

NBER WORKING PAPER SERIES

EXPECTATIONS AND FORECASTS FROM  
BUSINESS OUTLOOK SURVEYS

Victor Zarnowitz

Working Paper No. 845

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge MA 02138

January 1982

Financial support from the National Science Foundation and aid of the Graduate School of Business of the University of Chicago and NBER are gratefully acknowledged. I am greatly indebted to Louis Lambros for helpful suggestions and efficient research assistance. The research reported here is part of the NBER's research program in Economic Fluctuations. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

Expectations and Forecasts from Business Outlook Surveys

ABSTRACT

Each quarter since 1968 the National Bureau of Economic Research, in collaboration with the American Statistical Association, has been collecting a large amount of information on the record of forecasting in the U. S. economy. This paper is a progress report on a comprehensive study of the distribution of individual predictions from these surveys. It covers forecasts of quarterly developments in the year ahead for six variables representing inflation, real growth, unemployment, percentage changes in GNP and spending on consumer durables, and business inventory investment. The 79 respondents who participated in at least 12 of the 42 surveys covered constitute a broadly based and diversified group of experts and agents, mostly from the world of corporate business and finance--executives, analysts, economic consultants, also some government and academic forecasters. The data are in certain respects uniquely rich.

The first part of the paper reviews briefly the models of economic expectations and discusses the potential and problems of using survey data for testing these models. The second part offers a comparative analysis of the individual prediction series from the NBER-ASA as well as some earlier surveys. There are gains from combining predictions from different sources, e.g., the group mean forecasts are on the average over time more accurate than most of the corresponding sets of individual forecasts or expectations. But there is also a moderate degree of consistency in the relative performances of individual forecasters, some of whom score well above average with respect to several variables and predictive horizons.

The third section presents the distributions of an array of absolute accuracy measures for the survey respondents, regressions of actual on predicted values, and associated tests of bias and autocorrelation of error. The marginal forecast errors tend to increase, and the correlations between predictions and realizations tend to decrease, as the target quarter recedes into the future. The tests of the joint null hypothesis that the regressions have zero intercepts and unitary slope coefficients are very unfavorable to expectations of inflation, but they show the forecasts of the other variables generally in much better light. Inflation has been largely underestimated, with the predicted rates lagging behind the actual rates. On the other hand, real growth has been on the average overestimated. The incidence of autocorrelation in the prediction errors was also much higher for inflation than for the other variables.

A summary of findings is provided. The fifth and last section lists some additional questions raised by this study, to be dealt with in another paper.

Victor Zarnowitz  
Graduate School of Business  
The University of Chicago  
Chicago, Illinois 60637

ERRATA  
for  
EXPECTATIONS AND FORECASTS FROM  
BUSINESS OUTLOOK SURVEYS

Victor Zarnowitz  
NBER Working Paper No. 845

Page

- 11 Last line omitted. Add the following words at the bottom of the page:  
. . . as well as a measure of inflation. For each of these variables,  
nearly 400 . . .
- 22 Title: Should read CHART 2 instead of Chart 1.
- 31 First line: .4 to .6 for Q0 (not ".4 and .6").
- 40 Line 9: "not be known" (instead of "not known").
- 40 Footnote 28, line 4:  $R = n \sum_{k=1}^m \hat{r}_k \sim X_m^2$  ( $\hat{r}_k$  instead of  $k$ ).
- 40 Footnote 29, last line: add "for Q2 and Q3" after  $r_2$ .
- 43 Last para., line 2: note 28 (instead of note 26).
- 45 Line 5 from below: group mean (not "group, mean").

## I. Models, Tests and Data

Realizations of economic processes are singular, though often subject to different interpretations. In contrast, expectations are as a rule plural and mutually exclusive, referring as they do to sets of more or less probable outcomes of alternative courses of action and to the interplay of partially known past events and imagined future eventualities. The process of the formation of expectations is internal to their source and hidden to outsiders; it may be complex and need not be explicit. Even the results of the process are never readily observable; however, some can be and are reported through questionnaire surveys. These data pose many problems but so do the attempts to model the formation of economic expectations, that is, to infer the latter from postulated relations with observed (realization) variables.

### On Models of Expectations

Starting from deterministic formulae (simple moving averages, trend projections), the evolution of extrapolative models proceeded along two related lines, through recognition of (a) learning from errors and (b) the stochastic nature of most economic processes. The linear distributed-lag function generating one-period ahead predictions of a variable  $y$ ,

$$(1) \quad y_t^* = \sum_i^{\infty} \beta_i y_{t-i}, \quad \beta_i > 0, \quad \sum_i^{\infty} \beta_i = 1 + g$$

provided a general point of departure (where  $g$  is a trend rate of growth, which is zero in the stationary case). It is convenient (and, arguably, likely) that the  $\beta$ 's decline into the past from some point in time  $i > 1$ , so that most will be negligibly small. In particular, the popular trendless model

$$(2) \quad y_t^* = \sum_{i=1}^{\infty} \beta(1-\beta)^{i-1} y_{t-i}, \quad 0 < \beta < 1,$$

with geometrically declining  $\beta$  weights whose sum equals unity, requires the determination of only one parameter. It implies learning from ascertainable errors according to

$$(3) \quad y_t^* - y_{t-1}^* = \beta(y_{t-1} - y_{t-1}^*),$$

where under (over) estimation of  $y_{t-1}$  results in a partial upward (downward) revision of  $y_t^*$  relative to  $y_{t-1}^*$ .

Suppose that the series  $y$  was generated by a moving average process

$$(4) \quad y_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}, \quad \alpha_0 = 1,$$

where the  $\varepsilon$ 's represent independent random shocks or "white noise" (with zero mean and constant variance). Then the adaptive model (2) would produce optimal predictions of  $y$  in the special case of all  $\alpha_j$  being the same and equal to  $\beta$ , and the linear extrapolation (1) would do so under less restrictive conditions.<sup>1</sup>

The most recent and sophisticated methods of extrapolating the future values of a series from its past values derive from the analysis of stationary

---

<sup>1</sup>See Muth 1960; Nerlove and Wage 1964; Mincer 1969. Other learning models have been suggested, e.g.,

$$(5) \quad \Delta y_t^* = \gamma \Delta y_{t-1} \quad \text{and}$$

$$(6) \quad \Delta y_t^* = \delta(y_{t-1} - \bar{y}),$$

where  $0 < \gamma, \delta < 1$ ,  $\Delta y_{t-1} = y_{t-1} - y_{t-2}$ , and  $\bar{y}$  is a "normal" (perhaps a long-term average) value of the series  $y$ . Combining (3), (5), and (6) would make the revision term  $\Delta y_t^*$  a linear function of the adaptive, extrapolative, and regressive elements appearing, respectively, on the right-hand sides of the three equations. See Kane and Malkiel 1976; Jacobs and Jones 1980; Curtin 1980.

and homogeneous nonstationary stochastic processes.<sup>2</sup> When skillfully applied to sufficiently large amounts of good data, these techniques can be quite effective in practice. However, univariate times series models do not attempt to describe or explain the relationships expected to exist among different variables. In principle, knowledge of the true structure of such relations provides additional predictive information, which a properly specified econometric model can exploit but a purely extrapolative model cannot. In short, let  $X$  be the vector of those variables that influence  $y$  and let  $LX$  and  $Ly$  ( $L$  being a general lag operator) represent the appropriately weighted sums of the terms  $x_{t-i}$  and  $y_{t-j}$ , respectively. Then the model

$$(7) \quad y^* = f(LX, Ly)$$

would dominate even the best extrapolative model, since the latter would have as its arguments  $Ly$  alone.<sup>3</sup>

Economic expectations are to some extent autonomous in the sense of using information that is not contained in the past values of any regular time-series variable, e.g., news on current actions and plans of the government,

---

<sup>2</sup>This class of time series models includes the autoregressive (AR), moving average (MA), and mixed and integrated processes (ARMA and ARIMA). See Box and Jenkins 1970.

<sup>3</sup>Some but not all of the past effects on  $y$  of  $LX$  are accounted for indirectly in  $Ly$ . Implementation of (7) might take the form of a combined regression-cum-time-series model (or, if  $y^*$  and  $y$  are vectors, a multi-equation system including both econometric and time-series components). Such a model could provide a proper representation of a multivariate time series process, which is consistent with, but more informative than, a set of the corresponding univariate time-series models (Zellner and Palm 1974). Only in very special cases are purely extrapolative expectations "rational" in the sense that they employ all available information (Nelson 1975). It is well to remember, however, that this theoretical advantage of econometric models presupposes their validity and may be canceled by misspecification errors. Also, the incremental gains from using a correct econometric model instead of an efficient time-series model need not always be large enough to justify the costs which may be much larger for the former than for the latter (Feige and Pearce 1976).

strikes, international developments, opinion polls, etc. Market rewards motivate economic agents to use all information that can be acquired in a cost-effective way. The rational expectations hypothesis assumes that a sufficiently large number of agents know "how the world works," that is, recognize the structure of their economic environment and efficiently process all available and pertinent data. It is the so formed expectations that are decisive for what transpires in the market place and they are reflected in the equilibrating behavior of prices and other endogenous variables (Muth 1961; Poole 1976). Rational expectations contain only purely random (though possibly large) errors and thus satisfy

$$(8) \quad \mathcal{E}(y_t^* | I_{t-1}) = \mathcal{E}(y_t) ,$$

where  $\mathcal{E}$  is the expected value operator and  $I_{t-1}$ , the set of information available at the end of period  $(t-1)$  on which  $y^*$  is conditioned, includes the knowledge of the required models as well as the data.

Formulation (8) is too abstract to be useful in practice, but all attempts to work out the full implications of this approach in its strongest form confront a dilemma. Without specifying the contents of the information set  $I_{t-1}$ , the rationality of the corresponding expectations cannot be given a complete evaluation. But in most cases that matter (e.g., for the much studied expectations of inflation) it seems impossible to know what  $I_{t-1}$  contains. Economists do not agree on all the important features of their models, and insofar as the models contradict each other they surely cannot all

be properly specified.<sup>4</sup> It is difficult to accept the notion that the representative agent is free of the limitations of knowledge that are evident in experts' analysis of the economy. But consequences of incomplete information or deficient knowledge may be mistaken for departure from rational expectations.<sup>5</sup>

#### Tests of the Models and Survey Data

One implication of (8), namely that the expectational errors are unbiased,

$$(9) \quad \mathcal{E}(y_t^* - y_t) = 0 ,$$

can, of course, be tested without specifying  $I_{t-1}$ , providing that adequate data on  $y_t^*$  are available. To this end, the regression

$$(10) \quad y_t = a + by_t^* + u_t$$

is estimated to verify or falsify the joint hypothesis that  $a$  and  $b$  are not statistically different from 0 and 1, respectively. However, this is a weak test of rational expectations, since the latter imply an efficient use of all pertinent information, not just unbiasedness. A considerably stronger test would require, in addition to  $H_0: (a, b) = (0, 1)$ , that the errors

---

<sup>4</sup>The existing econometric models in effect summarize the average historical experience in the past few decades, and their fixed-weight equations attribute to economic agents the same behavior patterns under very different economic conditions and policy regimes. But such invariance is not really credible. For a serious criticism of such models from the point of view of the rational expectations hypothesis, see Lucas 1976. However, the construction of econometric models that would conform strictly to the rule of rational expectations runs into major identification problems.

<sup>5</sup>Some errors that appear to be "systematic" in retrospect could not have been detected and corrected on a current basis, as will be illustrated later in this paper. There is often much uncertainty about what is happening in the economy for lack of sufficiently long series of consistent and timely observations (Zarnowitz 1982).



that are themselves a part of  $I_{t-1}$  be essentially free of all avoidable systematic elements such as significant autocorrelations.

Much effort was spent in recent years on the collection and processing of expectational data from periodic surveys of various groups: consumers, corporate managers, business and financial analysts, economists. This work was motivated mainly by the prospect of obtaining useful tools for practical forecasting, but it is increasingly recognized that the surveys provide the principal source of direct measures of economic expectations. Since the models discussed above raise many questions and need to be tested, such measures have important analytical uses.

The basic survey data represent anticipations, intentions, forecasts or plans of individuals and teams (organizations), but studies have used only or mainly the time series of averages based on responses to a sequence of surveys by groups whose composition varies over time. This raises the possibility of serious aggregation errors due to the neglect of the cross-section and distributional aspects of the data: differences among the individuals and subgroups; sampling variation; consistency and representativeness of the employed averages.<sup>6</sup> This study will pay attention to some of these aspects.

At any time, some people will outperform others in anticipating the future, partly by chance and partly because of systematic factors such as greater incentives, more skills and knowledge, and access to better information. Those who succeed relatively often tend to reap market gains; the competitive game of economic prediction cannot be comprehended by treating expectations as if they were single-valued and universally shared. Modeling

---

<sup>6</sup>An early study which dealt with certain characteristics of the relation between aggregate and individual forecasts is Zarnowitz 1967, pp. 123-126. A recent analysis of disaggregated data from surveys of inflation forecasts is Figlewski and Wachtel 1981.

of expectations, therefore, should allow for their dispersion. In this context, it is important to distinguish between individual and market expectations. Prices in a market may incorporate all available information, even though price expectations of many, perhaps most, traders do not meet the rationality criterion. This will happen simply if some resourceful participants have their way in eliminating the unexploited profit opportunities in the given market.<sup>7</sup>

However, under uncertainty, quantity signals may be as important as price signals, particularly in areas of the economy other than the competitive auction markets with flexible prices. Economic agents are presumably most interested in the local variables (quantities as well as prices) which relate closely to their own activities. But aggregate measures such as GNP and components, changes in money, credit, price levels and interest rates, sensitive cyclical indicators, foreign trade and exchange rates, are also widely monitored and selectively used. For most of the macrovariables, market expectations are nonexistent or unobservable, but it is evident that numerous predictions are being regularly made and used throughout the economy. The outlook for the economy plays an essential role in shaping expectations of many decision-making units, including prominently large and diversified corporations. Macropredictions serve as important inputs to micropredictions.

