Volume Title: Business Cycles, Indicators and Forecasting

Volume Author/Editor: James H. Stock and Mark W. Watson, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-77488-0

Volume URL: http://www.nber.org/books/stoc93-1

Conference Date: May 3-4, 1991

Publication Date: January 1993

Chapter Title: Introduction to "Business Cycles, Indicators and Forecasting"

Chapter Author: James H. Stock, Mark W. Watson

Chapter URL: http://www.nber.org/chapters/c7188

Chapter pages in book: (p. 1 - 10)

Introduction

James H. Stock and Mark W. Watson

The provision of reliable economic forecasts has long been one of the principal challenges facing economists. This continues to be an important way in which the economics profession contributes to the operation of business and government. The past ten years have seen major advances in computational power, data availability, and methods for time-series and econometric analysis. These technological advances have in turn led to the development of new methods for economic forecasting with sufficient scope to provide serious competition to more traditional structural econometric models and to judgmental forecasts, at least for the primary economic aggregates such real GDP and the GDP deflator.

This volume consists of papers presented at a conference held at the offices of the National Bureau of Economic Research, Cambridge, Massachusetts, 3 and 4 May 1991. The purpose of the conference was to provide a forum for discussing and investigating new methods for economic forecasting and for the analysis of business cycles. This conference continues the long-standing involvement of the NBER with economic forecasting and the leading and coincident indicators. Indeed, it was in a 1938 *NBER Bulletin* that NBER researchers Arthur Burns and Wesley Mitchell initially proposed their system of coincident, leading, and lagging indicators (see Mitchell and Burns [1938] 1961). Under the guidance of Geoffrey Moore, Victor Zarnowitz, and their associates, Burns and Mitchell's proposal evolved into the system of economic indicators maintained today by the U.S. Department of Commerce.

The papers in this volume document several of the new macroeconomic forecasting techniques and models developed over the past ten years, compare

James H. Stock is professor of political economy at the Kennedy School of Government, Harvard University. Mark W. Watson is professor of economics at Northwestern University. Both are research associates of the National Bureau of Economic Research.

their performance to traditional econometric models, and propose methods for forecasting and time-series analysis. The volume starts with an exhaustive study by Victor Zarnowitz and Phillip Braun of the historical performance of economic forecasts. The data for their study are the responses to a quarterly survey of professional forecasters conducted by the NBER in conjunction with the American Statistical Association from 1968 to 1990. The forecasters surveyed used mostly traditional econometric models, augmented to various degrees by individual judgment. Because these data constitute actual forecasts, they provide a framework for evaluating historical forecast performance and for making hypothetical comparisons of actual forecasts to forecasts that would have been produced had other techniques been used.

The traditional focus of economic forecasting has been the forecasting of growth rates (in the case of the real GNP, e.g.) or levels (say, in the case of interest rates) of economic variables. However, nontechnical audiences—the business press, politicians, and business people—are often less interested in growth rate forecasts than in answers to simple questions, such as "Are we going to be in a recession next year?" or, "When will the current recession end?" To address such questions, the forecaster must construct conditional probability distributions of future paths of output or other key variables, a task made possible by the recent advances in computational power. The next group of papers (the papers by James H. Stock and Mark W. Watson, Ray C. Fair, and Christopher A. Sims) takes advantage of these computational advances to study different approaches to forecasting growth rates and discrete events, in particular, to forecasting whether the economy will be entering a recession. The specific models used for these forecasts are quite different: Stock and Watson use a monthly time-series model based on a range of leading economic indicators, Fair uses a structural econometric model, and Sims uses a reduced-form Bayesian vector autoregressive (VAR) system. The common element in the papers by Stock and Watson and by Fair is the use of stochastic simulation to produce probability forecasts of discrete recession/expansion events. All three papers study the performance of their forecasting systems during the 1990–91 recession and draw lessons from this experience.

The third group of papers consists of a pair of empirical studies of historical relations among specific economic time series. A number of recent studies have found that the spread between the short-term commercial paper rate and the rate on matched-maturity U.S. Treasury bills—the "paper-bill" spread—has had a strong predictive relation to aggregate economic activity. Benjamin M. Friedman and Kenneth N. Kuttner document this relation, summarize several hypotheses for why this predictive relation exists, and provide empirical evidence on the extent to which these hypotheses plausibly explain the power of this spread as a leading indicator.

