NBER WORKING PAPER SERIES

PREDICTING U.S. RECESSIONS:
FINANCIAL VARIABLES AS
LEADING INDICATORS

Arturo Estrella
Frederic S. Mishkin

Working Paper 5379

PREDICTING U.S. RECESSIONS:
FINANCIAL VARIABLES AS
LEADING INDICATORS

ABSTRACT

This article examines the performance of various financial variables as predictors of subsequent U.S. recessions. Series such as interest rates and spreads, stock prices, currencies, and monetary aggregates are evaluated singly and in comparison with other financial and non-financial indicators. The analysis focuses on out-of-sample performance from 1 to 8 quarters ahead. Results show that stock prices are useful with 1-2 quarter horizons, as are some well-known macroeconomic indicators. Beyond 2 quarters, the slope of the yield curve emerges as the clear choice, and typically performs better by itself out of sample than in conjunction with other variables.

Arturo Estrella
Federal Reserve Bank of New York
33 Liberty Street
New York, NY 10045

Frederic S. Mishkin
Federal Reserve Bank of New York
33 Liberty Street
New York, NY 10045
and NBER

1. Introduction

Financial variables, such as the prices of financial instruments, are commonly associated with expectations of future economic events. For example, long-term interest rates are frequently analyzed as weighted averages of expected future short-term interest rates. In this framework, spreads between rates of different maturities are interpreted as expectations of future rates corresponding to the period between the two maturities. Stock prices are similarly interpreted as expected discounted values of future dividend payments, which incorporates views regarding both the future profitability of the firm and future interest or discounting rates.

In this article, we examine the usefulness of various financial variables in predicting specifically whether or not the U.S. economy will be in a recession anywhere between 1 and 8 quarters ahead. The variables are examined by themselves and in some plausible combinations, and the results are compared with similar exercises involving more traditional macroeconomic indicators, including widely used indices of leading indicators and their component variables.

The present analysis differs in two important respects from earlier research which examines the usefulness of financial variables in predicting future macroeconomic outcomes.[1] First, we focus simply on predicting recessions, rather than quantitative measures of future economic activity. We believe that this is a useful exercise in that it addresses a question frequently posed by policy makers and market participants, and it sidesteps the problem of spurious accuracy associated with quantitative point estimates of, say, future real GDP growth.[2] Second, the primary criterion of predictive accuracy in this

1

article is out-of-sample performance, that is, accuracy in predictions for quarters beyond the period over which the model is estimated. In-sample performance can always be improved by introducing additional variable, but in the out-of-sample context, more is not necessarily better, as our results will show.

With the existence of large scale macroeconometric models and with the judicious predictions of knowledgeable market observers, why should we care about the indications of one or a few financial variables? Is such an approach too simplistic?

There are several important reasons why policy makers and market participants should look at a few well-chosen financial indicators. First, the indicators may be used to double-check both econometric and judgmental predictions. There is no question that forecasting with macroeconometric models can be quite helpful. Beyond the mere potential accuracy of the forecasts, such models allow the economic analyst to think about the causal relationships that may lead to a specific result, that is, to think about the structure of the economy. In many cases, the bottom line prediction is not the most interesting or useful part of the modeling exercise. Judgmental forecasts, although not necessarily based on strict statistical analysis, are also typically based on thinking about economic relationships and have similar benefits.

Nevertheless, a quick look at a financial indicator may flag some problem with the results of more involved approaches. If the model and the indicator agree, confidence in the model's results can be enhanced. If the indicator gives a different signal, the model is not necessarily wrong, but it may be worthwhile to review the assumptions and relationships that

led to the prediction. Of course, the significance one may attach to a particular indicator depends on its historical out-of-sample performance, which is the focus of this article.

A second reason for looking at simple financial indicators is the phenomenon that econometricians call "overfitting". Most econometric models forecast future activity through the use of some sort of statistical regression. Weighted sums of explanatory variables are constructed in order to maximize the predictive power of the sum over the sample period. Generally, the more variables one includes, the better the in-sample results will be. However, liberal inclusion of explanatory variables in the regression will not necessarily help, and frequently hurts, when extrapolating beyond the sample's end.

Intuitively, the reason is that even when a variable is not "truly" related to future economic activity (its "true" weight is zero), the estimation procedure is subject to error and may produce a non-zero weight. With this incorrect weight, the predictions of the model may be worse than if the specific variable had been left out altogether. The potential cost of leaving out a variable that belongs in the model has to be considered against the potential cost of including a variable that does not belong. Our results suggest that, in predicting recessions, especially with long horizons, the second type of cost is typically large.

A third reason for looking at financial indicators is that it is quick and simple. Of course, this presupposes that they are also accurate, otherwise it is a waste of (a little bit of) time. Our analysis should be helpful in determining which particular indicators are worth watching. An additional benefit of the analysis in this paper is that it provides a forecasted probability of a future recession, a probability that is of interest in its own right.

To preview the results, the analysis focuses on out-of-sample performance in predicting whether or not the economy will be in a recession between 1 and 8 quarters ahead. Stock prices are useful predictors, particularly 1 and 2 quarters ahead. This performance is comparable to that of some well-known macroeconomic indicators, such as the Commerce Department's index of leading indicators and its component series. Beyond 2 quarters, the slope of the yield curve emerges as the clear choice. It outperforms other indicators in one-on-one comparisons, and the addition of other variables is generally more likely to hurt at these longer horizons.

In the following section, we describe the basic model used to perform the predictive tests, as well as the criteria used to evaluate the results. Next, we list and explain the indicators that are included in the tests. We then proceed to discuss some in-sample results that are both illustrative and somewhat useful in model selection. Full in-sample results are contained in the appendix. The text continues with the out-of-sample results, which is the central part of the article, and concludes with a case study that shows how the indicators would be estimated and used in practice.

2. The basic model and criteria for evaluation of results

2.1 The model

In order to quantify the predictive power of the variables examined with respect to future recessions, we use a statistical regression technique. The particular form of the model used, the probit equation, is dictated by the fact that the variable being predicted takes on only two possible values. Specifically, the dependent variable in the model is

$$R_t = 1 \qquad \text{if the economy is in recession in quarter t}$$
$$\quad = 0 \qquad \text{otherwise.}$$

With this type of dependent variable, the usual linear regression model is inappropriate. However, the probit form is very analogous to the linear regression and the interpretation of its results is similar.

The probit equation we use has the general form

$$P(R_{t+k}=1) = F(\alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{2t} + ...) \, , \qquad \qquad \textbf{(1)}$$

where the α coefficients are statistically estimated and F is the normal cumulative distribution function (see box). A weighted sum of one or more explanatory variables (X_{it}), observed as of the end of quarter t, is used to predict whether at k quarters ahead R_{t+k} will be 1 or 0. The effect of applying the function F to the weighted sum is to convert the result into a probability that a recession will result in quarter t+k. A probability close to 1 indicates a strong prediction of a recession, whereas the opposite is true of a probability close to 0.

As an example, consider the results of one of the most successful models in the paper, the probit equation for a recession four quarters ahead estimated using only the spread

between the 10-year Treasury note and the 3-month Treasury bill (SPREAD) as an explanatory variable. The estimates using data from the first quarter of 1960 to the first quarter of 1995 are:

$$P(R_{t+4}=1) = F(-0.66 - 0.81\, QSPT_t) \ . \tag{2}$$

Table 1 shows what values of the spread variable, SPREAD, correspond to estimated probabilities of a recession four quarters in the future. It shows that the estimated probability of a recession four quarters ahead from this model is 10% when the spread averages 0.76 percentage points over the quarter, while the probability is 50% when the spread averages -0.82 percentage points, and is 90% when the spread averages -2.40 percentage points.

To see how a probit model like this would be used in practice, consider that in the third quarter of 1994, the spread averaged 2.74 percentage points, and the corresponding predicted probability of a recession in the third quarter of 1995 was only 0.2 percent, and indeed a recession did not materialize. In contrast, the spread averaged -2.18 percentage points in the first quarter of 1981, implying a probability of a recession of 86.5 percent four quarters later. As predicted, the first quarter of 1982 was in fact labeled as a recession quarter by the National Bureau of Economic Research (NBER).

6

Table 1

Estimated Recession Probabilities for Probit Model With SPREAD

$$P(R_{t+4}=1) = F(-0.66 - 0.81 \; SPREAD_t)$$

Recession Probability 4 Quarters Ahead (Percent)	Value of SPREAD (Percentage Points)
5	1.21
10	0.76
15	0.46
20	0.22
25	0.02
30	-0.17
40	-0.50
50	-0.82
60	-1.13
70	-1.46
80	-1.85
90	-2.40

2.2 Test criteria

In this article, we examine many variables with potential predictive power for recessions and we consider each variable with predictive horizons ranging from 1 to 8 quarters ahead. The volume of output generated by this type of analysis makes it important to be able to summarize the results in a meaningful way. Hence, we introduce a few summary measures of the predictive power of a given variable with a given horizon.

The principal measure is a pseudo R^2 developed in Estrella (1995), that is, a simple measure of goodness of fit that corresponds intuitively to the widely-used coefficient of determination, or R^2, in a standard linear regression. Although the absolute levels of this new measure may differ from standard measures proposed earlier in the literature, the ordering of alternative models produced by the different measures is consistent. For the in-sample results, the measure takes on values between 0 and 1. A value of this measure that is close to 0 indicates that the variable or variables in the model have little explanatory power, a value close to 1 indicates a very close fit, and intermediate values may be used to rank the models in terms of predictive power. When applied to out-of-sample results, there is not guarantee that the value of the pseudo R^2 will lie between 0 and 1, as is also the case in the standard linear regression. Nevertheless, the pseudo R^2 for out-of-sample results is useful as a simple measure of fit and is reported in the appropriate section.