Not surprisingly, professional business analysts and economists produce the bulk of the macroeconomic predictions, both for public and internal uses, and many of them participate in periodic business outlook surveys. It might be argued that these are forecasts of people who study the economy (experts), which are quite unlike the expectations of those who act in the economy

---

<sup>7</sup>In short, rational market reactions may coexist with a large element of individual "irrationality." For an early argument along these lines, see Becker 1962.

(agents). On the one hand, the experts are usually credited with more knowledge of the economy at large than the agents have. On the other hand, the experts are often charged with being less strongly motivated to predict optimally than the agents who are seen as having more at stake.

In practice, the distinction between agents and experts is at this point very blurred. The forecasters who respond to business outlook surveys act and are treated as "experts" but they certainly are also "agents" in their own rights. Indeed, many of them are influential agents who have passed critical market tests, as certified by their positions and by the rewards their forecasts and advice earn them in the business world. It can be presumed that, in general, they do have incentives to perform well and strive to do so.

In my view, therefore, it is appropriate that the results of business outlook surveys have received alternative interpretations in the literature. They are treated either as agents' expectations, e.g., in tests of whether they conform to the hypotheses of rational or adaptive expectations, or as experts' forecasts, e.g., in comparisons with predictions from particular econometric models.<sup>8</sup>

#### Problems of Observation and Measurement

An ideal survey would use a large, properly constructed random sample to insure that the respondents represent well the universe of those whose expectations count, and a system of rewards and penalties to insure that they have a stake in their responses. In the absence of market expectations, it would then be interesting to test the quality and analyze the properties of predictions from such a survey, and perhaps particularly those of the composite or weighted average forecasts.

---

<sup>8</sup>For examples and further references, see Theil 1965; Mincer 1969; Mincer and Zarnowitz 1969; Zarnowitz 1972, 1974, and 1979; McNees 1975 and 1978; Nelson 1975; Carlson 1977; Wachtel 1977; Pearce 1979; Figlewski and Wachtel 1981.

Unfortunately, the ideal surveys do not exist and the actual surveys may be far from ideal. If a survey yields inferior or biased predictions, it is possible that carelessness, poor information, or other failings of particular respondents are to blame, which should not be generalized.

As already noted, past studies of expectations concentrated on the performance of simple averages (means or medians) of the participants' responses. But a series of averages from small samples whose composition varies over time may lack consistency. A few outliers may cause large errors and even bias the results. Moreover, the individual data will at times inevitably contain errors of reporting and transcription, some of which at least could be detected and eliminated. Thus the survey data need to be carefully edited and interpreted. Neglect of data problems explains why some of the survey evaluations yielded mixed and contradictory results, very limited in both scope and applicability.<sup>9</sup>

The identification of the agents, which is a matter of common sense in many cases (e.g., a survey of potential car buyers), presents problems with regard to comprehensive phenomena that affect virtually everyone but mean different things to different people. Thus individuals with widely divergent consumption patterns may have quite disparate perceptions of inflation, reflecting the shifts in relative prices of the different bundles of

---

<sup>9</sup> A case in point is a long series of surveys of economic forecasters, conducted semiannually since 1947 by Joseph A. Livingston, a syndicated financial columnist. The predictions for CPI included in this survey have been used in several recent tests of the "rationality" of inflationary expectations before it turned out that the data need to be rather extensively reworked (Carlson 1977). Some of the subsequent tests showed the average expectations derived from the revised survey data in a relatively favorable light, but others, including the most recent and persuasive ones, do not. For a summary of the earlier studies, see Pearce 1979; for other criticisms and comparisons with indicators and consumer expectations of inflation, see Moore 1977 and Juster 1979. More conclusive results, based on a summary of the behavior of individual expectations, are presented in Figlewski and Wachtel 1981.

commodities they evaluate. The problem is altered but not resolved by a stricter definition of the variable in question. Suppose that instead of asking about expected changes in the "general level of prices of things you buy," a poll referred explicitly to the changes in the official consumer price index (CPI). This might well be more confusing than helpful in a broadly based inquiry, since not many people are interested in, and well informed on, the specific and technical matters involved in the construction of this index. This is one reason why presumably informed views on the prospective changes in specific macroeconomic variables (e.g., CPI) are collected from professional economic forecasters. But, in the absence of ideal surveys, time series of expectations reported by different groups of agents and experts need to be analyzed and compared (allowing for the difficulties of such comparisons), to expand the coverage and improve the chances of detecting any interesting regularities of predictive behavior. Groups, as individuals, have their particularities.

#### New Departures

Owing to the efforts of the National Bureau of Economic Research, in collaboration with the American Statistical Association, a large amount of systematic, quantitative and qualitative information has been assembled on the record of forecasting changes in the U. S. economy since 1968. Each quarter, the NBER examines the results of a survey questionnaire mailed by the ASA.<sup>10</sup> The survey reaches a broadly based and diversified group of persons who are regularly engaged in the analysis of current and prospective business

---

<sup>10</sup>The quarterly reports on each survey prepared by the NBER present and discuss mainly the medians of the responses. These reports have been published regularly by the NBER, first in Explorations in Economic Research and later (since 1977) in the NBER Reporter, and by the ASA, first in the American Statistician and later (since 1974) in AmStat News. On the origin of the survey and the design of the questionnaire, see Moore 1969 and Zarnowitz 1969a.

conditions. Most of the respondents are from the world of corporate business and finance but academic institutions, government, consulting firms, trade associations, and labor unions are also represented. In the past, the numerical predictions from each survey have covered eleven important macroeconomic variables (in 1981 the coverage has been substantially extended). The target periods always include the current and the next four quarters.

The NBER-ASA surveys provide unique data on the methods and assumptions used by the respondents, and on the probabilities they attach to alternative prospects concerning changes in output and price levels. So far only the overall results have been used in a number of studies;<sup>11</sup> the rich detail is yet to be processed and explored.

A comprehensive study of the individual predictions from these surveys is currently under way. This paper, a progress report on the early and still incomplete results of this investigation, addresses several questions. How accurate are the individual expectations relative to the group average predictions? How representative are the latter of the former? What is the evidence concerning the frequencies and significance of bias and autocorrelated errors? How do the findings vary for different variables and predictive horizons?

Unlike most recent studies of expectations which consider only inflation forecasts, this paper compares predictions for six variables: growth rates in real and nominal income, two expenditure components of GNP, unemployment rate,

---

<sup>11</sup>Charts showing the median, upper quartile, and lower quartile ASA-NBER survey forecasts have been featured quarterly in articles of current interest published in Economic Prospects, a report by the Commercial Credit Company (1972-73), and in Economic Outlook USA, a report by the Survey Research Center at the University of Michigan (since 1974). Evaluations of the median series and, in some cases, of dispersion and some other aspects of the survey forecasts include Christ 1975; Fair 1974; McNees 1974, 1975, 1976, 1979; Moore 1977; Su and Su 1975; and Zarnowitz 1969b, 1971, 1974, and 1979.

quarterly time series of forecasts are analyzed, consisting of sets of five series per source (for five target quarters). Given the large scope of the study, only the summary measures of error for the whole period covered are presented in this paper to keep its size manageable. A sequel will deal with the variations over time, the cross-sectional (survey-by-survey) results, and disaggregation by method.

A serious limitation of most empirical studies in this area is that they refer mainly to the 1970s, a relatively short period and a particular one in several respects.<sup>12</sup> Efforts are being made to construct longer series of predictions comparable to those from the ASA-NBER surveys with the aid of valuable data from a large group of business economists organized into the New York Forecasters' Club. A small segment of this information is used in this paper.

## II. Individual vs. Group Mean Predictions

### Data and Measures

The body of data used in this paper consists of 42 consecutive surveys conducted quarterly from 1968:4 through 1979:1. Altogether, the list of those who replied to any of the NBER-ASA survey questionnaires includes 172 names (which are treated confidentially). However, many individuals responded only once or a few times, and some decision had to be made on the minimum number of

---

<sup>12</sup>For at least a part of the 1960s, some influential forecasts by government agencies and econometric model builders are available, but earlier data are sparse, not easy to access, and often in need of much careful processing; they are mainly informal predictions from business sources and surveys (Zarnowitz 1967, 1979).

surveys that would qualify a participant for inclusion. It was set at 12, which still left as many as 79 individuals in the sample.<sup>13</sup>

Four of the variables covered have strong upward trends, and it is not their levels that are of major interest but rather their rates of change which reflect real growth and/or inflation. These are gross national product and consumer expenditures for durable goods, both in current dollars (GNP and CEDG); GNP in constant dollars (RGNP); and the GNP implicit price deflator (IPD). For these series, forecast errors are measured as differences, predicted minus actual percentage change.

The change in business inventories (CBI), a current-dollar series, is trendless, being already in first-difference form. The unemployment rate (UR) represents the percentage unemployed of the civilian labor force and is dominated by short-term, mainly cyclical movements, not a long-term trend. For these two variables, therefore, forecast errors are measured as differences, predicted level minus actual level.<sup>14</sup>

The "actual" values are not well defined for many economic variables, such as GNP and components, which are subject to several, often sizable, revisions. In this paper, they are represented by the last data available prior to the benchmark revisions of January 1976 and December 1980. These are

---

<sup>13</sup>As noted earlier, a careful proofreading of the survey questionnaire is needed to detect simple mistakes of calculation, copying, and typing which chance or neglect will always occasion in some replies. The voluminous NBER-ASA materials were submitted to such an audit with the aid of the computer and, where needed, an inspection of the original submissions. Although the number of mistakes that were thus identified turned out to be very small relative to the mass of the data, failure to eliminate them would have affected adversely the evaluation of several individual records.

<sup>14</sup>See Zarnowitz 1967, pp. 32-35, and 1979, p. 6; and McNees 1973, pp. 7-10, on the definitions, measures, and merits of level and change errors.



presumably the "best" of those estimates that are conceptually comparable to the corresponding survey predictions.<sup>15</sup>

Let the level error be defined as

$$(11) \quad E_{t+j} = P_{t+j} - A_{t+j}, \quad j = 0, 1, \dots, 4,$$

where  $P$  and  $A$  refer to the predicted and actual levels, respectively. The surveys have generally been taken in the first half of each quarter, at a time when the most recent data available would be preliminary estimates for the preceeding quarter ( $A_{t-1}^*$ ).<sup>16</sup> Consequently, the  $P_t$  figures for the quarter during which the survey was taken ( $j = 0$ ) are authentic, if short, ex ante forecasts whose span is approximately one quarter. In addition, each survey produces predictions for the next four quarters extending into the future ( $j = 1, \dots, 4$ ).<sup>17</sup>

The errors of percentage change forecasts are

$$(12) \quad e_{t+j} = \begin{cases} \left( \frac{P_t - A_{t-1}^*}{A_{t-1}^*} - \frac{A_t - A_{t-1}}{A_{t-1}} \right) 100, & \text{if } j = 0 \\ \left( \frac{P_{t+j} - P_{t+j-1}}{P_{t+j-1}} - \frac{A_{t+j} - A_{t+j-1}}{A_{t+j-1}} \right) 100, & \text{if } j = 1, \dots, 4. \end{cases}$$

---

<sup>15</sup>This procedure imposes on the forecaster the burden to predict future revisions that are assumed to remove observational errors. An alternative is to compare the forecasts with provisional data that are closer to those that were available to the forecaster. The most informative approach is one that integrates the analysis of data errors and of prediction errors, which would be a good task for another paper. On the role of revisions in economic measurement and prediction, see Cole 1969 and Zarnowitz 1979 and 1982.

<sup>16</sup>An exception is the unemployment rate series which is available monthly.

<sup>17</sup>All this applies to predictions of any  $i$ -th respondent for any  $m$ -th variable, so for simplicity all subscripts other than those referring to the target periods are dropped from the formulae in this and the following paragraph of the text. A subscript for the time when the forecast was made is also redundant, since it is always  $t$  (see also equation 12 and text below).

