

Federal Reserve Bank of Minneapolis
Research Department Staff Report 267

June 2000

Predicting Turning Points

Dan Chin*

Kenwood Capital Management LLC

John Geweke*

University of Iowa

Preston Miller*

Federal Reserve Bank of Minneapolis
and Congressional Budget Office

ABSTRACT

This paper presents a new method for predicting turning points. The paper formally defines a turning point; develops a probit model for estimating the probability of a turning point; and then examines both the in-sample and out-of-sample forecasting performance of the model. The model performs better than some other methods for predicting turning points.

*While working on this paper, all three authors were affiliated with the Federal Reserve Bank of Minneapolis. The authors thank Doug Hamilton for his helpful suggestions. The views expressed herein are those of the authors and not necessarily those of the Congressional Budget Office, Federal Reserve Bank of Minneapolis, or the Federal Reserve System.

In this paper, we seek to develop a reliable method to predict turning points in observed economic variables. Our method is designed so that it can be used day to day with the information that is available at the time. We illustrate our method with respect to a particular variable: the civilian unemployment rate (UR). Our method applies as well to numerous other economic variables.¹

Many readers may not be familiar with the existing literature on turning points. For them, our topic is apt to raise at least the following questions:

- Exactly what defines a turning point, and is the definition a meaningful concept?
- Of what value is the prediction of turning points?
- Is it necessary to develop new methods for predicting turning points instead of just adapting traditional time-series models?

In order to address these questions, we refer to Figure 1, which plots UR (seasonally adjusted) monthly from January 1948 to March 1999. The pattern of the unemployment rate over time might be informally characterized as swings up and swings down at intermediate frequencies (e.g., 1–10 years) and wiggles around the swings at higher frequencies (e.g., 1–3 months). This pattern also characterizes many other economic variables.

We define a turning-point date as occurring when a swing in one direction ends and a swing in the other direction begins. Later we provide a formal definition of turning points. Dates of turning points are plotted in Figure 1 as the vertical lines. The formally defined turning points seem to pass the eyeball test: the lines occur at dates that look like turning points and do not occur at dates that do not look like turning points. We believe the definition is meaningful if there are regularities among observed economic variables and our defined turning-point dates that allow the dates to be

predicted. We show that our model has success at predicting whether a turning point will occur in some specified, future time interval.

We argue that accurate prediction of turning points is valuable from two different perspectives. First, it is well documented that standard forecasting models, those that predict values of variables over monthly or quarterly frequencies, tend to make relatively large errors at dates close to turning points.² Thus, better forecasts of turning points have the potential to improve forecasts of standard models. Second, accurate prediction of turning points would be of interest to policymakers. For some time now, policymakers have been concerned either that the long recovery and expansion period in the United States will end or that the long period of low and diminishing inflation will give way to a period of rising inflation. Those concerns really relate to directional changes in trajectories of variables, which are the focus of turning-point analysis, and not to the particular values of variables at selected future dates, which are the focus of standard forecasting models.

Our method to predict turning points in UR exploits two important characteristics of the data, both of which are evident in Figure 1. One important characteristic is that the swings are asymmetric.³ Downward swings in UR tend to be relatively long and gradual, while upward swings tend to be relatively short and steep. On average, downward swings last roughly 51 months, and upward swings last roughly 23 months. The other important characteristic is that UR values at turning points appear nonstationary in the sense that both the values of UR at highs and the values of UR at lows tend to vary considerably over time.⁴ The highs have varied from roughly 6 percent to 11 percent, whereas the lows have varied from roughly 4 percent to 6 percent.

Although one could use a traditional time-series model to forecast turning points, we believe our method has several advantages. If we used, for example, a vector autoregression (VAR) as our time-series model, we would generate time paths of economic variables for each of many drawings from

the underlying stochastic process. Each generated path could then be checked to see whether it satisfies the definition of a turning point in a selected variable (e.g., UR) over some given time horizon.⁵ The fraction of generated paths satisfying the definition could then be taken to be the probability of a turning point over the given horizon.

Intuitively, the standard time-series approach to predicting turning points has three shortcomings. First, time-series models do not directly predict turning points—they predict sequences of variable values over the future. An unwritten rule of forecasting is that accuracy is enhanced by forecasting directly what is of interest—in this case, turning points. Second, standard time-series models in practice seldomly allow for asymmetric movements in downswings and upswings. Third, the models do not address the apparent nonstationary variations in values at turning points. Our method attempts to rectify all three shortcomings. We show that our approach to predicting turning points dominates in an important respect a time-series model that we construct for comparison.

There appear to be two direct approaches to predicting turning points. One uses leading indicators and the other uses probit models. We use the second approach.

The first direct approach to predicting turning points is to construct a regression model for the variable of interest. The model includes, as explanatory variables, leading indicators that are hypothesized to be effective in anticipating turning points. The explanatory variables are lagged, relative to the dependent variable, by an amount of time equal to the forecasting horizon. Models differ in both turning-point variables and explanatory variables. For instance, two of the best-known models, which predict recessions as defined by the NBER, are the index of leading indicators maintained by the Conference Board and the experimental leading indicator index of Stock and Watson (1989). In contrast, the Center for International Business Cycle Research (CIBCR) uses separate leading-indicator regression models to predict turning points in several observable economic

variables.⁶ All of these models have in common that they forecast values for the variable of interest and then calculate turning-point probabilities from the distributions of forecasts.⁷

The second direct approach is to construct a probit model, or some other dichotomous-variable model, as we do here. To our knowledge the only other published work of this kind is the model of Estrella and Mishkin (1998) for recessions. They consider only financial variables for their probit model specifications. As Estrella-Mishkin argue, a probit model is the natural one to use for the prediction of turning points.

Our work differs from the models previously used under either direct approach.⁸ Most of them either predict an unobservable "state of the economy" or a recession as defined subjectively after considerable delay by the NBER. We chose to focus on turning points in the unemployment rate because we wanted our method to be useful in real time. That requires that the variable of interest is reported with little delay and is not subject to much revision and that the turning points in the variable can be formally defined—ruling out recessions defined by both GDP and NBER.⁹

Our method, like the CIBCR, predicts probabilities associated with movements in observed economic variables. However, other existing models, whether leading-indicator or probit, do not deal with the asymmetries and nonstationarities we have described. And in contrast to Estrella-Mishkin, we use nonfinancial variables as well as financial variables in our probit specifications. We now describe how we construct our probit model for turning points in the unemployment rate.

I. A Heuristic Description of Our Method

The first step is to formally define turning points. We use analogous definitions for turning points from downward to upward swings and for those from up to down. Suppose the unemployment rate is in a downward swing. We say a turning point occurs in month t if two conditions are met. The first condition requires that the unemployment rate reach a low at month t , and the second

condition requires that it rise by a significant margin (s.m.) by month $t + 12$. In particular, with UR_t denoting the value of the unemployment rate at date t , if the unemployment rate has moved from an upward swing to a downward swing at the previous turning point, then we say a turning point occurs at month t if

1. $UR_{t+i} > (UR_{t-1} + UR_t + UR_{t+1})/3$ for $i = 1, \dots, 12$ and
2. $(UR_{t+11} + UR_{t+12} + UR_{t+13})/3 - (UR_{t-1} + UR_t + UR_{t+1})/3 > \text{s.m.} = 0.4$ percentage points.

(Note that we use three-month centered moving averages in both conditions of our definition in order to filter out high-frequency noise. Also, note that knowledge of the previous turning point predicates our two conditions.)

In Figure 1 the vertical lines depicting turning points are based on our formal definitions. Notice that when the unemployment rate flattens out in some periods, our definitions attach those periods to the previous swing. Thus, if the unemployment rate had been swinging down, it is not enough that it quit moving down; it actually has to begin swinging up before we call it a turning point. These, like all definitions of turning points, are somewhat arbitrary. The arbitrariness introduces an imprecision in turning-point dates that is likely to make the forecasting performance of our model worse.