As in the linear regression case, the pseudo R^2 is associated with a valid statistical test but, by itself, is not sufficient for statistical hypothesis testing. Therefore, we present two additional measures that are associated with valid statistical tests. One is analogous to the t-statistic for an individual variable in a linear regression and may be interpreted in a similar

way. As a rule of thumb, an absolute value of 2 or above tends to indicate statistical significance. Finally, we also indicate whether a formal statistical test indicates that a variable is significant at the 5 percent and 1 percent levels.

In the earlier numerical example in which the Treasury spread is used to predict a recession 4 quarters ahead, the pseudo R^2 is 0.296 and the t-statistic for the spread variable SPREAD is -4.99. This result is significant at the 1 percent level and is one of the strongest single-variable results in the article. The negative sign associated with the t-statistic indicates that the variable is inversely related to the probability of a recession. That is, as the spread becomes larger -- as the yield curve steepens -- the probability of a subsequent recession diminishes.

3. Indicators examined and data used

The primary focus of this paper is to test whether simple financial variables may be useful predictors of future recessions. Thus, many of the variables we examine are interest rates, interest rate spreads, stock price indices, and monetary aggregates, both nominal and real. However, in order to establish the usefulness of these results, it is necessary to compare them with models based on traditional macroeconomic indicators. We therefore also include as explanatory variables the Commerce Department's index of leading economic indicators and several of its component series, two experimental indices of leading indicators constructed by Stock and Watson (1989) in conjunction with the NBER, and also lagged growth in real GDP.

The macroeconomic indicators have an established performance record with regard to the prediction of real activity. That record is not always subjected to comparison tests, and most of the predictive lead times are not as long as users might prefer. Nevertheless, a key question is whether our financial series can provide useful predictions over and beyond the results obtainable with the traditional variables.

Another important consideration is the possible lag in the availability of the data for the explanatory variables. Some variables, such as interest rates and stock prices, are available essentially on a continuous basis and there is effectively no informational lag. In contrast, many monthly macroeconomic series are only available one or two months after the period covered by the data, and GDP has a lag of almost one full quarter. To place all the variables on an equal footing with regard to these lags, only observations actually available as of the end of a given quarter are assigned to that quarter.

9

The recession variable is constructed using the NBER dates. Table 2 contains the names and descriptions of the other series, as well as the informational lag used for each variable, in months. Interest rates and spreads are calculated on a quarterly average basis. For other variables, a quarterly growth rate is used with the lags indicated in the table.

The equations are estimated using quarterly data from the first quarter of 1959 to the first quarter of 1995. The precise starting date does not seem to be crucial. The date actually chosen maximizes the availability of comparable data for all series. Results using data for some series that are available further back in the 1950s are not appreciably different from those presented in the article. Even though most series are available on a monthly basis, the estimates in this paper are based on quarterly data for two basic reasons: monthly data are generally too noisy, which produces somewhat weaker results, and using quarterly data again guarantees comparability of all series. However, we have found that results with monthly data lead to similar conclusions on the usefulness of financial indicators.[3]

Table 2

Indicator Series and Their Information Lags

Series	Description	Information Lag (Months)
Interest rates and spreads:		
SPREAD	10 year-3 month Treasury spread	0
CPTB	CP-Treasury spread (6 months)	0
BILL	3-month T bill	0
BOND	10-year T bond	0
Stock prices:		
DJIA	Dow Jones industrials	0
NYSE	NYSE composite	0
SP500	S&P 500	0
Monetary aggregates:		
M0	Monetary base	1
M1	M1	1
M2	M2	1
M3	M3	1
RM0	Monetary base deflated by CPI	1
RM1	M1 deflated by CPI	1
RM2	M2 deflated by CPI	1
RM3	M3 deflated by CPI	1
Individual macro indicators:		
GDPG1	Growth in real GDP, previous quarter	3
NAPMC	Purchasing managers' survey	0
VP	Vendor performance	0
CORD	Contracts and orders for plant and equip.	1
HI	Housing permits	1
CEXP	Consumer expectations (MI)	0
TWD	Trade-weighted dollar	0
MORD	Change in manufacturers' unfilled durable orders	1
Indexes of leading indicators:		
LEAD	Commerce Dept. leading index	2
XLI	Stock-Watson leading index	1
XLI2	Stock-Watson leading index (new)	1

Note: Interest rates and spreads are quarterly average levels, other variables are quarterly growth rates.

4. In-sample results

In-sample results are based on equations estimated over the entire sample period. Their predictions or fitted values are then compared with the actual recession dates. As noted in section 2, three types of results are provided: a pseudo R^2, a t-statistic, and indicators of significance at the 5 and 1 percent levels. Since the focus of the article is out-of-sample prediction, only a few selected in-sample results are presented in the section. The full in-sample results are provided in the appendix.

The general strategy of the analysis is the following. The probit equation is estimated using each series in table 1. Since the Treasury spread produced consistently strong results across all horizons, equations are also run containing the spread and each one of the other variables in turn. Some of the main results are summarized in tables 3 and 4 (the full results are presented in tables A1 and A2 in the appendix). In addition, for a few variables, in-sample results indicate that a second lag of the variable may be significant. For those variables, two-lag models are estimated with and without the spread (see appendix tables A3 and A4).

Table 3 contains several of the variables that performed best in sample and for which representative patterns of significance may be identified. The leading indicators and GDP, for example, are clearly strong predictors in the very short run, with the significance generally declining within a year. The strong performer among these is the original Stock-

Table 3

Measures of Fit and t-Statistics for Probit Models
Variables by themselves -- IN sample

$$P(R_{t+k}=1) = F(\alpha_0 + \alpha_1 X_{1t})$$

k = Quarters Ahead

X_{1t} Variables		1	2	3	4	5	6	7	8
SPREAD	Pseudo R^2	0.071	0.211	0.271	0.296	0.256	0.149	0.078	0.031
	t-stat	-3.01**	-4.56**	-4.89**	-4.99**	-4.78**	-4.03**	-3.09**	-2.02*
CPTB	Pseudo R^2	0.103	0.061	0.026	0.001	0.001	0.001	0.008	0.01
	t-stat	3.41**	2.74**	1.91	0.44	-0.41	-0.32	-1.01	-1.08
RM0	Pseudo R^2	0.153	0.103	0.156	0.168	0.118	0.072	0.046	0.014
	t-stat	-4.1**	-3.53**	-4.19**	-4.27**	-3.71**	-2.98**	-2.4*	-1.35
NYSE	Pseudo R^2	0.174	0.133	0.08	0.043	0.003	0	0.004	0.03
	t-stat	-4.35**	-3.91**	-3.19**	-2.4*	-0.68	0.07	0.76	1.93
LEAD	Pseudo R^2	0.236	0.132	0.112	0.018	0.005	0	0.006	0.007
	t-stat	-4.86*	-4.06**	-3.77**	-1.58	-0.8	-0.1	0.88	0.94
XLI	Pseudo R^2	0.387	0.332	0.205	0.103	0.056	0.022	0.006	0.001
	t-stat	-5.2**	-5.57**	-4.99**	-3.72**	-2.78**	-1.75	-0.91	-0.39
XLI2	Pseudo R^2	0.239	0.091	0.059	0.002	0.008	0.011	0.012	0.017
	t-stat	-4.73**	-3.38**	-2.76**	-0.51	1	1.19	1.24	1.49
GDPG1	Pseydo R^2	0.261	0.16	0.093	0.008	0.002	0.008	0.007	0.015
	t-stat	-4.93*	-4.29**	-3.4**	-1.02	-0.5	-1.02	0.98	1.42

For each model, the first row shows the pseudo R^2 and the second row contains the t-statistic for that variable. Significance is determined using a likelihood ratio test.

*Significant at the 5 percent level.
**Significant at the 1 percent level.

Watson indicator. Among the financial variables, stock prices and the commercial paper spread exhibit a pattern similar to the indices, although the fit is generally not as good, particularly for the spread.

Since the commercial paper spread is the difference between two six-month rates, it is not surprising that its predictive power is at 1 and 2 quarter horizons.[4] Stock prices are more forward-looking, at least in principle. Stock prices may be interpreted as expected present values of future dividend streams. Although the discounting reduces the effective predictive horizon, the projections should be relatively long-term. This is confirmed empirically by the results for the NYSE index, which are significant up to 4 quarters.

The real monetary base performs very well within the first year, and its fit is remarkably consistent over quarters 1 through 4. This pattern confirms the traditional view that the effects of monetary policy become apparent with long lags. The caveat is that the frequently stated principle says that the lags are not just long but also variable, and that the variables in question are supposed to be nominal rather than real. The predictive performance of nominal money growth is uniformly poor, although it should be noted that the dependent variable -- the recession index -- is real.

Finally, the significance and fit of the yield curve spread increase up to the fourth quarter, in which they peak. Technically, this spread corresponds to a forward interest rate applicable from 3 months to 10 years into the future. Although, as with stock prices, there should be some implicit discounting and a corresponding reduction in effective horizon, the spread should therefore be very forward looking. Between quarters 2 and 6, the Treasury spread is more significant than the other variables, with the exception of the Stock-Watson

indicator in quarter 2. It should also be noted that the Stock-Watson indicator includes a yield curve spread as one of its constituent variables, from which it seems to derive much of its out-of-sample predictive power.[5]

When the yield curve spread is combined with the other variables, as in table 4, the results of the single-variable analysis are generally confirmed, although some interesting combinations result. On one hand, the significance of the yield curve is basically undiminished beyond the first 2 to 3 quarters. Even within that range, only the real monetary base undoes the significance of the spread at the 5 percent level, and then only 1 quarter ahead. On the other hand, the other variables remain strong within 2 to 3 quarters, with two exceptions. With the inclusion of the spread, both the commercial paper spread and the real base become insignificant beyond one quarter.

