NBER WORKING PAPER SERIES

ON THE ACCURACY AND PROPERTIES OF RECENT MACROECONOMIC FORECASTS

Victor Zarnowitz*
University of Chicago
and
National Bureau of Economic Research

Working Paper No. 229

NATIONAL BUREAU OF ECONOMIC RESEARCH, INC.
261 Madison Avenue
New York, N.Y. 10016

Revised February 1978

NOTE: This is an abbreviated version of a report on work in progress supported by grants from the National Science Foundation and the Earhart Foundation to the National Bureau of Economic Research.

*I am grateful to Dennis Bushe, Thomas Kutzen, and Robert Osterlund for assistance with research and preparation of this paper.

NBER Working Paper 229
Revised February 1978

ON THE ACCURACY AND PROPERTIES OF RECENT MACROECONOMIC FORECASTS

Abstract

The aim of this study is to contribute to the measurement and analysis of errors in economists' predictions of changes in aggregate income, output, and the price level. Small sample studies of forecasts can be instructive, but their limitations must be recognized. Compilation of consistent forecast records extending over longer periods of time is necessary to establish a reasonably reliable base for assessments of forecasting behavior and performance. Thus the historical record of post-World War II forecasts assembled in the 1960's by the NBER is here extended and updated.

The end-of-year predictions of annual percentage changes in GNP earn good marks for overall accuracy when judged according to realistic rather than ideal standards. Moreover, they are found to have improved significantly in the period since the early 1960's compared with the previous years after World War II.

The corresponding predictions for GNP in constant dollars (real growth) and the GNP implicit price index (inflation) are considerably poorer. The former suffer from large turning point errors, the latter from large underestimation errors. Indeed, forecasts of inflation are not much better than projections of the most recently observed inflation rates, and they lag behind the actual rates much like such projections. But the errors in forecasts of real growth are negatively correlated with the errors in forecasts of inflation, which helped to make the nominal GNP predictions more accurate.

Abstract

Forecasts for the year as a whole can be satisfactory when based on a good record for the first two quarters; they tend to be more accurate than forecasts with longer spans. An examination of the recent multiperiod predictions from well-known econometric models and business outlook surveys shows that the errors for real growth and inflation cumulated rapidly beyond the spans of 2 to 4 quarters. Previous studies have shown the cumulation to be as a rule less than proportional to the increase in the span, but in the period of recession and recovery 1973-75 the build-up of errors was much greater. Again the nominal GNP forecasts benefitted from offsetting errors as the rise in prices was heavily underestimated and the downturn in real activity was missed. Forecasters were generally unprepared for the concurrence of accelerating inflation and slowing, then declining output rates: they optimistically (and probably also from a lingering faith in a simple Phillips trade-off) kept anticipating less inflation and more growth.

At the present time, the predictive value of detailed multiperiod forecasts reaching out further than a few quarters ahead must be rather heavily discounted. No doubt, in periods less turbulent than the recent past the longer forecasts can be considerably more accurate, but this fair-weather argument is not very persuasive or helpful.

Victor Zarnowitz
Graduate School of Business
University of Chicago
Chicago, Illinois 60637
312/753-3615

ON THE ACCURACY AND PROPERTIES OF RECENT MACROECONOMIC FORECASTS

Victor Zarnowitz

I. On Some Uses and Limitations of Forecast Data

How and how well economists forecast, and how much their predictions help or hurt public and private decision making, are matters that ought to receive much attention of the profession. This is so not only because of their direct interest to the authors, users, and critics of the forecasts, but also because of their intrinsic but less evident academic interest. What is the practical applicability of economic analysis in this critical area? What is the quality of foresight and counsel that can be expected of responsible economists? These are broad questions which are not easy to answer, but they are basic and surely deserve to be tackled. This requires that we systematically confront forecasts as indications of how economists ex ante thought events are likely to unfold with ex post knowledge of what actually did happen and how. The aims, from the least to the most ambitious, are (1) to measure forecast errors, (2) to explain them, and (3) to learn how to reduce them in the future.

Success in forecasting may be occasional and fortuitous or intuitive, but progress in forecasting, to the extent it is possible, can only come from advances of science, not art or chance. It presupposes that sufficiently important and persistent regularities in economic processes and relationships exist and be properly identified and used. Learning processes are involved, which can be time-consuming and discontinuous, reflecting in part the shifts and discontinuities in the economic change itself, in part the inadequacies of measurement and analysis.

Data on economic forecasts generally cover short time periods. Long time series on consistent predictions simply do not exist. Few if any forecast sets are fully identified according to the many aspects and dimensions

that matter (source, target, timing, assumptions, data, models, and methods used), so that it is often difficult to determine what constitutes a suitable "sample" of forecasts of a given type. Moreover, few forecasters leave their models and techniques unchanged for long as they seek improvements and try to adapt to new developments in the economy. Hence, a particular forecaster's past record is often a highly uncertain basis for inferences how he will perform in the future.

Even more hazardous, if not irresponsible, are attempts to grade forecasters on the evidence of how well they predicted change in a particular short period, say, a year or a few years. Clearly, on any individual occasion some forecasters will be ahead of others by sheer chance or for some idiosyncratic reasons. Strong evidence of significant and stable differences over time would be required to rank the forecasting individuals, groups, or models with a modicum of confidence, and such evidence is essentially lacking (Zarnowitz, 1967, 1971; Christ; McNees, 1975).

The proliferation in recent years of multiperiod quarterly macroforecasts offers no substitute for long historical series. These are rich data containing much interesting material that certainly deserves to be carefully recorded and analyzed. However, such forecasts, and so their errors, tend to be internally correlated in at least two ways: (a) serially, within each sequence made from a given base period and (b) across the successive sequences, which overlap and thus refer partially to the same target period. Each multiperiod forecast is a joint product of the common information, technique, and judgment used, and each depends on previous forecasts of which it is to some extent a revision. Thus, errors in the data, models, procedures, and judgments, autocorrelated disturbances, and certain types of distributed lags are all likely to induce interdependencies within and between the multi-

period forecasts. The resulting complex correlation structures resist estimation, given the small samples of comparable predictions from any given source. Consequently, measures of average accuracy, bias, etc., calculated from such samples are difficult to interpret from the viewpoint of statistical inference (Spivey and Wrobleski).

Two conclusions are surely valid. First, small-sample studies of forecasts are still needed and can be instructive, but their limitations must be recognized. Second, it is necessary to compile and examine forecast records extending as far back in time as possible, so as to gain information, take a longer view of forecasting behavior and performance, and place the short records of recent predictions in a proper perspective. Historical data on post-World War II forecasts assembled in the 1960's by the National Bureau of Economic Research provide a good base here, which I was able to partially extend and update. Some preliminary results for annual forecasts of three variables are reported below.

II. The Record of Annual GNP Forecasts Since 1947

In the early post-World War II period, most forecasts were made near the end of the calendar year for the next year and most referred to GNP in current dollars. The evidence we have on such forecasts goes back to 1947 but is quite fragmentary for the late 1940's and early 1950's.

The period of transition from the war economy witnessed the largest errors on record in the GNP forecasts. Even after the 1945-46 predictions were shown to have greatly underestimated the then prevailing levels of economic activity (Klein, 1946), expectations of a business slump stubbornly persisted. One small, reputable group of private forecasters came up with an average prediction for 1947 of a 6 percent decline in GNP, whereas the actual change turned out to be a rise of about 11 percent. For 1948 the group predicted a

fractional decline but GNP instead advanced again at much the same surprisingly high rate. The failure of forecasts during these years was widespread, with but a few partial exceptions; the developments of the time could not be predicted well with estimates based on data and relationships for the 1930's and false analogies with the early post-World War I period. When a recession finally came late in 1948, it proved shorter than many had expected. A "consensus forecast" by more than 30 respondents polled in December 1948 anticipated well the decline of nearly 2% in GNP during 1949, but a year later the same group was wide off the mark in predicting a drop of 3.5% while GNP actually staged a strong comeback in 1950 with a rise exceeding 10%.

The evidence for the period 1953-76 is summarized in Table 1 in terms of comparisons between the predicted and the actual annual percentage changes. It is generally instructive to analyze forecast errors in terms of levels, absolute changes, and percentage changes but, if a choice must be made for succinctness, there are several good reasons for using percentage changes where technically appropriate, particularly for variables with strong trends. (1) What is predicted in the first place is change from the last known or estimated level, and percent changes often vary less with the levels and are more stable and comparable over time than dollar changes. (2) The percent change forecasts are apt to be less affected by data revisions. (3) Some important measures of predictive performance, such as correlations with actual values, are much more meaningful for change forecasts than for level forecasts. (4) It is the rates of growth in economic aggregates (income, output, prices) that are of main interest to analysts and policy makers.