In their paper, Francis X. Diebold, Glenn D. Rudebusch, and Daniel E. Sichel use the historical NBER business-cycle chronology to document the striking differences between prewar and postwar U.S. business-cycle dynamics. Before World War II, the probability of an expansion ending increased with each additional month, while the probability of a recession ending was almost independent of duration. After World War II, the opposite has been true. Diebold et al. also find that the characterizations of prewar European business-cycle durations are similar to those of the United States.

The papers in the final group have substantial methodological as well as empirical components. Danny Quah and Thomas J. Sargent develop techniques for analyzing dynamic factor models in high-dimensional systems. Previous applications of dynamic factor models have involved fewer than ten variables. However, recent computational advances and improved algorithms make it possible to contemplate applications to systems with many more variables. In their study, Quah and Sargent find that the dynamics of employment in over fifty sectors of the U.S. economy are well explained using only two unobservable dynamic factors.

Clive W. J. Granger, Timo Teräsvirta, and Heather Anderson develop a family of nonlinear time-series models for forecasting economic variables. They apply their techniques to the possibility of forecasting postwar growth in real GNP using a nonlinear model involving the index of leading indicators.

What follows is a more detailed summary of the individual papers.

Historical Performance of Economic Forecasters

From 1968 to 1990, the NBER and the American Statistical Association (ASA) collaborated in the quarterly collection of quantitative economic forecasts made by professional forecasters in private firms, academic institutions, government, and nonprofit organizations with an interest in economic forecasting. These forecasts covered a broad range of macroeconomic series, including income, production, consumption, investment, profits, government purchases, unemployment, inflation, and interest rates. These surveys provided a public service by facilitating the comparison of forecasts and the construction of composite forecasts. They also serve a scientific purpose: after twenty-two years of collection, the results from these surveys, along with the actual outcomes of the variables being forecast, provide an opportunity to study historical forecast performance and thereby learn how forecasts can be improved.

In their paper, "Twenty-two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance," Zarnowitz and Braun use these data to study the performance of professional forecasters over this period. The authors start by discussing the uses and history of economic forecasting and by providing a history of forecast appraisals. They then describe the content and evolution of the NBER-ASA survey and document the construction of the associated data base.

Zarnowitz and Braun then turn to an extensive analysis of the survey forecasts, leading to several main conclusions. (1) At any point in time, and for any economic variable, there is typically great dispersion across forecasts, which typically increases with the forecast horizon. (2) Macroeconomic variables differ greatly in the ease with which they are forecast: growth in real GNP and consumption were forecast better than inflation, residential investment, and changes in business inventories. (3) Perhaps surprisingly, over this period there were no large systematic improvements in forecast accuracy: although inflation forecast accuracy increased, the accuracy of real GNP forecasts decreased. (4) Combined forecasts, in the form of group mean forecasts, are generally more accurate than individual forecasts. Moreover, the group mean forecast for many variables, including real output, outperforms time-series models, in particular, variants of vector autoregressions, when the comparison is conducted in a simulated "real-time" setting (i.e., when the time-series models are estimated, and forecasts constructed, using only those data available contemporaneously to the forecasts made by the survey participants).

The Prediction of Recessions and Expansions

In "A Procedure for Predicting Recessions with Leading Indicators: Econometric Issues and Recent Experience," Stock and Watson describe an approach to forecasting recessions in the U.S. economy that they have been using since October 1988 and analyze its out-of-sample performance. Unlike the earlier literature in this area, which has focused on predicting turning points, the problem here is posed as one of forecasting a discrete variable indicating whether the economy is in a recession. Stock and Watson's approach is to define recessions and expansions as different patterns of economic activity so that whether the economy will be in a recession is equivalent to whether the path of overall economic activity falls in a recessionary or an expansionary pattern. With quantitative definitions for these two patterns, the probability that the economy is in a recession during a future month can then be computed by the stochastic simulation of a model that forecasts future economic activity.

The specific forecasting system used for the stochastic simulation is a reduced-form monthly time-series model, developed in earlier work, based on seven leading and four coincident economic indicators. This model was estimated using data from January 1959 through September 1988. Since then, it has been used to produce three indexes of overall economic activity on a monthly basis: an experimental coincident index (the XCI); an experimental leading index (the XLI), which is a forecast of the growth in the XCI over the subsequent six months; and an experimental recession index (the XRI), which estimates the probability that the economy will be in a recession six months hence.