Our definitions allow us to classify all dates in our sample period as either being unemployment rate turning points or not. For any given horizon of n months, we define our dependent variable UR_t^{TP} by

$$UR_t^{TP} = \begin{cases} 0 & \text{if no turning point occurs in months } t + 1, \dots, t + n \\ 1 & \text{if a turning point occurs in months } t + 1, \dots, t + n \end{cases}$$

In this paper, we report the results only for $n = 12$.¹⁰

Our next step is to estimate a relationship of the form:

$$P[UR_t^{TP} = 1] = \Phi(a\Delta UR_t + b\Delta X_t + c).$$

In this relationship the function $\Phi(z)$ is the probability that a standard normal random variable is less than the number z . For example, if z , the value of the linear combination, is 0.00, the probability is 50 percent; if it is 1.65, the probability is 90 percent; and if it is -1.96 , then the probability is 5 percent.

Our X vector includes two variables:¹¹ the manufacturing capacity utilization rate, CU , and the spread in percentage points between the monthly average of Moody's Aaa corporate bond rate and the monthly average of the 90-day Treasury bill rate, $R-r$. We address the nonstationarity of unemployment rate turning points by defining the Δ operator as the change in a variable since the last unemployment rate turning point.¹² Thus, if t_0 is the date of the last turning point prior to date t , ΔUR_t , ΔCU_t , and $\Delta(R-r)_t$ are the differences in the values of these variables between dates t and t_0 . We address the asymmetry of swings by estimating different values of a , b , and c depending on whether the previous turning point was a high or a low.¹³ Roughly speaking, the coefficients are estimated so that the model comes as close as possible on average to generating a 1 when a turning point does occur in the next 12 months and a 0 when it does not. Given the estimates \hat{a} , \hat{b} , and \hat{c} , the estimated probability that UR will have a turning point over the next 12 months is $\Phi(\hat{a}\Delta UR_t + \hat{b}_1\Delta CU_t + \hat{b}_2\Delta(R-r)_t + \hat{c})$.

Our results indicate that explanatory variables can have very different effects in upswings and downswings. Since our model applies the nonlinear transformation Φ to the explanatory variables, the estimated coefficients do not directly provide information about the magnitudes by which turning-point probabilities change when given explanatory variables change. So, we provide tables with that information for both upswings and downswings and show that comparable entries in the two tables can be quite different.

We judge the success of our probit models using three types of criteria: in-sample fit, out-of-sample performance relative to in-sample performance, and out-of-sample performance compared with that of a standard macroeconomic forecasting model. By all of our criteria, our probit models appear quite successful.

II. Our Method in More Detail

A. Estimating the Model

There are 581 monthly observations available to estimate the model, beginning in October 1949 and extending through February 1998. (We know that a turning point from downswing to upswing did not occur in the months previous to March 1999 because of condition 1 of our definition. Thus, February 1998 is the last date for which we can be confident that a turning point did not occur in the next 12 months.) Of those observations, 422 months correspond to the eight periods of declining unemployment in the sample. Given that explanatory variables move smoothly most of the time (as indicated in Figures 2A and 2B) and that there are relatively few turning points, there is a serious danger of overfitting in a model with four unknown parameters.

To increase the reliability of our estimates in this situation, we use an approach analogous to the one used in the Federal Reserve Bank of Minneapolis's quarterly forecasting model, which also has a large number of unknown parameters relative to the amount of information in the data (see the appendix for an intuitive discussion of the approach). We introduce this assumption by means of 32 artificial observations added to the original sample of 581—16 for the upswing model and 16 for the downswing model. For the upswing model there are 8 pairs of observations corresponding to the 2^3 combinations of high and low values of the explanatory variables (ΔUR , ΔCU , and $\Delta(R-r)$) observed during upswings. In each pair, there is one artificial observation of 0 and one of 1 for the dependent variable UR^{TP} . The artificial observations for the downswing model are set in the same way but use

extreme values of the explanatory variables during downswings. Thus, the artificial observations are centered on a model in which the explanatory variables have no systematic effect on turning-point outcomes. That shrinks the coefficient estimates toward zero and avoids the problem of overfitting.

The estimation method for the four parameters in each model is maximum likelihood (incorporating the 32 artificial observations), a standard procedure for a probit model. The parameter estimates themselves are not directly useful, however, since the effect of a change in a variable on the probability of a turning point depends on the value of the probability of the turning point itself. So in Table 1 we present the estimated effects of typical changes in the explanatory variables on the probability of turning points, beginning from a turning-point probability of 50 percent, together with 80 percent confidence intervals for those changes. (The confidence intervals are based on standard errors adjusted for serial correlation by applying the Newey-West (1987) technique to the first-order conditions of the maximum likelihood estimates.) For example, the first row of this table indicates that during a period of declining unemployment, a change in capacity utilization of 1 percentage point (say, from 85 percent to 86 percent) increases the probability of an upturn in unemployment over the next year by an estimated 1.90 percentage points. The 80 percent confidence interval for this effect is from 0.00 to 3.40 percentage points. (For a 1 percentage point decrease in capacity utilization in a period of declining unemployment, reverse the signs.) Maximum effects occur beginning at 50 percent probability in the probit model, and they are smaller beginning from other probabilities. For example, if the initial probability is 10 percent or 90 percent, all effects need to be multiplied by a factor of about 0.44. Thus, during a period of declining unemployment when the probability of an upturn is 10 percent, the estimated effect of a 1 percentage point increase in capacity utilization is 0.80 percentage points, and the corresponding confidence interval is from 0.00 to 1.50 percentage points.

B. Results of Model Estimation

For the unemployment turning-point models, the changes in capacity utilization (up in unemployment downswings, down in unemployment upswings) and interest rate spread (down in downswings, up in upswings) indicated in the second column of Table 1 are in the same direction as has occurred in every postwar episode. In Figure 3, we divide the Figure 2 graphs into periods of downswings and periods of upswings. For example, we can see in Figure 3A that capacity utilization has always been lower at the ends of periods of rising unemployment than it was in the beginning.

Four explanatory variables are statistically significant by comfortable margins: capacity utilization in expansionary periods, the interest rate spread in both upswings and downswings, and the unemployment rate during contractions. The sixth coefficient—on the unemployment rate itself during expansions—is statistically insignificant.

To understand the substantive implications of Table 1, note from Figure 3A that the median change in capacity utilization is down about 12 percent during UR upswings and up about 7 percent from UR downswings. The median change in the interest rate spread is up about 2 percent during UR upswings and down about 2 percent during downswings. Multiplying those changes by the estimated probability changes reported in Table 1, the effects of both variables on changes in the probability of a turning point over the course of a typical expansion or contraction are about the same: 20 percent to 25 percent during expansions, 15 percent to 20 percent during contractions. However, there is a little more uncertainty in the estimates for capacity utilization than in the estimates for the interest rate spread. The unemployment rate itself, with a typical change of about 3 percent over both expansions and contractions, is at least as important as changes in capacity utilization and the interest rate spread during contractions. During expansions its effect is less important, and even the direction cannot be well estimated.

C. How Well Does the Model Predict?

Over a typical episode of rising or falling unemployment, the probability of a turning point always begins near zero; it usually rises to well above 50 percent, and often close to 100 percent, within a year of the next turning point. Figure 4 indicates the model's actual experience in anticipating specific historical turning points. In the figure, the solid line indicates the turning-point probabilities, based on coefficients estimated using the entire sample from October 1949 through February 1998 and using the actual turning points to define the beginning of each episode of rising or falling unemployment. Of course, not all of this information would have been available in real time: the entire sample was not available to estimate the coefficients, and turning points are only determined with certainty 13 months after they have in fact occurred. In Figure 4, the dashed line indicates the turning-point probabilities based on the information actually available at the time: in each month, coefficient estimates use the sample ending 13 months prior to the most recent month that is a viable candidate for a turning point, or, in other words, when both conditions of our turning-point definition could still be satisfied. Thus, these out-of-sample probabilities (dashed line) differ substantially from the in-sample probabilities (solid line) for 13 months beyond each turning point. More than 13 months beyond, the probabilities are close, and they become closer in later years as the estimation sample underlying the dashed line grows closer to that underlying the solid line.