The results for the model that combines the yield curve with stock prices suggests that these two financial variables, which are readily and continuously available, form a very strong combination across all the horizons examined. The significance at the short end is enhanced by the inclusion of the stock index, and the significance at the long end driven largely by the spread.[6]

Table 4

Measures of Fit and t-Statistics for Probit Models
Variables with spread -- IN sample

$$P(R_{t+k}=1) = F(\alpha_0 + \alpha_1 X_{1t} + \alpha_2 SPREAD_t)$$

k = Quarters Ahead

X_{1t} Variables		1	2	3	4	5	6	7	8
SPREAD	Pseudo R^2	0.071	0.211	0.271	0.296	0.256	0.149	0.078	0.031
	t-stat	-3.01**	-4.56**	-4.89**	-4.99**	-4.78**	-4.03**	-3.09**	-2.02*
CPTB	Pseudo R^2	0.142	0.233	0.272	0.307	0.285	0.165	0.102	0.051
	t-stat	2.84**	1.63	0.47	-1.09	-1.75	-1.36	-1.67	-1.54
	t-stat sp	-2.23*	-4.19**	-4.73**	-5.02**	-4.91**	-4.17**	-3.35**	-2.33*
RM0	Pseudo R^2	0.154	0.213	0.283	0.309	0.258	0.151	0.081	0.033
	t-stat	-3.13**	-0.49	-1.22	-1.23	-0.38	-0.44	-0.61	-0.14
	t-stat sp	-0.48	-3.44**	-3.6**	-3.72**	-3.67**	-3.**	-2.12*	-1.57
NYSE	Pseudo R^2	0.223	0.32	0.321	0.314	0.261	0.159	0.096	0.083
	t-stat	-4.08**	-3.44**	-2.48*	-1.49	0.78	1.13	1.52	2.46*
	t-stat sp	-2.45*	-4.18**	-4.58**	-4.77**	-4.77**	-4.11**	-3.3**	-2.56*
LEAD	Pseudo R^2	0.256	0.283	0.331	0.296	0.265	0.16	0.106	0.054
	t-stat	-4.38**	-3.**	-2.72**	-0.08	0.99	1.14	1.83	1.62
	t-stat sp	-1.58	-3.89**	-4.38**	-4.87**	-4.6**	-4.04**	-3.37**	-2.41*
XLI	Pseudo R^2	0.43	0.35	0.298	0.297	0.274	0.179	0.106	0.047
	t-stat	-4.92**	-3.87**	-1.89	0.34	1.45	1.63	1.58	1.26
	t-stat sp	2.17*	-1.43	-3.12**	-4.17**	-4.09**	-3.8**	-3.26**	-2.34*
XLI2	Pseudo R^2	0.289	0.268	0.298	0.302	0.356	0.21	0.121	0.07
	t-stat	-4.51**	-2.66**	-1.87	0.86	2.95**	2.48*	2.09*	2.07*
	t-stat sp	-2.42*	-4.14**	-4.57**	-4.98	-4.56**	-4.25**	-3.44**	-2.5*
GDPG1	Pseudo R^2	0.313	0.358	0.357	0.297	0.258	0.151	0.094	0.058
	t-stat	-4.68**	-3.92**	-3.09**	-0.34	0.48	-0.43	1.46	1.8
	t-stat sp	-2.45*	-4.1**	-4.58**	-4.95**	-4.74**	-3.95**	-3.22**	-2.32*

"t-stat sp" indicates the t statistic for the SPREAD variable.

*Significant at the 5 percent level.
**Significant at the 1 percent level.

5. Out-of-sample results

The out-of-sample results are obtained in the following way. First, a given model is estimated using data from the beginning of the sample up to a given quarter, say the first quarter of 1970. Then these estimates are used to form projections, say 4 quarters ahead. In the example, the projection would apply to the first quarter of 1971. After adding one more quarter to the estimation period, the procedure is repeated. That is, data up to the second quarter of 1970 is used to make a projection for the second quarter of 1971. In this way, the procedure mimics what a statistical model would have predicted with the information available at any point in the past. Data that became available subsequent to the prediction date are used neither to estimate nor to predict.

This type of procedure leads to a fairer and more realistic test of the predictive abilities of the various models than the in-sample results. It nevertheless has several drawbacks. First, instead of one regression for the whole sample, as in the in-sample case, regressions must be run for each observation following the starting point. Second, the pseudo R^2, which is easily interpretable in sample, is not longer guaranteed to lie between 0 and 1. This is not a consequence of the probit form, but is also true of predictions based on linear regressions. Third, statistical tests of significance are no longer available in a strict sense.

We deal with these issues in the following ways. First, we let the computer crunch away, dealing with the problem of having to do a multitude of regressions. Second, we present only non-negative pseudo R^2s in the results. For regressions in which this measure

of fit is significant in sample, the out-of-sample analogue works also reasonably well; negative values are generally meaningless.

Finally, in place of the 5 and 1 percent significance tests, we use the following informal tests of improved predictive power. We first ask whether the pseudo R^2 increases with the addition of the last variable. Although this seems like an easy test to pass, it is actually not so easy to achieve with out-of-sample predictions, and we use it as an analogue to the in-sample 5 percent test. We also ask whether the pseudo R^2 increases to the extent that would be required for 1 percent significance in an in-sample test. This last criterion is used as an rough indicator of large improvements in fit, since the test with out-of-sample results is not strictly based on statistical theory.

The first data point for which predictions are made is the first quarter of 1971. Although an earlier date would have been possible based solely on sample size considerations, it is important to capture some recession observations in order to arrive at accurate parameter estimates. The 1960s were essentially an uninterrupted economic expansion, so the sample must extend to the early 1970s. Predictions are computed for the period up to the first quarter of 1995. The results are presented in tables 5 and 6 (and in tables A5 and A6 in the appendix).

Table 5 includes results for each of the variables in the data set by itself. The table in general exhibits patters similar to those described in the previous section, although a few of the results may be somewhat surprising. Variables that perform well, confirming expectations, are the yield curve spread, the Treasury bill, the real monetary base, stock prices, and the indices of leading indicators. Compared with the in-sample results, there is

Table 5

Measures of Fit for Probit Models
Variables by themselves -- OUT OF sample

$$P(R_{t+k}=1) = F(\alpha_0 + \alpha_1 X_{1t})$$

k = Quarters Ahead

X_{1t} Variable	1	2	3	4	5	6	7	8
SPREAD	0.072*	0.236**	0.328**	0.295**	0.155**	0.141**	--	--
CPTB	--	--	--	--	--	--	--	--
BILL	0.078**	0.101**	0.070*	--	0.004*	0.066*	0.018*	--
BOND	--	--	--	--	--	--	--	--
M0	--	--	--	--	--	--	--	--
M1	0.040*	--	--	--	--	--	--	--
M2	--	--	--	--	--	--	--	--
M3	--	--	--	--	--	--	0.006*	0.038*
RM0	0.157**	0.073*	--	0.176**	0.101**	0.097**	--	--
RM1	0.169**	0.048*	--	--	--	0.042*	--	--
RM2	0.129**	0.061*	--	--	--	--	--	--
RM3	0.093**	--	--	--	--	--	--	--
NYSE	0.161**	0.077**	0.075*	0.016*	--	--	--	0.028*
SP500	0.159**	0.073*	0.068*	0.018*	--	--	--	0.027*
DJIA	0.137**	0.036*	0.024*	0.009*	--	--	--	0.001*
NAPMC	0.195**	0.046*	0.005*	--	--	--	--	--
VP	0.095**	0.007*	0.015*	--	0.007*	0.002*	0.001*	--
CORD	0.078**	--	--	--	--	--	--	--
HI	0.105**	0.098**	0.205**	0.047*	--	--	--	--
CEXP	--	--	--	--	--	--	--	--
TWD	--	--	--	--	--	--	--	--

Table 5 (continued)

MORD	--	--	--	--	--	--	--	--
LEAD	0.121**	--	--	--	--	--	--	--
XLI	0.324**	0.141**	--	0.015*	0.067*	0.016*	--	--
XLI2	0.196**	0.028*	--	--	--	--	--	--
GDPG1	0.234**	0.060*	--	--	--	--	--	--

For each model, the pseudo R^2 is shown, followed by informal indicators of significance. "--" indicates a negative value.

*Additional variable improves fit.
**Improvement in fit corresponds to 1 percent significance.

some deterioration in the performance of these variables, both in terms of accuracy and length of the predictive horizon. Nevertheless, the same basic patterns emerge for most of these predictors as in the in-sample case.

For a few variables, the deterioration in performance is substantial. For example, the commercial paper spread (CPTB), which was highly significant for 1 and 2 quarters in sample, has a negative pseudo R^2 for every predictive horizon out of sample. The Treasury bond rate follows a similar pattern, although it is clear that it enhances the predictive power of the bill rate as they combine in the yield curve spread (SPREAD). The Commerce Department leading indicators also have significantly diminished predictive power as compared with the in-sample results. The original Stock-Watson index outperforms the other leading indicators, particularly 1 quarter ahead.

As in the in-sample results, the yield curve spread tends to dominate the results starting with the 2-quarter ahead predictions. Although predictive power at 7 and 8 quarters is absent, the results for 2 and 3 quarters are actually stronger than in sample. No other single variable exhibits this kind of performance, including the traditional macro indicators. Thus, we proceed to add the yield curve spread to the other variables, as we did in sample.

