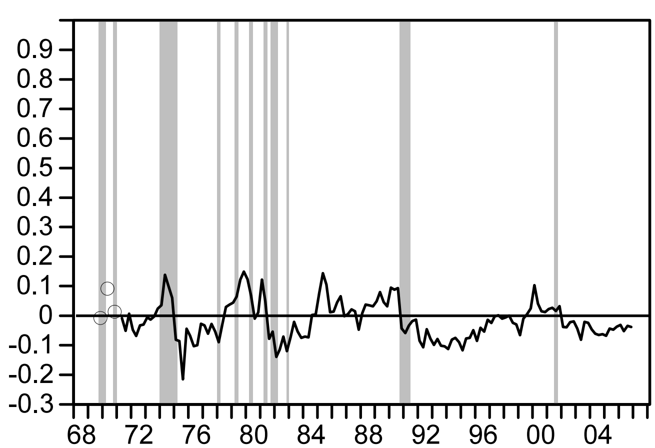
Two-quarter-ahead forecasts



Three-quarter-ahead forecasts



Four-quarter-ahead forecasts



Legend:

— SPF reported probabilities minus Naive recession probabilities

Note:

Shaded bars denote R1 recession quarters.

Figure 6
SPF and Naive Probability Forecast Errors

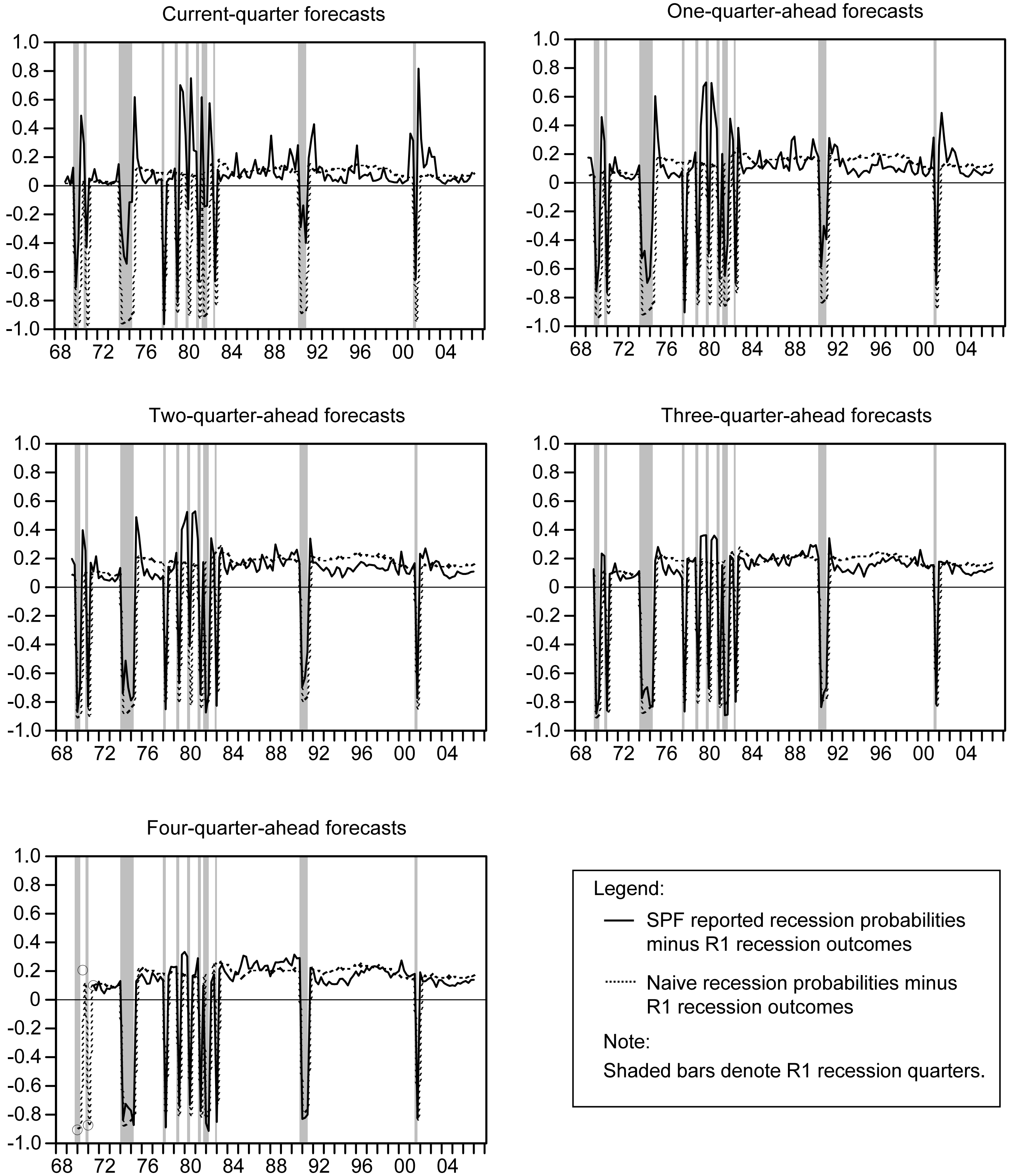


Figure 7
Differences Between Yield Spread and SPF Probabilities

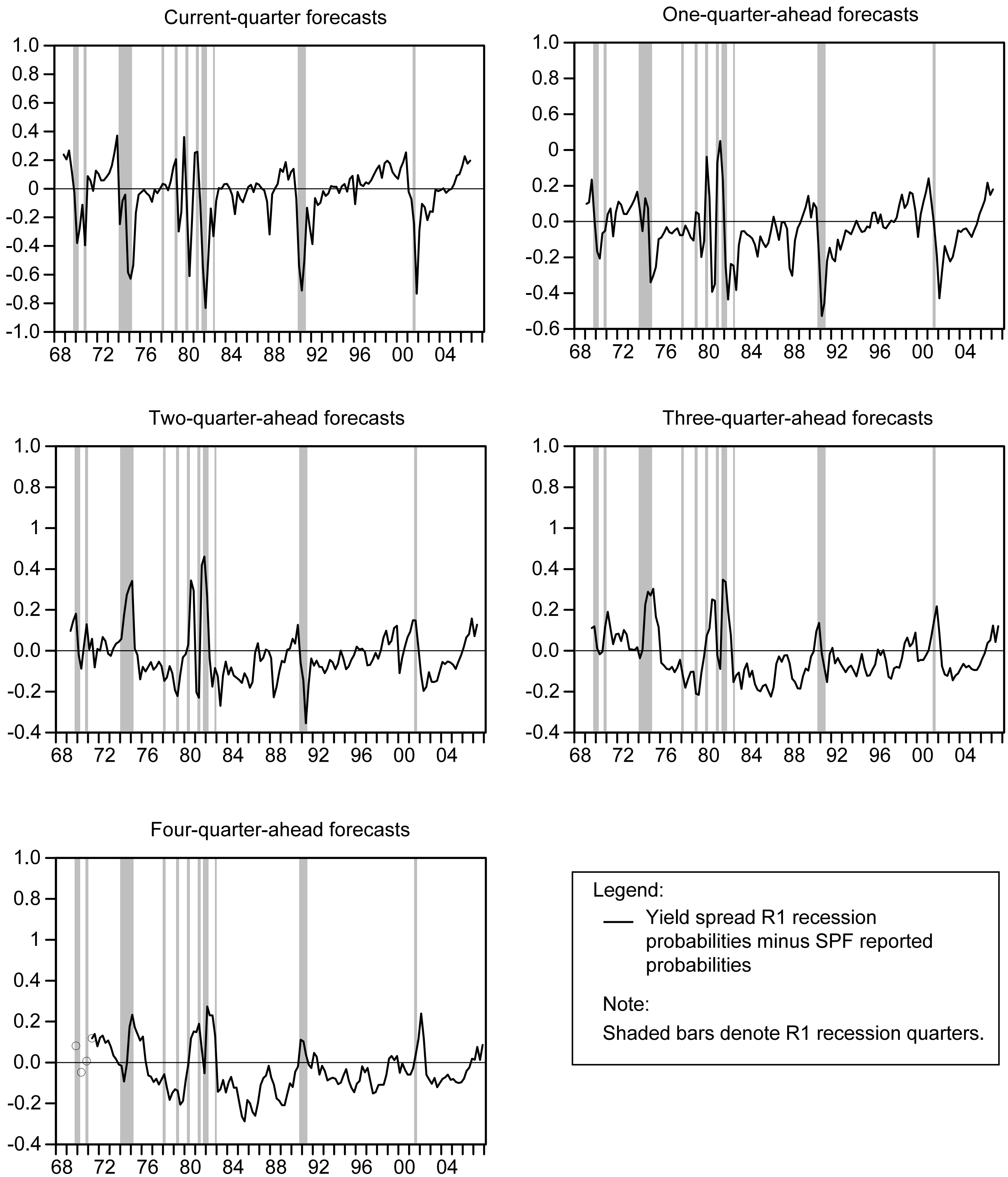


Figure 8
SPF and Yield Spread Probability Forecast Errors