Thus  $e_t$  contains an error due to the discrepancy between  $A_{t-1}^*$  and  $A_{t-1}$ . This difference is a measurement error, but it is usually akin to a very short forecast error, since  $A_{t-1}^*$  is in most cases an extrapolation based on incomplete data. For  $j > 0$ , the base of each percentage change forecast is itself a forecast, namely that of the level in the preceeding quarter ( $P_{t+j-1}$ ). The differences between the successive levels predicted in a chain of forecasts made at time  $t$ ,  $P_{t+j} - P_{t+j-1}$ , are implicit predictions of changes over the successive subperiods covered. Note that each of these marginal or "intraforecast" change forecasts covers a single quarterly interval, so the target periods do not overlap.<sup>18</sup>

#### Comparisons with Group Averages

Chart 1 shows the distributions of ratios of root mean square errors,  $M_i/M_{gi}$ , where  $M_i$  refers to the expectations by the  $i$ -th individual and  $M_{gi}$  to the corresponding group average, i.e., the mean of all the individual predictions that covered the same period as that included in  $M_i$ .<sup>19</sup> There are 30 distributions, one for each of the variables and target quarters (Q0, . . . , Q4) covered, and every one of them is skewed to the right. Thus, only minorities of the individuals had ratios of less than 1, that is, outperformed the group

---

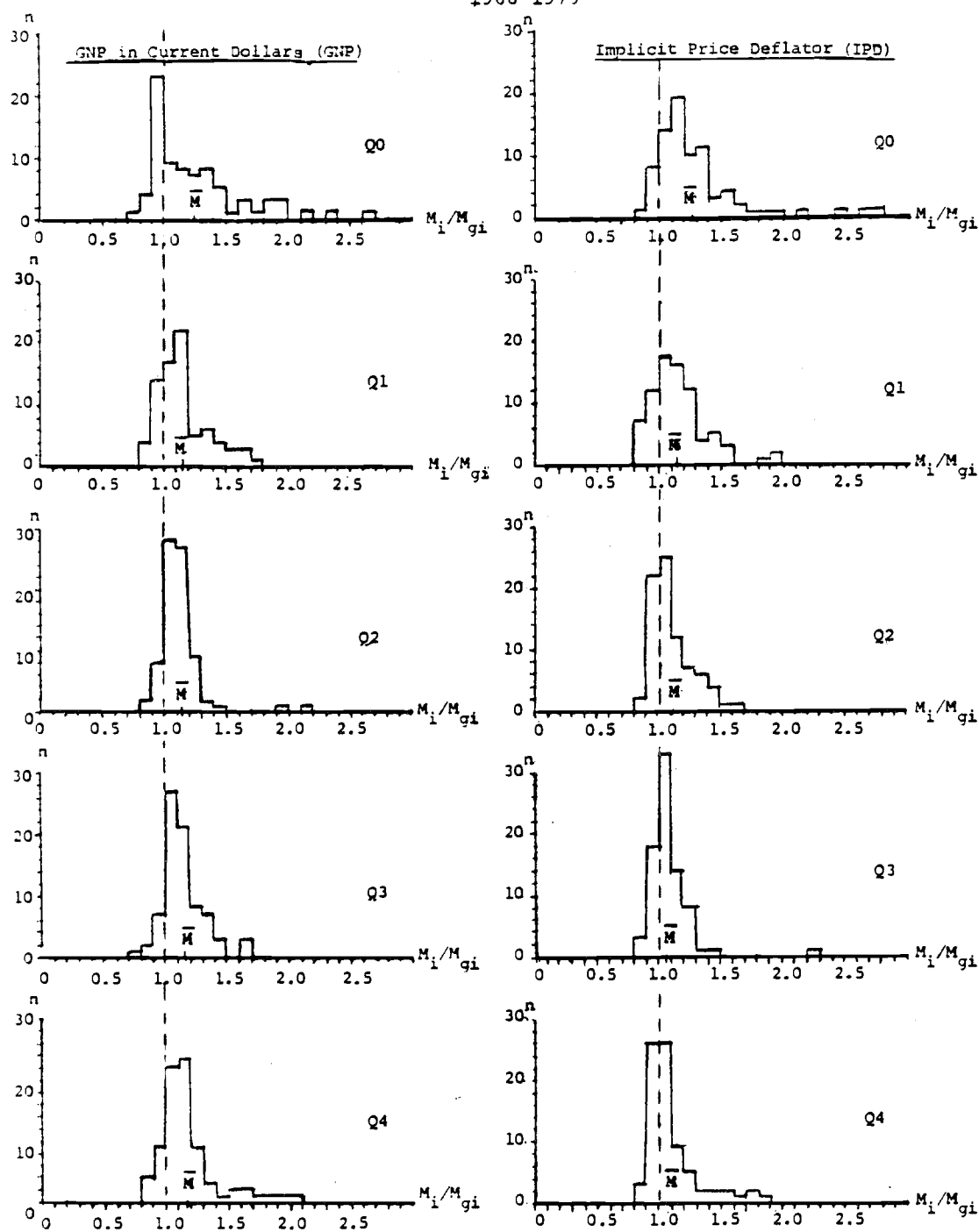
<sup>18</sup>The targets are changes over successive quarters (0-1, 1-2, . . .). In contrast, forecasts of average changes over increasing spans (0-1, 0-2, . . .) involve overlapping target periods and their errors are therefore necessarily intercorrelated. See Zarnowitz 1967, pp. 64-70.

<sup>19</sup>For level forecasts (UR and CBI),  $M_i = \sqrt{\frac{1}{n} \sum_t^n (E_{it} - \bar{E}_i)^2}$ , for percentage change forecasts,  $M_i = \sqrt{\frac{1}{n} \sum_t^n (e_{it} - \bar{e}_{it})^2}$ . The  $M_{gi}$  measures are computed in the same way from time series of errors of the corresponding group mean forecasts.

CHART 1

SEVENTY-NINE INDIVIDUAL FORECASTS OF MULTIPERIOD CHANGES IN  
SIX AGGREGATE VARIABLES, COMPARISONS WITH GROUP MEAN FORECASTS,

1968-1979



NOTE: In each graph, the ratio of root mean square errors ( $M_i/M_{gi}$ ) is plotted horizontally, the number of forecasters in each class (n) is plotted vertically. Total n equals 79 in each case, except for CEDG (75). The number of surveys covered is 42 for Q0, 41 for Q1, 40 for Q2, 39 for Q3, and 33 for Q4. See text, equations 6 and 7 and note 15, for the definitions of the symbols and measures used. The points labeled  $\bar{M}$  on the horizontal axes locate the mean  $M_i/M_{gi}$  ratio in each case.

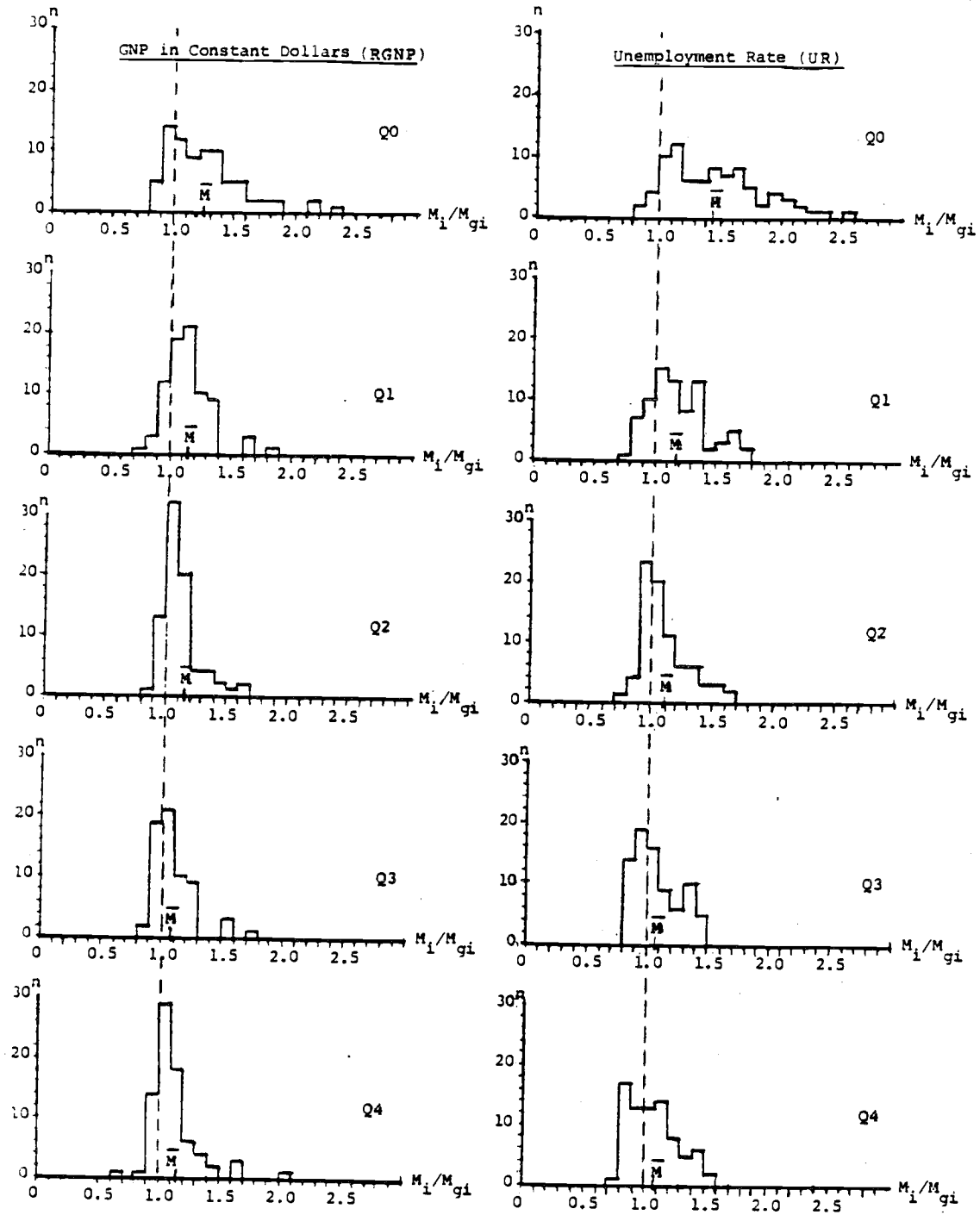
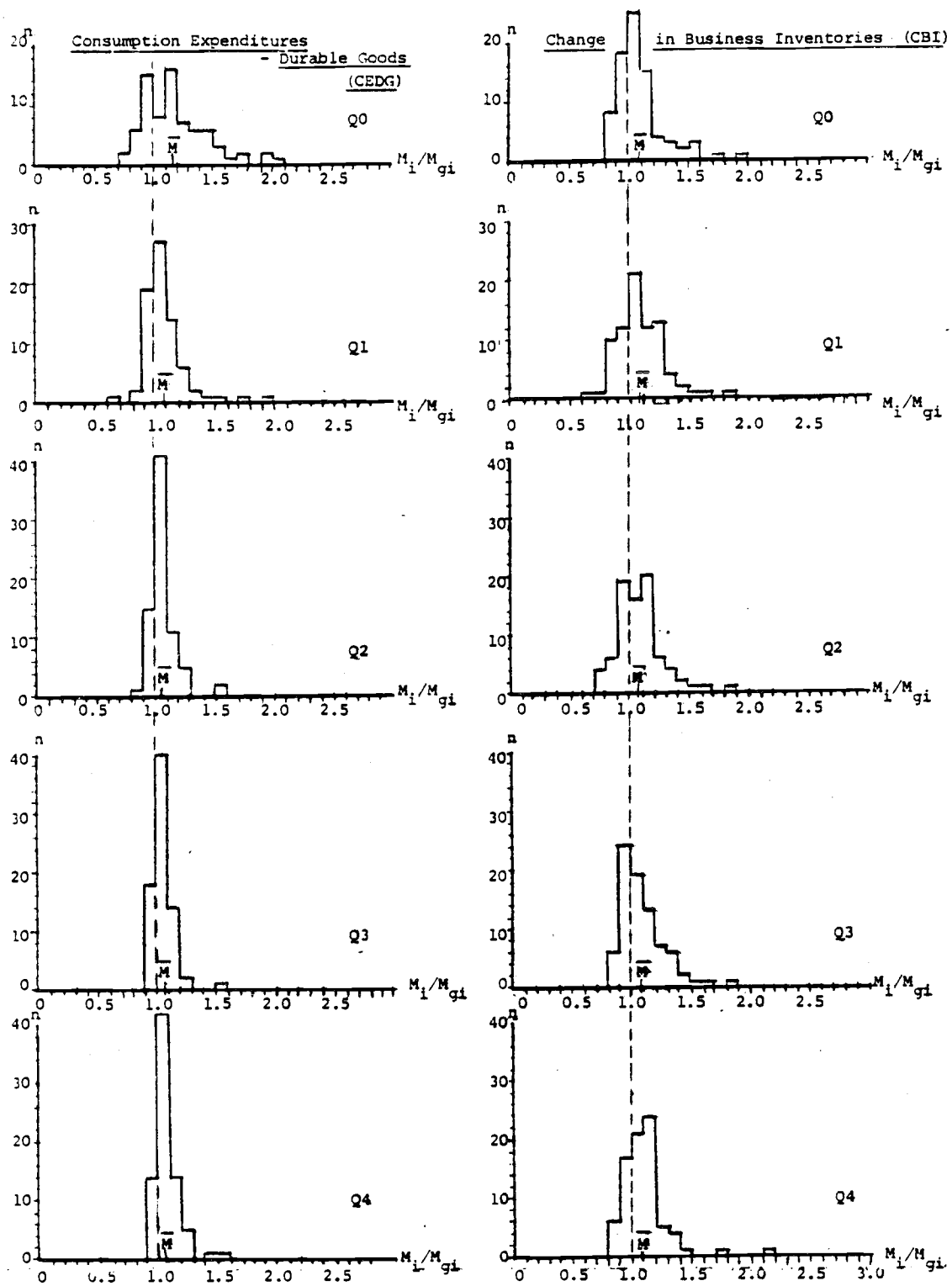
CHART 1  
(Continued)

CHART 1  
(Concluded)

averages over time. The best (lowest) ratios fall between 0.7 and 0.9, the worst (highest) between 1.4 and 2.2. The means of the ratios (marked  $\bar{M}$ ) are all located to the right of the unity (broken vertical) lines. The histograms seem to get tighter and also, often, less skewed for the more distant quarters (i.e., as one moves down the chart, from Q0 to Q4, for each variable).

Table 1 shows that the mean ratios are remarkably close: when rounded, all but nine of the 30 statistics are 1.1. The higher mean ratios, ranging from 1.2 to 1.4, refer to the shortest predictions, for Q0 and, less so, for Q1. The standard deviations of the  $M_i/M_{qi}$  ratios tend to decrease strongly with the distance to the target quarter, from Q0 to Q3.<sup>20</sup> An exception is CBI, where the horizon of the expectations apparently does not matter much (all the means are approximately 1.1 and the decline in the dispersion of the ratios is very small).

The proportions of the better-than-average forecasters ( $M_i/M_{qi} < 1$ ) vary strongly with the target quarter for some variables, much less so for others. Thus for UR the range is 8 to 42 percent, for CBI it is only 29 to 38 percent. Averaged across Q0-Q4, the figures fall between 20 percent for GNP and 33 percent for CBI (see the last section of Table 1).