The forecasts are made late in the year $t - 1$ or, in a few cases, very early in the target year t; typically, the forecasters know the official estimates for the first three quarters but not for the last quarter of the

TABLE 1

SUMMARY MEASURES OF ERROR FOR ANNUAL PREDICTIONS OF PERCENTAGE CHANGE IN GNP, 1953-76

Line	Period and No. of Years Covered	Livingston Survey, Mean Forecast[a] (1)	Selected Private Forecasts, Mean[b] (2)	N.Y. Forecasters Club, Mean Forecast[c] (3)	ASA/NBER Survey, Median Forecast[d] (4)	Economic Report of the President[e] (5)	Michigan Model[f] (6)	Wharton Model[g] (7)	Extrapolations: Last Change[h] (8)	Extrapolations: Average Change[i] (9)	Actual: Preliminary[j] (10)	Actual: Revised[k] (11)
					Mean Absolute Percentage Change, Predicted and Actual							
1	1953-76(24)	5.7	6.0								6.6	6.9
2	1956-63 (8)	4.3	4.7	4.0							5.0	5.3
3	1963-76(14)	7.1	7.2			7.6	7.2	7.6			7.9	8.2
4	1969-76 (8)	8.1	8.0		8.1	8.6	8.3	8.2			8.4	8.8
					Mean Absolute Error, in Percentage Points							
5	1953-76(24)	1.6	1.2						3.1	2.3	0.5	
6	1956-63 (8)	1.7	1.4	1.7					3.4	1.9	0.5	
7	1963-76(14)	1.1	0.9			0.9	1.3	0.8	2.3	1.8	0.3	
8	1969-76 (8)	1.0	0.6		1.0	0.8	1.0	0.9	2.6	2.0	0.3	
					Mean Error, in Percentage Points							
9	1953-76(24)	-1.0	-0.7						-0.3	-0.1	-0.5	
10	1956-63 (8)	-0.4	-0.1	-0.8					0.2	-0.4	-0.3	
11	1963-76(14)	-0.8	-0.6			-0.2	-0.5	-0.3	-0.3	-0.6	-0.3	
12	1969-76 (8)	-0.3	-0.4		-0.3	0.2	-0.1	-0.2	-0.3	-0.5	-0.3	
					Squared Correlation (r^2) Between Predicted and Actual Change							
13	1953-76(24)	.739	.792						.012	.054		
14	1956-63 (8)	.634	.497	.563					.155	.038		
15	1963-76(14)	.717	.791			.752	.603	.689	.006	.079		
16	1969-76 (8)	.780	.875		.768	.834	.746	.669	.002	.000		

Notes to Table 1

[a]Based on surveys conducted by Joseph A. Livingston, syndicated columnist, now with the Philadelphia Inquirer, and published in the Philadelphia Bulletin and American Banker. Of the semiannual surveys, only the end-of-year ones are used here; they typically cover answers to a questionnaire mailed in November and appear in a "Business Outlook" column late in December. The participants in these surveys, listed at the end of the Bulletin columns, varied in number between 44 and 62.

[b]Average of end-of-year annual GNP forecasts from the following sources: (1) Fortune magazine ("Business Roundup"); (2) Harris Trust and Savings Bank; (3) IBM Economic Research Department; (4) National Securities and Research Corporation; (5) NICB (now Conference Board) Economic Forum; (6) Robert W. Paterson, University of Missouri; (7) Prudential Insurance Company of America; (8) UCLA Business Forecasting Project. The earliest of these predictions were made in October, the latest in January. Most but not all of the forecasts in each of these eight sets are available in published form; those for the period ending in 1969 were analyzed in NBER studies of economic forecasting (Zarnowitz, 1967, 1972, 1974).

[c]Of the semiannual forecasts of this group, only the end-of-year ones are included. The group mean forecasts used here cover individual predictions varying in number between 31 and 39. These data, too, were analyzed in NBER studies (see ref. in note b), but no forecasts were collected for the period after 1963. The predictions for 1956-58 were made in October, those for 1959-63 in December.

[d]Source: Quarterly releases by the American Statistical Association and the National Bureau of Economic Research, published in the ASA AmStat News and the NBER Explorations in Economic Research. Median forecasts from the November surveys only are used. The membership in these surveys varied between 45 and 84. See Zarnowitz, 1969, and V. and J. Su, 1975.

[e]Forecasts by the Council of Economic Advisers (CEA) as stated in the Economic Report (usually as the midpoint in a relatively narrow range). As a rule, the Economic Report appears in January. For some earlier studies of these forecasts, see Moore, 1969 and 1977; Zarnowitz, 1972; Fellner, 1976; McNees, 1977.

[f]Published ex ante forecasts from the Research Seminar in Quantitative Economics (RSQE) of the University of Michigan. Based on several working models (see Suits, 1962; Hymans and Shapiro, 1970 and 1974). The forecasts are those released in connection with the University of Michigan annual "Conference on the Economic Outlook," occurring usually in November (in 1974 and 1975, December).

[g]Source: Wharton Economic Newsletter, Econometric Forecasting Unit, Wharton School of Finance & Commerce, University of Pennsylvania. Forecasts based on a series of versions of Wharton models. See Evans and Klein, 1968; Evans, Klein, and Saito, 1972; McCarthy, 1972; Duggal, Klein, and McCarthy, 1974. The forecasts here covered are dated in November or (as in 1971 and 1973-75) in December.

Notes to Table 1 (contd.)

[h]Assumes that next year's percentage change will be the same as that of the previous year. The actual changes used are those based on the preliminary estimates explained in note j below.

[i]Assumes that next year's percentage change will be the same as the average percentage change in the four previous years. On the actual changes used, see note j.

[j]Based on the first official estimates following the year for which the forecast was made.

[k]Based on current data taken from U.S. Department of Commerce, Bureau of Economic Analysis, Handbook of Cyclical Indicators, A Supplement to the Business Conditions Digest, May 1977.

/ r is negative.

year t - 1. The actual changes used to compute the errors are based on the first official estimates for the year t published early in the following (t + 1) year. These are provisional values which are themselves partly near-term predictions, and subsequent revisions indicate that the errors in the early data are by no means negligible (cols. 10-11). On the average, without regard to sign, these revisions are about one-third the size of the forecast errors (lines 5-8). The errors are computed by subtracting the actual from the predicted (or estimated) changes, and they are predominantly negative, which shows that both forecasts and the provisional figures strongly tend to understate the changes in GNP (lines 9-12). By far most of these underestimated changes are increases (for a review of similar findings of earlier studies, see Zarnowitz, 1972).

Table 1 discloses a substantial correspondence between the forecasts and the realizations. The predicted changes approximate the actual ones well in each period covered, the averages of the former being generally less than one percentage point smaller than the averages of the latter (lines 1-4). Where the mean actual changes increased (as from 5% p.a. in 1956-63 to 8% p.a. in 1963-76), so did the mean predicted changes; moreover, the discrepancies between the two diminished in the latter years. The forecasts are in all cases considerably more accurate than the naive model which assumes that next year's percentage change will be the same as that of the previous year and more accurate than the--somewhat less naive--trend extrapolation model which projects the average percentage change of the four previous years. Collectively, the mean absolute error of forecasts is less than half that of the first naive model (lines 5-8, col. 8), and the ratio of the two declines from 0.47 in 1956-63 to 0.43 in 1963-76 and 0.34 in 1969-76. The corresponding ratios for comparisons of the forecasts with the four-year moving average ("trend") extrapolations (col. 9) are 0.84, 0.56, and 0.44.

The average error measures are important but they fall far short of telling the whole story. Measures of correlation (which unfortunately are often omitted from forecast evaluations) are needed to show how well the predicted changes have tracked the actual changes over time. The r coefficients for the forecasts covered in Table 1 are all positive and significant; the r^2 statistics generally exceed 0.5 and, for the more recent periods, average 0.7 or higher (lines 13-16, cols. 1-7). In contrast, the corresponding coefficients for the extrapolations (cols. 8 and 9) are zero or near-zero (where larger, r is negative).

Because sufficiently long and consistent annual time-series data for GNP are not available, no attempt was made here to test the forecasts against higher standards provided by more effective extrapolation methods such as the autoregressive integrated moving average (ARIMA) models. However, recent comparisons of quarterly forecasts with such models show the forecasts to be on the average more accurate (Hirsch, Grimm, and Narasimham; Christ; Spivey and Wrobleski), and I would expect this to be a fortiori true for the annual forecasts and particularly with respect to the correlations with the actual values.

The evidence supports the conclusion that the end-of-year forecasts of current-dollar GNP next year had a reasonably satisfactory record of accuracy since 1953. Indeed, in comparisons with earlier forecasts (Sapir; Okun; Zarnowitz, 1967), that record improved considerably in the 1960's and even in the 1970's, a turbulent period presumed to have been particularly difficult to forecast.

It must be noted that our collection is certainly no random sample, including as it does the official Administration forecasts and several of the most reputable and influential sets of private predictions by business and academic economists (see notes to Table 1). It is also true that our data

and measures have some shortcomings that must not be overlooked. In particular, the estimates of the current position (ECP) which the forecaster actually used as the starting point or base are not always reported. In some cases, therefore, the base values had to be imputed, which was done using data as of the (precise or approximated) date of the forecast plus such information as was available on how the forecaster derived his ECP's on other occasions. The imputations, even if carefully made, undoubtedly contain some errors. However, these errors are definitely not such as to invalidate the broad conclusions of this paper.[1]

More detailed inferences concerning the relative accuracy of the different forecast sets covered cannot be drawn from these results. One reason is that the forecasts differ appreciably with regard to their precise dates, and it is known from previous research that the earlier predictions have a significant advantage over the later ones (Zarnowitz, 1967; McNees, 1975). It is relevant, however, to make the general observation that the average error and correlation measures do not show large, consistent differences among the forecast sets being compared. This is in agreement with earlier findings which together strongly suggest that the search for a consistently superior forecaster is about as promising as the search for the philosophers' stone (Zarnowitz, 1971; McNees, 1973, 1975, 1976; Christ).