These indexes performed well from 1988 through the summer of 1990—for example, in June, 1990, the XLI model forecast a 0.4 percent (annual rate) decline in the XCI from June through September when in fact the decline was only slightly greater, 0.8 percent. However, the XLI and the XRI failed to forecast the sharp declines of October and November 1990. Even so, the short-horizon recession probabilities produced by the model performed relatively well during this episode. After investigating a variety of possible explanations for the forecast failure, Stock and Watson conclude that the main source was the failure of the individual leading indicators included in the model to forecast the sharp drop in aggregate growth, either individually or collectively. In short, the XLI and the XRI relied on financial variables during a recession that, unlike the previous three recessions, was not associated with a particularly tight monetary policy.

This poor forecasting record entering the 1990 recession is typical of a wide range of economic indicators. Of a broad set of forty-five coincident and leading indicators, Stock and Watson find that almost all performed poorly during this episode; even the best had large forecast errors by historical standards, and, moreover, they performed relatively poorly in the recessions of the 1970s and early 1980s. This in turn suggests that there was only limited room for improvement in the performance of the recession forecasts.

In "Estimating Event Probabilities from Macroeconomic Models Using Stochastic Simulation," Fair also considers the problem of forecasting recessions. In contrast to Stock and Watson's reduced-form approach based on monthly data, Fair studies recession forecasts using his quarterly structural econometric model, a nonlinear dynamic simultaneous equation system consisting of 30 stochastic equations, 98 identities, and 179 estimated coefficients. Because his system has exogenous variables, he can study recession forecasts that incorporate three different types of uncertainty: uncertainty in the future path of the endogenous variables, given future values of the exogenous variables and the model coefficients; uncertainty in the exogenous variables, given the model coefficients; and uncertainty (arising from estimation error) about the model coefficients themselves.

Fair considers business-cycle events defined in terms of real GNP, and he examines three alternative discrete events: (1) at least two consecutive quarters of negative growth in real GNP during the next five quarters; (2) at least two quarters of negative real GNP growth during the next five quarters; and (3) at least two quarters of the next five having inflation exceeding 7 percent at an annual rate. Because the event "two consecutive quarters of negative real GNP growth" is a conventional, if sometimes inaccurate, rule-of-thumb definition of a recession, the first of these events corresponds to a recession occurring sometime during the next five quarters.

On a computational level, Fair's approach is to draw a set of exogenous variables, disturbances, or coefficients, depending on which of the three types of simulation is being performed, and to use these to compute a stream of forecasts over 1990:I–1991:I. The fraction of times that the forecast registers the indicated type of event yields the probability of that event occurring.

Fair's initial focus is the five quarters 1990:I–1991:I. Even though his mean forecasts predict positive growth in four of the five quarters, the probabilities of the two contraction events are rather high: for the full stochastic simulation (with uncertainty arising from endogenous variables, exogenous variables, and coefficients), the probability of two consecutive declines in GNP approaches 40 percent.

Fair then turns to the more computationally demanding task of computing a sequence of event probabilities over the period 1954:I–1990:I and compares the event forecasts produced by his model to those produced by a "naive" model, a univariate autoregression for GNP. For each of the two recession events, Fair's model (with endogenous variable and coefficient uncertainty, using future values of the exogenous variables) outperforms the naive model using conventional probability scores. Overall, these results are encouraging and suggest pursuing further work using stochastic simulation to predict discrete events.

One of the most important advances in forecasting methodology during the 1980s was the development and refinement of small multivariate time-series forecasting models, in particular, vector autoregressions (VARs). Since first introducing VARs to economists, Christopher Sims and his students have pursued a research program aimed in part at improving the forecasts made by VARs. A key aspect of this program has been the ongoing production of quarterly forecasts from a Bayesian VAR. This model was originally developed and maintained by Robert Litterman at the Federal Reserve Bank of Minneapolis. Sims took over the responsibility for this model and its forecasts in 1987.

The paper by Sims in this volume—"A Nine-Variable Probabilistic Macroeconomic Forecasting Model"—documents the current version of this model and summarizes the changes that have been made to it over the years. Sims then provides various measures of the model's performance, both in sample and out of sample. The version of the model currently in use incorporates nine variables and can be thought of as a third-generation VAR. Because a nine-variable, five-lag VAR would have a very large number of coefficient estimates (forty-five regression coefficients per equation, plus a constant), unrestricted estimation of this system would result in imprecise coefficient estimates and thus a good chance of poor out-of-sample performance. Sims uses Bayesian techniques to restrict the otherwise large number of parameters. Sims has also modified Litterman's original model to incorporate two deviations from the standard linear/Gaussian framework, conditional heteroskedasticity and nonnormal errors. Also, the model has been modified to permit cointegration among the variables. These modifications and the priors incorporated into the model are documented in Sims's paper.

