Volume Title: An Appraisal of Short-Term Economic Forecasts

Volume Author/Editor: Victor Zarnowitz

Volume Publisher: NBER

Volume ISBN: 0-87014-445-6

Volume URL: http://www.nber.org/books/zarn67-1

Publication Date: 1967

Chapter Title: Time Span of Forecasts and Predictive Accuracy

Chapter Author: Victor Zarnowitz

Chapter URL: http://www.nber.org/chapters/c0940

Chapter pages in book: (p. 60 - 82)

5

TIME SPAN OF FORECASTS AND PREDICTIVE ACCURACY

AVERAGE ERRORS OF FORECASTS WITH DIFFERENT SPANS: GNP AND INDUSTRIAL PRODUCTION

Five of our forecast sets provide chains of predictions made at a given date for two or more successive periods, e.g., for the first and second halves, or the four quarters, of the coming year. These data have been used to analyze the performance of forecasts over different intervals between the current base and the future target (which is called the *span* of the forecast).

The evidence shows clearly that the average errors of short-term forecasts increase as the span increases. Table 9 demonstrates this for predictions of changes in GNP and industrial production, expressed in units of the predicted series.[1] But the decreasing accuracy of longer-span forecasts could also be shown in terms of the level errors or the relative change errors (predicted relative change minus actual relative change). The progression to larger errors appears in all summary measures (means, standard deviations, and root mean square errors) and in forecasts of variables with quite different characteristics.

The increase in the mean absolute errors with the extension of the predictive span is in general steady and substantial, though in some cases it weakens considerably at the longer end of the range, between a year and eighteen months (Table 9, lines 1–5 and 11–15). There can be no doubt that the predictions are considerably more accurate, in absolute terms, over the next three months than over the next six, and more accurate again over six months than over nine or twelve.

[1] The error of predicted change is $E_{\Delta(t+m)} = P_{\Delta(t+m)} - \Delta A_{t+m}$, where m is the span of the forecast. See footnote 1, Chapter 4, which refers to the case where $m = 1$ year.

Averages taken with regard to sign (lines 6–10 and 16–20) are negative for all GNP and industrial production forecasts with spans of six months or more, except for set G. The absolute values of these arithmetic means tend to increase with the span of forecast in all cases.

TABLE 9

Average Errors in Forecasts of Changes in GNP and Industrial Production for Spans from Three to Eighteen Months, 1947-63

		Span of Forecast (Months)					
Line	Forecast Set[a]	Three (1)	Six (2)	Nine (3)	Twelve (4)	Fifteen (5)	Eighteen (6)
	Gross National Product: Mean Absolute Errors (billion dollars)						
1.	A		10.4		17.6		
2.	E		11.9		20.1		
3.	C[b]	5.0	8.3	9.6	12.2	15.1	
4.	D	5.0	9.8	10.0	11.0		
5.	G[c]	5.7	8.8	13.0	14.8	16.2	16.6
	Gross National Product: Mean Errors (billion dollars)						
6.	A		-5.8		-11.9		
7.	E		-6.6		-13.4		
8.	C[b]	-1.7	-3.0	-3.4	-3.9	-10.1	
9.	D	2.0	-1.2	-0.1	-4.0		
10.	G[c]	4.9	6.8	9.7	10.0	13.8	13.6
	Industrial Production: Mean Absolute Errors (1947-49 = 100)						
11.	A		5.9		6.6		
12.	E		5.2		7.6		
13.	C[b]	2.1	4.3	5.0	5.7	7.1	
14.	D	2.8	5.6	6.5	5.7		
15.	G	2.7	4.9	6.7	8.1	8.0	7.8
	Industrial Production: Mean Errors (1947-49-100)						
16.	A		-1.4		-1.8		
17.	E		-0.7		-2.0		
18.	C[b]	0.1	-1.5	-2.6	-2.5	-4.4	
19.	D	0.4	-1.0	-1.3	-1.9		
20.	G	2.0	2.9	3.9	4.3	5.0	5.1

Notes to Table 9

[a]The years covered and numbers of observations per span (in parentheses) are as follows: A: 1947-49, 1955-56, 1958-63 (11); C: 1958-63 (22, 20, 19, 13, and 7 for spans of three to fifteen months, respectively); D: 1959-63 (9); E: 1956, 1960-63 (5) for GNP; E: 1951-63 (13) for industrial production; G: 1955-63 (16) for GNP; G: 1956-63 (13) for industrial production.

[b]The entries on this line are not strictly comparable because some of the forecast chains are "incomplete," so that the number of observations for different spans varies (see note a). However, adjustments were made to allow for this by deleting selected observations, and the resulting differences were not large. The progression to larger errors associated with the increase in span was found to be characteristic of these forecasts regardless of the adjustments.

[c]Based on forecasts in constant prices, as reported.

TYPES OF ERROR IN MULTIPERIOD FORECASTS

The minus signs of the arithmetic means in Table 9 reflect in large measure facts that are already familiar: most of the time the economy moved upward and forecasters expected it to do just that, but the predicted increases were often smaller than those that materialized on the record. However, a closer analysis of these forecasts discloses great variation in the type of error. The relative frequency of overestimates here is high, higher than in the annual forecasts, as shown by Table 10 which can be compared with Table 7. The frequency of directional errors is, as would be expected, much larger than in the annual forecasts (fewer directional changes occur in the annual than in the semiannual and quarterly figures, and fewer are also predicted).

In one set of forecasts, G, overestimates are actually more numerous than underestimates. These are highly optimistic predictions which show a tendency to overstate increases and to continue upward missing the decline. In view of the different characteristics of the forecasts in set G, the results for this set are shown separately in Table 10; the other sets, which have more in common, however, are handled as a group on this occasion.