As can be seen in Figure 4, the prediction record over periods when unemployment is rising—that is, the prediction of the end of a recession—is about the same throughout the sample. The probability rises rapidly from near zero to over 50 percent in every case, sometimes to over 90 percent. The prediction record over periods when unemployment is falling—that is, the prediction of the end of an expansion—is poor at the start of the sample but generally improves over time. For example, the probability of the unemployment upturns of June 1953 and March 1957 was never more

than 45 percent before their occurrence, whereas the upturns since the mid-1970s have all been anticipated, with probabilities exceeding 60 percent during part or all of the year preceding their occurrence. The record for the in-sample and out-of-sample forecasts is similar in that regard.

We compare measures of the in-sample and out-of-sample forecasting performance of the probit model in Table 2. Our error measures are referred to as Type I and Type II. A Type I error occurs if a turning point occurs within the next 12 months but the model's estimated probability of a turning point is less than 50 percent. A Type II error occurs if a turning point does not occur within the next 12 months but the model's estimated probability of a turning point is greater than 50 percent. Thus, a Type I error is a failure to predict a turning point when one occurs, or a "false negative," and a Type II error is a prediction of a turning point when none occurs, or a "false positive."

We believe that Type I errors are the most important for comparing in-sample to out-of-sample performance or for comparing performance across models. We could have just looked at the total number of incorrect predictions, but that is likely to understate the differences among models. The reason is that the vast majority of observations have no turning points within the next 12 months. A naive model, assuming, say, that the probability was always zero of a turning point occurring within the next 12 months, would then be incorrect for a fairly small minority of the time. This same naive model would also make no Type II errors, by definition. So any model that seriously attempts to predict turning points is likely to have a higher Type II error rate than does a naive model that never predicts a turning point. The discriminating test between alternative forecasts is then with respect to Type I errors.¹⁴ In fact, a main objective of turning-point models is to reduce the large Type I errors of standard macroeconomic models, which seldom predict turning points whether they in fact occur or not.

In constructing the error statistics, we excluded observations within 13 months following the last turning point. That, of course, removes a large number of errors from the probit model's out-of-

sample forecasts. Although the practice obviously makes the out-of-sample forecasts look better, we think that it is justified. In real time, we do not know whether a turning point has occurred until 13 months after the date. Thus, in real time, it is useful to know that the probability of a turning point occurring is high even if one has in fact occurred. The interpretation would be that the probability is high of a turning point occurring around the current date.

Judging from a comparison of the appropriate columns of Table 2, it appears that the probit model overall does as well out-of-sample as it does in-sample. Somewhat surprisingly, the Type I errors are even smaller out-of-sample than in-sample. Similarly, the model seems to do about as well in upswings as in downswings with one exception.

The exception to the probit model's good forecasting performance, both in-sample and out-of-sample, is its record of forecasting a turning point when the unemployment rate is swinging up and none in fact occurs—a Type II error. It turns out that all of these Type II errors in periods of upswings occurred from 1979 to 1982, when the FOMC changed its monetary policy procedures. As Miller and Roberds (1991) showed, the change in the FOMC's procedures resulted in changed relationships of interest rates to other economic variables. Thus, since the probit model relies on the existence of a stable relationship between the interest rate spread and real economic activity, it is not surprising that the model's forecasts deteriorated in a period when that relationship broke down.

D. Comparison with Other Models

We first compared out-of-sample turning-point predictions in our model with those from VAR models. We constructed a VAR model analogous to the probit models used in this study, incorporating capacity utilization, the interest rate spread, and the unemployment rate. The definitions of turning points described above were applied to the simulated paths, and probabilities were equated to the share that satisfied the definitions. The model used 13 lags and simple Bayesian priors for the

parameters. In comparing the last two columns of Table 2, it is clear that the VAR does far worse than our probit model in failing to predict turning points when they in fact occur. That is, the Type I error rate for the VAR is almost 10 times that for our probit model. Moreover, the Type I error rate is higher for the VAR whether in periods of unemployment downswings or upswings.

The Table 2 comparison also reveals that the VAR has a somewhat smaller Type II error rate than does the probit model. Some of that superiority with respect to Type II errors just reflects the breakdown of the probit model in the 1979–1982 period. However, most of the superiority of the VAR with respect to Type II errors merely reflects that, like the naive model, it seldom yields a probability of a turning point above 50 percent, as Figure 5 shows.

We next compared predictions from our model with those from leading indicators. Two of the most commonly used indexes are the index of leading indicators maintained by the Conference Board and the experimental leading indicator index of Stock and Watson (1989). A comparison of those indicators in anticipating recessions (as defined by the NBER) is presented in Estrella and Mishkin (1998). For a one-year forecasting horizon, neither of the indicators ever produced more than a 50 percent probability of recession within the one year preceding any of the four NBER recessions that have occurred since 1973. The only time recession probability exceeded 50 percent with either index was late in the 1973:IV to 1975:I recession. The 1990:IV to 1991:I recession, in particular, was entirely unanticipated by either index. For the other recessions, the Stock and Watson index performed somewhat better than the Conference Board index, giving high probabilities of recession before recessions occurred and lower probabilities when they did not occur.

Finally, we compared predictions from our model with those of a probit model of Estrella and Mishkin. Beginning from a much longer list of financial variables, Estrella and Mishkin identify the same interest rate spread used in our probit model and the New York stock exchange index, as the best for a probit model intended to forecast NBER-determined recessions. They report out-of-sample

recession probabilities for 100 quarters, from 1971 through 1995. Using the same 50 percent probability criterion applied in Table 2, their model predicts incorrectly in 6 out of 14 quarters of contractions and in 5 out of 86 quarters of expansions. Those figures are comparable with our probit model. However, their model never assigned a probability of more than 25 percent to the Gulf War recession, whereas our model assigned a probability of more than 50 percent to that turning point.

Estrella and Mishkin also report goodness-of-fit measures for their models, analogous to R^2 statistics for regression. The highest value for any of their models is .33. For the unemployment turning points the goodness-of-fit measure in our model is .29 for the entire sample period and .44 for the 1959–1995 sample period used by Estrella and Mishkin. The improvement relative to Estrella and Mishkin is consistent with the better forecasting track record of our model. The higher values for the 1959–1995 sample period are consistent with the poorer tracking of turning points before 1975 than after in our models, which was discussed earlier.

Conclusion

We have argued that prediction of turning points could help in standard economic forecasting and policymaking. Our work strongly suggests that special-purpose models can improve upon standard economic models for the purpose of predicting turning points. The special-purpose models have the properties that they directly predict the object of interest and account for asymmetries in swings and for nonstationarities in turning-point values. However, we believe our work leaves unanswered two important questions that warrant further research.

The first question is, Can direct prediction of turning points be shown in practice to improve standard economic forecasts? Our analysis suggests that it may be fruitful to estimate two-regime standard economic models, one for upswings and one for downswings. That two-regime standard model would then be combined with a turning-point model that indicates the probability of being in

one regime versus the other, period by period. Until this is done, it is only conjecture to suggest that turning-point prediction can improve standard economic forecasts.

The second question is, Are the processes for turning points of different important variables related? Our analysis suggests that there are regularities in relationships among variables at intermediate (e.g., business cycle) frequencies. It would be worthwhile to explore, for instance, whether inflation and unemployment turning points are related. In particular, one could inquire whether turning points in unemployment and inflation should be modeled as a joint process. That is far different from analyzing correlations at high frequencies between the two variables, as is commonly done in the literature.¹⁵ Until this is done, policymakers' expressed concern that a continued downswing in unemployment will lead to an upswing in inflation cannot be summarily dismissed. That is because their concern is about intermediate-frequency trajectories and not about high-frequency correlations.