When the yield curve spread is added to each of the variables in table 5, the effects are quite dramatic, as seen in table 6. One important feature of the table is that, with very few exceptions, additional predictive power is absent beyond 1 quarter when variables are combined with the yield curve spread. Of course, the variables that do not perform well by themselves remain poor predictors. What is noteworthy, however, is that some variables that

Table 6

Measures of Fit for Probit Models
Variables with spread -- OUT OF sample

$$P(R_{t+k}=1) = F(\alpha_0+\alpha_1 X_{1t}+\alpha_2 SPREAD_t)$$

k = Quarters Ahead

X_{1t} Variable	1	2	3	4	5	6	7	8
SPREAD	0.072	0.236	0.328	0.295	0.155	0.141	--	--
CPTB	--	--	--	0.153	0.105	0.140	--	--
BILL	0.046	0.101	0.046	0.145	0.095	0.064	--	--
BOND	0.046	0.101	0.046	0.145	0.095	0.064	--	--
M0	0.059	0.223	0.230	0.157	--	0.118	--	--
M1	0.078*	0.211	0.249	0.230	0.110	--	--	--
M2	--	--	0.000	0.207	0.114	0.127	0.002*	--
M3	0.018	--	--	--	0.081	0.141	--	--
RM0	0.127*	0.176	--	0.171	--	0.114	--	--
RM1	0.106*	0.199	--	0.201	0.128	0.089	--	--
RM2	0.010	0.131	--	0.225	0.148	0.130	--	--
RM3	0.083*	0.031	--	0.181	0.161*	0.134	--	--
NYSE	0.208**	0.316**	0.367*	0.274	0.161*	0.120	--	--
SP500	0.205**	0.314**	0.359*	0.277	0.161*	0.133	--	--
DJIA	0.172**	0.248*	0.318	0.292	0.153	0.079	--	--
NAPMC	0.205**	0.222	0.265	0.233	--	0.090	--	--
VP	0.128*	0.212	0.306	0.256	0.190*	0.193*	--	--
CORD	0.127*	0.224	0.322	0.279	0.170*	0.148*	--	--
HI	0.114*	0.237*	0.400*	0.254	0.126	0.137	--	--
CEXP	--	0.034	--	0.244	--	--	--	--
TWD	--	0.005	--	0.073**	--	--	--	--

Table 6 (continued)

MORD	--	--	0.030	0.115	--	0.016	--	--
LEAD	0.079*	--	0.149	0.254	0.121	0.081	--	--
XLI	--	0.136	0.015	0.192	--	--	--	--
XLI2	0.252**	0.270*	0.311	0.139	--	--	--	--
GDPG1	0.285**	0.265*	0.059	0.241	0.112	0.011	--	--

"--" indcates a nagative value.
*Additional variable improves fit.
**Improvement in fit corresponds to 1 percent significance.

do extremely well by themselves, such as the real monetary base and the original Stock-Watson index, are almost completely overshadowed by the spread.

As noted earlier, the Stock-Watson index is partly based on the spread, indicating that out-of-sample there is little additional information in that measure. It is more difficult to find a direct link to the reduced significance of the real base, although the empirical results are almost equally striking. More generally, the lesson from this table is that parsimonious models work best out-of-sample. A combination based on two variables, even variables that are good individual predictors, tends to do worse than each variable on its own.[7]

It is clear from table 6 that the only variables that truly and consistently enhance the out-of-sample predictive power of the yield curve beyond 1 quarter are the stock price indices. With horizons of 1, 2, 3 and 5 quarters, the results are better with either of the broader market indices, namely NYSE and SP500. Even for 4 and 6 quarters, the reduction in predictive fit is not that large.

We may draw a couple of additional conclusions. First, stock prices provide information that is not contained in the yield curve spread and which is useful in predicting future recessions. Second, a simple model containing these two variables is about the best that can be constructed from financial variables for out-of-sample prediction. Again, it generally pays to be parsimonious. For example, adding GDP growth to the yield curve spread and the NYSE index increases the fit of the 1 quarter prediction dramatically to 0.433, as compared with 0.285 without GDP. However, for every other horizon, the results are much worse in the 3 variable case.

17

6. Case study: an application of the approach

Predicting the future is a tricky business. An good example of what may happen is provided by the experience with the Stock-Watson (1989) experimental index of leading indicators. In a very useful piece of post mortem analysis, Watson (1991) and Stock and Watson (1992) describe and analyze the disappointing performance of their indicator in predicting the 1990-91 recession.

We have shown in this article how out-of-sample performance may deteriorate significantly with the use of too many explanatory variables. In different ways, the leading indicators of both the Commerce Department and Stock-Watson (1989) are susceptible to this type of overfitting problem. The Commerce Department measure is based on movements in 11 individual variables, which are combined in a weighted average. The Stock-Watson (1989) indicator uses a fairly complex modeling specification that includes 7 individual series, with several lags for each of the series.

In this section, we examine the performance of two parsimonious models -- SPREAD only and SPREAD with NYSE -- in forecasting the 1990-91 recession out-of-sample, and compare the results to the Commerce and Stock-Watson leading indicators. For each series, we use a forecasting horizon that roughly maximizes its performance; for the leading indicator series, we use 2 quarters, whereas we use 4 quarters for the simple models.

Before turning to the 1990-91 results, consider the earlier performance of the series. All four variables had been fairly reliable until the late 1980s. Chart 1, for example, shows the recession probabilities implied by the Commerce (LEAD) and Stock-Watson (XLI) indicators from 1971 to 1987. Both series produce larger signals approximately consistent

18

with the actual recessions, but the Stock-Watson measure is superior in timing and accuracy. The representation of their results is somewhat different from that in their paper. However, comparison of chart 1 with figure 4 of Watson (1991) reveals very similar patterns. The indications of the Commerce variable come too late, are more volatile, and are too high in early 1985.

Chart 2 shows the corresponding probabilities using the yield curve spread (SPREAD) and the combined spread and stock index (NYSE) models. In general, the results are fairly accurate, even though the signal in 1973-74 comes a bit late. The two models produce similar results, and both are very accurate in non-recessionary periods.

In the 1990-91 recession, the predictive power of the two leading indicator series broke down, as shown in chart 1. Stock and Watson have documented how their indicator surged too early, declined, and gave a feeble signal within the recession. Our chart shows pretty much the same pattern. The Commerce indicator again was worse. It gave a very strong signal after the recession, and two somewhat strong signals before it, but essentially missed the recessionary quarters.

In this recession, the simple models outperformed both leading indicators. Chart 2 shows that the spread by itself was quite informative. It surged a bit prematurely, but less so than the Stock-Watson measure, and the signal was weaker than in some earlier recessions.[8] Nevertheless, it provided a clear signal that continued to rise into the onset of the recession. The addition of the NYSE index improves the results somewhat in that the probabilities have a similar peak in the recession, but are not as strong before the recession starts.

The lessons of these out-of-sample forecasting exercises, particularly in 1990-91, suggest that the simple financial variable models compare favorably with the more complex leading indicators. The results illustrate the dangers of overfitting and the potential benefits of using simple financial variables as indicators. The results are all the more impressive in that forecasting horizon for the financial variables -- 4 quarters -- is twice as long as that of the leading indicators, for which even 2 quarters may be a bit of a stretch.

Notes to charts 1 and 2

- The probabilities in these charts are based on out-of-sample forecasts either 2 or 4 quarters ahead. The probability shown is a forecast for the contemporaneous quarter, based on data from either 2 or 4 quarters earlier.

- The predictive horizon is chosen to suit the particular variable. For the Commerce and Stock-Watson leading indicators, we use 2 quarters, whereas for the yield curve spread (with and without the stock index), we use 4 quarters.

- The model labeled "SPREAD+NYSE" includes the yield curve spread and the stock market index as separate explanatory variables.