As longer spans are taken, increases usually become more, and decreases less, numerous. The number of turning-point errors tends to decline, since fewer directional changes occur and fewer are expected over longer spans. In particular, the frequency of such errors is reduced

TABLE 10

Forecasts of GNP and Industrial Production with Different Spans: Distribution by Type of Change and Type of Error, 1947-63

Line	Forecast Set[b]	Type of Change[c]	Number of Forecasts of Changes (all spans)[a]: Total (1)	Underestimates (2)	Overestimates (3)	Turning-Point Errors (4)	Not Classified[d] (5)
			Gross National Product				
1.	A,C,D,E	Increases	135	67	47	17	4
2.	A,C,D,E	Decreases	14	4	2	7	1
3.	G	Increases	76	16	55	4	1
4.	G	Decreases	20	0	0	20	0
			Industrial Production				
5.	A,C,D,E	Increases	127	57	44	19	7
6.	A.C.D.E	Decreases	37	11	3	9	14
7.	G	Increases	50	18	29	3	0
8.	G	Decreases	28	2	3	19	4

[a]Underestimates: predicted change is less than actual change. Overestimates: predicted change exceeds actual change. Turning-point errors: sign of predicted change differs from sign of actual change.

[b]This table covers the same sets of forecasts as Table 9. For detail on periods and numbers of observations, see Table 9, note a.

[c]Increases and decreases refer to the direction of changes in the actual values (first estimates for the given series).

[d]Line 1: Includes three forecasts of no change and one case of numerical agreement between predicted and reported change.

Line 2: Perfect numerical agreement between predicted and reported change.

Line 3: Prediction of no change.

Line 5: Includes three forecasts of no change, three instances of zero reported change (associated with forecasts of increases), and one case of numerical agreement between predicted and actual change.

Line 6: Includes seven forecasts of no change, six instances of zero reported change (associated with forecasts of decreases), and one case of numerical agreement between predicted and actual change.

Line 8: Includes three forecasts of no change and one instance of zero reported change (associated with a forecast of decrease).

because there are fewer *declines* to be overlooked. Underestimates can be expected to gain with the increase in span, since we have already observed that the larger rises are particularly apt to be understated and since, the further away his target, the more cautious the forecaster is likely to be. Just as uncertainty grows with the span, so does the forecaster's "caution," which leads him to avoid predicting large changes in general and to discount the more distant changes.[2]

Comparisons for all forecast sets and spans confirm that the longer the forecast, the smaller the share of turning-point errors and the larger the share of underestimates. Table 11 summarizes the evidence for all sets except G by contrasting the shortest (three- and six-month) with the longer (nine-, twelve-, and fifteen-month) forecasts. The percentages of underestimates are about 44 for the short and 55 for the longer forecasts of GNP. For industrial production, the corresponding figures are 41 and 53 per cent. The proportions of overestimates are essentially stable and those of turning-point errors drop sharply for the longer forecasts.[3]

Overestimates rather than underestimates prevail among those forecasts in set G which are associated with increases in GNP and industrial production. The forecasts for periods of decreases show here a very high concentration of turning-point errors; most of the errors of this type are due to failure to foresee downturns. As the spans increase, both over- and underestimates gain at the expense of turning-point errors, but the shifts in the distribution of errors by type are not large. In short, predictions from this source had a tendency to be overoptimistic regardless of span.[4]

INTRAFORECAST CHANGES

Consider a chain of forecasts made at the time $t = 0$ for several future periods, say, the next four quarters. The predictions refer to levels at

[2] Although this argument seems plausible and in line with the notions of many forecasters, it is by no means conclusive. Greater uncertainty should show itself in lower probabilities being attached to forecasts rather than in biased forecasts. Regrettably, such probabilities are typically not stated by forecasters, who content themselves simply with point predictions.

[3] The latter fall from about 24 to 10 per cent for GNP and from 26 to 13 per cent for industrial production. These figures refer to all forecasts covered in Table 11, but the results for the dominant category of increases are similar.

[4] The proportions of under- and overestimates in the GNP forecasts of set G are: 13 and 52 per cent for the short, and 19 and 61 per cent for the longer predictions, respectively. For industrial production, the percentages are 25 and 42 for the short forecasts, 28 and 44 for the longer ones.

TABLE 11

Forecasts of GNP and Industrial Production with Spans of Under and Over Six Months: Distribution by Type of Change and Type of Error, 1947-63

Line	Span of Forecast[a] (months)	Type of Change[b]	Number of Forecasts of Changes: Total (1)	Under-estimates[c] (2)	Over-estimates[d] (3)	Turning-Point Errors[e] (4)
			Gross National Product			
1.	3 & 6	Increases	60	28	21	11
2.	3 & 6	Decreases	11	3	2	6
3.	9, 12, & 15	Increases	71	39	26	6
4.	9, 12, & 15	Decreases	2	1	0	1
			Industrial Production			
5.	3 & 6	Increases	52	22	20	10
6.	3 & 6	Decreases	16	6	2	8
7.	9, 12, & 15	Increases	68	35	24	9
8.	9, 12, & 15	Decreases	7	5	1	1

[a]This table covers forecasts A, C, D, and E; these are the same as in Table 10, except that the "not classified" observations are excluded here. Set G is not covered (see text).

[b]Increases and decreases refer to the direction of changes in actual values.

[c]Predicted change is less than actual change.

[d]Predicted change exceeds actual change.

[e]Sign of predicted change differs from sign of actual change.