As will be shown below, the average accuracy of forecasts varies considerably across the individuals, variables, and target periods. Highly volatile series such as CEDG and CBI are much more difficult to predict than relatively smooth, trend-dominated series such as GNP. In general, the uncertainty and difficulty (hence errors) of prediction tend to increase for the more distant future. The remarkable degree of standardization in the

---

<sup>20</sup>It should be noted that these comparisons are somewhat impaired by the fact that the measures for Q0, Q1, Q2, Q3, and Q4 refer to 42, 41, 40, 39, and 34 surveys, respectively. In particular, the relatively large figures for Q4 compared to those for Q3 probably reflect the drop in survey coverage.

TABLE 1  
MEANS AND STANDARD DEVIATIONS OF THE  $M_i/M_{gi}$  RATIOS,  
BY VARIABLE AND TARGET QUARTER, 1968-1979

	GNP	IPD	RGNP	UR	CEDG	CBI
	(1)	(2)	(3)	(4)	(5)	(6)
<b>MEANS</b>						
Q0	1.25	1.29	1.25	1.43	1.18	1.10
Q1	1.16	1.16	1.14	1.19	1.10	1.11
Q2	1.13	1.11	1.11	1.11	1.07	1.09
Q3	1.14	1.08	1.09	1.08	1.06	1.10
Q4	1.17	1.10	1.12	1.08	1.08	1.10
<b>STANDARD DEVIATIONS</b>						
Q0	.38	.36	.31	.36	.28	.20
Q1	.21	.23	.18	.24	.18	.21
Q2	.18	.16	.15	.19	.11	.19
Q3	.17	.18	.15	.18	.09	.18
Q4	.24	.20	.19	.19	.10	.19
<b>PERCENTAGE OF CASES WHERE <math>M_i/M_{gi} &lt; 1</math></b>						
Average	20	26	22	29	24	33
Range	13-35	11-37	18-27	8-42	19-31	29-38

Note: Based on quarterly ASA-NBER business outlook surveys 1968:4-1979:1. On coverage and symbols used, see text and Chart 1.

$M_i/M_{gi}$  ratios stands in sharp contrast to the diversity of the average accuracy measures for the individuals,  $M_i$ .

The advantage of the group means  $M_{gi}$  is the greatest for the nearest targets and it becomes less and less important as the predictions reach out further into the future. One may speculate that the individual forecasts for Q0 and Q1 contain more independent information than those for Q2-Q4, hence the gains from averaging are larger for the former than for the latter.<sup>21</sup> The abilities to predict CBI are particularly limited, even for the nearest quarters, so here the means and dispersion of the ratios  $M_i/M_{gi}$  depend little on the distance to the target quarter ( $j = 0, 1, \dots, 4$ ).

Similar results are obtained from another, earlier survey. Chart 2 shows the distributions of  $M_i/M_{gi}$  ratios for six-month and twelve-month forecasts of industrial production in 1947-63. The data come from those members of the New York Forecasters' Club who participated in at least five surveys. There are separate comparisons for the predictions of levels six and twelve months after the date of the survey ("0") and of absolute changes during the first and second six-month periods and the twelve-month period ahead (0-6, 6-12, and 0-12). The distributions are strongly skewed to the right, with most of the ratios falling between 1.0 and 1.4 and the classes below 0.8 almost

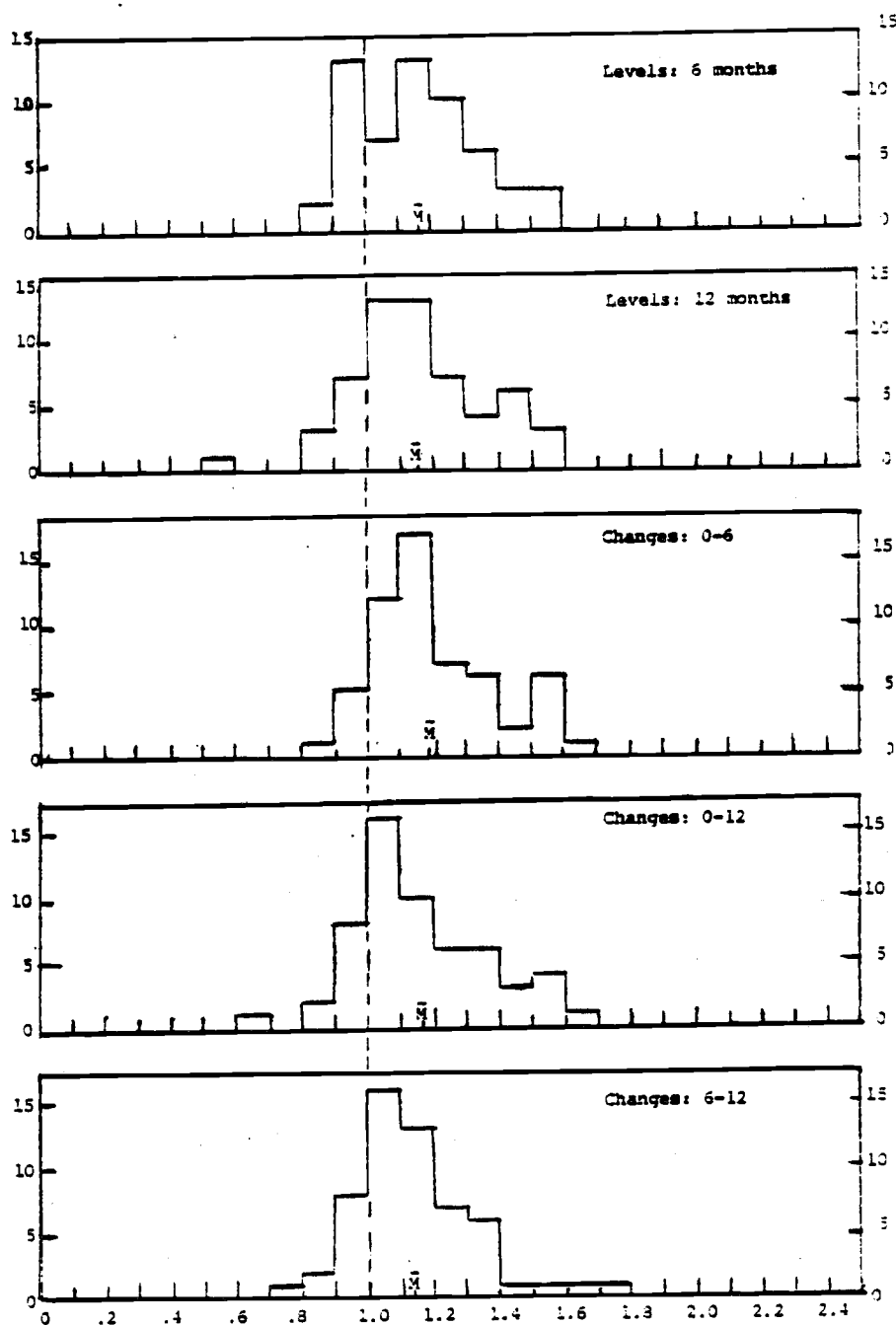
---

<sup>21</sup>The large means and standard deviations of the ratios for Q0 may be associated with the disparities in the quality of the current data available to different individuals. Although the survey questionnaire provides the most recent information on the values of the series to be predicted, some respondents choose to use different jump-off levels which may be more or less accurate. It is not quite clear why the figures for the shortest predictions of UR should be particularly high, as Table 1, column 4, shows them to be, but it is suggestive that this is the only variable covered for which monthly data are available. Some individuals are likely to lag behind the majority in absorbing these monthly data (and related weekly information on unemployment claims).



## CHART 1

Fifty-Seven Individual Forecasts of Levels of Changes of Industrial Production over Spans of Six and Twelve Months, Comparisons with Group Mean Forecasts, 1947-63.



NOTE: In each panel, the vertical scale represents the number of forecasters. The horizontal scale represents the ratios of root mean square errors for the individual forecasts and the corresponding group mean forecasts,  $M_i/M_{gi}$  (see text for further explanation). "0-6" refers to the change in the first six-month interval following the survey date; "6-12" to the second six-month interval; and "0-12" to the twelve-month interval. The points labeled " $\bar{M}$ " on the horizontal axes locate the mean  $M_i/M_{gi}$  ratio in each case.

empty. The average ratios ( $\bar{M}$ ) are all concentrated in the narrow range 1.1-1.2.<sup>22</sup>

These findings indicate that it is difficult for most individuals to predict consistently better than the group. Contemporaneous expectations for a given target may be distributed more or less symmetrically about their mean, but over time the individuals' positions within these distributions are likely to fluctuate. For most people, most of the time, the predictive record is spotty, with but transitory spells of relatively high accuracy. A series of group averages has the advantage that it is helped by the cancellation of individual errors of opposite sign.

#### Rank Tests of Predictive Consistency

Success in one class of predictions (say, for GNP in Q1) may or may not coincide with success in another class (say, GNP in Q4, or for IPD). If the degree of coincidence were very low (e.g., if very few people managed to "beat" the group mean in more than one class), then the success, being rather isolated, might be attributable more to chance than to better techniques or skills.

The NBER-ASA survey participants have been ranked according to the  $M_i/M_{gi}$  ratios for each of the variables and target quarters covered in Chart 2. The correlations among the resulting ranks could be either close to zero (indicating very little consistency in the relative performance of the forecasters across different variables or predictive spans) or significantly negative (those who succeed in one category tend to fail in another) or significantly positive (those who succeed in one category also tend to succeed in others).

---

<sup>22</sup>GNP forecasts made by members of the same group in the period 1956-63 (Zarnowitz 1967, pp. 123-126) tell much the same story. The number and identity of those who did better than the group mean varied for predictions with different spans and other characteristics, but the average success rate in these terms was no more than 19 percent.

The rank correlations are presented in Table 2, both across the variables for each target quarter (part I) and across target quarters for each variable (part II). All the correlations are positive and in general they appear to be significantly so (see note in the table). Thus there is some degree of consistency in the predictive performance of the individuals as revealed by their  $M_i/M_{gi}$  ranks.

People who predict relatively well the rates of change in nominal GNP also tend to do so for the rates of change in real GNP: the average rank correlation coefficient  $\bar{\rho}$  is 0.74 in this case. For variables that are not so closely related, the correlations are much lower (e.g.,  $\bar{\rho} = 0.23$  for CEDG and IPD, and also for CEDG and UR). However, only 15 of the 75 coefficients ( $\rho \neq 1$ ) in part I of the table are less than 0.2. The overall mean of the  $\rho$  statistics is 0.36.

For any of the variables, people who rank high (low) in predicting one quarter also tend to rank high (low) in predicting the next quarter. The  $\rho$ 's for Q0-Q1 average 0.61, those for Q1-Q2, Q2-Q3, and Q3-Q4 average 0.52-0.55 (see part II of Table 2). For non-adjointing target periods, the rank correlations are lower,  $\bar{\rho}$  being 0.40 where the distance is two quarters (Q0-Q2, Q1-Q3, and Q2-Q4) and 0.31 where it is three quarters (Q0-Q3 and Q1-Q4). The further apart the target periods, the less correlated are the values to be predicted, and the above results suggest that the ranking consistency declines correspondingly. But the reductions in the rank correlations vary considerably in size and regularity, being most pronounced for CEDG, least for GNP. When averaged over the quarters Q1-Q4, the  $\bar{\rho}$  coefficients are relatively low for CEDG, GNP, and RGNP (.27-.33) and high for IPD, UR, and CBI (.55-.66).

TABLE 2

RANK CORRELATIONS AMONG PARTICIPANTS IN ASA-NBER SURVEYS  
 ACCORDING TO RATIOS OF INDIVIDUAL TO GROUP ROOT MEAN  
 SQUARE ERRORS, MULTIPERIOD PREDICTIONS FOR SIX AGGREGATE  
 VARIABLES, 1968-1979

I. Across Variables, for Each Target Quarter <sup>a</sup>													
Q0							Q1						
	GNP	IPD	RGNP	UR	CEDG	CBI		GNP	IPD	RGNP	UR	CEDG	CBI
GNP	1.00						GNP	1.00					
IPD	.57	1.00					IPD	.23	1.00				
RGNP	.83	.65	1.00				RGNP	.77	.48	1.00			
UR	.42	.43	.39	1.00			UR	.20	.30	.40	1.00		
CEDG	.69	.56	.69	.50	1.00		CEDG	.43	.20	.50	.30	1.00	
CBI	.40	.41	.36	.21	.42	1.00	CBI	.39	.31	.38	.22	.38	1.00
Q2							Q3						
	GNP	IPD	RGNP	UR	CEDG	CBI		GNP	IPD	RGNP	UR	CEDG	CBI
GNP	1.00						GNP	1.00					
IPD	.35	1.00					IPD	.27	1.00				
RGNP	.66	.48	1.00				RGNP	.69	.33	1.00			
UR	.23	.48	.27	1.00			UR	.41	.44	.49	1.00		
CEDG	.27	.15	.21	.12	1.00		CEDG	.21	.05	.14	.19	1.00	
CBI	.41	.43	.32	.36	.31	1.00	CBI	.28	.42	.15	.37	.03	1.00
Q4							Average, Q0-Q4						
	GNP	IPD	RGNP	UR	CEDG	CBI		GNP	IPD	RGNP	UR	CEDG	CBI
GNP	1.00						GNP	1.00					
IPD	.44	1.00					IPD	.37	1.00				
RGNP	.76	.51	1.00				RGNP	.74	.49	1.00			
UR	.39	.49	.31	1.00			UR	.33	.45	.35	1.00		
CEDG	.18	.17	.17	.06	1.00		CEDG	.36	.23	.34	.23	1.00	
CBI	.49	.36	.38	.19	.21	1.00	CBI	.32	.39	.32	.27	.27	1.00

<sup>a</sup> These measures refer to 75 individuals who participated in at least 12 quarterly ASA-NBER business outlook surveys 1968:4-1979:1 and predicted all six variables covered. The symbols for the variables are identified in part II of the table. The symbols Q0-Q4 refer to the current and the four successive future quarters. The rank correlation coefficients shown are Spearman's  $\rho = 1 - [6\sum d^2 / (n^3 - n)]$  where  $d$  is the rank difference and  $n$  is the number in each ranking.