Chart 1

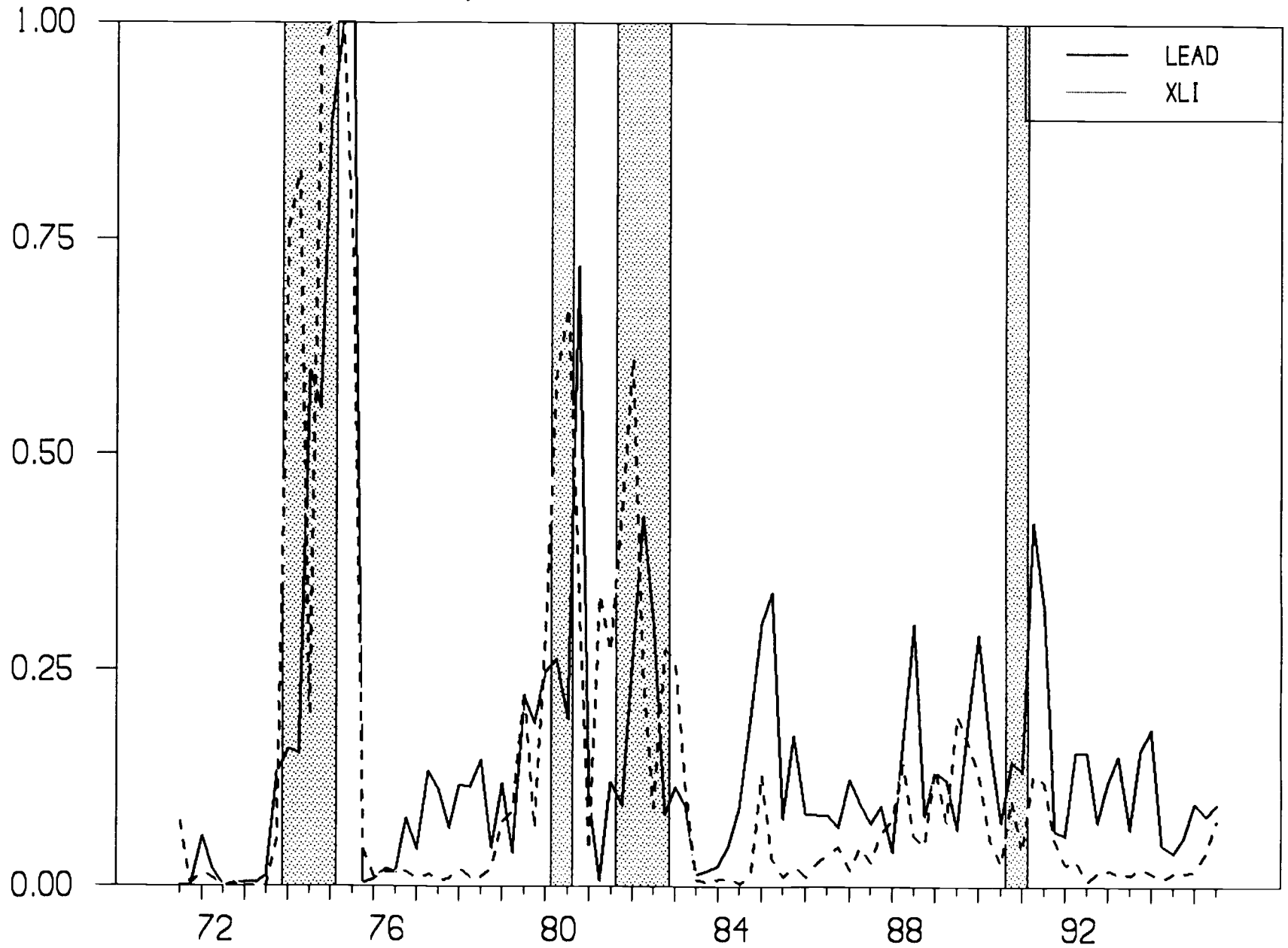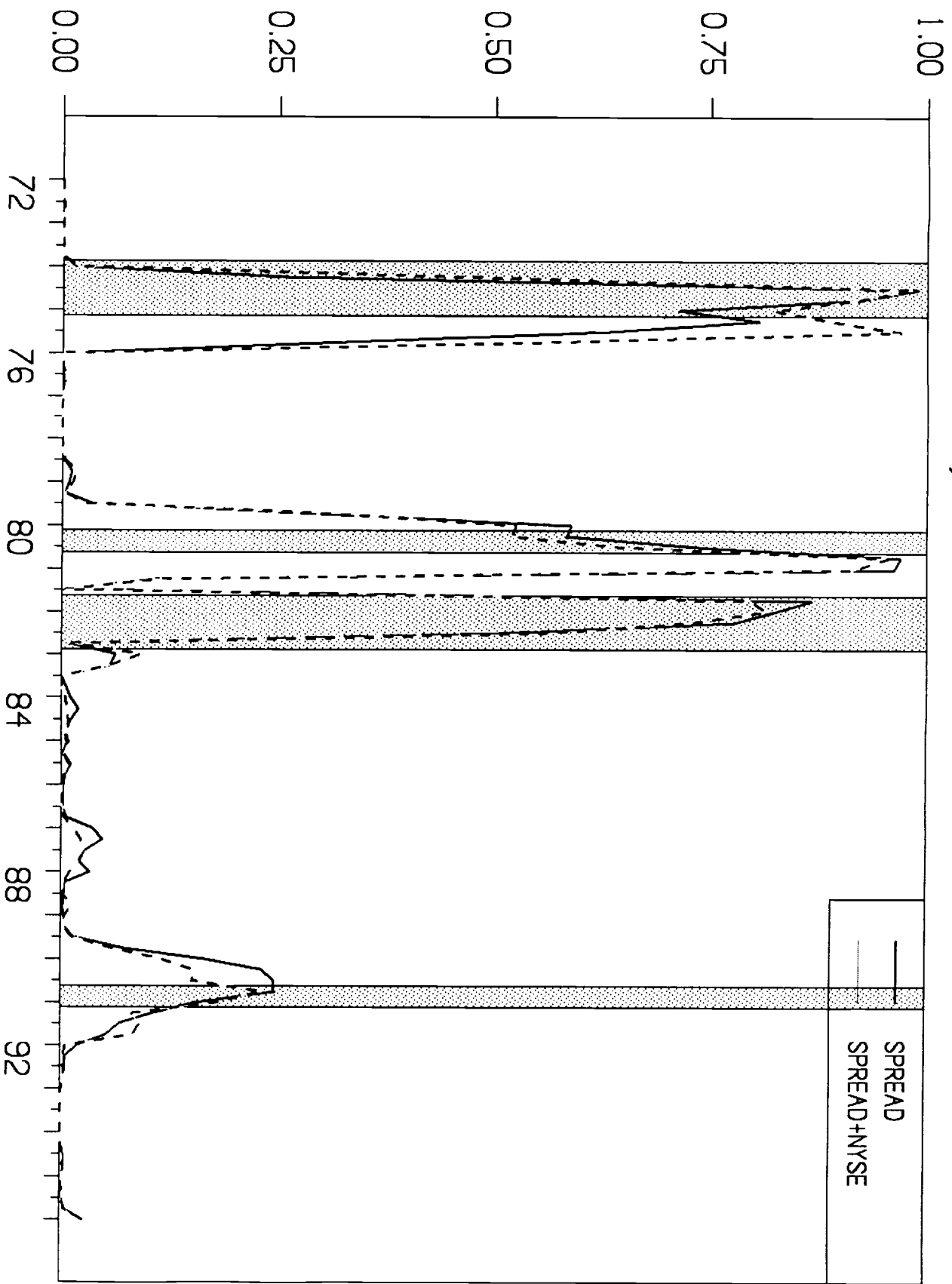## Probability of Recession, 2 Quarters Ahead

Legend:
- LEAD
- XLI

Chart 2

Probability of Recession, 4 Quarters Ahead

Legend:
— SPREAD
---- SPREAD+NYSE

7. Conclusions

This article has examined the performance of various financial variables in predicting future U.S. recessions, particularly out-of-sample. The results obtained using the yield curve spread and stock prices are encouraging and suggest that these measures can have a useful role in macroeconomic prediction. As stated earlier, we do not propose that these indicators supplant macroeconomic models and judgmental forecasts. Instead, we conclude that the financial variables can usefully supplement the models and other forecasts, and can be used as a quick reliable check of more elaborate predictions.

In the process of examining the predictive power of the variables, several general principles were clearly illustrated. First, overfitting is a serious problem in macroeconomic predictions. Even when only a few variables are used, the addition of a single variable or another lag of a variable can undermine the predictive power of a parsimonious model. Second, in-sample and out-of-sample performance can differ greatly. A good illustration is the 6-month commercial paper-Treasury bill spread, which does very well in sample for 1 and 2 quarters, but has no out-of-sample predictive power at any horizon.

A third principle is that it is important to determine the optimal out-of-sample horizon for each financial variable. For instance, the yield curve spread shows the best predictive performance across the range of horizons examined. For a 1 quarter horizon, however, even though it has some power, it is substantially outperformed by a number of other indicators, including the stock price indices, the Commerce and Stock-Watson leading indicators, and some of the Commerce indicator's components. Other than the yield curve, the indicators

we have studied tend to perform best with short horizons, although in some cases (for example, stock prices) the performance extends to 2 or 3 quarters.

As to specific conclusions, the yield curve spread and stock price indexes emerge as the most useful simple financial indicators. The may be observed individually over their respective primary horizons, or they may be combined to produce a very reliable simple model that would have provided clear indications of the last recession with a 4 quarter horizon.

8. Appendix: complete in-sample results and supplementary out-of-sample results

Some key in-sample results were presented in section 4. In this appendix, we include in-sample results for all the variables listed in table 1. Results for single variables are given in table A1, single variables with the yield curve spread in table A2, and variables with two lags, with and without the spread, in tables A3 and A4.

Out-of-sample results for models with two lags of the explanatory variables are reported in tables A5 and A6.

Table A1

Measures of Fit and t-Statistics for Probit Models
Variables by themselves -- IN sample

$$P(R_{t+k}=1) = F(\alpha_0 + \alpha_1 X_{1t})$$

k = Quarters Ahead

X_{1t} Variables		1	2	3	4	5	6	7	8
SPREAD	Pseudo R^2	0.071	0.211	0.271	0.296	0.256	0.149	0.078	0.031
	t-stat	-3.01**	-4.56**	-4.89**	-4.99**	-4.78**	-4.03**	-3.09**	-2.02*
CPTB	Pseudo R^2	0.103	0.061	0.026	0.001	0.001	0.001	0.008	0.01
	t-stat	3.41**	2.74**	1.91	0.44	-0.41	-0.32	-1.01	-1.08
BILL	Pseudo R^2	0.133	0.193	0.177	0.151	0.113	0.064	0.036	0.015
	t-stat	4.05**	4.62**	4.44**	4.16**	3.69**	2.9**	2.22*	1.42
BOND	Pseudo R^2	0.077	0.077	0.054	0.036	0.022	0.012	0.007	0.003
	t-stat	3.12**	3.11**	2.65**	2.19*	1.74	1.27	1.	0.62
M0	Pseudo R^2	0	0.001	0.001	0.002	0.006	0	0.001	0.024
	t-stat	-0.05	0.36	-0.33	-0.57	-0.87	0.03	0.34	1.74
M1	Pseudo R^2	0.052	0.021	0.03	0.004	0	0.001	0	0.005
	t-stat	-2.46*	-1.66	-1.95	-0.72	-0.16	-0.29	0	0.83
M2	Pseudo R^2	0.022	0.02	0.024	0.002	0.001	0.005	0.004	0.011
	t-stat	-1.67	-1.63	-1.76	-0.5	0.39	0.82	0.71	1.22
M3	Pseudo R^2	0.001	0.002	0.003	0	0.004	0.03	0.031	0.039
	t-stat	0.36	-0.48	-0.63	0.01	0.76	1.93	1.98*	2.16*
RM0	Pseudo R^2	0.153	0.103	0.156	0.168	0.118	0.072	0.046	0.014
	t-stat	-4.1**	-3.53**	-4.19**	-4.27**	-3.71**	-2.98**	-2.4*	-1.35
RM1	Pseudo R^2	0.209	0.12	0.154	0.092	0.041	0.038	0.023	0.009
	t-stat	-4.39**	-3.77**	-4.19**	-3.34**	-2.29*	-2.16*	-1.72	-1.07
RM2	Pseudo R^2	0.172	0.136	0.171	0.103	0.037	0.022	0.017	0.008
	t-stat	-4.17**	-3.9**	-4.21**	-3.45**	-2.17*	-1.69	-1.48	-1.06
RM3	Pseudo R^2	0.105	0.09	0.117	0.082	0.028	0.005	0.002	0.001
	t-stat	-3.53**	-3.33**	-3.69**	-3.19**	-1.93	-0.8	-0.49	-0.29
NYSE	Pseudo R^2	0.174	0.133	0.08	0.043	0.003	0	0.004	0.03
	t-stat	-4.35**	-3.91**	-3.19**	-2.4*	0.68	0.07	0.76	1.93