Notes

¹Our method applies to all economic variables that move with the business cycle. In an earlier version of this paper, we showed that it even applies to predicting turning points in consumer price inflation.

²See, for example, McNees (1992), pp. 29–30. He states, “History shows a close association between business-cycle turning points and the size of forecasting errors.” (p. 29)

³Asymmetry of cycles has been noted by many economists. See, for example, DeLong and Summers (1986), Hamilton (1989), Neftci (1984), Rothman (1998), and Sichel (1993).

⁴Nonstationarity of turning points is evident in the analysis of Staiger-Stock-Watson (1997) and Pesaran-Potter (1997).

⁵In practice, standard econometric models used this way have focused on predicting recessions, defined operationally as two consecutive quarters of negative real GDP growth. Nevertheless, in principle, they could be used as well to predict turning points, however defined, in other observed economic variables.

⁶We looked at various monthly reports put out by the CIBCR on its indicators.

⁷The variable being forecast in recession models is an unobserved index in a single common factor model.

⁸Our work is closely related to Hamilton’s. His model, like ours, has asymmetric swings down and up that correspond to a business cycle. However, Hamilton sought to identify dates of turning points, while we seek to predict them.

⁹Many analysts have noted the large revisions in GDP that can even reverse the sign of a quarterly change. Meanwhile, the NBER’s proclamation of a recession usually is based somewhat on subjective criteria and may not be issued until well over a year after the event.

¹⁰The shorter the forecast horizon, the lower the probability of a turning point in that

period. By repeating our procedure for horizons ranging from $n = 1$ to $n = 12$, one can compute a cumulative distribution indicating the probabilities of a turning point within the next n months as n increases from 1, 2, ..., 12.

¹¹We considered a few variables to be included in X . Our test for inclusion was that they significantly improve the fit of our probit model (in terms of number of correct hits and likelihood) in real-time simulations. We found that CU and $R-r$ were useful by this test and that other variables, such as time elapsed since the last turning point and the quit rate, did not improve the fit after CU and $R-r$ were included.

¹²This method is similar to the one used by Pesaran-Potter in their floor and ceiling model of output. It does not remove the nonstationarity of turning points altogether, but it does prevent any drift from accumulating. This method has not been used to form probabilities of future turning points, and it has not performed well in predicting periods of positive and negative output growth.

¹³We discuss the estimation method in the next section.

¹⁴In a sense, predicting turning points is similar to predicting tornadoes. A naive model that predicted no tornadoes ever occurring would be right most of the time. However, we would view such a model as a failure because it never gave any warning. In contrast, a naive model that always predicted a tornado occurring would be wrong a majority of the time. We might seek models that have good records at predicting tornadoes that do occur (small Type I errors) and then, from those, choose the ones with lowest Type II errors.

¹⁵See, for example, Staiger-Stock-Watson.

Appendix

The Bayesian Shrinkage Estimator in the Probit Model

This paper uses a Bayesian shrinkage estimator similar to the one applied in vector autoregressions to improve forecasting performance. In our model, as well as in VARs, the effect of shrinkage is to reduce variation at the cost of introducing some bias in the estimator. In both cases, the variance reduction is greater than the increase in bias, and so forecasts are improved.

The problem solved by the shrinkage estimator in our model is different than the one that occurs in a large VAR like the forecasting model at the Federal Reserve Bank of Minneapolis. In the VAR, there are many parameters relative to the number of observations, and the variance of the maximum likelihood estimator is high as a result. In our model there are only eight parameters, while in our estimation sample there are 581 monthly observations. But there are only 15 turning points. Consequently, the dependent variable UR_t^{TP} takes on long runs of “0” when there is no turning point in the next 12 months, each run followed by a run of 12 “1’s.” There are 16 runs of “0” and 15 of “1.” In addition, the explanatory variables are serially correlated. In this situation, the probit model can fit the data very well. In the extreme cases—which do not happen in our sample—the fit can be perfect.

A perfect fit in a probit model is very different from a perfect fit in a linear regression. To appreciate the problem, consider the simple example with hypothetical data presented in Figure A1. As in any probit model there are two possible outcomes: $d = 0$ and $d = 1$. In the hypothetical sample portrayed in Figure A1, whenever the single explanatory variable x is below x_6 , the outcome is $d = 0$, and whenever it is above x_7 , the outcome is $d = 1$. This situation is an example of the perfect classifier problem—so named because the ordering of x perfectly sorts the values of d into 0 and 1. Recall that in the probit model the probability of $d = 1$ is $\Phi(\beta_1 + \beta_2 x)$, and the probability of $d = 0$ is

$1 - \Phi(\beta_1 + \beta_2 x)$. The maximum likelihood estimator picks β_1 and β_2 to maximize the product of the probabilities of all of the observed d 's. For the example shown, a vertical line through any point between x_6 and x_7 —like \hat{x} —will make the probabilities of the observed d 's equal to 1, because $\Phi(-\infty) = 0$ and $\Phi(+\infty) = 1$. A more precise statement is that the probability of the observed outcomes increases continuously as the slope parameter β_2 approaches $+\infty$ and the intercept parameter β_1 approaches $-\infty$, in such a way that $\beta_1 + \beta_2 \hat{x}$ remains between 0 and 1.

When the situation is close to that shown in Figure A1, estimated values of slope coefficients are high in absolute value, but standard errors are also high. The reason is clear from the limiting case shown in Figure A1: the likelihood with $\beta_2 = 100$ and $\beta_1 = 0.5 - 100\hat{x}$ is almost indistinguishable from the likelihood with $\beta_2 = 1,000$ and $\beta_1 = 0.5 - 1,000\hat{x}$. Therefore, confidence intervals are very wide. With several explanatory variables, there are many more ways in which a perfect classifier could occur.

In most situations—certainly including the one studied in this paper—we do not believe there is a perfect classifier. Formally, we could set up a prior distribution to reflect this belief. Less formally but much more easily, we can set up an informal prior distribution, using artificial data, that conveys the belief that the perfect classifier sample is unlikely. In the Figure A1 example, those artificial observations are placed at the points labeled A , B , C , and D . The pair (A, B) corresponds to the smallest observed x and the outcomes $d = 0$ and $d = 1$. The pair (C, D) corresponds to the largest observed x and the outcomes $d = 0$ and $d = 1$. Based on these four artificial data points alone, the fitted probit model would be a horizontal line. As actual data are added, the slopes change in the direction indicated by the data. Eventually the data come to dominate, and the model is estimated consistently. The artificial data points keep the slope from becoming too steep when there are just a few data points.

The model estimated in the paper is more complex than the one shown in Figure A1, of course, because there are more explanatory variables. But the principle is the same: a few artificial observations greatly reduce the variability of the estimates and greatly improve forecasting, both in- and out-of-sample.

References

- CIBCR Indicators, Center for International Business Cycle Research, various monthly reports.
- DeLong, B. and L. Summers (1986): "Are Business Cycles Symmetrical?" in *American Business Cycle: Continuity and Change*, edited by R. Gordon. Chicago: University of Chicago Press, 166–79.
- Estrella, A. and F. Mishkin (1998): "Predicting U.S. Recessions: Financial Variables as Leading Indicators," *The Review of Economics and Statistics*, 80, 45–61.
- Hamilton, J. (1989): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357–84.
- McNees, S. (1992): "How Large Are Economic Forecast Errors," *New England Economic Review*, July/August, 25–42.
- Miller, P. and W. Roberds (1991): "The Quantitative Significance of the Lucas Critique," *Journal of Business and Economic Statistics*, 9, 361–87.
- Neftci, S. (1984): "Are Economic Time Series Asymmetric Over the Business Cycle?" *Journal of Political Economy*, 92, 307–28.
- Newey, W. and K. West (1987): "A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–08.
- Pesaran, M. and S. Potter (1997): "A Floor and Ceiling Model of U.S. Output," *Journal of Economic Dynamics and Control*, 21, 661–95.
- Rothman, P. (1998): "Forecasting Asymmetric Unemployment Rates," *Review of Economics and Statistics*, 80, 164–68.
- Sichel, D. (1993): "Business Cycle Asymmetry: A Deeper Look," *Economic Inquiry*, 31, 224–36.