A few further observations seem warranted. Although the forecasters included differ in many respects, even a detailed inspection reveals few sharp contrasts between their predictions for the same years. Of course,

---

[1] Other possible errors, also not critical, might arise from the fact that some of our forecasts, lacking directly reported annual predictions, are averages of forecasts for shorter periods within the coming year. This could cause some deviations from the span or target period intended by the forecasters (Carlson, 1977).

competent forecasters use common data and techniques, regularly interact, and are often similarly influenced by recent events and current attitudes and ways of thinking. The genuine ex ante forecasts here considered are all to a large extent "judgmental." Large doses of judgment enter, mostly helpfully, the forecasts derived with the aid of econometric models (see, e.g., Haitovsky, Treyz, and Su). This could well tend to reduce the dispersion among the corresponding prediction of this type; there is indeed some evidence that errors of ex ante forecasts with econometric models vary less than errors of ex post forecasts made without judgmental adjustments (Christ). At the same time, many so-called judgmental forecasters use partly some more or less explicit econometric equations or models, "outside" or "own" (Zarnowitz, 1971; V. and J. Su). While published forecasts by ranking practitioners are often developed with particular skill or care, group average forecasts benefit over time greatly from cancellations of individual errors of opposite sign (Zarnowitz, 1967 and 1972). At any given time, the deviations between corresponding forecasts from different sources are likely to be reduced by the working of these balancing factors. Thus, it is not surprising that forecasts for the same variable and target period tend to be similar. Indeed, the correlations between pairs of the forecast sets included in Table 1, computed for the four periods distinguished therein, are significantly higher than the correlations between predictions and realizations recorded on lines 13-16. The r^2 coefficients for eight pairs of the predicted percentage change series all exceed 0.8, and some are considerably higher.

Of the 110 observations comprised in our seven forecast sets, about 64 percent are underestimates and 34 percent are overestimates. By far most of the latter refer to years marked by economic recessions (1954, 1960, 1970, 1974) or slowdowns (1962, 1967). The provisional GNP values show but two

year-to-year declines in the period covered in Table 1: in 1954, which the forecasts overstated, and in 1958, which the forecasts missed (accounting for the only turning-point errors in this sample). Thus underestimation was limited to the increases in GNP; moreover, it was most pronounced when the increases were particularly large as in 1953, 1955-56, 1965-66, 1968-69, and 1973.

These results suggest the presence of "systematic" errors, but not in the sense of a bias that could have been readily escaped or corrected in advance. It seems difficult to discount them as merely another manifestation of the familiar tendency of forecasts to underestimate the observed changes (which, for series with random elements, is a property of even unbiased and efficient forecasts; Mincer and Zarnowitz, 1969; Hatanaka, 1975). What is underestimated here is the average annual rate of growth in a series which, as properly recognized by the forecasters, is trend-dominated and seldom declines from year to year. This outcome can be traced to the forecasters' tardy recognition of high-growth phases ("booms") and, increasingly, of inflation speedups, but it was also mitigated by their even tardier recognition of business recessions and slowdowns. Such movements are recurrent and not purely random; they have important, detectable regularities as shown by historical studies of business cycles; but they are also nonperiodic and indeed vary a great deal over time, so their predictability remains very limited. In any event, simple "learning from past errors" would not have been of much use here as the errors of these forecasts generally have zero or very low autocorrelations.

III. Annual Forecasts of Real GNP and the Price Level

It is difficult to obtain and verify consistent forecasts of GNP in constant dollars and the implicit price deflator (IPD) that would cover more than just the most recent period. Few business forecasters in the 1950's and

1960's made systematic efforts to decompose their predictions of current-dollar GNP into quantity and price elements. Of the forecasters with econometric models, who paid more attention to real GNP, only two (Michigan and Wharton) have longer records.[2]

Table 2 shows that the predicted changes in real GNP, taken without regard to sign, differed from the actual changes by less than one percentage point on the average (lines 1-3). The predicted changes tend to be smaller than the actual ones, except for the CEA forecasts (col. 3) where the reverse obtains. The mean absolute errors of the forecasts average a little over four tenths of those of the simple last-change extrapolations in 1959-67 and 1962-76, about one third in 1969-76 (lines 4-6, col. 6). Comparisons with extrapolations of the average percentage change of the four previous years give very similar results, except for 1959-67 where the forecast errors average about two thirds of the extrapolation errors (lines 4-6, col. 7).

Correlations between the predicted and actual changes are all significantly positive, and they too suggest some improvement in recent years: the r^2 coefficients for 1969-76 are higher than those for the earlier and longer periods (lines 10-12). It is interesting to observe that all but one of them exceed the corresponding coefficients for current-dollar GNP forecasts, particularly so for the predictions with the Wharton models and the ASA/NBER group medians (cf. Table 2, line 12, and Table 1, line 16). In contrast to the reasonably high correlations for the forecasts proper, those for the extrapolations (cols. 6-7) are here again extremely low or negative.

---

[2]Some of the econometric forecasts were released at more than one date near the end of the year, and in more than one version depending on the data used or policy assumptions made. In all but a few doubtful instances where somewhat arbitrary decisions had to be made, the forecasts chosen are those preferred by the forecaster or, lacking stated preferences, those which embodied assumptions most common to the forecasts made at the time.

TABLE 2

SUMMARY MEASURES OF ERROR FOR ANNUAL PREDICTIONS OF PERCENTAGE CHANGES IN REAL GNP, 1959-76

Line	Period and No. of Years Covered	Selected Private Forecasts, Mean[a] (1)	ASA/NBER Survey, Median Forecast (2)	Economic Report of the President[b] (3)	Michigan Model (4)	Wharton Model (5)	Extrapolations: Last Change (6)	Extrapolations: Average Change (7)	Actual: Preliminary (8)	Actual: Revised (9)
					Mean Absolute Percentage Change, Predicted and Actual					
1	1959-67 (9)	4.1			3.7				4.3(5.0)[c]	4.5
2	1962-76(15)			4.4	3.8				4.1	4.0
3	1969-76 (8)		3.3	4.0	3.5	3.2			3.6	3.3
					Mean Absolute Error, in Percentage Points					
4	1959-67 (9)	1.3			1.0		2.7	1.7	0.5	
5	1962-76(15)			1.1	1.4		2.8	2.6	0.4	
6	1969-76 (8)		1.0	1.2	1.6	0.9	3.7	3.6	0.3	
					Mean Error, in Percentage Points					
7	1959-67 (9)	-0.9			-0.5		-0.6	-1.1	-0.2	
8	1962-76(15)			0.6	0.2		-0.3	0.1	-0.1	
9	1969-76 (8)		0.7	0.8	0.8	0.5	-0.15	0.7	0.06	
					Squared Correlation (r^2) Between Predicted and Actual Change					
10	1959-67 (9)	.445			.531		.306[d]	.000		
11	1962-76(15)			.775	.617		.012	.049[d]		
12	1969-76 (8)		.936	.857	.709	.941	.001	.320[d]		

NOTE: For sources and explanations of the data used in columns 2-9, see footnotes d through k, respectively, in Table 1.

[a] Average of end-of-year annual forecasts of real GNP inferred from the forecasts of current-dollar GNP, the consumer price index (CPI) and the wholesale price index (WPI) from the following sources: (1) Harris Trust and Savings Bank; (2) National Securities and Research Corporation; (3) NICB (Conference Board) Economic Forum; (4) Robert W. Paterson, University of Missouri; (5) UCLA Business Forecasting Project. These forecasts were obtained by dividing the forecasts of GNP, as reported in current dollars, by the composite price level forecasts. The latter are weighted sums of the reported forecasts of CPI and WPI, the weights being .647 and .353, respectively (the first of these proportions represents the average ratio of consumption expenditures to GNP in the period 1953-64). For further detail and analysis of the individual forecasts in this set, see Zarnowitz, 1968. See also Table 1, fn. b.

[b] The forecasts for 1962, 1963, 1965, and 1968 must be inferred from statements in the Report; they are confirmed by the Council as approximately correct, though not in all cases precisely correct (Moore, 1977). The other forecasts are all based on figures given in the Report and so fully verified.

[c] The figure in parentheses is based on preliminary GNP figures deflated by weighted averages of the corresponding data for CPI and WPI (with weights as given in fn. a above). This series of "actual" values is comparable to the forecasts used in column 1 only.

[d] r is negative.

These summary measures, then, present the annual forecasts of real GNP in a generally favorable light. However, the accuracy of these forecasts varied greatly in different years, which at times impaired seriously their usefulness, and this does not show up in the summary. As suggested by the averages with regard to sign (lines 7-9), the usual tendency of forecasts to underestimate changes prevailed in the first half of the period 1959-76 but not in the second half. Actually, the errors varied considerably in each subperiod, primarily reflecting cyclical change and in particular the disturbing effects of missed downturns. Real GNP turned down in 1954, 1958, 1970, and 1974, but of the 10 predictions for these years which are available eight specified continued rises and only two succeeded in signaling declines. Again, and not surprisingly, nearly all of the significantly large overestimation errors refer to the years during which national output grew at relatively low or decreasing rates, and most of the larger underestimation errors refer to the years of high real growth rates.

It is of considerable interest to note that the turning-point errors are much larger than other errors (on the average about 2 1/2-3 times larger, for all forecasts in this collection). Thus, even though relatively few, these directional errors had a strong adverse impact on the overall accuracy of the real GNP forecasts, as indicated by the following tabulation.