TABLE 2  
(concluded)II. Across Target Quarters, for Each Variable<sup>b</sup>

GNP in Current Dollars (GNP)						Implicit Price Deflator (IPD)					
	Q0	Q1	Q2	Q3	Q4		Q0	Q1	Q2	Q3	Q4
Q0	1.00					Q0	1.00				
Q1	.51	1.00				Q1	.55	1.00			
Q2	.18	.19	1.00			Q2	.45	.68	1.00		
Q3	.14	.18	.40	1.00		Q3	.41	.54	.60	1.00	
Q4	.50	.32	.40	.47	1.00	Q4	.39	.52	.51	.62	1.00

GNP in Constant Dollars (RGNP)						Unemployment Rate (UR)					
	Q0	Q1	Q2	Q3	Q4		Q0	Q1	Q2	Q3	Q4
Q0	1.00					Q0	1.00				
Q1	.57	1.00				Q1	.64	1.00			
Q2	.33	.47	1.00			Q2	.38	.78	1.00		
Q3	.05	.19	.48	1.00		Q3	.32	.62	.85	1.00	
Q4	.38	.19	.33	.33	1.00	Q4	.27	.53	.75	.92	1.00

Consumer Expenditures-Durable Goods (CEDG)						Change in Business Inventories (CBI)					
	Q0	Q1	Q2	Q3	Q4		Q0	Q1	Q2	Q3	Q4
Q0	1.00					Q0	1.00				
Q1	.66	1.00				Q1	.70	1.00			
Q2	.53	.43	1.00			Q2	.63	.76	1.00		
Q3	.12	.07	.14	1.00		Q3	.51	.56	.81	1.00	
Q4	.14	.00	.04	.08	1.00	Q4	.57	.62	.63	.76	1.00

<sup>b</sup> These measures refer to the sample covered in Chart 2: 79 individuals for each of the variables except CEDG (80). The rank correlation coefficients are Spearman's  $\rho$ .

NOTE: For rankings without ties, the variance of  $\rho$  equal's  $\frac{n}{n-1}$  (Kendall 1948, p. 46). For  $n = 75$ , therefore, the standard error  $S_{\rho} = 1/\sqrt{0.74} = 0.1162$  (for  $n = 79$ ,  $S_{\rho} = 0.1125$ ; for  $n = 80$ ,  $S_{\rho} = 0.1132$ ). Hence, all entries  $\rho > 0.23$  in the table are significant at the 5% level, and all  $\rho \geq 0.20$  at the 10% level.

### III. Assessing the Forecasters' Performance

Because of discrepancies in time coverage, absolute accuracy measures are not strictly comparable across the individual participants in the surveys. (This was one of the reasons for the standardization by means of the relative accuracy measures  $M_i/M_{gi}$ .) However, we are dealing with numerous responses to a relatively large number of surveys, and there is no apparent reason for any significant bias due to missed observations.<sup>23</sup> With some caution, therefore, it should be instructive to examine the distributions of the statistics that sum up the records of the individuals. The corresponding measures for the group mean forecasts provide some further interesting comparisons.

#### Overall Accuracy

Table 3 shows the distributional statistics for the root mean square errors of the individuals (columns 1-5). With virtually no exceptions, the averages of the individual RMSE's are larger than the RSME's for the corresponding group mean forecasts (compare line by line the entries in columns 1 and 4 with their counterparts in column 6). In most cases, the medians are somewhat smaller than the means, which indicates some skewness to the right, that is, toward large RSME's. The measures for the group mean tend to be closer to the lower quartile than to the median of the distribution of the individual RMSE's (cf. columns 3, 4, and 6).

The more distant the target quarter, the larger tend to be the prediction errors, as demonstrated by the increases from Q0 through Q4 of the entries in

---

<sup>23</sup>Each of the 42 surveys has an adequate coverage; on the average, 43 participants with a standard deviation of 9. Each of the 79 individuals responded to at least 12 surveys; the mean is 23 with a standard deviation of 8. The distribution of the individuals among the periods covered appears to be dominated by random choice.

28  
TABLE 3

ROOT MEAN SQUARE ERRORS OF 79 INDIVIDUAL FORECASTS OF SIX  
AGGREGATE VARIABLES, MEASURES OF CENTRAL TENDENCY AND DISPERSION,  
1968-1979

Quarter Predicted <sup>a</sup>	Distributional Statistics for the Individual RMSE's <sup>b</sup>					Group Mean Forecast RMSE	
	Mean	Standard Deviation	Lower Quartile	Median	Upper Quartile		
	(1)	(2)	(3)	(4)	(5)		
		GNP in Current Dollars (GNP)					(6)
Q0	.77	.27	.58	.68	.92	.66	
Q1	.95	.26	.78	.92	1.05	.86	
Q2	1.06	.19	.96	1.07	1.15	.96	
Q3	1.10	.26	.92	1.10	1.23	.98	
Q4	1.12	.28	.94	1.08	1.22	.94	
		Implicit Price Deflator					
Q0	.55	.16	.45	.49	.60	.42	
Q1	.69	.16	.58	.66	.77	.59	
Q2	.79	.16	.69	.78	.87	.70	
Q3	.88	.19	.78	.86	.95	.77	
Q4	.98	.21	.86	.94	1.09	.88	
		GNP in Constant Dollars (RGNP)					
Q0	.85	.28	.67	.78	.96	.70	
Q1	1.09	.28	.91	1.03	1.26	.95	
Q2	1.24	.22	1.13	1.25	1.37	1.12	
Q3	1.39	.25	1.22	1.36	1.53	1.23	
Q4	1.46	.31	1.25	1.39	1.69	1.23	
		Unemployment Rate					
Q0	.22	.06	.17	.21	.26	.16	
Q1	.46	.11	.38	.44	.51	.41	
Q2	.71	.17	.60	.67	.81	.65	
Q3	.94	.23	.78	.91	1.09	.88	
Q4	1.04	.24	.88	1.00	1.19	.98	
		Consumer Expenditures--Durable Goods (CEDG)					
Q0	3.37	.96	2.64	3.10	3.85	2.87	
Q1	4.16	.78	3.79	4.13	4.60	3.77	
Q2	4.24	.71	3.82	4.22	4.68	4.04	
Q3	4.44	.65	4.12	4.41	4.78	4.09	
Q4	3.98	.69	3.47	3.99	4.47	3.64	
		Change in Business Inventories (CBI)					
Q0	8.21	2.65	6.70	8.10	9.61	8.07	
Q1	9.17	3.06	7.06	8.89	11.86	9.11	
Q2	10.42	3.36	8.22	10.08	12.87	9.79	
Q3	10.99	3.32	9.16	11.12	13.10	10.08	
Q4	11.22	3.00	9.13	11.14	13.12	10.80	

Notes to Table 3

<sup>a</sup>Q0 denotes the current quarter, Q1 the following (first future) quarter, etc. The number of the surveys covered is 42 for Q0, 41 for Q1, 40 for Q2, 39 for Q3, and 33 for Q4.

<sup>b</sup>These measures refer to the sample covered in Chart 1 (75 individuals forecast CEDG, 79 each of the other variables). See text, eqs. 6 and 7 and note 15, for the definitions of the symbols and measures used (RSME = root mean square error).

SOURCE: Quarterly ASA-NBER business outlook surveys; 1968:4-1979:1. Minimum number of surveys covered by any individual is 12.

columns 1 and 3-6. However, the increases tend to taper off: the individuals on the average predict Q0 substantially better than Q1, and Q1 still noticeably better than Q2, but their ability to anticipate Q3 is not much less limited than their ability to anticipate Q2, and the same applies even more to Q4 vs. Q3. In short, these measures suggest that the RMSE's tend to approach asymptotically a high plateau at the more distant target quarters. Note that these results apply to the marginal prediction errors for each successive quarter (i.e., changes 0-1, 1-2, . . .). To the extent that such errors are positively correlated, their cumulation will produce much greater increases in the average prediction errors for changes over increasing, overlapping spans (0-1, 0-2, . . .).<sup>24</sup>

The RMSE's for CEDG and CBI are particularly large and their Q0-Q4 differences are relatively small and irregular; those for UR, in contrast, are very low for Q0 and display relatively large and regular increases for the successively more distant quarters. Measures of dispersion for these distributions (standard deviations in column 2 and interquartile ranges implied by columns 3 and 5) increase but weekly and irregularly with the index  $j$  for most variables.

#### Correlations of Predicted with Actual Values

Table 4 shows that the correlations between predictions and realizations decline strongly and consistently with the lengthening horizon. The squared

---

<sup>24</sup>The buildup of average prediction errors with increasing spans is a general phenomenon to be expected and is well documented in forecast evaluations. However, some evidence for earlier periods has shown marginal errors varying narrowly and irregularly over the range of several quarters ahead, without any systematic upward drift (Zarnowitz 1967, pp. 64-72, and 1979, pp. 18-19; McNees 1973, pp. 24-25). The present results may differ because of the nature of the period covered (and Zarnowitz 1979 provides some support for this hypothesis), but they also inspire more confidence than those of other studies, being based on much larger samples of better controlled data.



TABLE 4

**CORRELATIONS BETWEEN PREDICTED AND ACTUAL VALUES OF SIX AGGREGATE VARIABLES,  
79 INDIVIDUAL FORECASTS, SELECTED DISTRIBUTIONAL STATISTICS, 1968-1979**

Quarter Predicted	Coefficients of Squared Correlation for the Individuals					Group Mean Forecast $\bar{r}^2$
	Percentage equal to or exceeding					
	<u>0.1</u> (1)	<u>0.3</u> (2)	<u>0.5</u> (3)	<u>Mean</u> (4)	<u>Standard Deviation</u> (5)	
<b>GNP in Current Dollars (GNP)</b>						
Q0	91	77	56	.49	.22	.63
Q1	76	34	9	.24	.16	.31
Q2	53	8	1	.13	.12	.18
Q3	39	11	4	.12	.13	.16
Q4	34	13	1	.12	.15	.18
<b>Implicit Price Deflator (IPD)</b>						
Q0	92	78	48	.45	.20	.64
Q1	80	43	13	.28	.18	.35
Q2	54	11	1	.14	.12	.17
Q3	32	5	0	.10	.10	.12
Q4	37	5	0	.10	.10	.08
<b>GNP in Constant Dollars (RGNP)</b>						
Q0	99	92	75	.60	.18	.75
Q1	89	72	20	.38	.17	.48
Q2	70	15	4	.18	.13	.25
Q3	37	6	1	.10	.11	.10
Q4	38	11	3	.12	.15	.16
<b>Unemployment Rate (UR)</b>						
Q0	100	100	100	.97	.02	.99
Q1	100	100	100	.86	.06	.91
Q2	100	100	90	.68	.12	.75
Q3	99	81	53	.48	.17	.53
Q4	85	56	18	.32	.19	.27
<b>Consumer Expenditures--Durable Goods (CEDG)</b>						
Q0	84	61	39	.40	.25	.63
Q1	35	13	3	.11	.15	.13
Q2	16	3	0	.05	.07	.01
Q3	17	1	0	.05	.07	.01
Q4	32	1	0	.08	.08	.13
<b>Change in Business Inventories (CBI)</b>						
Q0	84	61	26	.36	.21	.55
Q1	76	51	21	.31	.22	.51
Q2	71	39	12	.25	.19	.41
Q3	62	26	9	.20	.18	.40
Q4	66	30	8	.21	.17	.35

NOTE: The correlations are between the corresponding series of actual and predicted percentage changes for GNP, IPD, RGNP, and CEDG, and between the actual and predicted levels for UR and CBI.  $\bar{r}^2$  denotes the squared coefficient of correlation, corrected for the degrees of freedom. For source, see Table 3. Minimum number of surveys covered by any individual is 12. Number of individuals covered is 79, except for CEDG (75).

correlation coefficients average about .4 and .6 for Q0 and .1 to .2 for Q4, except for the unemployment rate where they are much higher, exceeding .9 for Q0 and .3 for Q4 (column 4). This reflects the fact that UR is a relatively smooth series, since levels rather than changes are used in this case. The dispersion of the correlation coefficients across the individuals declines as the distance to the target quarter increases, again except for UR, where the opposite happens (column 5; cf. also columns 1-3).

The  $\bar{r}^2$  coefficients for the group mean forecast are consistently, and often substantially, higher than the averages of the squared correlations for the individuals (compare columns 4 and 6). They, too, decline sharply and regularly for Q0 through Q4, the drop being most pronounced for IPD, least for CBI (column 6).

#### Mean Errors

Table 5 shows that almost all survey members underestimated inflation, i.e., had negative mean errors in their expectations of the rates of change in IPD during the period 1968-79 (columns 1-2). On the average, these statistics increase strongly with the predictive horizon, from Q0-Q4 (columns 3-4). In contrast, real growth as measured by the rates of change in RGNP was predominantly overestimated in this period, increasingly so for the more distant future. The underestimates of the price component and the overestimates of the quantity component tend to cancel each other in the predictions of rates of change in current-dollar GNP; most of the mean errors are here negative but very small (the overall averages are close to zero).<sup>25</sup> The signs of the

---

<sup>25</sup>Negative correlations between errors in predicting real growth and inflation have long been observed (Zarnowitz 1979; pp. 15-16). They are not necessarily indicative of poor forecasts but are difficult to reconcile with the positive short-term effects on output of unanticipated inflation as implied by the recent model of an "expectations-augmented Phillips curve."