Sims next examines the performance of his model. The early VARs had produced good forecasts of real variables, but their forecasts of inflation were substantially worse than forecasts from traditional structural econometric equations. According to Sims's evidence, his subsequent model modifications improved the inflation forecasts without deteriorating the real forecasts. The final evidence examined in Sims's paper is the performance of the model in the 1990 recession. Like almost all (unadjusted) formal quantitative economic models, this VAR failed to forecast the negative GNP growth in the fourth quarter of 1990 and the first quarter of 1991. Sims concludes by exploring the lessons of this episode for future work.

Historical Empirical Studies

A series of recent papers has shown that the difference between interest rates on commercial paper and U.S. Treasury bills—the "paper-bill spread"—has, for the past three decades, exhibited a systematic relation to subsequent fluctuations of real economic activity. Friedman and Kuttner's paper "Why Does the Paper-Bill Spread Predict Real Economic Activity?" documents the empirical facts about this spread as a leading economic indicator and studies various economic reasons why this spread has such a strong historical forecasting record.

Friedman and Kuttner start by documenting the value of this spread as a predictor of economic activity. Commercial paper represents the unsecured, discounted short-term (up to 270 days) liability of either nonfinancial business corporations or financial intermediaries. The paper-bill spread outperforms any other interest rate or any monetary aggregate as a predictor of output. In contrast to the monetary aggregates, the authors argue, this spread clearly forecasts real rather than nominal economic activity; it predicts nominal magnitudes only to the extent that nominal magnitudes reflect real ones. In his discussion of Friedman and Kuttner's paper, Ben S. Bernanke presents additional evidence concerning the striking predictive performance of this spread as a leading economic indicator.

Friedman and Kuttner turn next to a description of several factors that can account for the levels of and changes in the spread. One explanation of the mean level is the difference in tax treatments between commercial paper and Treasury bills when the interest is received by entities domiciled in states or municipalities with an income tax; the authors calculate that an effective state/municipal tax rate of 8.1 percent would suffice to explain the spread between six-month commercial paper and six-month Treasury bills. A second factor in the spread is that commercial paper is subject to potential default by private obligors, a factor that is exacerbated by the junior standing of commercial paper as unsecured debt. A third factor underlying this spread is the greater liquidity of the Treasury-bill market than the commercial paper market. Although the total value of commercial paper outstanding in 1989 was $579 billion, as recently as 1960 the volume outstanding was only $6.5 billion. In contrast, the U.S. Treasury market has been well developed throughout the postwar period, with a total value outstanding of $482 billion in 1990. This growth of the commercial paper market during this period, along with legal restrictions on the use of commercial paper, raises the possibility that the commercial paper market had substantially less liquidity than the Treasury-bill market for much of this episode. To quantify these factors, Friedman and Kuttner provide a decomposition of changes in the level of the paper-bill spread among changes in the level of interest rates (as suggested by the tax and default arguments), changes in quality as measured by the P2-P1 commercial paper premium, and residual, unexplained changes. They find that all three components are statistically and economically large.

Having documented these historical relations, Friedman and Kuttner study three hypotheses about why this spread predicts aggregate economic activity. The first concerns changes in perceptions of default risk: a widening of the spread reflects an increasing fear of a downturn, business failures, and concomitant defaults on debt. The second hypothesis is that the paper-bill spread is an indicator of monetary policy. The third hypothesis emphasizes changes in borrowers' cash flows: to the extent that borrowers' cash flows vary cyclically, borrowing requirements might rise toward the end of an expansion (because of constant costs in the face of declining sales), with the result that the increasing spread would reflect an increasing supply of commercial paper and an increasing commercial paper rate. When these hypotheses are studied empirically using a more structural approach based on imperfect substitutability between commercial paper and Treasury bills, empirical support is found for each of these three hypotheses.