Staiger, D., J. Stock, and M. Watson (1997): “The NAIRU, Unemployment and Monetary Policy,”

Journal of Economic Perspectives, 11, 33–49.

Stock, J. and M. Watson (1989): “New Indexes of Coincident and Leading Indicators,” in *NBER*

Macroeconomic Annual, edited by O. Blanchard and S. Fischer, 351–94.

Table 1: Model Estimates of Turning-Point Probabilities*

Expansionary periods: Unemployment declining			
In ...	A change of ...	Implies a change in probability of contraction in the next year of ...	80% confidence interval for this change
Capacity utilization	+1.0%	+1.9%	+0.0% to +3.4%
Interest rate spread	-1.0%	+11.9%	+6.5% to +17.0%
Unemployment rate	-1.0%	+0.4%	-5.1% to +5.8%
Time	+1 year	+2.7%	-0.1% to +6.3%
Contractionary periods: Unemployment increasing			
In ...	A change of ...	Implies a change in probability of expansion in the next year of ...	80% confidence interval for this change
Capacity utilization	-1.0%	+1.7%	+0.0% to +3.1%
Interest rate spread	+1.0%	+5.6%	+2.3% to +8.9%
Unemployment rate	+1.0%	+7.9%	+2.8% to +12.8%
Time	+1 year	+1.6%	-6.8% to +9.9%

*When probability of the indicated change in the next year begins at 50%. For other probabilities, the change will be smaller: e.g., when the probability of change begins at 25% or 75%, the probability changes (and the endpoints of the confidence intervals) are smaller by a factor of .80; when the probability of change begins at 10% or 90%, the factor is .44.

Table 2: Comparison of Forecast Errors
(Percentage of incorrect observations)

Forecast Errors	(Number of Observations)	Probit Model		BVAR Model
		In-Sample	Out-of-Sample	
Type I: Turning point occurs within next 12 months and prob < 50%				
All periods	(52)	13.46	5.77	55.77
Downswing to upswing	(24)	29.17	12.50	100.00
Upswing to downswing	(28)	0.00	0.00	17.86
Type II: Turning point does not occur within next 12 months and prob > 50%				
All periods	(157)	8.28	15.29	5.73
Downswing to upswing	(141)	4.96	5.67	1.42
Upswing to downswing	(16)	37.50	100.00	43.75