$t + m$ where $m = 3, 6, 9, 12$ (months) and to changes over the intervals with increasing length, 0–3, 0–6, 0–9, and 0–12. Taking the differences between either the successive level or the successive change predictions (the results are the same), one gets the "intraforecast changes." These differences within a given chained forecast represent implicit predictions of changes over the successive *subperiods* covered. For example, in the four-quarter case, there are, in addition to the first change from the base (0–3), three intraforecast changes, 3–6, 6–9, and 9–12. They relate to slices of the future that are of equal length but are increasingly distant from the present. The sum of the errors of such marginal changes equals the error of the total change predicted over the entire span of the chain (here 0–12).[5]

Since this is so for each forecast chain, the well-known basic propositions about the (arithmetic) means and variances of sums apply here. Obviously, the sum of the mean errors of all subperiod changes must equal the mean error of the total change predicted per chain. Furthermore, the variance of the total change errors equals the sum of variances of the component change errors, plus the covariances. If the component errors were independent, the covariance terms would vanish. If these errors are correlated instead, which seems more likely, then the covariances will be nonzero, having the signs of the correlations among the errors.[6]

Unless the errors for the subperiods (the intraforecast changes) are *negatively* correlated to a sufficiently high degree,[7] they must clearly be cumulative. Since high negative correlations are unlikely, summary measures will probably show the errors of predicted changes to be larger, the greater the span: a longer forecast covers more of the subperiods over which the errors cumulate. Here, then, is another technical

[5] Every intraforecast change has the form $\Delta P_{ij} = P_{t+j} - P_{t+i}$, where i and j are two successive values assumed by the span m; e.g., the expression $(P_{t+6} - P_{t+3})$ would denote the difference between the simultaneously made predictions for two quarters and one quarter ahead. The error of ΔP_{ij} equals $E_{\Delta ij} = E_{t+j} - E_{t+i}$. The total change error $E_{\Delta(t+m)} = \Sigma E_{\Delta ij}$.

For brevity, let us illustrate these relations for a chain of only two forecasts for six and twelve months, and use simple self-explanatory symbols and subscripts, without the delta signs for changes. Then,

$$P_{0\text{-}12} = P_{0\text{-}6} + P_{6\text{-}12} \text{ and } E_{0\text{-}12} = E_{0\text{-}6} + E_{6\text{-}12} = (P_{0\text{-}6} - A_{0\text{-}6}) + (P_{6\text{-}12} - A_{6\text{-}12}).$$

[6] Returning to the simple example in footnote 5, let us adopt the shorthand subscripts 1 and 2 for the subperiods (0–6) and (6–12), respectively; the total period (0–12) can then be denoted by the subscript (1 + 2). In the usual symbols, the relation between the mean errors is

$$\bar{E}_{1+2} = \bar{E}_1 + \bar{E}_2,$$

and the relation between the variances is

$$\text{Var } (E_{1+2}) = \text{var } (E_1) + \text{var } (E_2) + 2 \text{ cov } (E_1, E_2) = S_{E_1}{}^2 + S_{E_2}{}^2 + 2rS_{E_1}S_{E_2},$$

where cov is covariance, var $= S^2$ is variance, and r is the coefficient of correlation between the errors E_1 and E_2.

[7] For var (E_{1+2}) to be less than var (E_1), the correlation between E_1 and E_2 must be such that (1) $r < 0$ and (2) $|r| > S_{E_2}/2S_{E_1}$. An analogous condition for var $(E_{1+2}) <$ var (E_2) is also readily derived from the relation given in footnote 6.

TABLE 12

Average Errors of Intraforecast Changes for GNP and Industrial Production, 1947-63

Line	Forecast Set[a]	0-3 (1)	0-6 (2)	3-6 (3)	6-9 (4)	6-12 (5)	9-12 (6)	12-15 (7)	15-18 (8)
		Interval of Predicted Change (months)[b]							
	Gross National Product: Mean Absolute Errors (billion dollars)								
1.	A		10.4			11.9			
2.	E		11.9			10.4			
3.	C[c]	5.0		5.4	4.7		4.8	4.6	
4.	D	5.0		5.1	6.0		5.6		
5.	G[d]	5.7		4.3	6.5		4.5	6.6	5.0
	Gross National Product: Mean Errors (billion dollars)								
6.	A		−5.8			−6.1			
7.	E		−6.6			−6.8			
8.	C[c]	−1.7		−1.6	−0.6		0.8	−2.3	
9.	D	2.0		−3.3	1.1		−3.9		
10.	G[d]	4.9		1.9	2.9		0.3	3.8	−0.2
	GNP: Correlations of Predicted with Actual Changes								
11.	A		.518			.094			
12.	E		.421			.136			
13.	C[c]	.658		.217	.164		.397	.215	
14.	D	.234		.264	.143		−.072		
15.	G[d]	.649		.471	.243		.282	.197	−.016
	Industrial Production: Mean Absolute Errors (1947-49=100)								
16.	A		5.9			3.2			
17.	E		5.2			6.0			
18.	C[c]	2.1		3.5	3.4		3.3	2.4	
19.	D	2.8		4.5	3.1		3.6		
20.	G	2.7		3.5	4.1		3.9	4.0	3.6
	Industrial Production: Mean Errors (1947-49=100)								
21.	A		−1.4			−0.4			
22.	E		−0.7			−1.3			
23.	C[c]	0.1		−1.5	−0.9		1.6	0.7	
24.	D	0.4		−1.4	−0.3		−0.7		
25.	G	2.0		0.9	1.0		0.4	0.7	0.1

(continued)

TABLE 12 *(concluded)*

Line	Forecast Set[a]	Interval of Predicted Change (months)[b] 0-3 (1)	0-6 (2)	3-6 (3)	6-9 (4)	6-12 (5)	9-12 (6)	12-15 (7)	15-18 (8)
	Industrial Production: Correlations of Predicted with Actual Changes								
26.	A		.486			.576			
27.	E		.192			.434			
28.	C[c]	.800		.309	.020		.335	.500	
29.	D	.559		-.314	.553		-.179		
30.	G[d]	.831		.402	.183		.211	.329	.325

[a]This table covers the same forecast sets as Table 9. For detail on periods and numbers of observations, see Table 9, note a.