SP500	Pseudo R²	0.169	0.134	0.079	0.043	0.003	0.001	0.007	0.031
	t-stat	-4.31**	-3.93**	-3.17**	-2.38*	-0.59	0.29	0.97	1.97*
DJIA	Pseudo R²	0.131	0.102	0.065	0.05	0.003	0	0.003	0.013
	t-stat	-3.92**	-3.54**	-2.89**	-2.55*	-0.65	0.2	0.68	1.32
NAPMC	Pseudo R²	0.151	0.04	0.049	0.025	0.006	0.001	0	0.004
	t-stat	-3.82**	-2.27*	-2.51*	-1.8	0.87	-0.3	0.12	0.71
VP	Pseudo R²	0.074	0.013	0.014	0.006	0.016	0.012	0.006	0.008
	t-stat	-2.95**	-1.33	-1.41	-0.9	1.43	1.26	0.9	1.02
CORD	Pseudo R²	0.084	0.027	0.001	0.001	0.002	0.002	0.003	0
	t-stat	-3.16**	-1.89	-0.27	-0.45	0.53	0.52	0.62	-0.03
HI	Pseudo R²	0.086	0.085	0.171	0.056	0.014	0.003	0	0
	t-stat	-3.23**	-3.22**	-4.31**	-2.65**	-1.39	-0.62	-0.25	-0.07
CEXP	Pseudo R²	0.03	0.047	0.024	0.039	0.001	0	0.005	0
	t-stat	-1.94	-2.4*	-1.73	-2.19*	-0.42	-0.06	-0.8	0
TWD	Pseudo R²	0.007	0.015	0.006	0.005	0.011	0.003	0	0.003
	t-stat	0.91	1.29	0.8	0.7	1.08	0.59	0.19	-0.55
MORD	Pseudo R²	0.016	0	0.014	0.044	0.045	0.029	0.016	0.001
	t-stat	-1.46	0.08	1.36	2.37*	2.39*	1.95	1.44	0.37
LEAD	Pseudo R²	0.236	0.132	0.112	0.018	0.005	0	0.006	0.007
	t-stat	-4.86**	-4.06**	-3.77**	-1.58	-0.8	-0.1	0.88	0.94
XLI	Pseudo R²	0.387	0.332	0.205	0.103	0.056	0.022	0.006	0.001
	t-stat	-5.2**	-5.57**	-4.99**	-3.72**	-2.78**	1.75	-0.91	-0.39
XLI2	Pseudo R²	0.239	0.091	0.059	0.002	0.008	0.011	0.012	0.017
	t-stat	-4.73**	-3.38**	-2.76**	-0.51	1	1.19	1.24	1.49
GDPG1	Pseudo R²	0.261	0.16	0.093	0.008	0.002	0.008	0.007	0.015
	t-stat	-4.93**	-4.29**	-3.4**	-1.02	-0.5	-1.02	0.98	1.42

For each model, the first row shows the pseudo R^2 and the second row contains the t-statistic for that variable. Significance is determined using a likelihood ratio test.

*Significant at the 5 percent level.
**Significant at the 1 percent level.

Table A2

Measures of Fit and t-Statistics for Probit Models
Variables with spread -- IN sample

$$P(R_{t+k}=1) = F(\alpha_0 + \alpha_1 X_{1t} + \alpha_2 SPREAD_t)$$

k = Quarters Ahead

X_{1t} Variables		1	2	3	4	5	6	7	8
CPTB	Pseudo R^2	0.142	0.233	0.272	0.307	0.285	0.165	0.102	0.051
	t-stat	-2.84**	1.63	0.47	-1.09	-1.75	-1.36	-1.67	-1.54
	t-stat sp	-2.23*	-4.19**	-4.73**	-5.02**	-4.91**	-4.17**	-3.35**	-2.33*
BILL	Pseudo R^2	0.145	0.271	0.305	0.314	0.263	0.153	0.081	0.033
	t-stat	3.02**	2.68**	2.05*	1.46	0.9	0.72	0.67	0.47
	t-stat sp	-1.25	-2.9**	-3.51**	-3.79**	-3.64**	-3.09**	-2.33**	-1.53
BOND	Pseudo R^2	0.145	0.271	0.305	0.314	0.263	0.153	0.081	0.033
	t-stat	3.02**	2.68**	2.05*	1.46	0.9	0.72	0.67	0.47
	t-stat sp	-2.92**	-4.4**	-4.71**	-4.79**	-4.57**	-3.89**	-3.**	-1.98*
M0	Pseudo R^2	0.072	0.216	0.271	0.298	0.257	0.152	0.083	0.066
	t-stat	0.37	0.84	-0.21	-0.48	-0.34	0.6	0.78	2.02*
	t-stat sp	-3.05**	-4.65**	-4.77**	-4.81**	-4.73**	-4.1**	-3.19**	-2.32*
M1	Pseudo R^2	0.105	0.215	0.282	0.298	0.272	0.155	0.083	0.047
	t-stat	-.2*	-0.69	-1.2	0.43	1.38	0.86	0.85	1.41
	t-stat sp	-2.61**	-4.43**	-4.67**	-5.05**	-4.94**	-4.13**	-3.21**	-2.34*
M2	Pseudo R^2	0.093	0.231	0.295	0.297	0.281	0.177	0.093	0.055
	t-stat	-1.66	-1.55	-1.72	0.3	1.74	1.87	1.4	1.69
	t-stat sp	-3.01**	-4.55**	-4.82**	-4.93**	-4.84**	-4.15**	-3.23**	-2.33*
M3	Pseudo R^2	0.076	0.223	0.29	0.299	0.259	0.18	0.11	0.074
	t-stat	-0.8	-1.25	-1.55	-0.6	0.57	1.91	1.97*	2.2
	t-stat sp	-3.1**	-4.65**	-4.96**	-5.**	-4.78**	-4.02**	-3.08**	-2.11*
RM0	Pseudo R^2	0.154	0.213	0.283	0.309	0.258	0.151	0.081	0.033
	t-stat	-3.13**	-0.49	-1.22	-1.23	-0.38	-0.44	-0.61	-0.14
	t-stat sp	-0.48	-3.44**	-3.6**	-3.72**	-3.67**	-3.**	-2.12*	-1.57
RM1	Pseudo R^2	0.21	0.226	0.294	0.297	0.27	0.15	0.078	0.033
	t-stat	-3.67**	-1.39	-1.74	-0.18	1.29	0.15	-0.07	-0.01
	t-stat sp	-0.39	-3.4**	-3.76**	-4.26**	-4.26**	-3.5**	-2.61**	-1.76
RM2	Pseudo R^2	0.183	0.242	0.304	0.298	0.271	0.153	0.078	0.033
	t-stat	-3.48**	-1.96*	-2.01*	-0.4	1.34	0.72	0.15	-0.07
	t-stat sp	-1.16	-3.26*	-3.48**	-3.87**	-3.98**	-3.52**	-2.66**	-1.75

RM3	Pseudo R²	0.141	0.239	0.306	0.306	0.258	0.157	0.083	0.035
	t-stat	-2.93**	-1.88	-2.07*	-1.1	0.5	0.99	0.8	0.55
	t-stat sp	-2.15*	-3.82**	-4.06**	-4.19**	-4.2**	-3.79**	-3.02**	-2.06*
NYSE	Pseudo R²	0.223	0.32	0.321	0.314	0.261	0.159	0.096	0.083
	t-stat	-4.08**	-3.44*	-2.48*	-1.49	0.78	1.13	1.52	2.46*
	t-stat sp	-2.45*	-4.18**	-4.58**	-4.77**	-4.77**	-4.11**	-3.3**	-2.56*
SP500	Pseudo R²	0.217	0.319	0.319	0.312	0.264	0.165	0.103	0.085
	t-stat	-4.03**	-3.45**	-2.43*	-1.42	0.96	1.43	1.78	2.52*
	t-stat sp	-2.44*	-4.17**	-4.57**	-4.77**	-4.77**	-4.15**	-3.35**	-2.59**
DJIA	Pseudo R²	0.177	0.282	0.303	0.315	0.266	0.166	0.098	0.063
	t-stat	-3.56**	-2.93**	-2.02*	-1.51	1.1	1.46	1.59	1.95
	t-stat sp	-2.39*	-4.19**	-4.6**	-4.75**	-4.73**	-4.14**	-3.32**	-2.49*
NAPMC	Pseudo R²	0.182	0.217	0.28	0.297	0.343	0.156	0.086	0.046
	t-stat	-3.35**	-0.87	-1.07	-0.19	2.86**	0.94	1.03	1.33
	t-stat sp	-1.98*	-4.23**	-4.58**	-4.83**	-4.92**	-4.06**	-3.21**	-2.32*
VP	Pseudo R²	0.119	0.211	0.271	0.297	0.342	0.194	0.099	0.051
	t-stat	-2.38*	-0.11	-0.09	0.24	2.91**	2.24*	1.59	1.5
	t-stat sp	-2.39*	-4.46**	-4.8**	-4.96**	-4.71**	-4.22**	-3.31**	-2.33*
CORD	Pseudo R²	0.144	0.233	0.271	0.296	0.273	0.156	0.087	0.033
	t-stat	-2.93**	-1.63	0.35	0.1	1.43	0.95	1.06	0.23
	t-stat sp	-2.75**	-4.47**	-4.89**	-4.99**	-4.82**	-4.08**	-3.19**	-2.07*
HI	Pseudo R²	0.113	0.225	0.326	0.297	0.272	0.163	0.086	0.037
	t-stat	-2.29*	-1.32	-2.53*	-0.28	1.36	1.34	1.04	0.78
	t-stat sp	-1.88	-3.82**	-3.85**	-4.56**	-4.6**	-4.05**	-3.19**	-2.18*
CEXP	Pseudo R²	0.088	0.245	0.283	0.322	0.261	0.154	0.078	0.034
	t-stat	-1.45	-2.04*	-1.23	-1.76	0.8	0.81	-0.19	0.49
	t-stat sp	-2.73**	-4.39**	-4.78**	-4.85**	-4.76**	-4.04**	-3.01**	-2.12*
TWD	Pseudo R²	0.111	0.288	0.358	0.374	0.318	0.206	0.121	0.063
	t-stat	0.45	0.2	-0.53	-0.77	0	-0.24	-0.45	-0.97
	t-stat sp	-3.26**	-4.51**	-4.68**	-4.69**	-4.63**	-4.09**	-3.32**	-2.42*
MORD	Pseudo R²	0.143	0.258	0.282	0.296	0.257	0.15	0.079	0.034
	t-stat	-3.02**	-2.42*	-1.23	0.05	0.35	0.38	0.37	-0.35
	t-stat sp	-3.92**	-4.86**	-4.87**	-4.74**	-4.48**	-3.69**	-2.81**	-2.06*
LEAD	Pseudo R²	0.256	0.283	0.331	0.296	0.265	0.16	0.106	0.054
	t-stat	-4.38**	-3.*	-2.72**	-0.08	0.99	1.14	1.83	1.62
	t-stat sp	-1.58	-3.89**	-4.38**	-4.87**	-4.6**	-4.04**	-3.37**	-2.41*
XLI	Pseudo R²	0.43	0.35	0.298	0.297	0.274	0.179	0.106	0.047
	t-stat	-4.92**	-3.87**	-1.89	0.34	1.45	1.63	1.58	1.26
	t-stat sp	2.17*	-1.43	-3.12**	-4.17**	-4.09**	-3.8**	-3.26**	-2.34*