Figure 1: Unemployment Rate

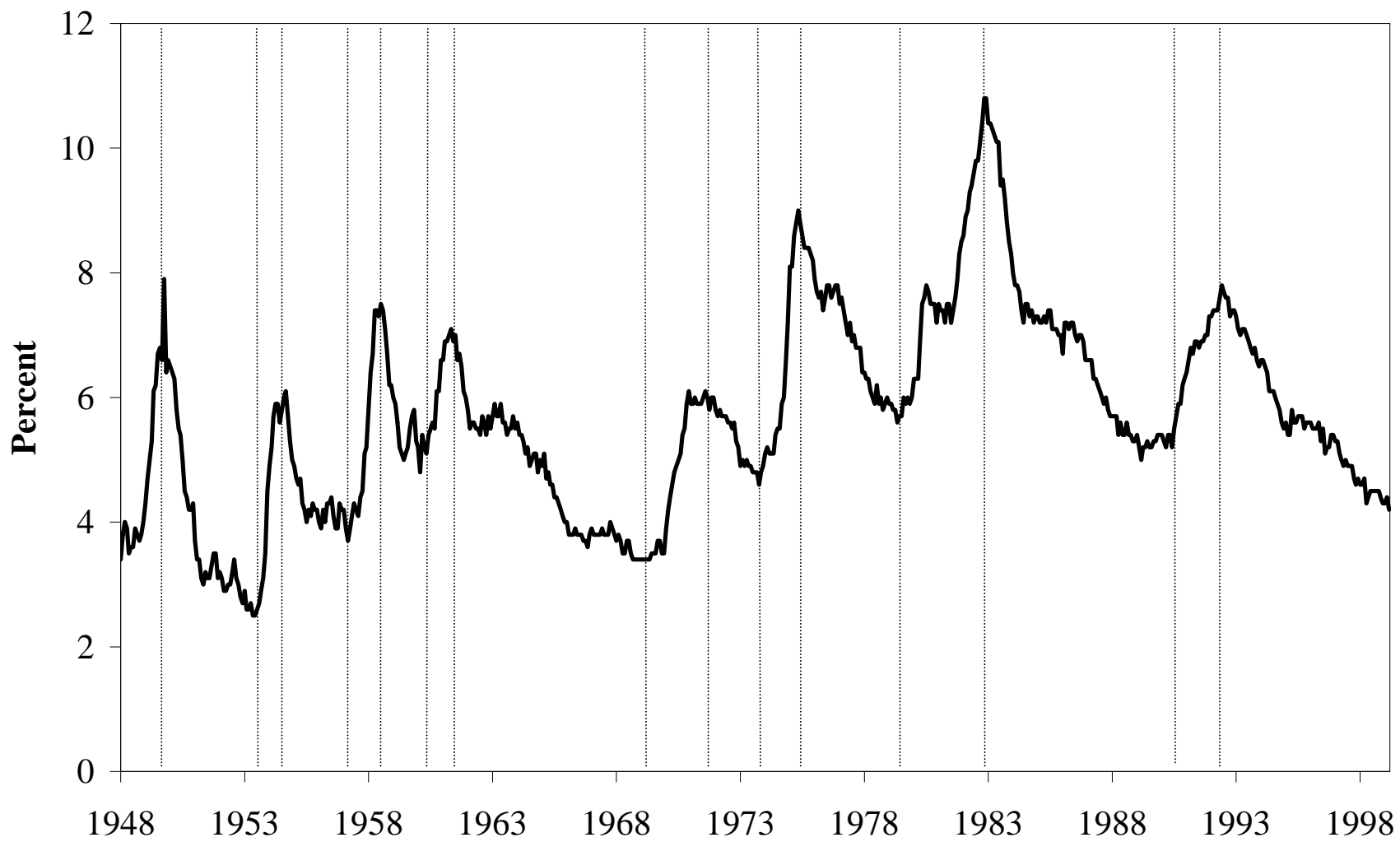


Figure 2A: Capacity Utilization

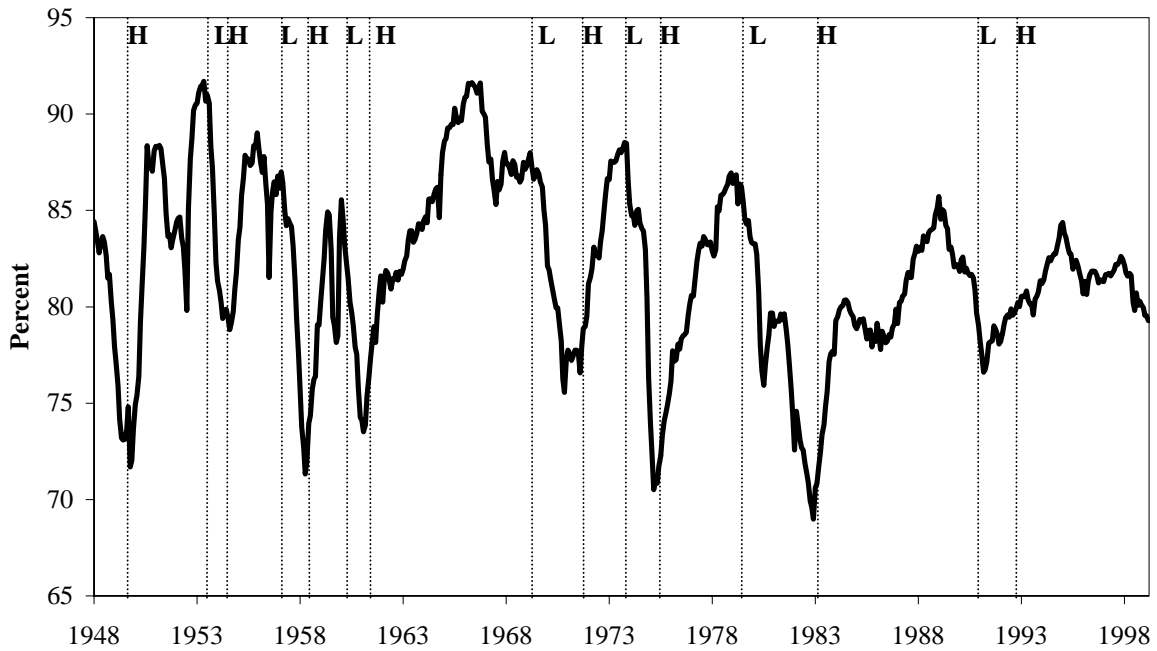
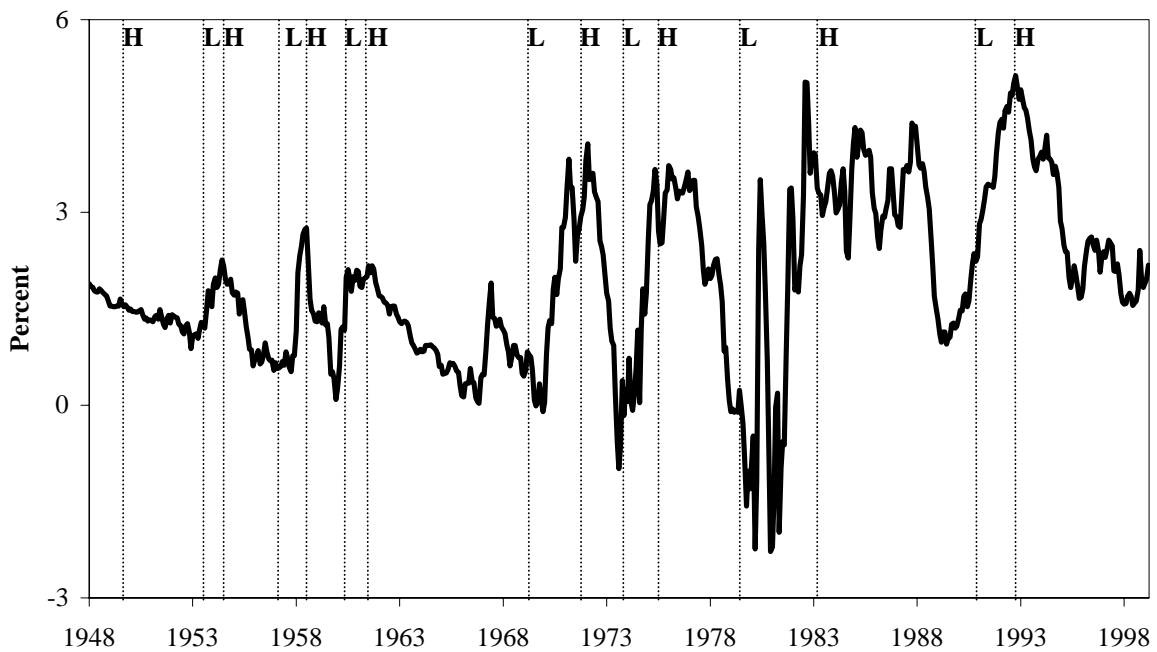
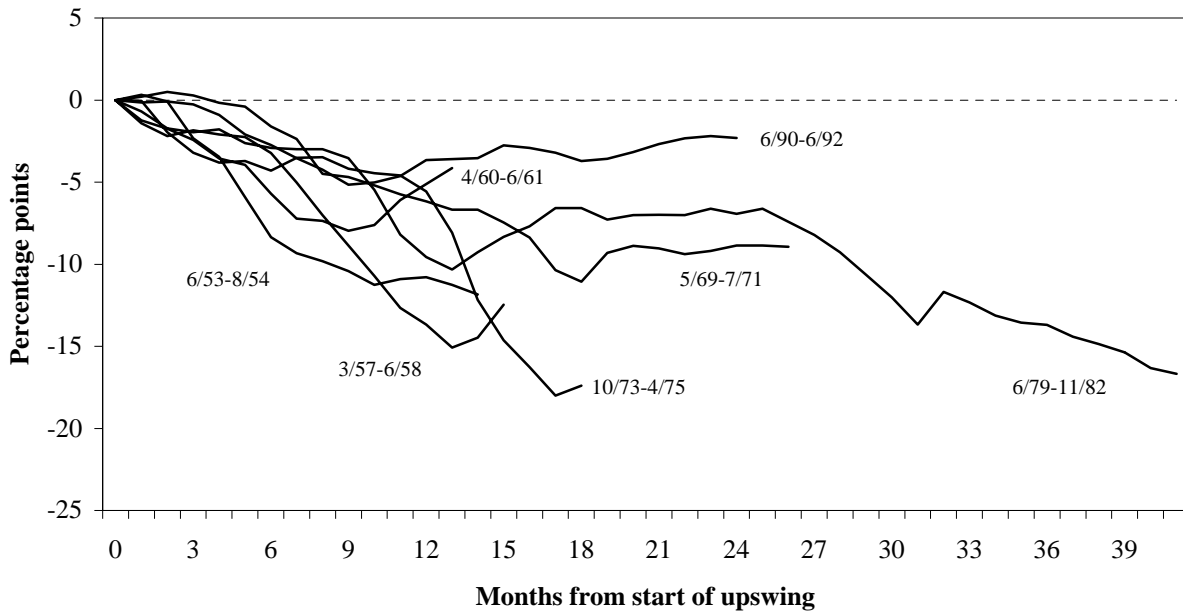