	Number	Mean Absolute Error, % Points	Percent of Total Absolute Error
Underestimation errors	33	1.12	46.8
Overestimation errors	21	0.92	24.4
Turning-point errors	8	2.85	28.8

This evidence contradicts the argument that turning-point errors matter little because they are few and far between (cf. Samuelson). But the argument goes further to say that such few large errors are the necessary (and small) price to pay for the avoidance of many large errors "between turning points" by means of optimal estimation procedures such as least squares. However, it is not clear that these procedures imply more than that the variance of the predicted changes must be less than that of the actual changes (and progressively declining as the forecast span is lengthened). The inevitability (indeed, even the existence) of a trade-off between errors at major turning points and other errors has never been demonstrated, and it would seem a counsel of despair for the forecasters to accept it. Prediction of cyclical turns in such series as real GNP, though certainly difficult, is not necessarily impossible, particularly on an annual basis (note the good record in forecasting troughs). In sum, there are indeed strong reasons for makers and users of economic forecasts to give a great deal of attention to turning-point errors. Actually, most of them realize this, as shown by the widespread practice of analyzing such errors (Hickman, ed., 1972; studies in I.E.R., 1974-75). However, there is certainly much need for improvement here, and room for some new initiatives (e.g., on how to use current signals from leading indicators, see Vaccara and Zarnowitz).

The worst single year for the predictions covered in Table 2 was 1974, on the eve of which forecasters across the field missed the onset of a serious recession. This, plus the smaller turning-point errors for 1970, are the main reasons for the rise in the average errors of these forecasts in 1969-76 compared with the earlier years. But the rise in the absolute errors was not large, and there was no decline in accuracy as measured by the criteria of comparisons with extrapolations and correlations of predicted with actual changes (Table 2, cols. 3 and 4). Limited evidence from one longer series of

forecasts suggests that real GNP was predicted with similar average errors in the two 8-year periods 1953-60 and 1969-76, with much smaller errors in the relatively quiet years 1961-68.

Although the forecasts of real GNP are about as good relative to our simple extrapolative benchmark models as are the forecasts of GNP in current dollars, they are less accurate in terms of comparisons of the errors with the actual percentage changes to be predicted. The point is that the extrapolations perform substantially better for nominal GNP than for real GNP. This can be shown by dividing the error of extrapolation into the size of the actual change, without regard to sign, which gives the following overall ratios for the X1 (last change) and X2 (average change) models:

GNP--X1, 0.44; X2, 0.30 Real GNP--X1, 0.78; X2, 0.68 .

These results accord with expectations, since the growth rates in constant-dollar GNP varied considerably more than those in current-dollar GNP. The ratios of forecast error to extrapolation error average about 0.4 when X1 is the standard, 0.5 to 0.6 when X2 is, and the results are much the same for either variable.

Table 3 surveys the performance of forecasts of percentage changes in the price level (IPD) that match the real GNP predictions covered in Table 2. On the average, the predicted inflation rates fall short of the actual ones by fractions of one percentage point (lines 1-3). The 1959-67 forecast sets are less accurate than simple last-change extrapolations (line 4), and the other sets outperform the naive models by relatively small margins, much less than those observed for the GNP series. The naive models work comparatively well here, with errors averaging about 3/10 of the actual changes in IPD. Projections of the last change are in this case better than those of the average

TABLE 3

SUMMARY MEASURES OF ERROR FOR ANNUAL PREDICTIONS OF PERCENTAGE CHANGES IN THE PRICE LEVEL, 1959-76

Line	Period and No. of Years Covered	Selected Private Forecasts, Mean[a] (1)	ASA/NBER Survey, Median Forecast (2)	Economic Report of the President[b] (3)	Michigan Model (4)	Wharton Model (5)	Extrapolations: Last Change (6)	Extrapolations: Average Change (7)	Actual: Preliminary (8)	Actual: Revised (9)
		Mean Absolute Percentage Change, Predicted and Actual								
1	1959-67 (9)	1.5			1.9				1.9(1.4)[c]	2.0
2	1962-76(15)			3.7	3.8				4.2	4.5
3	1969-76 (8)		4.9	5.3	5.0	5.3			5.9	6.2
		Mean Absolute Error, in Percentage Points								
4	1959-67 (9)	0.6			0.7		0.3	0.7	0.3	
5	1962-76(15)			1.0	1.0		1.3	1.4	0.4	
6	1969-76 (8)		1.3	1.4	1.4	1.4	2.0	2.1	0.4	
		Mean Error, in Percentage Points								
7	1959-67 (9)	0.2			0		-0.1	0.04	-0.3	
8	1962-76(15)			-0.5	-0.5		-0.2	-1.0	-0.3	
9	1969-76 (8)		-0.9	-0.6	-0.9	-0.6	-0.2	-1.2	-0.4	
		Squared Correlation (r^2) Between Predicted and Actual Change								
10	1959-67 (9)	.389			.424		.365	.068[d]		
11	1962-76(15)			.768	.682		.536	.508		
12	1969-76 (8)		.526	.581	.454	.604	.166	.059		

NOTE: For sources and explanations of the data used in columns 2-9, see footnotes d through k, respectively, in Table 1.

[a] Average of end-of-year annual forecasts of the composite price level (a weighted sum of forecasts of CPI and WPI). See Table 2, fn. a, on the weights used and sources.

[b] See Table 2, fn. b.

[c] The figure in parentheses is based on weighted averages of data for CPI and WPI. This series of actuals for the composite price level is comparable to the forecasts used in column 1 only (cf. Table 2, fn. c).

[d] r is negative.

change (cols. 6-7), which is the reverse of the situation for GNP in both current and constant dollars. The forecasts underestimated strongly (much more than the last-change extrapolations) the average inflation since 1961 (lines 7-9). The predicted and actual percentage changes in the price level are all positively correlated, but the correlations for 1969-76 are generally lower than their counterparts for GNP and, still more so, for real GNP (lines 10-12).

Forecasts of inflation often have much in common with projections of the last observed rate of inflation. To illustrate, correlations between the errors of these forecasts and the errors of the corresponding extrapolations produce the following r^2 coefficients: Michigan, 1959-76: 0.51; CEA, 1962-76: 0.78; ASA/NBER, 1969-76: 0.95; Wharton, 1969-76: 0.80. For growth rates in real GNP, the correlations between forecast errors and extrapolation errors are also positive but throughout lower, in most cases much lower. These results are not surprising and they have a positive aspect inasmuch as forecasts should be closer to extrapolations of a given type in those cases where such extrapolations are more effective (for an elaboration, see Mincer and Zarnowitz). However, our comparisons are constrained to naive models which presumably do not represent high standards for economic forecasting. In particular, price-level forecasts that are highly correlated with last-change extrapolations must share the property of the latter to lag a year behind the actual rates of inflation. Indeed, the correlations between the predicted changes and the previous year's actual changes are all positive and high: the r^2 coefficients for the four sets of IPD forecasts listed earlier in this paragraph are 0.76, 0.87, 0.81, and 0.72, respectively.

The annual percentage changes in real GNP are inversely related to those in IPD and positively related to those in current-dollar GNP, while the last

two variables do not show a strong or stable association. The relationships between the predicted changes generally parallel the actual ones. This is illustrated by the r^2 coefficients tabulated below (for symbols, see Table 4).

	1962-76			1969-76		
	Actual	Michigan	CEA	Actual	Michigan	CEA
RGNP-IPD	.567(-)	.328(-)	.528(-)	.646(-)	.472(-)	.651(-)
RGNP-GNP	.297	.210	.222	.644	.464	.491
IPD-GNP	.020	.217	.068	.085(-)	.004	.022(-)

The errors of the forecasts are similarly interrelated. Table 4 demonstrates a pervasive pattern of negative correlation between errors in forecasting real growth and inflation (col. 1). The tendency for these errors to be offsetting, which benefits the forecasts of GNP in current dollars, is most strongly in evidence for the more recent years. When forecasters overestimated real growth, or missed a downturn and projected continued growth instead, they typically also underestimated inflation, as in 1969-71 and 1973-74. Underprediction of real growth occurred in 1972 and 1975-76 in combination with overprediction of inflation.

These observations, which have some precedents (Zarnowitz, 1969; Moore, 1969, 1977), are consistent with a view of the world in which nominal GNP changes are predicted directly and relatively well, but their division into real and price changes continues to pose great problems. Many forecasters may agree with that view in general terms, and some subscribe to models consistent with it (a specific example might be the St. Louis model in which the dollar change in total GNP expenditure is determined mainly by the dollar change in a measure of money stock). However, most macroeconometric models, including the two sets covered here, have separate aggregate real demand, output, and price level equations, and it is not at all clear why they should predict GNP

TABLE 4

CORRELATIONS BETWEEN ERRORS OF FORECASTS OF PERCENTAGE CHANGES IN NOMINAL GNP, REAL GNP, AND IPD, 1962-76

Line	Source of Forecast	Squared Correlation (r^2) Between Forecast Errors for RGNP and IPD (1)	for RGNP and GNP (2)	for IPD and GNP (3)
		1962-76 (15 years)		
1	Economic Report (CEA)	.297(-)	.359	.114
2	Michigan model	.494(-)	.429	.006
		1969-76 (8 years)		
3	Economic Report (CEA)	.677(-)	.004	.259
4	Michigan model	.684(-)	.209	.014
5	Wharton model	.340(-)	.036	.466
6	ASA/NBER survey, median	.524(-)	.013	.351

NOTE: The symbols RGNP, IPD, and GNP denote real GNP, the implicit price deflator, and nominal GNP, respectively. The correlations (r) are positive, except where the sign (-) following the r^2 coefficient indicates that r is negative.

better in current than in constant dollars. In fact, some studies of the recent performance of quarterly models arrive at the opposite conclusion, namely that the results for real GNP are better than those for nominal GNP because of deficient price forecasts (Duggal, Klein, and McCarthy; Eckstein, Green, and Sinai). The available evidence seems too limited and too mixed to permit any conclusive generalizations on this point. But it is interesting to observe that the importance of output errors vs. price errors may vary with changes in the relative roles of real vs. nominal factors and disturbances: in the 1970's the errors of the GNP forecasts were for the most part better correlated with the IPD errors than with the RGNP errors, whereas in the 1960's the contrary situation obtained (Table 4, cols. 2-3).