[b]The current (base) period t is marked 0, so that 0-3 denotes the three-month interval between t and $(t + 3)$; 3-6, the three-month interval between $(t + 3)$ and $(t + 6)$; etc.

[c]The entries on this line are not strictly comparable because some of the forecast chains are "incomplete," so that the number of observations for different spans varies. See Table 9, note b.

[d]Based on forecasts in constant prices, as reported.

"explanation" of the already familiar inverse relationship between the span and the accuracy of forecasts.

Do the errors of the intraforecast changes tend to increase with the distance from the present? One might think that the change during, say, the first quarter in the chain (0–3) should be easier to predict than the change during the second quarter (3–6), and so on. However, in our data, errors of the implicit forecasts of such successive changes show no systematic increases. The absolute averages in Table 12 (lines 1–5 and 16–20) sometimes rise and sometimes decline; the differences among these figures are neither regular nor large.[8]

A recurrent bias in the chained forecasts could contribute to such results. In a simple hypothetical case, let the present level (assumed

[8] Other summary measures such as root mean square errors lead to similar conclusions. All these measures reflect both the central tendency and the dispersion of errors. The relations involving means and variances, which were set out earlier in this section, have clear implications for the mean square errors (it will be recalled that $M^2 = \bar{E}^2 + S_E{}^2$).

to be correctly estimated) be 100 and the predictions for two successive future periods be 103 and 106. If the actual levels turn out to be 104 and 108, then the errors of the two predictions are -1 and -2, increasing absolutely with the span. But the errors of the two intraforecast changes are the same (each being -1).[9]

As already noted, elements of persistent biases exist in the forecasts used here. Underestimation is common in all sets, except G in which overestimation prevails; these facts underlie the mean errors in Table 12 (lines 6–10 and 21–25). But this is by no means the whole story. The signs of the intraforecast change errors do vary in some chains. In most of these cases, the errors still cumulate because the negative correlation among them is not sufficiently high and they do not quite cancel each other out; and sometimes they do increase absolutely with the distance from the present.

It may be significant that intraforecast change errors showed more frequent and more sustained increases in the periods in which cyclical turning points occurred than at other times when movement continued in the same direction. In the latter sequences, errors were virtually always smaller and their variation was usually more irregular than in the former, as illustrated by the following tabulation of mean absolute errors (in billion dollars) for two sets of GNP forecasts (corresponding to lines 4 and 5 of Table 12):

	Interval of Intraforecast Change (months)					
	0–3	3–6	6–9	9–12	12–15	15–18
Forecasts D						
Periods with turning point (TP)	4.6	5.5	7.3	9.5		
Periods without TP	5.2	4.9	5.3	3.6		
Forecasts G						
Periods with TP	6.1	5.2	6.0	6.2	7.6	7.4
Periods without TP	3.9	2.9	5.3	2.3	6.2	2.5

The correlations between predicted and actual changes are in most instances higher for the very near future than for the more distant intervals, though irregularities and contrary cases do not appear uncommon (lines 11–15 and 26–30). If the decline in correlation were actually the rule, this would confirm the presumption that the forecasters do, in one sense, "know more" about the near future (say,

[9] For the first subperiod, the error is (103–100) – (104–100) = –1; for the second, it is (104–103) – (108–106) = –1.

0–3 months) than about the more remote future (say, 12–15 months). It would also give additional support to the hypothesis that it is the bias that accounts mainly for the similarity of typical intraforecast change errors for the different intervals.

MULTIPERIOD FORECASTS INTERPRETED AS RATE-OF-GROWTH PREDICTIONS

A rough indication that the intraforecast change errors do not generally increase with the distance to the target interval could have been obtained earlier (from Table 9). Reading across that table, one can see that the average errors usually increase less than proportionately to the extension of the span. The errors of twelve-month forecasts are, on the whole, less than twice as large as the errors of six-month forecasts and less than four times as large as those of three-month forecasts. These observations suggest another way of looking at multiperiod forecasts.[10] Suppose that what forecasters really try to do is to predict average rates of growth. Under this assumption, one would want to compute errors by (1) taking differences between the predicted and the actual percentage changes and (2) expressing these differences on a per-unit-of-time basis.[11]

When the mean absolute errors of percentage changes are divided by the length of span, it appears that they become *smaller* the longer the forecast. For example, the figures for the six-month forecasts are 3.14 (percentage points) for set A and 2.38 for set E, while the corresponding figures for the twelve-month forecasts, when divided by two, are 2.81 and 2.06, respectively. The results for quarterly forecast chains also suggest such declines. Thus the figures for set G, obtained by dividing the error measures for spans of one to six quarters by 1, 2, . . . 6, are 1.22, .94, .93, .80, .70, and .60 percentage points.

This may seem puzzling indeed: Are we to infer that the longer forecasts are after all better, not worse, than the very short ones? At this point, it becomes important to consider what the *effective* span of a

[10] I am indebted to Victor Fuchs and Geoffrey Moore of the National Bureau for comments which prompted the approach described here, although they are in no way responsible for the interpretations made.