XLI2	Pseudo R^2	0.289	0.268	0.298	0.302	0.356	0.21	0.121	0.07
	t-stat	-4.51**	-2.66*	-1.87	0.86	2.95**	2.48*	2.09*	2.07*
	t-stat sp	-2.42*	-4.14**	-4.57**	-4.98**	-4.56**	-4.25**	-3.44**	-2.5*
GDPG1	Pseudo R^2	0.313	0.358	0.357	0.297	0.258	0.151	0.094	0.058
	t-stat	-4.68**	-3.92**	-3.09**	-0.34	0.48	-0.43	1.46	1.8
	t-stat sp	-2.45*	-4.1**	-4.58**	-4.95**	-4.74**	-3.95**	-3.22**	-2.32*

*Significant at the 5 percent level.
**Significant at the 1 percent level.

Table A3

Measures of Fit and t-Statistics for Probit Models
Variables with two lags -- IN sample

$$P(R_{t+k}=1) \;=\; F(\alpha_0+\alpha_1 X_{1t}+\alpha_2 X_{1t-1})$$

k = Quarters Ahead

X_{1t} Variables		1	2	3	4	5	6	7	8
BILL	Pseudo R^2	0.217	0.194	0.177	0.158	0.152	0.084	0.067	0.047
	t-stat X_{1t}	-1.62	1.49	1.85	2.31*	2.88**	2.42*	2.52*	2.22*
	t-stat X_{1t-1}	2.8**	0.34	-0.12	-0.92	-1.99*	-1.57	-1.95	-1.77
BOND	Pseudo R^2	0.078	0.112	0.089	0.072	0.064	0.023	0.031	0.034
	t-stat X_{1t}	0.37	2.63**	2.54*	2.45*	2.51*	1.43	1.9	2.02*
	t-stat X_{1t-1}	0.33	-2.06*	-2.07*	-2.08*	-2.21*	-1.2	-1.73	-1.86
RM1	Pseudo R^2	0.241	0.203	0.181	0.097	0.057	0.044	0.027	0.011
	t-stat X_{1t}	-3.51**	-2.45*	-3.22**	-2.6**	-1.6	-1.66	-1.56	-0.7
	t-stat X_{1t-1}	-2.14*	-3.31**	-1.9	-0.8	-1.46	-0.94	-0.32	-0.62
NYSE	Pseudo R^2	0.256	0.175	0.1	0.043	0.004	0.004	0.03	0.033
	t-stat X_{1t}	-3.67**	-3.33**	-2.7**	-2.3*	-0.72	-0.12	0.26	1.8
	t-stat X_{1t-1}	-3.12**	-2.34*	-1.66	0.02	0.026	0.77	1.84	0.55
SP500	Pseudo R^2	0.253	0.175	0.099	0.043	0.004	0.007	0.032	0.033
	t-stat X_{1t}	-3.63**	-3.36**	-2.69**	-2.31*	-0.69	0.04	0.47	1.84
	t-stat X_{1t-1}	-3.16**	-2.31*	-1.64	0.12	0.46	0.93	1.82	0.41
DJIA	Pseudo R^2	0.198	0.139	0.094	0.05	0.004	0.003	0.014	0.014
	t-stat X_{1t}	-3.38**	-3.04**	-2.4	-2.48	-0.73	0.03	0.26	1.35
	t-stat X_{1t-1}	-2.87**	-2.19*	-1.95	0.04	0.37	0.65	1.22	-0.24
NAPMC	Pseudo R^2	0.186	0.089	0.069	0.034	0.006	0.001	0.004	0.004
	t-stat X_{1t}	-3.6**	-2.22*	-2.34*	-1.92	0.84	-0.29	0.26	0.72
	t-stat X_{1t-1}	-2.13*	-2.51*	-1.62	1.1	-0.19	0.1	0.72	0.31

*Significant at the 5 percent level.
**Significant at the 1 percent level.

Table A4

Measures of Fit and t-Statistics for Probit Models
Variables with two lags plus spread -- IN sample

$$P(R_{t+k}=1) = F(\alpha_0+\alpha_1 X_{1t}+\alpha_2 X_{1t-1}+\alpha_3 SPREAD_t)$$

k = Quarters Ahead

X_{1t} Variables		1	2	3	4	5	6	7	8
BILL	Pseudo R^2	0.291	0.299	0.335	0.328	0.265	0.154	0.093	0.052
	t-stat X_{1t}	-2.69**	-0.82	-1.07	-0.76	0.78	0.56	1.37	1.56
	t-stat X_{1t-1}	3.39*	1.87	1.92	1.33	-0.53	-0.34	-1.22	-1.38
	t-stat sp	-2.77*	-3.23**	-3.63**	-3.6**	-3.25**	-2.79**	-1.79	-0.82
BOND	Pseudo R^2	0.17	0.271	0.308	0.32	0.263	0.157	0.084	0.048
	t-stat X_{1t}	-1.21	0.43	-0.28	-0.66	-0.04	-0.58	0.68	1.34
	t-stat X_{1t-1}	1.79	0.04	0.63	0.91	0.2	0.72	-0.57	-1.22
	t-stat sp	-3.22**	-3.92**	-4.12**	-4.18**	-4.1**	-3.71**	-2.55*	-1.37
RM1	Pseudo R^2	0.241	0.289	0.306	0.297	0.275	0.151	0.082	0.034
	t-stat X_{1t}	-3.14**	-0.6	-1.19	-0.06	1.41	0.27	-0.17	0.14
	t-stat X_{1t-1}	-2.08*	-2.85**	-1.24	-0.25	-0.75	-0.44	-0.01	-0.41
	t-stat sp	-0.02	-3.03**	-3.58**	-4.23**	-4.18**	-3.43**	-2.59**	-1.71
NYSE	Pseudo R^2	0.296	0.354	0.337	0.316	0.264	0.169	0.137	0.086
	t-stat X_{1t}	-3.4**	-2.75**	-1.79	-1.53	0.58	0.79	0.91	2.31*
	t-stat X_{1t-1}	-2.95**	-2.07*	-1.44	0.4	0.54	1.09	2.12*	0.64
	t-stat sp	-2.19*	-4.05**	-4.49**	-4.8**	-4.77**	-4.14**	-3.43**	-2.56*
SP500	Pseudo R^2	0.292	0.353	0.335	0.315	0.269	0.178	0.144	0.088
	t-stat X_{1t}	-3.35**	-2.75**	-1.71	-1.51	0.68	1.02	1.15	2.38*
	t-stat X_{1t-1}	-2.99**	-2.05*	-1.44	0.52	0.78	1.26	2.12	0.49
	t-stat sp	-2.18*	-4.05**	-4.48**	-4.8**	-4.77**	-4.17**	-3.47**	-2.59**
DJIA	Pseudo R^2	0.236	0.313	0.332	0.317	0.272	0.176	0.12	0.064
	t-stat X_{1t}	-3.**	-2.24*	-1.26	-1.62	0.85	1.15	1.06	1.96*
	t-stat X_{1t-1}	-2.72**	-1.95	-1.88	0.54	0.86	1.08	1.59	-0.15
	t-stat sp	-2.17*	-4.06**	-4.49**	-4.79**	-4.71**	-4.15**	-3.42**	-2.48*
NAPMC	Pseudo R^2	0.209	0.267	0.308	0.309	0.346	0.157	0.098	0.047
	t-stat X_{1t}	-3.15**	-0.68	-0.71	-0.39	2.86**	0.96	1.24	1.35
	t-stat X_{1t-1}	-1.84	-2.38*	-1.85	1.18	0.56	0.22	0.89	0.43
	t-stat sp	-1.7	-4.04**	-4.44**	-4.77**	-4.58**	-4.05**	-3.29**	-2.33*

*Significant at the 5 percent level.
**Significant at the 1 percent level.

Table A5

Measures of Fit and t-Statistics for Probit Models
Variables by themselves -- 2 lags -- OUT OF sample

$$P(R_{t+k}=1) = F(\alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{1t-1})$$

k = Quarters Ahead

X_{1t} Variable	1	2	3	4	5	6	7	8
BILL	0.173**	0.007*	0.027*	--	--	--	--	--
BOND	--	--	--	--	--	--	--	--
RM1	0.173**	--	--	--	--	--	--	--
NYSE	0.220**	0.125**	0.069*	--	--	--	0.001*	0.024*
SP500	0.218**	0.121**	0.067*	0.000	--	--	0.110*	0.022*
DJIA	0.153**	0.062*	0.033*	--	--	--	--	--
NAPMC	0.234**	0.020*	--	--	--	--	--	--

For each model, the pseudo R^2 is shown, followed by informal indicators of significance.