IV. Quarterly Multiperiod Forecasts, 1970-75: An Overall Appraisal

Here we have space only for a summary of some early results from a study in progress. The forecasts and actual data are used in the same form as before, but they now refer to overlapping sequences of quarters, not simply to a series of successive years. Our materials cover 22 quarters from 1970:3 through 1975:4, a period for which forecasts from several new sources are available. First estimates for the preceding year, taken from the data prior to the 1976 benchmark revision of the national income accounts, serve as comparable realizations. The full version of the study will include also comparisons with the revised data in an integrated treatment of forecast errors and measurement errors. Adjustments of the forecasts for base revisions, used in some forecast evaluations, are regarded as questionable and are avoided.

The mean absolute errors of GNP forecasts are close to one percentage point (like the annual forecasts, see Table 1) for two quarters ahead, and about half of that or less for one quarter ahead. Over longer spans, the MAE rise more or less steadily by increments varying from 0.3 to 0.5 of one percentage point for each additional quarter; they approach and exceed 2 percentage points for 4-quarter and 5-quarter spans, and 3 percentage points for 7-quarter and 8-quarter spans, respectively (Table 5, lines 1-3). Consistent with earlier findings and interpretations for various types of multiperiod forecasts (see, e.g., Zarnowitz, 1967, pp. 60-72), the MAE increase less than in proportion to the extension of the span. The errors in forecasts of percentage changes expressed on a per-unit-of-time basis (roughly, errors divided by the length of the effective span) neither rise nor decline systematically as the forecast reaches further into the future. The same applies to the errors of the implicit predictions of changes during the successive

TABLE 5

SUMMARY MEASURES OF ERROR FOR QUARTERLY MULTIPERIOD PREDICTIONS OF PERCENTAGE CHANGE IN GNP, 1970-75

Line	Forecast Set[a]	Span of Forecast in Quarters[b]							
		One (1)	Two (2)	Three (3)	Four (4)	Five (5)	Six (6)	Seven (7)	Eight (8)
		Mean Absolute Error, in Percentage Points (MAE)[c]							
1	Chase	.42	1.03	1.32	1.68	2.22	2.73	3.19	3.49
2	DRI	.53	1.04	1.43	1.94	2.43	2.69	2.95	2.80
3	GE	.42	.95	1.34	1.71	2.19	2.59	2.88	3.25
		Mean Error, in Percentage Points (ME)[d]							
4	Chase	.01	.04	.02	.08	-.14	-.66	-1.48	-2.34
5	DRI	-.01	.11	.05	.11	.01	-.42	-1.12	-1.69
6	GE	-.14	-.15	-.30	-.15	-.15	-.44	-.95	-1.68
		Squared Correlation (r^2) Between Predicted and Actual Change							
7	Chase	.752	.451	.107	.058	.127	.134	.179	.293
8	DRI	.632	.469	.069	.000*	.008	.102	.249	.600
9	GE	.753	.577	.284	.159	.132	.180	.227	.225
		Theil's Inequality Coefficient (U)[e]							
10	Chase	.241	.287	.294	.268	.236	.218	.198	.189
11	DRI	.284	.292	.299	.295	.272	.218	.182	.148
12	GE	.245	.260	.260	.243	.233	.204	.181	.172

[a] Chase: Chase Econometric Associates, Inc.; DRI: Data Resources, Inc.; GE: MAPCAST group at the General Electric Company. The forecast data are those used and described in S. K. McNees, 1975, 1976. Chase and DRI are "early-quarter forecasters," while GE is a "late-quarter forecaster" (for the release dates, see McNees, 1976, p. 41).

[b] Number of forecasts covered (n) for spans 1 to 7, respectively: 22, 21, 20, 19, 18, 17, and 16 (for each of the sets). For span 8, the number is 15 (Chase), 14 (DRI), and 12 (GE).

[c] Defining the predicted change and the actual change (for the given set, variable, period, and span) as P_t and A_t, respectively, $MAE = \frac{1}{n} \Sigma |e_t|$, where $e_t = P_t - A_t$.

[d] $ME = \frac{1}{n} \Sigma e_t$. [e] $U = \sqrt{\Sigma e_t^2 / \Sigma A_t^2}$, separately for each span. * r is negative.

single quarters covered; it is the cumulation of these intraforecast ("marginal") change errors that technically accounts for the tendency of errors in the total predicted changes to grow with the span.[3]

Where both forecasts and realizations refer to increases (as they do most of the time by far in the case of GNP), errors of positive sign denote overestimation of actual change. The mean errors in Table 5, lines 4-6, are predominantly small and positive, except for the longer spans where some of them are large and negative. As will be shown below, these averages conceal large errors of opposite sign in the forecasts for some of the different economic phases of the period 1970-75.

The r^2 coefficients for the correlations between the predicted and actual changes in GNP exceed 0.6 or 0.7 for one quarter ahead (like the annual forecasts) and exceed 0.4 or 0.5 for two quarters ahead. They are much smaller for the longer spans, mostly in the 0.1-0.25 range, in a few cases near zero (lines 7-9).

Theil's inequality coefficients generally fall between 0.2 and 0.3 (lines 10-12). This indicates that these forecasts are all much better than a naive model extrapolating the last recorded percentage change (for which $U = 1$). That model, it should be noted, is but a minimal standard for economic forecasts. Interestingly, the U coefficients do not increase with the forecast span; in fact, they decline slightly below .2 for the longest spans.

The next two tables have the same format as Table 5, which facilitates presentation and comparisons of these measures. Real GNP forecasts have MAE (in percentage points) rising from 0.5-0.6 for one quarter ahead to 5-6 for

---

[3]Note that fewer observations are available for the longer spans (Table 5, fn. b). This reduces the comparability of the measures reported for the different spans, but does not eliminate it.

TABLE 6

SUMMARY MEASURES OF ERROR FOR QUARTERLY MULTIPERIOD PREDICTIONS OF PERCENTAGE CHANGE IN REAL GNP, 1970-75

Line	Forecast Set[a]	Span of Forecast in Quarters[b]							
		One (1)	Two (2)	Three (3)	Four (4)	Five (5)	Six (6)	Seven (7)	Eight (8)
		Mean Absolute Error, in Percentage Points (MAE)[c]							
1	Chase	.51	1.11	1.81	2.46	3.29	4.19	4.95	5.31
2	DRI	.61	1.37	2.08	2.75	3.52	4.15	4.78	5.58
3	GE	.50	1.20	1.75	2.15	2.80	3.80	4.76	5.15
		Mean Error, in Percentage Points (ME)[d]							
4	Chase	.17	.51	.92	1.46	1.98	2.38	2.63	2.82
5	DRI	.26	.77	1.20	1.82	2.59	3.16	3.66	4.72
6	GE	.00	.22	.36	.95	1.53	2.09	2.46	2.58
		Squared Correlation (r^2) Between Predicted and Actual Change							
7	Chase	.839	.817	.727	.703	.733	.710	.604	.596
8	DRI	.793	.745	.598	.584	.785	.827	.741	.638
9	GE	.808	.741	.677	.607	.772	.764	.661	.662
		Theil's Inequality Coefficient (U)[e]							
10	Chase	.433	.502	.607	.673	.711	.741	.758	.741
11	DRI	.504	.622	.721	.769	.781	.774	.774	.836
12	GE	.427	.498	.548	.606	.627	.676	.714	.694

Footnotes a through e: See the corresponding footnotes in Table 5.

eight quarters ahead, that is, somewhat more than in proportion to the measured span (Table 6, lines 1-3). The errors for the two shortest spans are not much larger than those for GNP in current dollars, but the errors for the longest spans are 50 to 100 percent larger. The unusually rapid build-up of the MAE can be traced in large part to turning point errors. In quarterly multiperiod forecasting, turning points are more frequent and more difficult to predict than in annual forecasting, but the errors associated with them matter much more yet: here, missing a turn often means that a whole chain of predictions for the subsequent observations is badly off.

The mean errors of these forecasts are all positive, which is largely due to the effects of missing or underestimating the declines in real GNP during the recession (Table 6, lines 4-6). The ME also cumulate continuously and rapidly here, quite unlike those for the nominal GNP forecasts. On the other hand, the r^2 coefficients are rather surprisingly high in Table 6, lines 7-9, much above the corresponding figures for GNP in Table 5, particularly for spans of 3-8 quarters. Relative to the size of the actual changes, however, the real GNP errors are much larger than the current-dollar GNP errors: the inequality coefficients rise from .4-.5 to .7-.8 (Table 6, lines 10-12).

The MAE of forecasts of inflation in terms of the GNP implicit price deflator are like those of the GNP forecasts for the shortest spans--0.5 or less one quarter ahead, approximately 1 percentage point two quarters ahead--but they cumulate rapidly, especially for the longest spans (Table 7, lines 1-3). The figures for the eight-quarter-ahead predictions are here more than 12 times as large as those for the one-quarter-ahead predictions. This exceptionally strong build-up of errors reflects a progression of underestimates of the inflation rates, rising more than in proportion to the span extensions (Table 7, lines 4-6).