[11] This amounts to dividing the errors in forecasts of percentage changes by the length of span. Compounding is ignored to simplify matters; given the shortness of the forecasts and our present purpose, this should not be a cause of any significant errors.

forecast is, since the above calculations depend critically on assumptions regarding the relative spans—that, e.g., 0–12 actually represents twice the distance involved in 0–6, etc. But recall that the position at the time "0," that is, on the date of forecast, is itself not known as a rule, but estimated (predicted) with an error.[12] Suppose then that, to account for this, the forecast of the first interval in the chain is treated as if its span were three months longer—0–9 instead of 0–6, for example. The longer forecast, say, 0–12, must be treated accordingly, i.e., as 0–15. In semiannual units, the two spans are no longer to be represented by 1 and 2 but rather by 1.5 and 2.5, and the divisors in our calculations must be changed accordingly.[13]

When the spans are thus recomputed, the following figures are obtained for the mean absolute errors of the implicit forecasts of rates of change: [14]

Span of Forecast (months)		*GNP (billion dollars): Forecast Sets*				*Industrial Production (1947–49 = 100): Forecast Sets*			
"Apparent"	"Effective"	A	E	D	G	A	E	D	G
3	6			.48	.61			.85	.89
6	9	2.09	1.59	.64	.63	2.94	2.48	1.13	1.08
9	12			.49	.70			.97	1.10
12	15	2.25	1.65	.43	.64	2.01	2.13	.70	1.08
15	18				.58				.89
18	21				.51				.75

These measures show no definite pattern of dependence upon the span of forecast. The adjustment for the fact that the recent past and present must in part also be predicted (because of the lag of information) has removed the previously observed tendency for the mean absolute errors of rate-of-change forecasts to get smaller as the predictions

[12] As shown earlier, subtracting that error algebraically from the error of the level forecast yields the implicit error of predicted change. But this is a matter of defining the latter measure; it does not imply that errors of changes are independent of those of the base (ECP). The following argument assumes dependence between the two error categories in that the imperfection of knowledge about the current position acts as if to lengthen the effective span of forecast.

[13] A similar adjustment for quarterly spans would lead to the use of divisors 2, 3, 4, and 5 (instead of 1, 2, 3, and 4) for the forecasts 0–3, 0–6, 0–9, and 0–12, respectively.

[14] Measures for different forecast sets are not comparable here, not only because of differences in periods covered but also because some of the figures refer to rates per six-month intervals (A,E), while others refer to rates per quarter. The only comparisons intended are those between spans for a given forecast set.

grow longer.[15] The differences between the summary error measures thus obtained are mostly small and irregular. This evidence is consistent with the idea that projection of a certain rate of growth over a sequence of periods often served as a basic device in the construction of these multiperiod forecasts. And this conception fits in, too, with our earlier finding that the intraforecast change errors do not vary systematically with the distance from the present.

To sum up, in this chapter we have developed some tentative and partial explanations of how an increase in the predictive span influences the marginal as well as the average accuracy of forecasts. They include the notions of sustained biases, differentiation between sequences with and without turning points, and modified rate-of-growth projections. These are not competing hypotheses, though they have not yet been integrated. Further evidence and analysis will be necessary to arrive at a more definitive explanation.

DIRECTIONAL AND TURNING-POINT ERRORS IN MULTIPERIOD FORECASTS

Appraisals of turning-point errors in forecasts of sequences of short intervals present some complications that do not arise in the annual data. In the latter, virtually all turns are associated with business cycle reversals; in series with shorter unit periods, there are some additional directional changes superimposed sporadically on the continuing trend-cycle developments. Such changes often reflect events caused by exogenous forces, which could hardly have been foreseen, and it is at least questionable whether they should be treated on a par with the major cyclical turns.

Also, in the series of short unit periods, runs of negative as well as positive signs occur, i.e., sequences of decreases as well as increases. The corresponding forecast chains can likewise contain runs of either sign. Comparisons of these signs will yield the frequencies of all directional errors, whether or not they are connected with actual or predicted turning points. A distinction can be made here between directional errors and turning-point errors, where the latter represent a subset of

[15] In other words, these errors decline with the increase in the "apparent" spans but not with the increase in the "effective" spans (see the tabulation in the text above).

the former. For the annual aggregates, on the other hand, this distinction has virtually no practical significance, since the two categories coincide.[16]

Table 13 presents the frequencies of directional errors for the five sets of multiperiod forecasts of GNP that are used throughout this chapter. These measures refer to the intraforecast changes; comparisons by span are of less interest in this context.[17]

Most of the observed semiannual or quarterly changes in GNP are increases reflecting the trend, though decreases (and, consequently, turning points) are considerably more frequent here than in the annual data, as would be expected. There were relatively few errors of predicting a decline for a period which actually saw a rise in GNP (see columns 2 and 3). Forecasters are, of course, aware of the economy's growth and tend to predict increases most of the time. Indeed, the predominant type of directional error comes about by missing a downturn, that is, predicting a rise for a period in which a decline actually occurred (columns 4 and 5). In two of the sets, all decreases were missed as none at all were predicted (lines 12–19).

The percentages of directional errors are computed by counting all instances of divergent signs of actual and predicted changes and relating the results to the corresponding observation totals. Even though the latter include the large category of actual increases that coincide

[16] Decreases in such series as GNP or industrial production have never lasted as long as two years in the postwar period, nor were they expected to. Only in 1946 and 1947 were declines predicted twice in a row by some forecasters (these would both be classified as "false signals" of peaks).

[17] Scores computed according to the agreement in direction between actual and predicted changes classified by span may be quite sensitive to errors in trend estimation. It is indeed possible (though not likely) that such scores would be poor solely because of misjudgment of the trend. Thus, in the example below, the peak was correctly predicted and dated(*). The intraforecast changes (ΔP) agree in sign with the actual changes (ΔA) in each successive period. Nevertheless, changes predicted for increasing spans, from t to $t + 1$, $t + 2$, etc. (which amount to cumulations of ΔP) disagree in sign with the corresponding actual changes in two of the four intervals.