TABLE 5

DISTRIBUTIONS OF MEAN ERRORS IN 79 INDIVIDUAL FORECASTS  
FORECASTS OF SIX AGGREGATE VARIABLES, 1968-1979

Quarter Forecast Predicted	Mean Errors for the Individuals				Group Mean	
	Under-	Over-	Standard		Mean	Standard Devia-
	estimates, percent (1)	estimates, percent (2)	Mean (3)	Deviation (4)	Error (5)	tion of Errors (6)
<b>GNP in Current Dollars (GNP)</b>						
Q0	71	29	-.12	.21	-.12	.66
Q1	63	37	-.07	.20	-.08	.87
Q2	76	24	-.13	.19	-.14	.96
Q3	73	27	-.13	.21	-.16	.98
Q4	62	38	-.08	.29	-.14	.94
<b>Implicit Price Deflator (IPD)</b>						
Q0	89	11	-.16	.14	-.16	.39
Q1	96	4	-.30	.17	-.30	.52
Q2	98	2	-.39	.18	-.39	.59
Q3	99	1	-.49	.17	-.47	.62
Q4	99	1	-.61	.21	-.59	.66
<b>GNP in Constant Dollars (RGNP)</b>						
Q0	34	66	.04	.24	.04	.70
Q1	11	89	.23	.22	.22	.94
Q2	14	86	.26	.23	.25	1.11
Q3	10	90	.35	.25	.30	1.21
Q4	2	98	.53	.31	.45	1.17
<b>Unemployment Rate (UR)</b>						
Q0	14	86	.04	.05	.05	.16
Q1	47	53	-.01	.11	.01	.41
Q2	80	20	-.12	.17	-.09	.65
Q3	92	8	-.29	.23	-.23	.85
Q4	86	14	-.32	.27	-.24	.96
<b>Consumer Expenditures--Durable Goods (CEDG)</b>						
Q0	91	9	-.92	.67	-.94	2.74
Q1	64	36	-.36	.70	-.39	3.80
Q2	65	35	-.27	.77	-.27	4.08
Q3	52	48	.03	.76	-.23	4.14
Q4	41	59	.14	.80	.02	3.70
<b>Change in Business Inventories (CBI)</b>						
Q0	95	5	-2.40	1.95	-2.36	7.81
Q1	85	15	-1.88	2.28	-1.96	9.01
Q2	71	29	-1.39	2.82	-1.62	9.78
Q3	69	31	-1.10	3.10	-1.20	10.13
Q4	73	27	-1.85	2.80	-1.70	10.83

NOTE: The mean errors of percentage change forecasts  $\bar{e}$  are used for GNP, IPD, RGNP, and CEDG; the mean errors of level forecasts  $\bar{E}$  are used for UR and CBI. Underestimation is indicated when  $\bar{e} < 0$  or  $\bar{E} < 0$ ; overestimation when  $\bar{e} > 0$  or  $\bar{E} > 0$  (for any individual and target quarter). See text and eqs. and above for detail. Source and average as in Tables 3 and 4.

errors are mixed for the forecasts of the unemployment rate in Q0 and Q1, while for Q2-Q4 underpredictions prevail, which is consistent with the overprediction of real growth. The rates of change in spending on consumer durables are underestimated in the short expectations but not in the longer ones. Negative mean errors prevail in the forecasts of business inventory investment.

The series of group expectations have, of course, the same mean errors as the aggregates of individual expectations, except for minor discrepancies due to rounding (cf. columns 3 and 5), but the standard deviations of their errors are about 3 to 4 times higher than the corresponding statistics for the individuals (cf. columns 4 and 6). Further, the standard deviations of the group mean forecast errors show a strong and general tendency to rise with the distance to the target quarter (i.e., with the index  $j$ ).

In sum, there is the familiar tendency toward underestimation of change in most of the forecasts, with the important exception of RGNP. After a decade of relatively stable and high rates of growth, the 1970s gave rise to a novel phenomenon commonly called stagflation and an unexpectedly serious recession: these facts should go far in explaining the average overestimation of real growth observed in our data. For IPD and RGNP, but not for the other variables, the mean errors increase in absolute value with  $j$ .

#### Regression Statistics and Tests of Bias

Regressions of the actual on the predicted values have been computed for all individuals and the group means. For any variable, then,

$$(13) \quad A_{jt} = a_{ij} + b_{ij}P_{ijt} + u_{ijt}, \quad j = 0, 1, \dots, 4,$$

where  $i$  denotes the  $i$ -th forecaster (the group mean being included as a particular case) and  $j$  denotes the time distance of the target quarter

from  $t$ .<sup>26</sup> The sample least-squares estimates  $a$  and  $b$  (the subscripts may be dropped for simplicity) lend themselves to statistical tests of the joint null hypothesis that the true (population) parameters of the relation between  $A$  and  $P$  are  $\alpha = 0$  and  $\beta = 1$ . A sufficiently high  $F$ -ratio refutes that hypothesis, suggesting that the forecast contains some systematic errors.

The results of these tests do not show any regular dependence on the time distance  $j$ , so their representation can be greatly simplified without loss of substance by aggregation across the quarters Q0-Q4. Table 6 sums up the evidence from a very large sample (2,350 regressions). What stands out clearly is the contrast between the predictions of inflation and those of the other variables. For IPD, almost half of the  $F$ -ratios are significant at the 5% level and about two-thirds at the 10% level, which indicates a disturbing frequency of apparently biased or inefficient predictions. Elsewhere no more than six to twelve percent of the  $F$ -ratios are significant at the 5% level and 14 to 20 percent at the 10% level; and at the levels of 0.5%, 1%, and 2.5% some of the proportions are low enough to be attributable entirely to chance. The relatively good results for UR, CEDG, and CBI deserve to be noted.

Table 7 shows the detailed results of the regressions for the group mean predictions. Most of the intercepts  $a$  are small fractions, with signs about evenly mixed (column 1). All of the slope coefficients  $b$  are positive and most are not far away from 1.0 (column 2). However, in several cases the absolute values of  $a$  deviate significantly from zero and the values of  $b$  deviate significantly from unity, as determined by the  $F$  and  $t$  statistics (columns 3-5). Of the thirty  $F$ -ratios, four are significant at the 1% level, nine at the 5% level, and fourteen at the 10% level. The tests reject

---

<sup>26</sup>In this simplified notation, the distinction between levels and percentage changes (see equations (11) and (12) and text above) is disregarded.

TABLE 6

Tests of  $H_0: \alpha = 0$  and  $\beta = 1$  for Forecasts by Regular Participants  
in the ASA-NBER Business Outlook Surveys, 1968-1979

Variable <sup>a</sup>	Coverage		Percentage of Forecasts with F-ratios <sup>d</sup> that are significant at the level of:				
	No. of Indivi- duals <sup>b</sup>	No. of Forecast Series <sup>c</sup>	$\frac{1}{2}$ of 1%	1%	2½%	5%	10%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GNP	79	395	2.3	3.0	7.1	12.2	19.8
IPD	79	395	10.9	18.0	30.1	47.3	66.3
RGNP	79	395	1.0	2.0	5.3	10.9	19.0
UR	79	395	0.2	0.5	2.8	6.1	14.4
CEDG	75	375	0.8	2.4	2.9	7.7	15.7
CBI	79	395	0	0.8	3.0	8.2	17.5

<sup>a</sup>The symbols for the variables are identified in Table 5.

<sup>b</sup>Includes those individuals who participated in more than 12 surveys.

<sup>c</sup>Contains each individual's predictions for five target quarters (Q0-Q4). Base of the entries to the right (columns 3-7).

<sup>d</sup>Refers to the F statistics for testing the joint null hypothesis that  $a_{ij}$  and  $b_{ij}$  in eq. (13) are not statistically different from zero and one, respectively. See text.

SOURCE: Quarterly ASA-NBER surveys, 1968:4-1979:1.

TABLE 7  
TESTS OF  $H_0: \alpha = 0$  and  $\beta = 1$  AND MEAN SQUARE ERROR COMPONENTS,  
GROUP MEAN<sup>0</sup> FORECASTS FROM THE ASA-NBER BUSINESS OUTLOOK SURVEYS,  
1968-1979

Quarter Predicted	Regression Estimates		F-ratio for $\alpha = 0$ $\beta = 1$	t-tests for		Components of MSE (percent)		
	a	b		$\alpha = 0$	$\beta = 1$	MC	SE	RV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<b>GNP in Current Dollars (GNP)</b>								
Q0	-.42	1.26	1.97	-1.21	1.63	3.2	6.6	90.1
Q1	.02	1.03	.18	.03	.12	0.9	0.0	99.1
Q2	.27	.94	.42	.41	-.21	2.1	0.1	97.8
Q3	-.01	1.08	.51	-.02	.22	2.6	0.1	97.2
Q4	-.76	1.42	.67	-.70	.83	2.1	2.2	95.7
<b>Implicit Price Deflator (IPD)</b>								
Q0	-.08	1.17	3.61	-.35	1.14	14.8	3.2	81.9
Q1	.37	.95	5.11 <sup>#</sup>	1.18	-.22	26.1	0.1	73.8
Q2	.72	.73	6.27 <sup>†</sup>	1.90 <sup>§</sup>	-.91	30.8	2.2	67.0
Q3	.85	.67	7.31 <sup>†</sup>	2.01 <sup>§</sup>	-.95	37.1	2.4	60.5
Q4	.98	.66	7.21 <sup>†</sup>	1.72 <sup>§</sup>	-.71	44.9	1.6	53.5
<b>GNP in Constant Dollars (RGNP)</b>								
Q0	-.36	1.44	4.63 <sup>#</sup>	-2.37	3.02 <sup>*</sup>	0.3	22.8	76.9
Q1	-.63	1.50	2.83 <sup>§</sup>	-2.38	1.91 <sup>§</sup>	5.2	9.3	85.5
Q2	-.67	1.49	1.60	-1.65	1.17	4.8	3.6	91.6
Q3	-.58	1.30	1.26	-1.02	.52	6.1	0.7	93.2
Q4	-1.30	1.88	2.70 <sup>§</sup>	-1.67	1.14	13.3	4.2	82.6
<b>Unemployment Rate (UR)</b>								
Q0	-.05	1.00	1.85	-.47	.03	9.2	0.0	90.6
Q1	-.02	1.00	.01	-.07	.05	0.0	0.0	100.0
Q2	.30	.96	.45	.56	-.40	2.0	0.4	97.6
Q3	1.11	.85	1.97	1.41	-1.13	7.2	3.4	89.3
Q4	2.25	.65	2.49 <sup>§</sup>	1.95 <sup>§</sup>	-1.77 <sup>§</sup>	6.0	10.1	83.9
<b>Consumer Expenditures—Durable Goods (CEDG)</b>								
Q0	.08	1.63	5.81 <sup>*</sup>	.14	2.72 <sup>#</sup>	10.6	18.4	70.9
Q1	.03	1.20	.30	.03	.44	1.0	0.5	98.5
Q2	.66	.79	.14	.49	-.33	0.5	0.3	99.3
Q3	.51	.85	.08	.32	-.19	0.3	0.1	99.6
Q4	-2.53	2.39	.93	-1.29	1.36	0.0	6.0	94.0
<b>Change in Business Inventories (CBI)</b>								
Q0	-.50	1.39	3.36	-.25	1.81 <sup>§</sup>	8.6	8.2	83.2
Q0	-3.94	1.71	3.82	-1.38	2.41 <sup>#</sup>	4.6	14.9	80.4
Q2	-5.40	1.80	2.91 <sup>§</sup>	-1.51	2.19	2.7	12.6	84.5
Q3	-7.43	1.95	2.87 <sup>§</sup>	1.80 <sup>§</sup>	2.28	1.4	14.1	84.5
Q4	-7.59	1.95	2.25	-1.46	1.93 <sup>§</sup>	2.5	12.0	85.5

NOTE: See text and eqs. 13 and 14 for the explanation of the symbols and tests used. All measures refer to the means of predictions by those respondents to the quarterly ASA-NBER surveys, 1968:4-1979:1, who participated in at least 12 surveys.

<sup>†</sup>Significant at the level of  $1/2$  of one percent.

<sup>\*</sup>Significant at the 1 percent level.

<sup>#</sup>Significant at the  $2 1/2$  percent level.

<sup>§</sup>Significant at the 5 percent level.

<sup>§</sup>Significant at the 10 percent level.

the joint hypothesis that  $\alpha = 0$  and  $\beta = 1$  most strongly for the predictions of inflation in Q2 and beyond, and the  $t$  statistics suggest that this is attributable mainly to the mean bias ( $\alpha > 0$ ). Other, generally weaker, rejections are indicated by the  $F$ -tests for the IPD-inflation in Q0 and Q1, and also for the other variables (except GNP) in at least one or two target quarters. In most of these instances, the  $t$ -ratios suggest inefficiency in the sense of  $\beta > 1$ .

The mean square errors of the group forecasts for each variable have been decomposed according to the general formula

$$(14) \quad M_j = \bar{E}_j^2 + (1 - b)^2 S_{P_j}^2 + S_{u_j}^2, \quad j = 0, 1, \dots, 4$$

where  $\bar{E}_j$  is the mean error and  $S_{P_j}$  and  $S_{u_j}$  denote the standard deviations of the forecasts and the residual disturbances from the regressions of  $A_{jt}$  on  $P_{jt}$ , respectively. The three terms on the right-hand side of (14) may be labeled the mean component, slope component, and residual component (MC, SC, and RV).<sup>27</sup> These estimates, expressed in percent of the corresponding mean square errors, are listed in columns 6-8 of Table 7. Given  $M_j$ , the larger its random component RV and the smaller both MC and SC, the better.