TABLE 7

SUMMARY MEASURES OF ERROR FOR QUARTERLY MULTIPERIOD PREDICTIONS OF PERCENTAGE CHANGE IN THE PRICE LEVEL, 1970-75

Line	Forecast Set[a]	Span of Forecast, in Quarters[b]							
		One (1)	Two (2)	Three (3)	Four (4)	Five (5)	Six (6)	Seven (7)	Eight (8)
		Mean Absolute Error, in Percentage Points (MAE)[c]							
1	Chase	.39	1.02	1.64	2.29	2.98	3.87	4.88	5.69
2	DRI	.54	1.11	1.69	2.37	3.05	4.04	5.17	6.78
3	GE	.39	.90	1.49	1.96	2.37	3.06	4.08	4.79
		Mean Error, in Percentage Points (ME)[d]							
4	Chase	-.15	-.49	-.96	-1.50	-2.33	-3.31	-4.42	-5.48
5	DRI	-.27	-.70	-1.22	-1.85	-2.82	-3.88	-5.12	-6.78
6	GE	-.12	-.36	-.70	-1.20	-1.85	-2.78	-3.76	-4.57
		Squared Correlation (r^2) Between Predicted and Actual Change							
7	Chase	.600	.440	.394	.287	.246	.233	.320	.381
8	DRI	.478	.426	.412	.346	.401	.398	.384	.371
9	GE	.657	.633	.508	.440	.438	.457	.524	.676
		Theil's Inequality Coefficient (U)[e]							
10	Chase	.311	.358	.377	.410	.438	.462	.480	.496
11	DRI	.375	.388	.397	.422	.444	.475	.508	.540
12	GE	.284	.286	.326	.354	.369	.395	.416	.410

Footnotes a through e: See the corresponding footnotes to Table 5.

The r^2 coefficients for the IPD inflation forecasts are generally higher than those for the forecasts of percentage change in nominal GNP (except for a few short predictions) but throughout lower than the corresponding statistics for the real growth forecasts. They range from .23 to .66 and tend to decrease as the spans lengthen (Table 7, lines 7-9). The U coefficients (lines 10-12) are close to .2 for the shorter spans and close to .5 for the longest; they are thus higher than their counterparts for the current-dollar GNP forecasts but lower than those for the real GNP forecasts.

The quantity and price ingredients of the GNP forecasts show a pattern of offsetting errors in the quarterly as well as annual data. The mean errors of real growth predictions are all positive, those of inflation predictions all negative, and these statistics, matched by source and span, have similar absolute values for most of the shorter forecasts As a result, the ME of the current-dollar GNP forecasts for spans 1-4 are as a rule positive but very small for Chase and DRI, negative but small for GE. The negative ME of the inflation forecasts outweigh the positive ME of the real growth forecasts in spans 5-8, so that the ME of the percentage change forecasts for GNP are predominantly negative and substantial (cf. lines 4-6 in Tables 5, 6, and 7).

The surveyed accuracy measures do not show any of the forecasters to be consistently superior to the others. They do favor GE over Chase and DRI in most instances, but by modest margins and in a way that can be explained by an advantage in timing: the GE forecasts are issued late in each quarter, the others early.

The results reported in this study should and will be carefully compared with those of other evaluations of the same forecasts, but the task is still to be completed. The error measures used here differ in several respects from those used by others, and additional computations are required to allow for these differences.

V. Quarterly Multiperiod Forecasts: An Analysis by Subperiods

The period 1970:3-1975:4, although short, was unusually varied and marked by major disturbances and drastic changes in the economy's course. It is useful to divide it into the following parts, as suggested by the contemporary business-cycle and inflationary developments.

I. 1970:3-1973:1. End of the mild 1970 recession followed by an expansion that accelerated in 1972, with relatively stable inflation.

II. 1973:1-1973:4. Slower real growth and a sharp inflation speedup (materials shortages, run-ups in commodity prices, oil embargo).

III. 1973:4-1975:1. Recession, severe in its last two quarters, accompanied first by a further rise and then by a downturn in the rate of inflation.

IV. 1975:1-1975:4. Sharp upturn and the initial recovery phase, with a further decline in inflation.

One question is whether forecasts that originated in these four subperiods show significantly different characteristics and performance. The other is whether forecasts for these subperiods (i.e., those that aimed at the corresponding groups of target quarters) are so differentiated. It turns out that the answers to both questions are definitely yes.

To illustrate the first point, the expansion phase I produced forecasts that underestimated growth in dollar GNP mainly because they underestimated inflation. The percentage changes in real GNP were partly underpredicted, partly (in some longer forecasts) overpredicted, but whether negative or positive the ME of these forecasts were small. In general, the record of the

forecasts that were made during the period I was good in terms of both the ME and the MAE figures, even for the long spans. In contrast, the slowdown phase II produced real growth predictions with very large positive ME and inflation forecasts with very large negative ME (underestimation errors). These errors balanced each other so that the ME for the nominal GNP predictions were moderate (and mostly negative, except for the longest forecasts). The recession phase III gave rise to even larger positive mean errors in the real growth forecasts as the declines were repeatedly missed and, when finally recognized, underestimated. These errors were larger absolutely than the negative errors on the price side, which reflected a continuing underestimation of inflation, so that the predictions of the growth rates in nominal GNP had consistently positive ME in the subperiod III.

The above summary is based on charts (not reproduced here) which show the average errors (MAE and ME) by span and by subperiod in which the forecasts originated.[4] These charts look very similar for such different models as Chase and DRI: they show in each case the same striking differences between the forecasts made in subperiods I, II, and III. The suggested inference is that concurrent predictions from different sources and models have common patterns such that their errors depend strongly and similarly on the characteristics of the time of their origin.

In a second exercise, the forecasts were assigned to the four subperiods according to their target quarters, not their base quarters, as illustrated In Charts 1-3. Here the samples are partitioned differently, hence the resulting patterns diverge from those obtained on the first plan, but the

---

[4] No averages for phase IV are used on this basis, since they contain too few observations in the truncated sample.

conclusion is analogous: the type and size of forecast errors depend critically on the economic properties of the target periods vis-à-vis those of the periods of origin. Forecasters perform best when the two periods are alike, belonging to the same already recognized phase, e.g., a continuing expansion as in 1971-72 (most of subperiod I). They perform worst when the target falls into a new phase, particularly when the latter departs sharply from the currently established pattern (forecasts made in subperiods II and III, and those for subperiods III and IV provide many examples, particularly in the long-span categories). Such period characteristics are much more important determinants of forecast errors than are any differences among the forecasters.

Chart 1 shows that both Chase and DRI persistently underestimated the percentage changes in GNP for subperiods I and, much more strongly, II. Both forecasters overestimated the changes in their short forecasts for subperiod III and underestimated them in their long forecasts for the same phase. Overestimates prevailed in all forecasts for the last phase covered, IV, and here the average errors behave in an unusual fashion, first increasing and then decreasing with the lengthening span. This is due to offsets between the real growth forecasts with positive ME and the inflation forecasts with negative ME (see Charts 2 and 3).

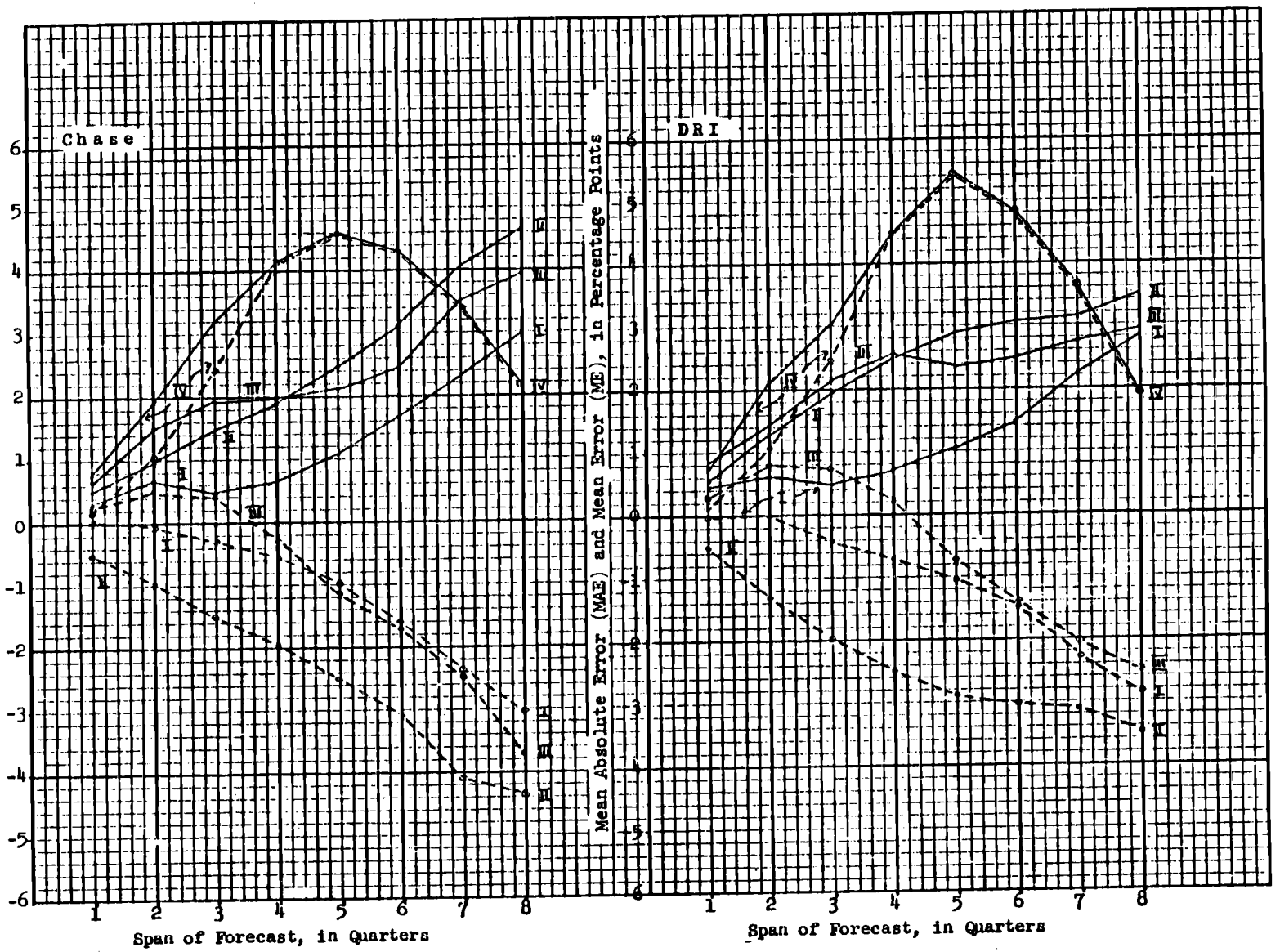
Chart 2, which covers the real growth forecasts, shows underestimates dominating the errors for I and II, much larger MAE and positive ME for III and IV. The huge average errors for the two latter phases derive mainly from the forecasters' failure to predict the declines in real GNP.[5] The long-span errors for the 1975 recovery (IV) are strikingly large here.