	Values of		*Successive Changes*		*Changes by Span (cumulated)*	
	P	A	ΔP	ΔA	Cum ΔP	Cum ΔA
t	100	100				
$t+1$	103*	105*	+3	+5	+3	+5
$t+2$	101	104	−2	−1	+1	+4
$t+3$	99	102	−2	−2	−1	+2
$t+4$	98	101	−1	−1	−2	+1

TABLE 13

Frequency of Directional Errors in Forecasts of Semiannual and Quarterly Changes in GNP, 1947-63

Line	Interval of Predicted Change (months)[a]	Total Number of Observations (1)	Number of Observed Increases That Were Predicted (2)	Number of Observed Increases That Were Not Predicted (3)	Number of Observed Decreases That Were Predicted (4)	Number of Observed Decreases That Were Not Predicted (5)	Percentage of Directional Errors[b] (6)
			Forecast Set A: 1947-49, 1955-56, 1958-63				
1.	0-6	11	7	2	2	0	18.2
2.	6-12	11	6	2	1	2	36.4
			Forecast Set C: 1959-63[c]				
3.	0-3	22	15	4	2	1	22.7
4.	3-6	20	14	3	1	2	25.0
5.	6-9	19[d]	15	1	0	2	16.7
6.	9-12	13[d]	10	0	0	2	16.7
7.	12-15	7[d]	6	0	0	0	0
			Forecast Set D: 1959-63				
8.	0-3	9	5	1	1	2	33.3
9.	3-6	9	7	1	0	1	22.2
10.	6-9	9	6	0	0	3	33.3
11.	9-12	9[d]	6	2	0	0	25.0
			Forecast Set E: 1956, 1960-63				
12.	0-6	5	2	3	0	0	60.0
13.	6-12	5	3	1	0	1	40.0
			Forecast Set G: 1955-63[e]				
14.	0-3	16	9	1	0	6	43.8
15.	3-6	16	13	1	0	2	18.8
16.	6-9	16	9	1	0	6	43.8
17.	9-12	16	14	0	0	2	12.5
18.	12-15	16	10	0	0	6	37.5
19.	15-18	16	14	0	0	2	12.5

Notes to Table 13

[a]The current (base) period t is marked 0, so that 0-3 denotes the three-month interval between t and $(t + 3)$; 3-6, the three-month interval between $(t + 3)$ and $(t + 6)$; etc.

[b]Sum of entries in columns 3 and 5, divided by the corresponding entry in line 1 and multiplied by 100. See also note d.

[c]The entries for this set are not strictly comparable because some of the forecast chains are "incomplete," so that the number of observations for different spans varies. See Table 9, note b.

[d]Includes one instance of no change in the actual value matched with a predicted increase; this observation was excluded from the computed percentage of directional errors in column 6.

[e]Based on forecasts in constant prices, as reported.

with predicted increases, the percentages are on the whole large (half of them exceed 25 per cent, see column 6).[18]

To identify directional errors, it is sufficient to compare the signs of actual and predicted changes in any given interval. However, turning points occur when the signs of changes in two successive periods differ; hence, in dealing with errors in forecasting turns, it is necessary to compare *sequences* of signs. A multiperiod forecast may contain directional errors which reflect a previous turning-point error but are not themselves associated with any turning point. It may also contain directional changes which correct previous errors and result in a realignment rather than in a divergence of signs. And it may even contain opposite sequences in the actual and predicted figures, e.g., a "peak" in the former matched by a "trough" in the latter.[19]

[18] Comparisons by span yield similar results here, but, as already noted, they are less appropriate for the study of directional errors than these measures which refer to changes in successive, nonoverlapping intervals.

[19] To illustrate, let the sign sequences be as follows:

Interval (months)	0–3	3–6	6–9	9–12	12–15	15–18
Sign of actual change	+	+	+	+	–	+
Sign of predicted change	+	–	–	+	+	–

Here there is a turning-point error in the interval 3–6, which is of the "false signal" type. In 6–9, there is a directional error but no new turning-point error. Then the signs of the forecast sequence change again, while those of the actual values do not, but this merely corrects the previous errors and restores the directional agreement between forecasts and realizations (in 9–12). In the interval 12–15, there is another turning-point error, this time of a "missed turn" type. Finally, there is an example of opposite errors in the transition from 12–15 to 15–18: a "trough" in the actual and a "peak" in the predicted values.

To define turning-point errors, they should (a) be associated with changes in sign of either the actual or the predicted change or both, and (b) result in a directional disagreement between the actual and predicted change. This definition excludes (1) the repeated directional error (which does not follow directly upon any sign changes) and (2) the "corrective" directional change (which results in an agreement of signs).[20]

Table 14 presents a count of turning-point errors thus defined (columns 2–6) and, separately, of cases in the two special categories just described (columns 7–9). The errors are classified into "missed turns" and "false signals," as in Table 8. But there is a difficulty with this distinction in the case of opposite turns in predictions and realizations. Such exceptional errors belong, in a sense, to both of these categories and are treated accordingly in Table 14 (see note f). Other decision rules that seemed reasonable were adopted to handle the cases of no change in either actual or predicted values.[21]

The performance of the multiperiod forecasts of GNP in regard to turning points appears to be poor indeed, according to Table 14. The dates of a few turns were correctly predicted, mainly over the shorter spans of three or six months, in three of the forecasts sets. In the more distant intervals, virtually all recorded turns were missed and virtually all predicted turns proved to be false signals. There is little point in presenting the percentages of turns missed and falsely predicted (the $\bar{E}_{T1}$ and $\bar{E}_{T2}$ measures, as in Table 8, columns 8–9); it is enough to observe the predominance of zero entries in the count of the "correctly predicted" turning points (Table 14, column 4). Two of the sets show no correct turning-point forecasts at all (lines 12–19).

Before accepting the verdict implied in these findings, we should point out that our method, in effect, assigns failure marks to all turning-point predictions that did not identify exactly the date of the

[20] It should be noted that (2) can occur in actual as well as predicted sequences. For example, let the signs of the actual changes be + − + and those of the predicted changes + + +. There are two directional changes in the actual values but we count only one turning-point error and assign it to the second interval. The change from the second to the third interval restores the agreement of signs. (See footnote 19 for an example of a "corrective" directional change in the predicted sequence.)