In "Further Evidence on Business-Cycle Duration Dependence," Diebold et al. use formal statistical techniques to take a new look at an idea found in popular discussions of the business cycle: that business cycles exhibit duration dependence. That is, the probability that an expansion or a recession will end depends on the length of that expansion or recession. The authors' previous research on duration dependence in the U.S. business cycle found evidence of substantial differences between the prewar and the postwar business cycle: during the postwar period, contractions exhibit duration dependence, but expansions do not, while the opposite is true during the prewar period. This paper extends this line of research to France, Germany, and Great Britain.

The analysis of duration dependence in business cycles is made difficult by the small number of observations of recessions or expansions contained in even one century of data. Thus, techniques for the analysis of duration dependence appropriate for large samples—the estimation of nonparametric or semiparametric hazard models—are inapplicable here because they require too many observations. Instead, Diebold et al. employ a quadratic hazard model that is parsimonious yet flexible enough to allow the nonmonotone hazards that might be found in business-cycle data.

The application of this quadratic hazard model to the U.S. business-cycle chronology confirms the authors' earlier findings, obtained using a simpler hazard specification, about the differences between duration dependence during the prewar and postwar periods. For example, they find that the hazard rate for postwar recessions rises from .07 to .29 over the course of twelve months.

Their results for France, Germany, and Great Britain indicate that prewar expansions exhibit positive duration dependence in all three countries and that in none of the countries do prewar contractions exhibit positive duration dependence. There is also evidence for duration dependence in prewar whole cycles in these three countries, which the authors attribute to the positive duration dependence of the expansion phase. Overall, these results are qualitatively the same for the United States, which leads the authors to suggest that, during the prewar period, there were substantial similarities across countries in business-cycle dynamics.

Methods for Analyzing Economic Time Series

Much of the aggregate economic data of primary interest to economic forecasters has disaggregated components. For example, the U.S. Bureau of Labor Statistics reports total private employment and employment disaggregated by industry. However, the richness provided by these disaggregated data has largely been ignored in many recent developments in the area of economic forecasting. From the point of view of economic theory, a study of the comovements of these data might elucidate the extent to which different sectors respond to aggregate shocks and might even help identify the number of separate aggregate shocks to the economy. From the point of view of economic forecasting, the use of these data might result in better measures of these different aggregate shocks, which could in turn be used to improve aggregate forecasts. However, the very richness of the data—the large number of disaggregated sectors—has posed a technical barrier to the simultaneous modeling of these comovements.

In "A Dynamic Index Model for Large Cross Sections," Quah and Sargent embark on a project to model simultaneously the comovements of a large number of disaggregated series. They examine dynamic factor models, in which the comovements of the series are presumed to arise from a reduced number of factors. These factors can affect different series with different lags and dynamic specifications. Because of computational limitations, these models have in the past been fit to small systems, for example, with four time series. The main technical advance in the paper is the development of procedures, based on the "EM" algorithm, for the fitting of these models to a large panel of time series.

In their empirical application, Quah and Sargent examine the comovements in U.S. employment in fifty-four industries over the period 1948–89. Their striking finding is that a large fraction of the variation in employment can be described by two common factors. Their results demonstrate that the construction of such large-scale dynamic factor models is both feasible and potentially valuable both for forecasting and for data description purposes.

Most statistical analysis of economic times series is done using linear models. However, economic theory typically predicts linear relations only as special cases; more often, the processes are likely to be nonlinear, in the sense that optimal forecasts will involve nonlinear rather than linear functions of the observed variables. In "Modeling Nonlinearity over the Business Cycle," Granger et al. outline a family of nonlinear time-series models and tests that might usefully be applied to economic data. Their main focus is on smooth-transition regression models, which allow for regression coefficients to take on two values and to shift between these two values in a continuous way.

The empirical focus of their paper is the relation between GNP and the Department of Commerce's index of leading indicators. Their objective is to ascertain whether a nonlinear model provides better forecasts of real GNP than a linear model does, in particular, whether a nonlinear model would have predicted the onset of the 1990 recession better than a linear one would have. Overall, in this application, the results are mixed: although formal statistical tests provide some evidence of nonlinearity, and although the nonlinear model provides quite different forecasts than the linear model, neither model performed particularly well in the slow-growth period leading to the recession and in the sharp contraction during the autumn of 1990.

Reference

Mitchell, W. C., and A. B. Burns. [1938] 1961. Statistical indicators of cyclical revivals. Reprinted in *Business cycle indicators*, ed. G. H. Moore. Princeton, N.J.: Princeton University Press.