*Additional variable improves fit (two lags jointly).
**Improvement in fit (two lags jointly) corresponds to 1 percent significance.

Table A6

Measures of Fit and t-Statistics for Probit Models
Variables by themselves -- 2 lags -- OUT OF sample

$$P(R_{t+k}=1) = F(\alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{1t-1} + \alpha_3 SPREAD_t)$$

k = Quarters Ahead

X_{1t} Variable	1	2	3	4	5	6	7	8
BILL	0.237**	0.094	0.054	0.097	--	--	--	--
BOND	0.039	0.047	0.037	0.078	--	--	--	--
RM1	--	--	--	0.159	--	--	--	--
NYSE	0.247**	0.347**	0.354*	0.250	--	--	--	--
SP500	0.245**	0.342*	0.350*	0.249	--	0.017	--	--
DJIA	0.169*	0.263*	0.328	0.276	--	--	--	--
NAPMC	0.233**	0.233	0.170	--	--	0.077	--	--

*Additional variable improves fit (two lags jointly).
**Improvement in fit (two lags jointly) corresponds to 1 percent significance.

<u>Box: The Probit Equation</u>

A probit equation is in many ways similar to an ordinary linear regression. The key difference is that the dependent variable only takes on the values 0 and 1. This requires that the right hand side of the equation be transformed so that it takes on values between 0 and 1, which can then be interpreted as probabilities of the dependent variable being 1.

In the text of this article, the dependent variable in the model is

$R_t = 1$ if the economy is in recession in quarter t

$ = 0$ otherwise.

A standard linear regression with one explanatory variable X would be of the form

$$R_{t+k} = \alpha_0 + \alpha_1 X_t + \varepsilon_t \; .$$

The probit equation we use has the form

$$P(R_{t+k}=1) = F(\alpha_0 + \alpha_1 X_t) \; ,$$

where the α coefficients are statistically estimated and F is the normal cumulative distribution function. The equation is estimated using the method of maximum likelihood, which, stated heuristically, maximizes the probability that the model is correct given the functional form and the observed values of the variables. Statistical tests on the coefficients are performed using likelihood ratio tests, which have asymptotic chi-squared distributions. For the technical details of these procedures, including a precise definition of the likelihood function, see Maddala (1983).

The measure of fit used, that is, the pseudo R^2, is calculated from a formula proposed in Estrella (1995). It is similar to measures proposed earlier in the literature, but corresponds more closely to the standard linear R^2 when its values are away from the endpoints 0 and 1. The formula used is

$$pseudo \ R^2 = 1 - \left(\frac{\log L_u}{\log L_c}\right)^{-\frac{2}{n}\log L_c},$$

where L_u is the value of the likelihood of the estimated model and L_c is the value of a model containing only the constant term α_0.

1. Papers that examine the predictability of future real activity include Harvey (1988), Laurent (1988, 1989), Estrella and Hardouvelis (1990, 1991), Chen (1991), Hu (1993), Bomhoff (1994), Davis and Henry (1994), Plosser and Rouwenhorst (1994), Barran et al. (1995), Davis and Fagan (1995), Estrella and Mishkin (1995). Papers that examine the predictability of future inflation include Mishkin (1990a, 1990b, 1991) and Jorion and Mishkin (1991).

2. Stock and Watson (1989, 1992) and Watson (1991) also focus on predicting recessions. Boldin (1994), in an alternative approach, models recessions using a regime-switching formulation.

3. The equations discussed in the text were also run using monthly data for the same period. Qualitatively, the results were the same: variables were ranked in the same order whether the data were monthly or quarterly. The fit, however, as measured by the pseudo R^2, was better with the quarterly data in the vast majority of cases. This pattern held for both in-sample and out-of-sample results, with exceptions generally found only for a few variables with horizons of 1 or 2 quarters.

4. Evidence of the predictive power of this variable has been provided by Stock and Watson (1989) and Friedman and Kuttner (1993), among others.

5. Stock and Watson use the 10-year minus 1-year Treasury rate spread. Other financial variables in their model are the commercial paper minus Treasury bill spread (CPTB in this article), the trade-weighted value of the dollar (TWD), and the 10-year Treasury rate (BOND). The remaining variables are housing permits (HI), manufacturers' unfilled orders for durable goods (MORD), and the number of people working part-time in nonagricultural industries because of slack work (not included here).

6. Since the dependent variable has only two values, it seems plausible to focus on yield curve inversions, that is, on cases where the SPREAD is negative. This variable was also examined, but the results are inferior to those for the SPREAD itself, and are insignificant when the SPREAD is included. We also tested a lagged dependent variable and the time (number of quarters) since the last recession. These variables were significant with maximum horizons of 2 and 1 quarters, respectively. However, this performance is not useful in practice, since the recession dates (and hence the recession variable) are available only with very long lags, possibly a year or more (see Boldin (1994)).

7. This principle also applies to multiple lags of an explanatory variable, as suggested by the results of tables A5 and A6 in the appendix.

8. The signal provided by the yield curve SPREAD in the last recession seems weak, but this weakness should be interpreted in relation to the strong signals in the 1980-1981 recessions. In the early 1980s, interest rate cycles exhibited unusually broad ranges. As

27

compared with the rest of the post-war period, steep yield curves were steeper and downward sloping curves were more negative. As a result, the signals produced by the yield curve per se were more extreme in both directions. Since the probit approach of this paper compresses one side of the interest rate cycle (large positive values of the SPREAD) to probabilities close to zero, the increase in the range of variation looks simply like an increase in the size of the signal in the early 1980s. This explanation may be confirmed by examining probit results that include earlier recessions in the post-war period (see, e.g., Estrella and Hardouvelis (1991)). In principle, these changes in the range of variation in the spread may be modeled econometrically, but the development of such a model is a topic for a different paper.

28

References

Barran, Fernando, Virginie Coudert and Benoit Mojon, 1995, "Interest Rates, Banking Spreads and Credit Supply: The Real Effects", Centre D'Etudes Prospectives et D'Informations Internationales, Working Paper No. 95-01 (March).

Boldin, Michael D., 1994, "Dating Turning Points in the Business Cycle", <u>Journal of Business</u>, 67:1.

Bomhoff, Eduard J., 1994, <u>Financial Forecasting for Business and Economics</u>, Academic Press.

Chen, Nai-Fu, 1991, "Financial Investment Opportunities and the Macroeconomy", <u>Journal of Finance</u>, 46:2 (June).

Davis, E. Philip and Gabriel Fagan, 1995, "Indicator Properties of Financial Spreads in the EU: Evidence from Aggregate Union Data", European Monetary Institute working paper.

Davis, E. Philip and S.G.B. Henry, 1994, "The Use of Financial Spreads as Indicator Variables: Evidence for the United Kingdom and Germany, <u>IMF Staff Papers</u>, 41:3 (September).

Estrella, Arturo, 1995, "Measures of Fit with Dichotomous Dependent Variables: Critical Review and a New Proposal", Federal Reserve Bank of New York Research Paper.

Estrella, Arturo and Gikas Hardouvelis, 1990, "Possible Roles of the Yield Curve in Monetary Analysis", in Intermediate Targets and Indicators for Monetary Policy, Federal Reserve Bank of New York.

Estrella, Arturo and Gikas Hardouvelis, 1991, "The Term Structure as a Predictor of Real Economic Activity", Journal of Finance, 46:2 (June).

Estrella, Arturo and Frederic S. Mishkin, 1995, "The Term Structure of Interest Rates and its Role in Monetary Policy for the European Central Bank", Federal Reserve Bank of New York working paper.

Friedman, Benjamin and Kenneth Kuttner, 1993, "Does the Paper-Bill Spread Predict Real Economic Activity?", in Business Cycles, Indicators, and Forecasting, University of Chicago Press.

Harvey, Campbell, 1988, "The Real Term Structure and Consumption Growth", Journal of Financial Economics, 22.

Hu, Zuliu, 1993, "The Yield Curve and Real Activity", IMF Staff Papers, 40:4 (December).

Jorion, Phillippe and Frederic S. Mishkin, 1991, "A Multi-Country Comparison of Term Structure Forecasts at Long Horizons," Journal of Financial Economics, 29, (January): 59-80.

Laurent, Robert, 1988, "An Interest Rate-Based Indicator of Monetary Policy", Federal Reserve Bank of Chicago Economic Perspectives, 12, January/February.

Laurent, Robert, 1989, "Testing the Spread", Federal Reserve Bank of Chicago Economic Perspectives, 13, July/August.

Maddala, G.S., 1983, Limited-Dependent and Qualitative Variables in Econometrics, Cambridge University Press.

Mishkin, Frederic S., 1990a, "What Does the Term Structure Tell Us About Future Inflation?" Journal of Monetary Economics 25 (January): 77-95.

Mishkin, Frederic S., 1990b, "The Information in the Longer-Maturity Term Structure About Future Inflation," Quarterly Journal of Economics, 55, (August):815-28.

Mishkin, Frederic S., 1991, "A Multi-Country Study of the Information in the Term Structure About Future Inflation," Journal of International Money and Finance, 19, (March): 2-22.

Plosser, Charles I. and K. Geert Rouwenhorst, 1994, "International Term Structures and Real Economic Growth", <u>Journal of Monetary Economics</u>, 33.

Stock, James and Mark Watson, 1989, "New Indexes of Coincident and Leading Indicators", in Blanchard, Olivier and Stanley Fischer, eds. <u>NBER Macroeconomic Annual</u>, 4.

Stock, James and Mark Watson, 1992, "A Procedure for Predicting Recessions with Leading Indicators: Econometric Issues and Recent Performance", Federal Reserve Bank of Chicago Working Paper WP-92-7, April.

Watson, Mark, 1991, "Using Econometric Models to Predict Recessions", <u>Federal Reserve Bank of Chicago Economic Perspectives</u>, 15:6, November/December.