[21] The main rule consists in the distinction between configurations such as + 0 − or − 0 +, which do constitute turning points, and configurations such as + 0 + or − 0 −, which do not.

TABLE 14

Frequency of Turning Points and Errors in Forecasts of Semiannual and Quarterly Changes in GNP, 1947-63

Line	Interval of Predicted Change (months)[a]	Total Number of Observations (1)	Number of Turning Points Counted[b]					No. of Turning Points Excluded from Count[c]		Number of Repeated Directional Errors[d] (9)
			Observed (TT+TN) (2)	Predicted (TT+NT) (3)	Correctly Predicted (TT) (4)	Missed (TN) (5)	Falsely Predicted (NT) (6)	Observed (7)	Predicted (8)	
					Forecast Set A: 1947-49, 1955-56, 1958-63					
1.	0-6	11	3	4	3	0	1	0	1	0
2.	6-12	11	3	2	1	2	1	0	0	1
					Forecast Set C: 1959-63[e]					
3.	0-3	22	4	4	2	2	2	0	1	1
4.	3-6	20	4	1	0	4	1	0	5	0
5.	6-9	19	3	0	0	3	0	1	3	0
6.	9-12	13	2	0	0	2	0	1	1	0
7.	12-15	7	0	0	0	0	0	1	0	0
					Forecast Set D: 1959-63					
8.	0-3	9	2	1	0	2	1	0	0	0
9.	3-6	9	1	1	1	0	0	1	0	1
10.	6-9	9	2	0	0	2	0	0	0	1
11.	9-12	9	1	2	0	1[f]	2[f]	1	0	0

(continued)

TABLE 14 *(concluded)*

Line	Interval of Predicted Change (months)[a]	Total Number of Obser-vations (1)	Number of Turning Points Counted[b]: Observed $(TT+TN)$ (2)	Predicted $(TT+NT)$ (3)	Correctly Predicted (TT) (4)	Missed (TN) (5)	Falsely Predicted (NT) (6)	No. of Turning Points Excluded from Count[c]: Observed (7)	Predicted (8)	Number of Repeated Directional Errors[d] (9)
					Forecast Set E: 1956, 1960-63					
12.	0-6	5	0	0	0	0	0	1	0	0
13.	6-12	5	1	1	0	1	1	0	0	0
					Forecast Set G: 1955-63[g]					
14.	0-3	16	4	1	0	4[f]	1[f]	0	1	2
15.	3-6	16	0	1	0	0	1	2	0	2
16.	6-9	16	4	0	0	4	0	0	0	3
17.	9-12	16	0	0	0	0	0	3	1	2
18.	12-15	16	4	0	0	4	0	0	0	2
19.	15-18	16	1	0	0	1	0	2	0	1

[a]The base period t is marked 0, so that 0-3 denotes the three-month interval between t and $(t+3)$, etc. The table covers the same forecast sets as previous tables in this chapter.

[b]Excludes the directional changes listed in columns 7 and 8. For symbols used, see text above.

[c]Directional changes that correct previous turning-point errors and result in agreement of signs of actual and predicted changes. See text and footnotes 19 and 20.

[d]Includes instances of divergent signs which have not changed from the preceding period. See text and footnote 19.

[e]The entries for this set are not strictly comparable because of differences in the number of available observations. See Table 9, note b.

[f]Includes one instance of opposite sequences of signs in the actual and predicted figures. See text and footnote 19.

[g]Based on forecasts in constant prices, as reported.

event. Now the objection could be raised that the requirement of exact dating of turning points in the chained (multiperiod) forecasts is too demanding. Moreover, some of the recorded directional changes have been short and shallow; in fact, some would be reversed by later data revisions. Failure to foresee such movements should not be considered a significant error, and in fact smoothing them out could under some circumstances be desirable.

The latter argument has a certain validity. Some of the directional changes involved were in fact reversed in the following quarter, and the agreement in sign of actual and predicted changes was often restored as promptly. An indication of this is provided where the "corrective" changes are frequent and the repeated directional errors rare (as in set C, lines 3–7, columns 7–9). Elsewhere, however, more sustained errors are in evidence (see, in particular, the high frequencies of repeated errors in lines 14–19, column 9). One can hardly dismiss all the errors recorded in Tables 13 and 14 on the ground that the changes to which they refer were all short and small.[22]

Recognizing that errors of misdating should not be judged too severely,[23] we have examined the importance of such errors by treating each multiperiod forecast as a single pattern and waiving the requirement of exact dating. For example, if a four-quarter chain forecast featured a rise followed by a decline (say, + + − −), and a peak did actually occur within the next twelve months but sooner than foreseen (+ − − −), then the forecast would be considered a correct prediction of that peak.

The outcome of this analysis can be simply stated: misdating was relatively unimportant as a source of errors in the forecasts under discussion. What happened in the great majority of turning-point errors was that the forecaster either (1) predicted a continuous development and missed a reversal, or (2) predicted a reversal which did not occur. In fact, despite the relaxation of the requirement of exact dating, com-

22 A direct, detailed examination of the differences among the movements involved should be rewarding when richer materials become available, but it is not required to bear out the above statement.

23 Suppose that a forecaster correctly predicted that a turn would occur the following year but misdated the turn by one quarter. This might be no mean achievement under the circumstances, yet one of the four interquarter changes in the given year would be marked wrong, yielding an error score of 25 per cent.

parisons of the sign patterns need not be more favorable to forecasters than the measures given in Tables 13 and 14.[24]

If the periods covered by the successive chain forecasts overlap, as in the case of forecasts C, D, and G, which are issued twice or four times in a year, the same error can reappear and be counted more than once. Again, however, this factor does not alter our main conclusion, which can now be stated with considerable assurance. Despite the possible mitigating circumstances that were considered, the record, whichever way one looks at it, simply does *not* indicate any significant ability of forecasters to predict a turning point several months ahead.[25]

SECTORAL FORECASTS FOR DIFFERENT SPANS

Forecasts of GNP components in set C were analyzed for spans varying from three to fifteen months. The results, based on chain predictions made in each quarter between 1958 and 1963, are summarized in Table 15.