---

[5]The change errors, $P_t - A_t$ (see Table 5 for the symbols), are positive where $P_t > 0$ and $A_t < 0$, and also where $P_t < 0$ and $A_t < 0$ but $|P_t| < |A_t|$. These cases dominate in Chart 2 the results for both the recession phase III and the recovery phase IV. Although real GNP reached a trough in 1975:1 and increased thereafter, the actual changes over longer spans ending in 1975 are negative; that is, real GNP was lower during period IV than in 1973 (II) and during most of 1974 (III).

CHART 1

QUARTERLY MULTIPERIOD FORECASTS OF PERCENTAGE CHANGES IN GNP,
AVERAGE ERRORS BY SUBPERIOD AND SPAN, TWO MODELS 1970-75

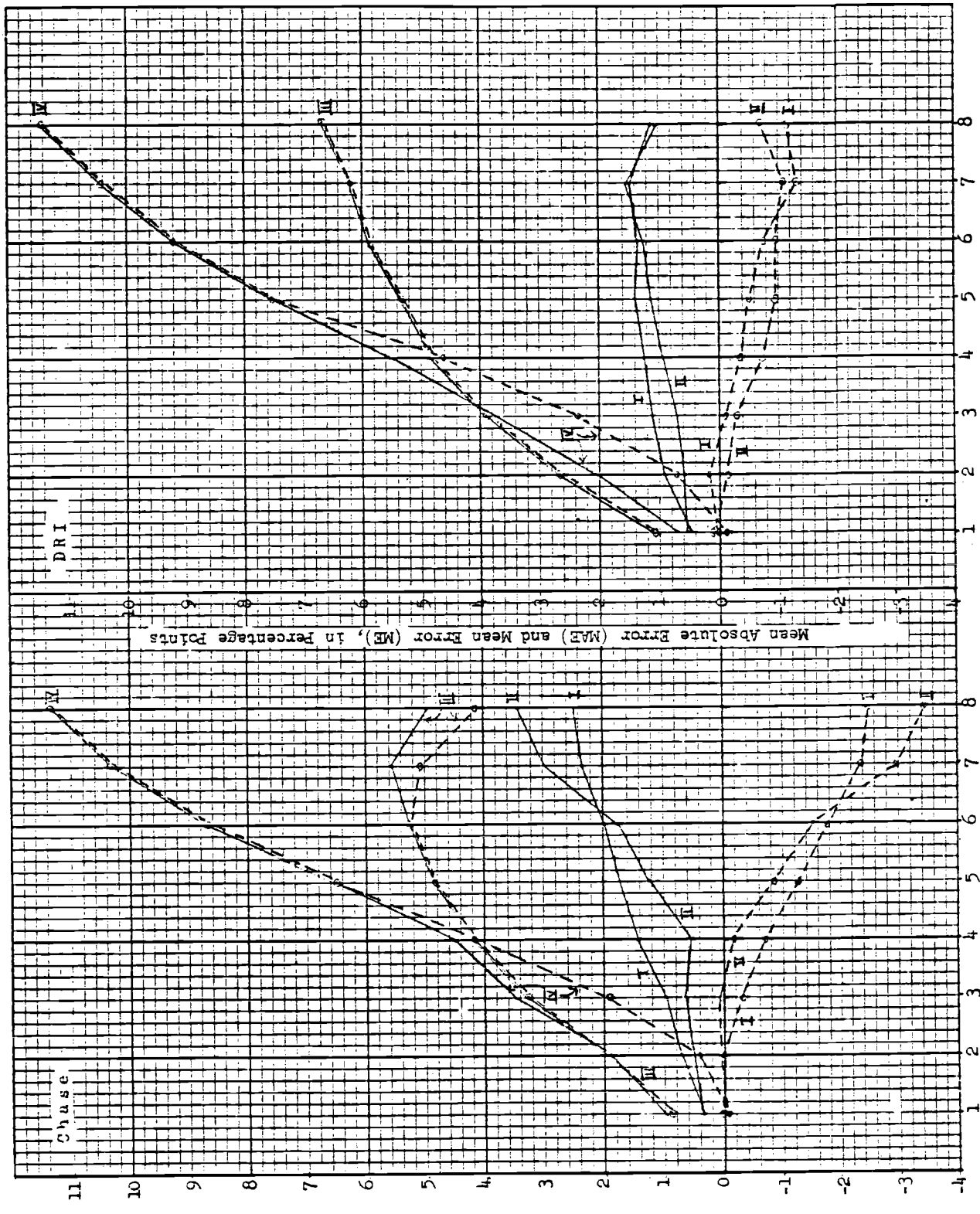


Legend:

I	1970:3-1973:1	n = 11, 10, ..., 4 for spans 1, 2, ..., 8 quarters, respectively.
II	1973:1-1973:4	n = 4 for each span.
III	1973:4-1975:1	n = 6 for each span.
IV	1975:1-1975:4	n = 4 for each span.
Solid lines	——•——	refer to mean absolute errors (MAE).
Broken lines	---•---	refer to mean errors (ME).

Chart 2

QUARTERLY MULTIPERIOD FORECASTS OF PERCENTAGE CHANGES IN REAL GNP,
AVERAGE ERRORS BY SUBPERIOD AND SPAN, TWO MODELS, 1970-75



Legend: See Chart 1 (p. 31).

The inflation forecasts also had very large average errors in the recession period (III), as demonstrated in Chart 3. Their errors were relatively small in subperiod I, considerably larger and on balance all negative in II; in neither phase did they increase strongly with the span. The short forecasts for phase IV had small errors; in several cases their means are positive, indicating overestimation of inflation rates that just began to decline. The long forecasts for this last subperiod, however, had very large underestimation errors.

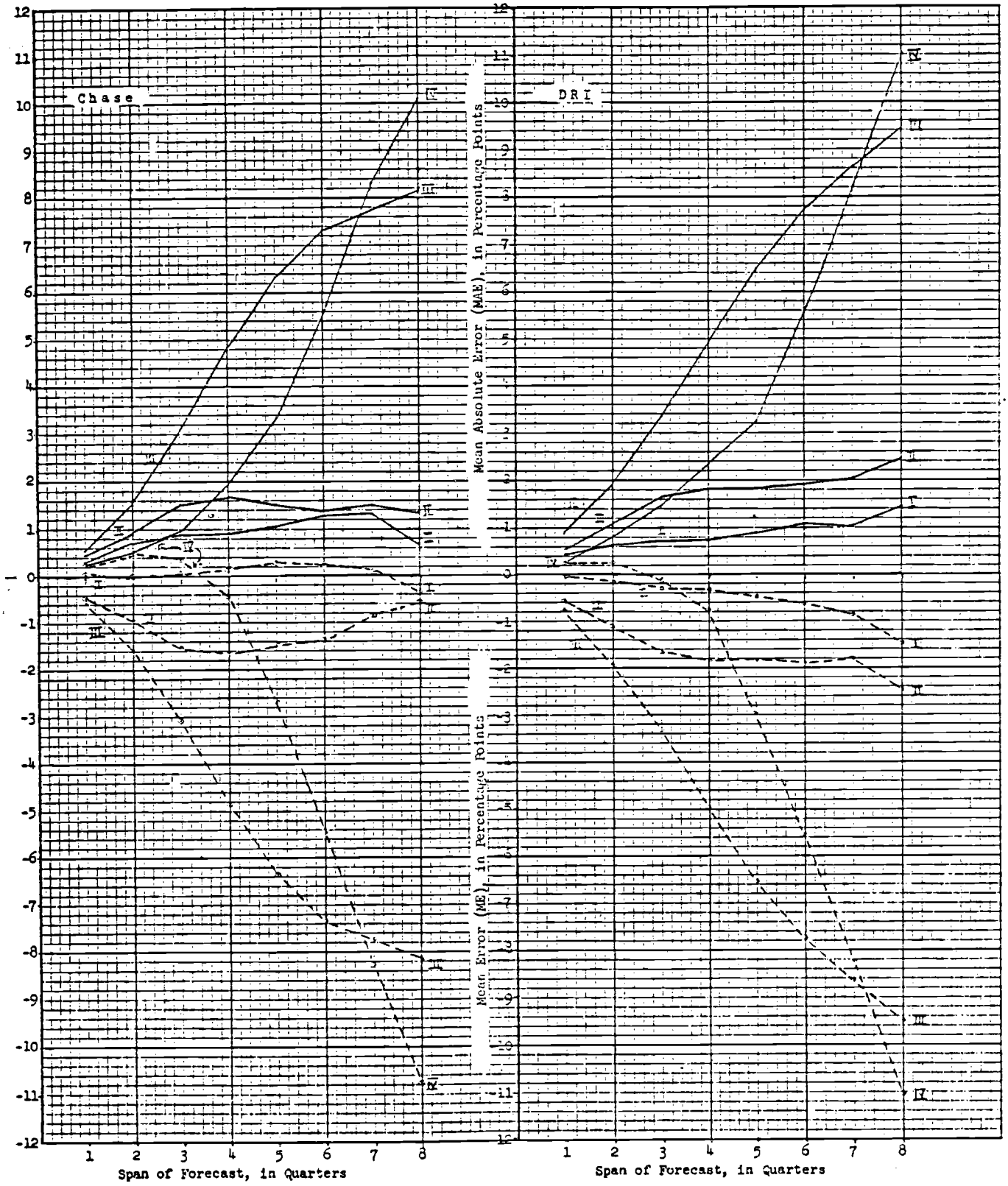
VI. Concluding Observations

The end-of-year forecasts of annual percentage changes in GNP earn good marks for overall accuracy when judged according to realistic rather than ideal standards. Moreover, they are found to have improved in the period since the early 1960's compared with the previous years after World War II.