Where the results of the  $H_0$  tests are favorable, the figures for MC and SC are very low, as would be expected: the best examples are found in the estimates for GNP, UR, and (except for Q0) CEDG. For IPD, in contrast to all other variables, the MC percentages are very high, rising from 15 for Q0 to 45 for Q4 and averaging 31 (the averages elsewhere range from two to six percent). The estimates for SC are on the whole smaller than those for MC, but they are high for some of the Q0 predictions, probably because of errors

---

<sup>27</sup>See Theil, 1965, p. 38, and Mincer and Zarnowitz 1969, pp. 10-11.



in the jump-off figures (column 7). RV accounts for 67 percent of the mean square errors of the inflation forecasts on the average across the target quarters (the figures decline sharply from 82% for Q0 to 54% for Q4), whereas the other RV estimates generally exceed 80 or even 90 percent (column 8).

#### Tests of Autocorrelation of Errors

The summary measures of accuracy and bias are informative but they tell only a part of the story; it is advisable to examine further the properties of the time series of errors and ask whether autocorrelations exist that could be exploited to improve the predictions (Granger and Newbold 1973).

The Box-Pierce statistic (R) serves as a convenient test of the presence of autocorrelations in the errors. In the present context, it is written as

$$(15) \quad R_j = \frac{n(n+2)}{n-k} \sum_{k>j}^6 \hat{r}_k,$$

where  $\hat{r}_k$  is a sample estimate of an autocorrelation coefficient for the lag k:

$$(16) \quad \hat{r}_k = \frac{n \sum_t^{n-k} (E_{jt} - \bar{E}_j)(E_{jt+k} - \bar{E}_j)}{(n-k) \sum_t^n (E_{jt} - \bar{E}_j)^2}.$$

Here  $\hat{r}_k$  is defined for level errors but the same formula, with  $e$  replacing  $E$  throughout, applies to percentage change errors. The autocorrelation lags are restricted to the range of one to six quarters because the available error series are short. Table 8 covers all predictions by those 18-20 individuals who participated in more than 12 consecutive surveys: a total of 452 error series whose length varies from 13 to 33 and averages 19 quarters.

TABLE 8

Chi-Square Tests of Autocorrelations of Errors in Forecasts by Selected Participants in the ASA-NBER Business Outlook Surveys, 1968-1979.

Variable <sup>a</sup>	Coverage		Percentage of Forecasts with $R_j$ coefficients <sup>d</sup> that are Significant at the level of				
	No. of Indivi- duals <sup>b</sup>	No. of Forecast Series <sup>c</sup>	$\frac{1}{2}$ of 1%	1%	2½%	5%	10%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GNP	19	75	13.3	14.7	18.7	22.7	37.3
IPD	20	80	23.8	28.8	46.2	52.5	67.5
RGNP	19	75	2.7	2.7	9.3	12.0	16.0
UR	18	71	23.9	25.4	32.4	39.4	46.5
CEDG	18	71	2.8	2.8	4.2	8.5	21.1
CBI	20	80	7.5	11.2	17.5	31.2	40.0

<sup>a</sup>The symbols for the variables are identified in Table 5.

<sup>b</sup>Includes those individuals who participated in more than 12 consecutive surveys.

<sup>c</sup>Contains each individual's predictions for four target quarters, Q0-Q3. (The few observations available for Q4 are excluded. For GNP, RGNP, UR, and CEDG, the number of predictions for Q3 is one less than that shown in Column 1.) Base of the entries to the right (columns 3-7).

<sup>d</sup>Refers to Box-Pierce statistics as defined in eq. (15), with 6-j degrees of freedom. See text.

SOURCE: Quarterly ASA-NBER surveys, 1968:4-1979:1.

The estimates refer to the series for  $j = 0, \dots, 3$ ; the Q4 errors, for which the samples are smaller, are not included.

The Box-Pierce statistic is approximately distributed as chi-square, in this case with  $6-j$  degrees of freedom.<sup>28</sup> The omission of  $\hat{r}_k$  for  $k < j$  from (10) reflects the fact that the information available at time  $t$ , when a survey is taken, includes the errors of past predictions through the previous quarter ( $t - 1$ ) but does not include the errors of the predictions made currently for Q0, Q1, etc. For example, the errors of the Q0 forecasts will not be known until a quarter later, hence they are not yet available to the forecasts for Q1, Q2, and Q3, which are all made at the same time as those for Q0.<sup>29</sup>

Overall, the frequencies of significant autocorrelations of errors are high: 65, 127, and 174 at the 1%, 5%, and 10% levels, respectively (that is, 14, 28, and 38 percent of the total of 452 series). For GNP, RGNP, and CEDG it is the Q0 errors that show the highest proportions of the significant  $R_j$  statistics, but for IPD, CBI, and UR it is the Q2 or Q3 errors. On the whole,

---

<sup>28</sup>If the errors formed random uncorrelated sequences, the  $\{\hat{r}_k\}$  would themselves be uncorrelated and would have variances equal to  $\frac{n-k}{n(n+2)}$ . For large values of  $n$  and relatively small number  $m$  of the autocorrelations included in  $R$ , the variances approximate  $\frac{1}{n}$  and  $R = n \sum_{k=1}^m \hat{r}_k^2 \sim \chi_m^2$ . In view of the small size of the available samples, it seemed advisable to avoid the approximations. See Box and Pierce 1970.

<sup>29</sup>Thus it is not only desirable, but also, at least in principle, possible for  $r_1$  to be reduced to a level not significantly different from zero for the Q0 predictions; but the lack of current knowledge of the most recent errors makes it difficult to accomplish the same for Q1 and any more distant quarters. This argument applies generally to  $r_k$  for  $k \leq j$  and here specifically to  $r_1$  for Q1, Q2, and Q3;  $r_2$  and  $r_3$  for Q3.

**TABLE 9**  
Autocorrelations of Error in the Group Mean Forecasts of Six  
Aggregate Variables, ASA-NBER Business Outlook Surveys, 1968-1979

Line	Quarter Predicted <sup>a</sup>	Estimated Autocorrelation Coefficients <sup>b</sup>												Box-Pierce Statistic <sup>c</sup>  R <sub>J</sub> (13)
		$\hat{r}_1$	$\hat{r}_2$	$\hat{r}_3$	$\hat{r}_4$	$\hat{r}_5$	$\hat{r}_6$	$\hat{r}_7$	$\hat{r}_8$	$\hat{r}_9$	$\hat{r}_{10}$	$\hat{r}_{11}$	$\hat{r}_{12}$	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	
GNP in Current Dollars (GNP)														
1	Q0	-.19	-.17	-.05	-.04	-.11	.26	.21	-.37	-.13	-.02	.22	-.10	21.39 <sup>*</sup>
2	Q1	-.01	-.20	-.10	-.05	-.03	.21	.15	-.32	-.26	-.01	.31	-.02	21.10 <sup>*</sup>
3	Q2	.02	-.18	-.02	.02	.06	.15	.13	-.31	-.28	.01	.23	-.16	16.89 <sup>†</sup>
4	Q3	.02	-.19	-.02	.06	.03	.16	.12	-.37	-.23	.02	.23	-.19	17.86 <sup>†</sup>
Implicit Price Deflator (IPD)														
5	Q0	.35	.33	.27	.02	-.04	-.22	-.45	-.28	-.55	-.35	-.19	-.35	64.90 <sup>†</sup>
6	Q1	.48	.28	.28	.16	-.07	-.25	-.46	-.48	-.51	-.31	-.27	-.47	73.47 <sup>†</sup>
7	Q2	.62	.41	.32	.16	-.08	-.30	-.53	-.54	-.59	-.44	-.38	-.45	93.46 <sup>†</sup>
8	Q3	.69	.51	.36	.17	-.09	-.36	-.53	-.58	-.64	-.52	-.47	-.42	105.32 <sup>†</sup>
GNP in Constant Dollars (RGNP)														
9	Q0	-.10	.01	-.03	-.13	-.17	.19	.04	-.27	-.21	-.08	.07	-.12	12.75
10	Q1	.06	-.06	-.03	-.09	-.14	.07	.01	-.32	-.24	-.07	.13	-.02	11.89
11	Q2	.17	-.00	.07	-.01	-.06	.00	-.02	-.32	-.32	-.10	.01	-.15	13.58 <sup>*</sup>
12	Q3	.22	.07	.10	.03	-.06	-.01	-.07	-.38	-.31	-.16	-.01	-.19	17.41 <sup>*</sup>
Unemployment Rate (UR)														
13	Q0	.23	-.17	.02	.01	-.07	-.03	.21	-.02	-.28	-.23	.00	.10	14.30
14	Q1	.56	-.03	-.22	-.20	-.08	.07	.20	.01	-.28	-.25	-.15	-.06	16.51
15	Q2	.68	.14	-.19	-.20	-.10	.04	.12	-.01	-.22	-.26	-.18	-.11	14.40
16	Q3	.77	.33	-.01	-.12	-.07	.01	-.02	-.07	-.20	-.26	-.25	-.19	13.56
Consumer Expenditures—Durable Goods (CEDG)														
17	Q0	-.24	-.11	-.20	.19	-.08	.10	-.22	.07	-.47	.22	.06	.23	29.06 <sup>†</sup>
18	Q1	-.28	-.12	-.15	.12	-.04	.10	-.18	.11	-.38	.21	-.01	.18	17.84 <sup>†</sup>
19	Q2	-.25	-.15	-.13	.11	.03	.04	-.08	.03	-.34	.10	-.01	.13	10.35
20	Q3	-.25	-.11	-.13	.15	-.02	.05	-.13	.06	-.34	.09	-.06	.14	10.47
Change in Business Inventories (CBI)														
21	Q0	.22	.01	-.06	-.03	-.17	-.08	.02	-.26	-.08	-.19	-.18	-.30	17.56
22	Q1	.41	-.02	-.12	-.08	-.14	-.05	.06	-.11	-.17	-.20	-.24	-.21	15.53 <sup>*</sup>
23	Q2	.42	.08	-.09	-.11	-.17	-.04	.07	-.17	-.20	-.33	-.17	-.19	17.09 <sup>†</sup>
24	Q3	.48	.11	-.05	-.10	-.17	-.09	-.04	-.22	-.24	-.31	-.18	-.17	17.66 <sup>†</sup>

<sup>a</sup>Q0 denotes the current quarter, Q1 the following (first future) quarter, etc.

<sup>b</sup>See equation (11) and text.

<sup>c</sup>See equation (10) and text.

<sup>†</sup>Significant at the level of  $1/2$  of one percent.

<sup>\*</sup>Significant at the 5 percent level.

<sup>§</sup>Significant at the 10 percent level.

SOURCE: Quarterly ASA-NBER surveys, 1968:4-1979:1. Group mean of forecasts by those individuals who participated in at least 12 surveys.

the distribution of these statistics does not depend strongly on the target quarter.<sup>30</sup>

Table 8, which uses aggregation across the quarters Q0-Q3, brings out strong differences between the autocorrelations of errors in forecasts of different variables. Once more it is the predictions of inflation that have definitely the worst record, with more than half of the forecast series showing  $R_j$  coefficients significant at the 5% level. By the same criterion, 39, 31, and 23 percent of the forecasts of unemployment, inventory investment, and GNP, respectively, have autocorrelated errors, as shown in column 6 of the table. The evidence for the predictions of real GNP and consumer expenditures on durable goods is much more favorable, the corresponding proportions here being twelve and eight percent. In general, the forecasters do considerably worse on these tests than on the bias tests (note that the entries in Table 8 tend to be much higher than their counterparts in Table 6).

Finally, Table 9 presents the sample estimates of the autocorrelation function for the errors in the ASA-NBER group mean forecasts. The coefficients  $\hat{r}_k$  are computed to the formula (16), but here the series are long enough to permit the autocorrelation lag  $k$  to vary from one to twelve quarters.<sup>31</sup>

---

<sup>30</sup>Thus, of the 174  $R_j$  coefficients that are significant at the 10% level, Q0 errors account for 29 percent, Q1 errors for 23 percent, Q2 errors for 24 percent, and Q3 errors for 24 percent. The corresponding proportions for the 127 coefficients that are significant at the 5% level are very similar (29, 21, 24, and 26 percent).

<sup>31</sup>Rather than from one to six quarters, the range used for the shorter individual error series (see text and Table 8 above). Accordingly, the Box-Pierce statistics listed in Table 9 differ from those defined in eq. (15) in that the sum  $\sum_{k>j}^6 \hat{r}_k$  is now replaced by  $\sum_{k>j}^{12} \hat{r}_k$ . All estimates refer to the series for  $j = 0, \dots, 3$ .

The Box-Pierce statistics  $R_j$  for the inflation (IPD) errors stand out for being very large and highly significant (at the level of  $1/2$  of one percent). They increase markedly with  $j$ , the distance to the target quarter, reflecting the rising absolute values  $|\hat{r}_k|$ , which can be seen by reading down the columns in lines 5-8. The  $\hat{r}_k$  coefficients are initially positive and declining as  $k$  rises from one to four, but then they turn negative and large as  $k$  rises from five to twelve. In short, they fail to approach zero even for the largest  $k$ 's.

For the other variables, there is much less evidence of autocorrelated errors. The  $\hat{r}_k$  coefficients tend to be much smaller absolutely, particularly for the larger values of  $k$ , and they do not show any comparable patterns in signs as related to  $k$  or magnitudes as related to  $j$ . The  $R_j$  statistics are significant at the 5% or 10% levels for the GNP errors and in a few instances for some of the other variables. The most favorable results in this respect are those for RGNP and UR, but those for CEDG and CBI are not much worse (column 13).