This table demonstrates the pervasive tendency of errors to increase with the length of the forecast. There are only a few, rather slight exceptions to the rule that the mean absolute errors or the root mean square errors grow steadily larger as the predictions reach further into the future. Evidently, the rule applies to variables with very different characteristics.

Again, as in Table 9 for GNP and industrial production, the errors in Table 15 increase less than proportionately to the indicated length of span. If the entries in column 2 were divided by two, those in column 3 by three, etc., the errors thus computed in the rates of change

[24] This is so because the measures given in Table 13 assign a common error to the chain forecast as a whole, whereas those in Table 14 result in many separate scores, one for each different future interval, and usually only some of these scores will reflect the error. In fact, comparing the sign patterns often yields results that are particularly, and sometimes unfairly, negative. The following are some of the results for the GNP forecasts (recall that *N* denotes the absence and *T* the presence of a turning point; the first symbol refers to the actual and the second to the predicted values).

Forecast Set	*No. of Chain Forecasts*	*NN*	*TT*	*TN*	*NT*	*Opposite Turns in P_t and A_t*
A	11	5	2	2	2	
D	9	2	1	3	2	1
G	16	4		11	1	

[25] A parallel study of the record of industrial production forecasts also leads to this conclusion.

TABLE 15

Forecasts of Eight Selected Components of GNP: Summary Measures of Error over Spans from Three to Fifteen Months, 1958-63[a]

Line	Predicted Variable	Span of Forecast (months)				
		Three (1)	Six (2)	Nine (3)	Twelve (4)	Fifteen (5)
		Mean Absolute Errors (percentage points)				
1.	Personal consumption expenditures	0.62	0.95	1.24	1.50	1.82
2.	Consumer durables	4.03	4.44	4.45	5.03	7.83
3.	Gross private domestic investment	4.75	9.42	9.69	10.76	14.23
4.	Plant and equipment outlays	2.40	3.77	5.21	6.75	9.13
5.	Residential construction	4.20	6.64	6.52	5.91	5.53
6.	Government expenditures	1.15	1.41	1.68	2.24	2.37
		Mean Absolute Errors (billion dollars)				
7.	Net change in inventories	2.14	4.38	4.35	5.14	5.44
8.	Net foreign balance	0.91	1.18	1.07	1.24	1.46
		Mean Errors (percentage points)				
9.	Personal consumption expenditures	−0.30	−0.41	−0.46	−0.53	−1.10
10.	Consumer durables	−1.55	−0.52	−1.20	−0.38	−4.99
11.	Gross private domestic investment	−0.32	−1.92	−2.61	−1.76	−7.73
12.	Plant and equipment outlays	0.07	−1.23	−3.13	−3.50	−6.04
13.	Residential construction	0.24	−1.89	−4.06	−4.29	−5.08
14.	Government expenditures	−0.26	−0.53	−0.85	−1.49	−1.58
		Mean Errors (billion dollars)				
15.	Net change in inventories	−0.40	−0.14	0.90	1.52	−1.07
16.	Net foreign balance	−0.22	−0.31	−0.51	−0.46	−1.46

[a]For consumer durables (lines 2 and 10), forecasts cover the period 1961-63. The numbers of observations for spans of 3 to 15 months are 13, 11, 10, 7, and 4. For net foreign balance (lines 8 and 16), forecasts cover the period 1958-63. The numbers of observations for spans of 3 to 15 months are 19, 17, 16, 10, and 5. For the remainder, forecasts also cover the period 1958-63. The numbers of observations are 21, 20, 19, 13, and 7 for spans of 3, 6, 9, 12, and 15 months, respectively. The numbers vary because some of the forecast chains do not include all spans (see Table 9, note b).

would systematically and strongly decrease. But, as pointed out before, this procedure is incorrect inasmuch as it disregards the fact that forecasters must in effect predict a little backward in time as well as forward. When, to make a broad allowance for this, each span is treated as if it were one quarter longer (that is, each divisor is increased by one; see pages 5–18 above), the resulting mean absolute errors of rates of change (in percentage points) show much smaller and more irregular differences, though they still tend to decline with the span in some cases.[26]

	Effective Span of Forecast (months)				
	Six	Nine	Twelve	Fifteen	Eighteen
Pers. consumption expenditures	0.31	0.32	0.31	0.30	0.31
Consumer durables	2.02	1.48	1.11	1.01	1.30
Gross priv. dom. investment	2.37	3.14	2.42	2.15	2.37
Plant and equipm. expenditures	1.20	1.25	1.30	1.35	1.52
Resid. construction	2.10	2.21	1.63	1.18	0.92
Government expenditures	0.58	0.47	0.42	0.45	0.40

The mean arithmetic errors are again found to be predominantly negative (Table 15, lines 9–16). They also tend to increase absolutely with the span of the forecasts, though this association is considerably less regular than the others.

[26] The "effective" span of forecast is obtained by adding 3 to the "apparent" span (listed in Table 15). The following figures for mean absolute errors of rates of change (in percentage points) illustrate the results of using the divisors for the apparent rather than those for the effective spans:

	Apparent Span of Forecast (months)				
	Three	Six	Nine	Twelve	Fifteen
Pers. consumption expenditures	0.62	0.48	0.41	0.38	0.36
Consumer durables	4.03	2.22	1.48	1.26	1.57
Gross private dom. investment	4.75	4.71	3.23	2.69	2.85
Government expenditures	1.15	0.70	0.56	0.56	0.47