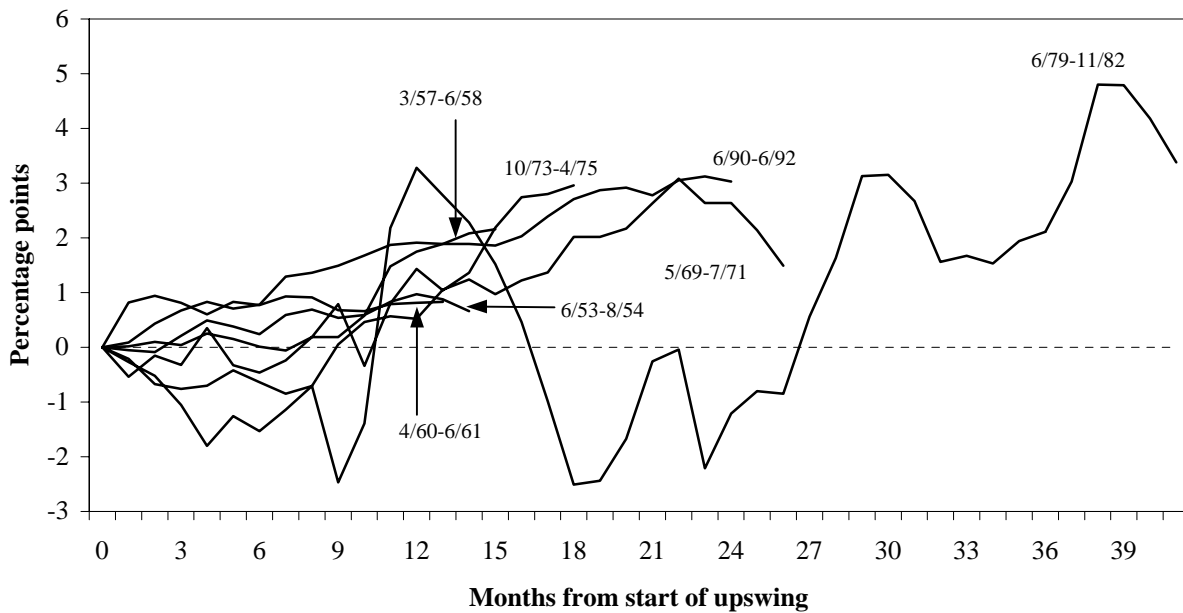
Figure 2B: Interest Rate Spread



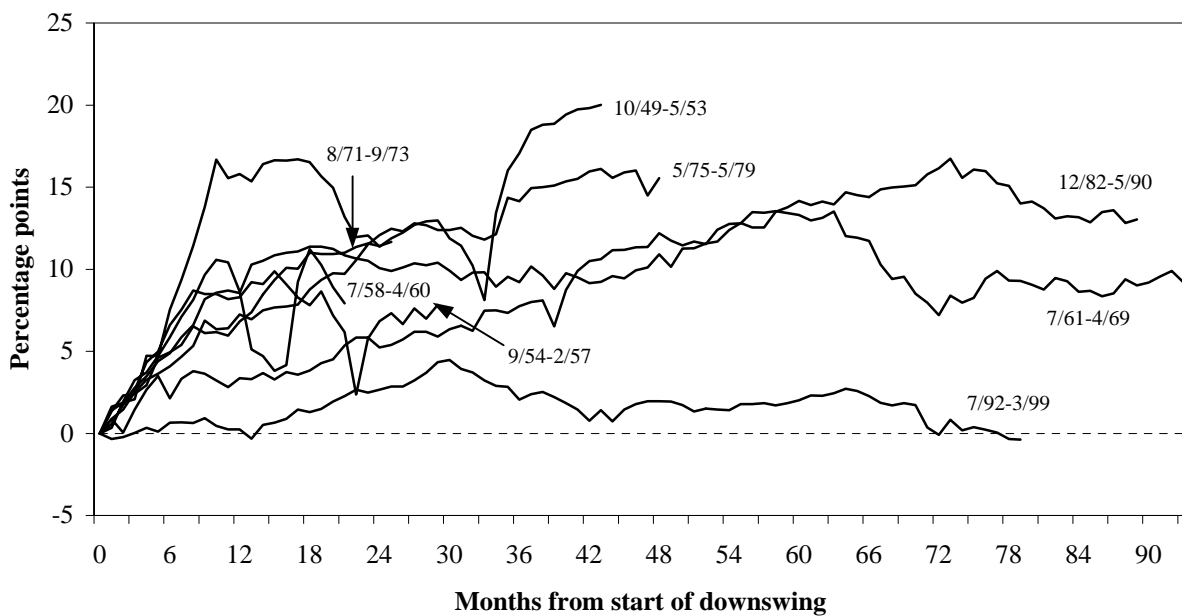
**Figure 3A: Change in Capacity Utilization
During Unemployment Rate Upswings**