Since the  $\hat{r}_k$  should be normally distributed random variables with mean zero and variance slightly less than  $1/n$  (see note 26 above), another test may be applied to these estimates to see how many of them fall outside range of  $\pm 2$  standard deviations from zero. In Table 8, these would be the cases where  $|\hat{r}_k|$  exceeds 0.25. Most of the autocorrelations for the inflation errors, but relatively few of those for the other variables, are larger than 0.25.<sup>32</sup>

---

<sup>32</sup>In fact, a large number of the  $r$  coefficients are probably not significantly different from zero on this test. It is intriguing that most of the large autocorrelations are found in columns 8-10 of the table, that is, for  $k$  of 8, 9, or 10 quarters. (The low-order autocorrelations for  $k \leq j$  are high throughout for IPD and in about half the cases for UR and CBI as well, but, as already noted, the errors involved are not part of the information that is available on the current basis.)

#### IV. Summary

In this section, the main results of the study are summed up in several points. It should be stressed that some of them may be particular to the period covered, 1968-79 (as noted earlier in a few specific instances).

1. A very strong finding of this study, applying to all variables and target quarters covered, is that the mean predictions from a series of surveys are on the average over time more accurate than most of the corresponding sets of individual forecasts or expectations. The minorities that did succeed in outperforming the group averages vary in size and composition for the different variables and predictive horizons, but typically represent one-fifth to one-third of the total number of individuals surveyed.

2. Rank correlations among the respondents according to the ratios of their root mean square errors to those of the corresponding group averages ( $M_i/M_{gi}$ ) are positive for all variables and target quarters, and in most cases significantly so. This suggests that a moderate degree of consistency exists in the relative performances of a sufficient number of the survey members, even though most people have but transitory spells of above-average accuracy (most of the  $M_i/M_{gi}$  ratios exceed 1.0). It remains to be seen whether weighted combinations of selected forecasts from the group would yield significantly large and persistent gains in accuracy, but our results do not rule out this possibility.

3. There is a reasonably well articulated tendency for the errors to increase in absolute size with the time distance to the target quarter, from Q0 to Q4, but by decreasing margins. This shows up both in the summary measures for the group forecasts and in the distributions of the corresponding statistics for the individual forecasts. Correlations between predictions and realizations decline steadily as the target quarter recedes into the future.

The  $\bar{r}^2$  coefficients for the group mean predictions exceed considerably their average counterparts for the individuals.

4. The mean errors are predominantly negative for all nominal variables covered, which reflects mainly large underestimates of inflation. They tend to be positive for real growth and negative for the unemployment rate. They are small for the percentage change in GNP as the underestimates of the IPD inflation cancel the smaller overestimates of growth in RGNP. For inflation and real growth, but not the other variables, the mean errors increase in absolute value between Q0 and Q4.

5. The tests of the joint null hypothesis that the regressions of actual on predicted values have zero intercepts and unitary slope coefficients are quite unfavorable to expectations of inflation. Of the individual F ratios, 18 percent exceed the 1% significance point, nearly half exceed the 5% point, and two-thirds exceed the 10% point. The F tests for the group mean forecasts confirm the bias. These findings are consistent with other evidence that in the past decade anticipated rates of inflation have generally erred on the low side, tending to lag behind the actual rates much like adaptive extrapolations would (Zarnowitz 1979; McNees 1981; Figlewski and Wachtel 1981).

6. For the other variables covered, these weak tests of rationality show the survey predictions generally in much better light. Here the F-tests reject at most three percent of the forecasts at the 1% level, six to twelve percent at the 5% level, and 14 to 20 percent at the 10% level. In many cases, particularly for the group, mean forecasts,  $H_0: (\alpha, \beta) = (0, 1)$  cannot be rejected.

7. Efficient use of information in the development of expectations includes continuous checking on and learning from past mistakes to the extent that this can be done in a timely manner. The process would tend to eliminate



systematic elements such as autocorrelations in the prediction errors. But the errors in the forecasts of the rates of change in IPD contain highly significant autocorrelations  $\hat{r}_k$  for both short and relatively long lags  $k$ . The Box-Pierce statistics are significant at the 1%, 5%, and 10% levels in slightly more than one-fourth, half, and two-thirds of the cases, respectively; and they are even significant at the level of  $1/2$  of one percent for the group mean forecasts of inflation.

8. The incidence of autocorrelations is relatively low in errors of the predictions for real growth and spending on consumer durables, as shown by the chi-square tests of both the individual series and the group means. For nominal GNP growth, the unemployment rate, and inventory investment, the statistics relating to the individual forecasts are much less favorable (though still considerably better than those for inflation), but the group mean predictions show little evidence of serially correlated errors.

#### **V. More Questions and Perspectives on Further Research**

The evidence that inflation has on the whole been poorly anticipated in the late 1960s and 1970s is extensive and substantial. I conclude that, in the absence of any empirically convincing challenge to it, this evidence should be taken seriously. But why have the expectations of inflation been so inaccurate and biased? Why so much worse than the forecasts of other important aggregate variables from the same sources? Did the great disturbances of this period make predicting inflation uniquely difficult?

A study of the behavior over time of cross-sectional data from the successive business outlook surveys can complement the overall accuracy analysis presented in this paper and make a contribution toward answering these important questions. What are the parameters of the error distributions

from each survey, by variable and target quarter, and how do they vary in periods with different economic characteristics (level of business activity, intensity of inflation)?

In addition, comparisons with autoregressive-moving average extrapolations will shed more light on the relative accuracy and efficiency of the survey predictions. The time-series models to be used will be estimated from the same data that were available to the participants in the successive surveys. To this end, all revisions of the relevant data have been recorded and stored in machine-readable form.

The recent marketing successes of several econometric service bureaus have received considerable publicity, and forecasts from such sources as Chase, DRI, and Wharton are considered to have gained much influence, at least in the corporate sector. Where do these forecasts fall in the spectrum of expectations provided by the business outlook surveys? Is there any evidence of their superiority that past studies (McNees 1979; Zarnowitz 1979) failed to detect? Can their influence be documented? These questions will be investigated with the aid of our compilation of the individual predictions by business respondents to the ASA-NBER surveys. In this context, it should be particularly helpful to classify the participants by their most favored forecasting methods (according to their own rankings provided in the surveys). Earlier work based on rather fragmentary survey data revealed no systematic differences in accuracy between the groups resulting from such classifications (Zarnowitz 1971; Su and Su 1975), but an examination of the comprehensive evidence that is now available may or may not confirm these findings and should prove instructive in either case.

## REFERENCES

- Becker, Gary S. 1962. "Irrational Behavior and Economic Theory," Journal of Political Economy 70, no. 1 (February): 1-13.
- Box, G.E.P. and Jenkins, G.M. 1970. Time Series Analysis, Forecasting and Control. San Francisco: Holden-Day.
- Box, G.E.P. and Pierce, David A. 1970. "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models," Journal of the American Statistical Association 65, no. 332 (December): 1509-1526.
- Carlson, John A. 1977. "A Study of Price Forecasts," Annals of Economic and Social Measurement 6, no. 1 (Winter)" 27-56.
- Christ, Carl F. 1975. "Judging the Performance of Econometric Models of the U. S. Economy," International Economic Review 16, no. 1 (February): 54-74.
- Cole, Rosanne, 1969. Errors in Provisional Estimates of Gross National Product. New York: National Bureau of Economic Research.
- Curtin, Richard T. 1980. "Consumer Perceptions of Inflation and Expectations." ISR Working Paper.
- Fair, Ray C. 1974. "An Evaluation of a Short-Run Forecasting Model," International Economic Review 15, no. 2 (June): 285-303.
- Feige, Edgar L. Pearce, and Pearce, Douglas K. 1976. "Economically Rational Expectations: Are Innovations in the Rate of Inflation Independent of Innovations in Measures of Monetary and Fiscal Policy?" Journal of Political Economy 84, no. 3 (June): 499-522.
- Figlewski, Stephen, and Wachtel, Paul. 1981. "The Formation of Inflationary Expectations," Review of Economics and Statistics (February): 1-10.
- Jacobs, Rodney L., and Jones, Robert A. 1980. "Price Expectations in the United States: 1947-75," American Economic Review (June): 269-277.

- Juster, F. Thomas. 1979. Statement presented in Hearings before the Task Force on Inflation of the Committee on the Budget, House of Representatives, 96th Congress, First Session (June), Washington, D. C.: U. S. Government Printing Office, pp. 214-226.
- Kane, Edward J., and Malkiel, Burton G. 1976. "Autoregressive and Nonautoregressive Elements in Cross-section Forecasts of Inflation," Econometrica (January): 1-16.
- Kendall, Maurice G. 1948. Rank Correlation Methods. London: C. Griffin.
- Lucas, Robert E., Jr. 1976. "Econometric Policy Evaluation: A Critique," in K. Brunner and A. H. Meltzer, eds., The Phillips Curve and Labor Markets, Carnegie-Rochester Conference Series no. 1. New York: North-Holland, pp. 19-46.
- McNees, Stephen K. 1973. "The Predictive Accuracy of Econometric Forecasts," New England Economic Review (September/October): 3-27.
- McNees, Stephen K. 1974. "How Accurate Are Economic Forecasts?" New England Economic Review (November/December): 2-19.
- McNees, Stephen K. 1975. "An Evaluation of Economic Forecasts," New England Economic Review (November/December): 25-29.
- McNees, Stephen K. 1976. "An Evaluation of Economic Forecasts: Extension and Update," New England Economic Review (September/October): 30-44.
- McNees, Stephen K. 1978. "The Rationality of Economic Forecasts," American Economic Review 68, no. 2 (May): 301-305.
- McNees, Stephen K. 1979. "The Forecasting Record for the 1970s," New England Economic Review (September/October): 33-53.
- Mincer, Jacob. 1969. "Models of Adaptive Forecasting," in J. Mincer, ed., Economic Forecasts and Expectations. New York: National Bureau of Economic Research: 81-111.
- Mincer, Jacob, and Zarnowitz, Victor. 1969. "The Evaluation of Economic Forecasts," in J. Mincer, ed., Economic Forecasts and Expectations. New York: National Bureau of Economic Research: 3-46.

- Moore, Geoffrey H. 1969. "Forecasting Short Term Economic Change," Journal of the American Statistical Association 64, no. 325 (March): 1-22.
- Moore, Geoffrey H. 1977. "Lessons of the 1973-1976 Recession and Recovery," in William Fellner, ed., Contemporary Economic Problems 1977. Washington, D.C.: American Enterprise Institute: 156-158.
- Muth, John F. 1960. "Optimal Properties of Exponentially Weighted Forecasts," Journal of the American Statistical Association 55, no. 290 (June): 299-306.
- Muth, John F. 1961. "Rational Expectations and the Theory of Price Movements," Econometrica 29, no. 3 (July): 315-335.
- Nelson, Charles R. 1975. "Rational Expectations and the Predictive Efficiency of Economic Models," Journal of Business 48, no. 3 (July): 331-343.
- Nerlove, Mark and Wage, S. 1964. "On the Optimality of Adaptive Forecasting," Management Science 10, no. 2: 207-224.
- Pearce, Douglas K. 1979. "Comparing Survey and Rational Measures of Expected Inflation: Forecast Performance and Interest Rate Effects," Journal of Money, Credit and Banking 11, no. 4 (November): 447-456.
- Poole, William. 1976. "Rational Expectations in the Macro Model," Brookings Papers on Economic Activity 2: 463-514.
- Su, Vincent, and Su, Josephine. 1975. "An Evaluation of the ASA/NBER Business Outlook Survey Forecasts," Explorations in Economic Research 2, no. 4 (Fall): 588-618.
- Wachtel, Paul. 1977. "Survey Measures of Expected Inflation and Their Potential Usefulness," in Joel Popkin, ed., Analysis of Inflation: 1965-1974, Cambridge, Mass.: Ballinger Publishing Company for the National Bureau of Economic Research, pp. 361-395.
- Theil, Henri. 1965. Economic Forecasts and Policy. Amsterdam: North-Holland (first edition 1958).
- Zarnowitz, Victor. 1967. An Appraisal of Short-Term Economic Forecasts. New York: National Bureau of Economic Research.

- Zarnowitz, Victor. 1969a. "The New ASA-NBER Survey of Forecasts by Economic Statisticians," American Statistician (February): 12-16.
- Zarnowitz, Victor. 1969b. "The ASA-NBER Quarterly Survey of the Economic Outlook: An Early Appraisal," 1969 Proceedings of the Business and Economic Statistics Section. Washington, D.C.: American Statistical Association: 241-249.
- Zarnowitz, Victor. 1971. "New Plans and Results of Research in Economic Forecasting," Fifty-fifth Annual Report. New York: National Bureau of Economic Research: 53-70.
- Zarnowitz, Victor. 1972. "Forecasting Economic Conditions: The Record and the Prospect," in V. Zarnowitz, ed., The Business Cycle Today. New York: National Bureau of Economic Research: 183-239.
- Zarnowitz, Victor. 1974. "How Accurate Have the Forecasts Been?" in William F. Butler, Robert A. Kavesh, and Robert B. Platt, eds., Methods and Techniques of Business Forecasting. Englewood Cliffs, NJ: Prentice-Hall: 565-596.
- Zarnowitz, Victor. 1979. "An Analysis of Annual and Multiperiod Quarterly Forecasts of Aggregate Income, Output, and the Price Level," Journal of Business 52, no. 1 (January): 1-33.
- Zarnowitz, Victor. 1982. "On Functions, Quality, and Timeliness of Economic Information," Journal of Business 55, no. 1 (January), forthcoming.
- Zellner, Arnold, and Palm, Franz. 1974. "Time Series Analysis and Simultaneous Equation Econometric Models," Journal of Econometrics 2 (May): 17-54.