The corresponding forecasts for GNP in constant dollars (real growth) and IPD (inflation) are weaker. The former suffer from large turning-point errors, the latter from large underestimation errors. But the errors in forecasts of real growth are negatively correlated with the errors in forecasts of inflation, which helped to make the nominal GNP predictions more accurate. In recent times, these correlations were connected with the unexpected concurrence of accelerating inflation and slowing, then declining output rates: optimistically, and probably also from a lingering faith in a simple Phillips trade-off, forecasters kept anticipating less inflation and more growth. But in the late 1950's and early 1960's, it was the relative stability of the price level that caused widespread surprises and offsetting errors resulted from the opposite combination of overestimates of inflation and underestimates of real growth.

CHART 3

QUARTERLY MULTIPERIOD FORECASTS OF PERCENTAGE CHANGES IN THE IMPLICIT PRICE DEFLATOR, AVERAGE ERRORS BY SUBPERIOD AND SPAN, TWO MODELS, 1970-75



Legend: See Chart 1 (p. 31).

Forecasts of inflation are not much better than projections of the most recently observed inflation rates, and they lag behind the actual rates much like such projections. The deficiency of price-level forecasts, documented in this and other studies, surely impairs the general ability of economists to analyze the prospects for the economy. Improvements will require major advances in our knowledge, presumably through research based on carefully worked out data (abstract speculation abounds but good information and observation are rare in this area).

The favorable record of annual GNP predictions does not imply that forecasters can perform well the more difficult task of predicting quarterly changes in GNP within the year ahead or even beyond it. Forecasts for the year as a whole can be satisfactory when based on a good record for the first two quarters; they tend to be more accurate than forecasts with longer spans.[6] An examination of the recent multiperiod predictions shows that the errors for real GNP and IPD cumulated rapidly beyond the spans of 2 to 4 quarters. Previous studies have shown the cumulation to be as a rule less than proportional to the increase in the span, but in this period the build-up of errors was much greater than usual. No doubt, in less turbulent times the longer forecasts can be considerably more accurate, but this fair-weather argument is not very persuasive or helpful. At the present time, the predictive value of detailed forecasts reaching out further than a few quarters ahead must be rather heavily discounted. Again, what is critical here is theoretical analysis and empirical research that would lead to improvements in our ability to

---

[6] Also, errors of predictions for the individual parts of the year at times offset each other to some degree (Zarnowitz, 1967; McNees, 1973, 1974). These gains from aggregation over time resemble those from aggregation over sectors (GNP is predicted with smaller average errors of relative change than are most of its components; see Zarnowitz, 1967, 1972; Fromm and Klein, 1976).

predict broad movements in the price level and business-cycle turning points. Despite setbacks, there is still no reason to give up moderate hopes for an ultimate advance on these fronts.

References

J. A. Carlson, "A Study of Price Forecasts," Annals of Economic and Social Measurement, Winter 1977.

C. F. Christ, "Judging the Performance of Econometric Models of the U.S. Economy," International Economic Review (IER), February 1975, pp. 54-74.

V. G. Duggal, L. R. Klein, and M. D. McCarthy, "The Wharton Model Mark III: A Modern IS-LM Construct," IER, October 1974, pp. 572-594.

O. Eckstein, E. W. Green, and A. Sinai, "The Data Resources Model: Uses, Structure, and Analysis of the U.S. Economy," IER, October 1974, pp. 595-615.

M. K. Evans and L. R. Klein, The Wharton Econometric Forecasting Model, University of Pennsylvania, Philadelphia, 1968.

M. K. Evans, L. R. Klein, and M. Saito, "Short-Run Prediction and Long-Run Simulation of the Wharton Model," in Econometric Models of Cyclical Behavior, pp. 139-191 (see Hickman, 1972).

W. Fellner, Towards a Reconstruction of Macroeconomics, American Enterprise Institute, Washington, D.C., 1976, pp. 118-124.

G. Fromm and L. R. Klein, "The NBER/NSF Model Comparison Seminar: An Analysis of Results," Annals of Economic and Social Measurement, Winter 1976, pp. 4-5.

Y. Haitovsky, G. L. Treyz, and V. Su, Forecasts with Quarterly Macroeconometric Models, National Bureau of Economic Research (NBER), New York, 1974.

M. Hatanaka, "The Underestimation of Variations in the Forecast Series: A Note," IER, February 1975, pp. 151-160.

B. G. Hickman, ed., Econometric Models of Cyclical Behavior, Studies in Income and Wealth, no. 36, 2 volumes, NBER, New York, 1972.

A. A. Hirsch, B. T. Grimm, and G. V. L. Narasimham, "Some Multiplier and Error Characteristics of the BEA Quarterly Model," IER, October 1974, pp. 617-631.

S. H. Hymans and H. T. Shapiro, The DHL-III Quarterly Econometric Model of the U.S. Economy, RSQE, University of Michigan, Ann Arbor, 1970.

_____, "The Structure and Properties of the Michigan Quarterly Model of the U.S. Economy," IER, October 1974, pp. 632-653.

L. R. Klein, "A Post-Mortem on Transition Predictions of National Product, Journal of Political Economy, August 1946.

M. D. McCarthy, The Wharton Quarterly Econpmetric Forecasting Model--Mark III, University of Pennsylvania, Philadelphia, 1972.

S. K. McNees, "The Predictive Accuracy of Econometric Forecasts," New England Review, Federal Reserve Bank of Boston, September/October 1973.

_____, "How Accurate Are Economic Forecasts?" ibid., November/December 1974.

_____, "An Evaluation of Economic Forecasts," ibid., November/December 1975.

_____, "An Evaluation of Economic Forecasts: Extension and Update," ibid., September/October 1976.

_____, "An Assessment of the Council of Economic Advisers' Forecast of 1977," ibid., March/April 1977.

J. Mincer and V. Zarnowitz, "The Evaluation of Economic Forecasts," in Mincer, ed., Economic Forecasts and Expectations, NBER, New York, 1969, pp. 14-20.

G. H. Moore, "Forecasting Short-Term Economic Change," Journal of the American Statistical Association, March 1969, pp. 1-22.

_____, "The President's Economic Report: A Forecasting Record," National Bureau Reporter, NBER, New York, April 1977.

J. F. Muth, "Rational Expectations and the Theory of Price Movements," Econometrica, July 1961, pp. 315-335.

A. M. Okun, "A Review of Some Economic Forecasts for 1955-57," Journal of Business, July 1959, pp. 199-211.

P. A. Samuelson, "Optimality of Sluggish Predictors Under Ergodic Probabilities," IER, February 1976, pp. 1-7.

M. Sapir, "Review of Economic Forecasts for the Transition Period," Studies in Income and Wealth, vol. 11, NBER, New York, 1949, part IV.

W. A. Spivey and W. J. Wrobleski, "An Analysis of Forecast Performance," Proceedings of the Business and Economic Statistics Section, American Statistical Association, Washington, D.C. (forthcoming).

V. Su and J. Su, "An Evaluation of ASA/NBER Business Outlook Survey Forecasts," Explorations in Economic Research, NBER, New York, Fall 1975, pp. 588-618.

D. B. Suits, "Forecasting and Analysis with an Econometric Model," American Economic Review, March 1962, pp. 104-132.

B. N. Vaccara and V. Zarnowitz, "How Good Are the Leading Indicators?" ibid.

V. Zarnowitz, An Appraisal of Short-Term Economic Forecasts, NBER, New York, 1967.

_____, "An Evaluation of Price-Level Forecasts," 1968 Proceedings of the Business and Economic Statistics Section, American Statistical Association, Washington, D.C., 1969, pp. 284-289.

_____, "The ASA/NBER Survey of Forecasts by Economic Statisticians," The American Statistician, February 1969.

_____, "New Plans and Results of Research in Economic Forecasting," Fifty-first Annual Report, NBER, New York, 1971.

_____, "Forecasting Economic Conditions: The Record and the Prospect," in Zarnowitz, ed., The Business Cycle Today, NBER, New York, 1972.

_____, "How Accurate Have the Forecasts Been?" in W. F. Butler, R. A. Kavesh, and R. B. Platt, Methods and Techniques of Business Forecasting, Prentice-Hall, Englewood Cliffs, N.J., pp. 565-596.