**Figure 3B: Change in Interest Rate Spread
During Unemployment Rate Upswings**



**Figure 3C: Change in Capacity Utilization
During Unemployment Rate Downswings**



**Figure 3D: Change in Interest Rate Spread
During Unemployment Rate Downswings**

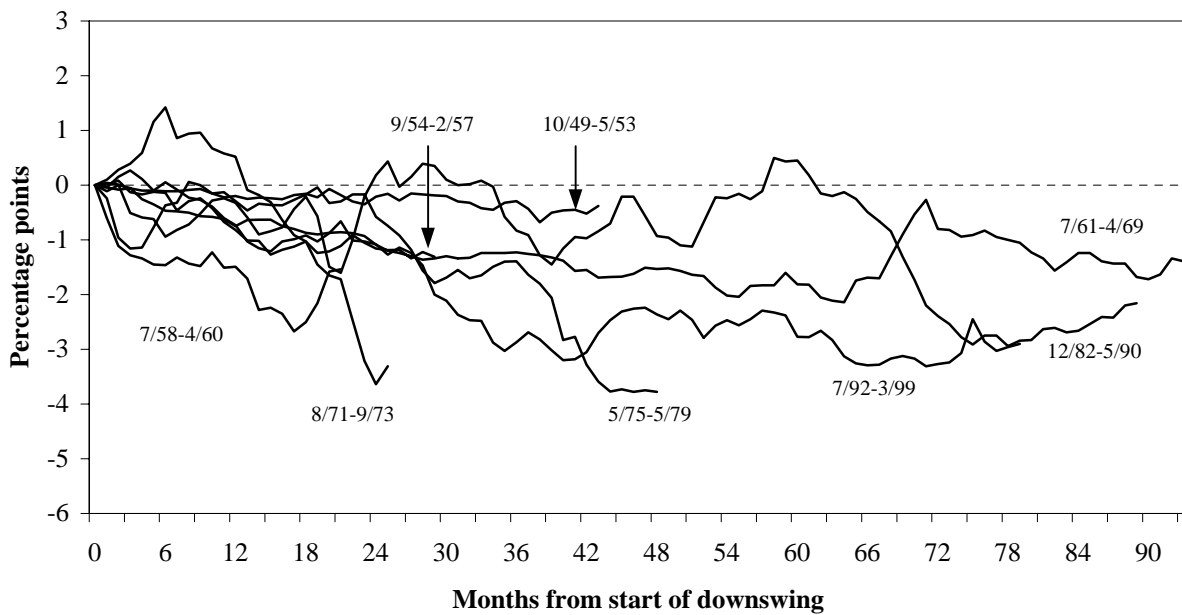


Figure 4: In-Sample Versus Out-of-Sample Turning-Point Probabilities

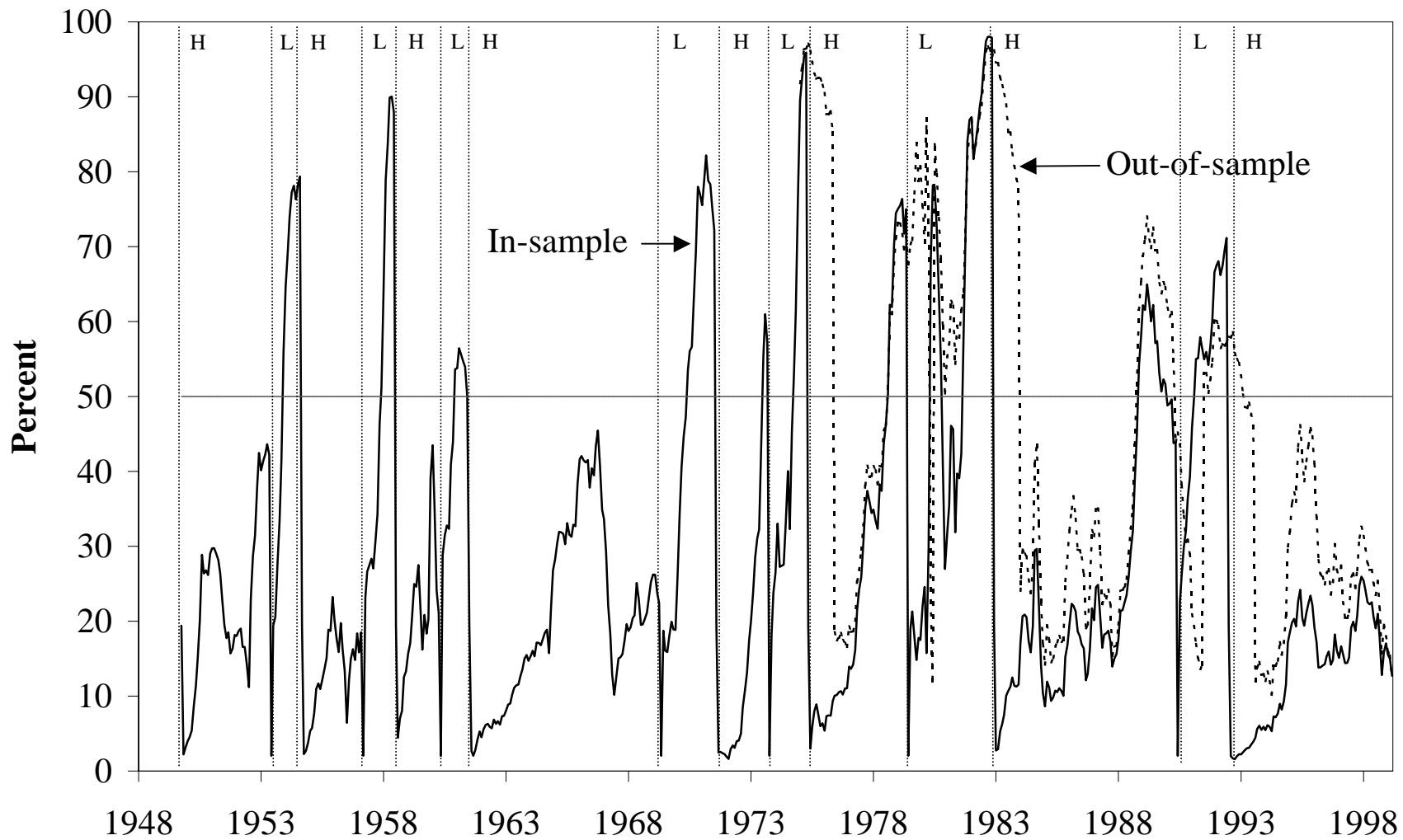


Figure 5: Out-of-Sample Turning-Point Probabilities

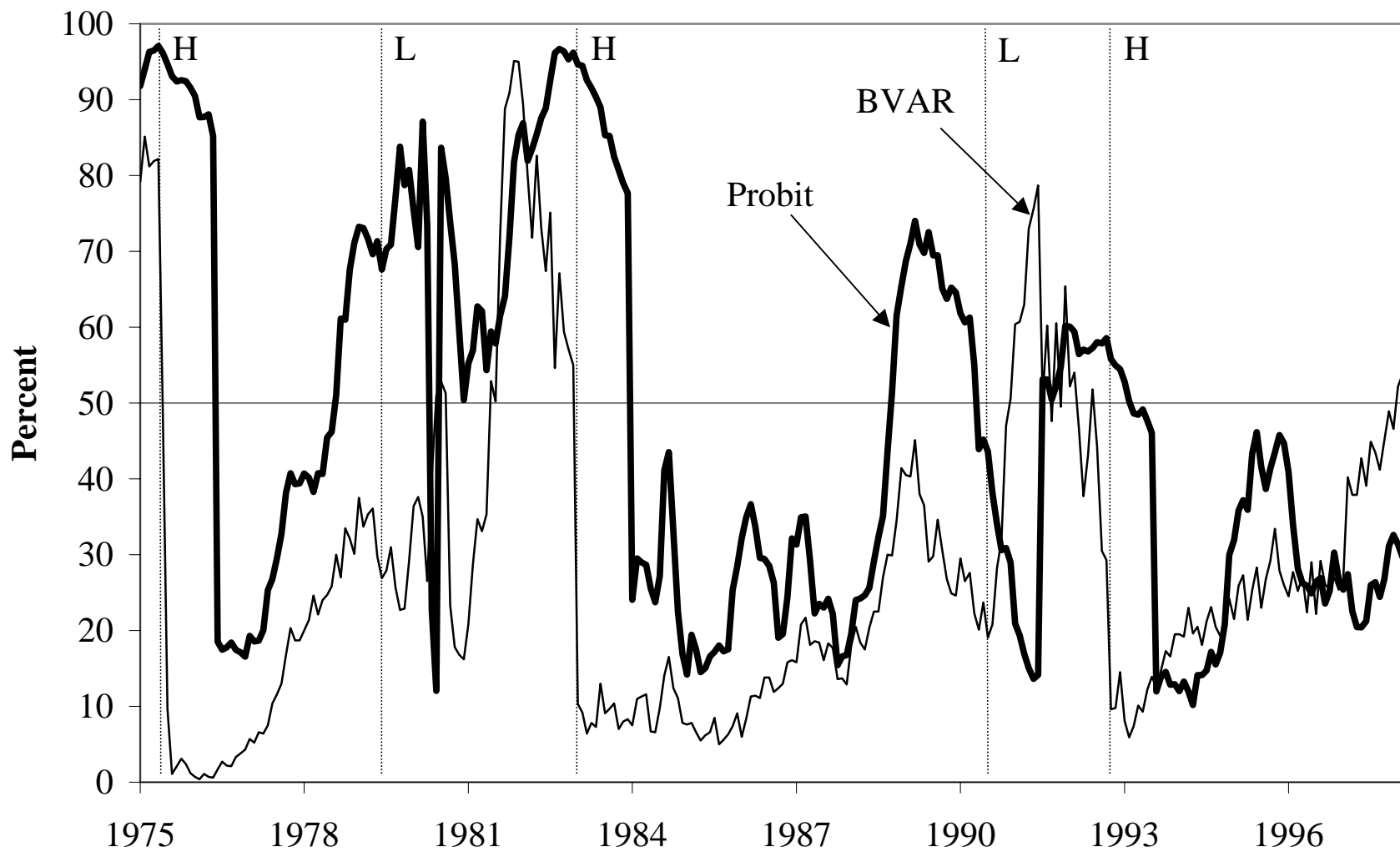
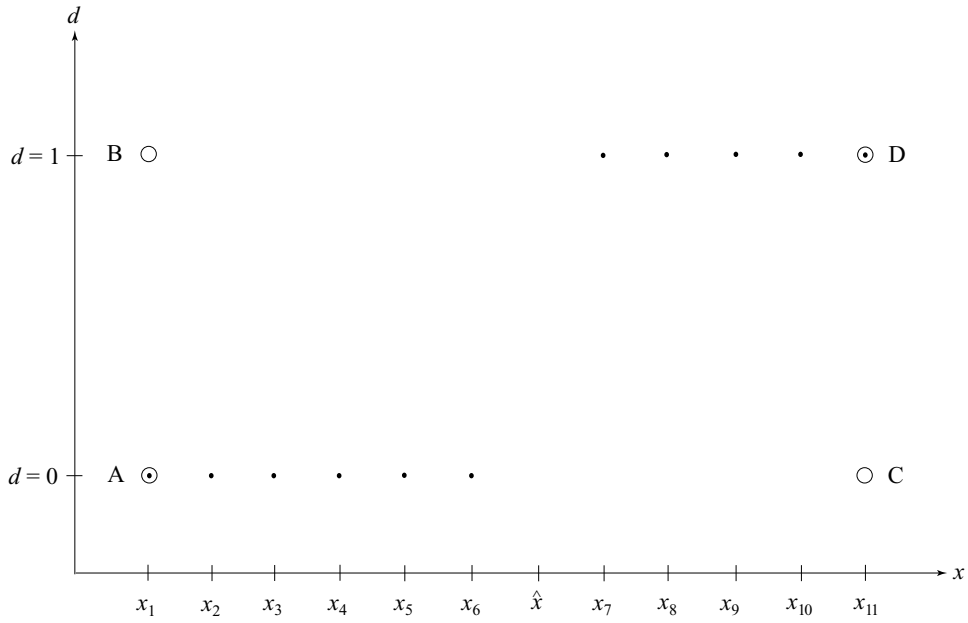


Figure A1: Adding Artificial Observations to the Probit Model



The figure portrays 11 actual observations, indicated by the dark points. The single explanatory variable is x , and the outcome variable is d . There are four artificial observations, at the centers of the circles labeled A, B, C, and D. The value \hat{x} is any point between x_6 and x_7 .