0401

SM/87/297

December 23, 1987

To: Members of the Executive Board

From: The Secretary

Subject: World Economic Outlook - Staff Studies - How Accurate is the World Economic Outlook? A Post Mortem on Short-Term Forecasting at the International Monetary Fund

The attached study is intended to serve as a background report for future discussions of the World Economic Outlook.

Mr. Larsen (ext. 4613) is available to answer technical or factual questions relating to this paper.

Att: (1)

Other Distribution:
Department Heads

INTERNATIONAL MONETARY FUND

World Economic Outlook: Staff Studies

How Accurate is the World Economic Outlook? A Post Mortem on Short-Term Forecasting at the International Monetary Fund

Prepared by M. J. Artis*

Approved by Jacob A. Frenkel

December 18, 1987

Contents

* Mr. Artis is professor at the University of Manchester. He prepared this report as a consultant to the Fund's Research Department. The author is grateful to officers of the IMF for inspiring and helping in the execution of this study, most particularly to Flemming Larsen; to John Llewellyn at the OECD for substantial assistance; to Peter Kenen for encouragement and example; to Eric Kiernan and Sue Duckworth for research assistance and to Hilary Thornton for secretarial services. None of the above bears any responsibility for opinions expressed.

Page

List of Tables

Page

List of Charts

How Accurate is the World Economic Outlook? A Post Mortem on Short-Term Forecasting at the International Monetary Fund

Recent developments in the sphere of international economic policy coordination produced an agreement at the May 1986 Tokyo Summit that the major countries should focus on a set of economic indicators as a means of strengthening the degree of co-operation in macroeconomic policy-making already in existence. The Fund was given the formal responsibility for carrying this suggestion forward. In the subsequent development of this idea (see, in particular, Crockett and Goldstein, 1987) emphasis has been given to a taxonomy of indicators of current economic developments, distinguishing those which are signals of policy posture from those which measure intermediate variables, and which in turn are distinguished from those measuring economic performance. Indicators may be used in a number of ways. On a rising scale of increasing international interdependence, they may provide individual countries with a check list of variables against which to monitor the short-run progress of their economies; they may provide information on the medium run sustainability of policies; and they may signal in a formal way the need for multilateral discussion of policies.

Because of the lags in the economic process, it is clear that indicators of current developments cannot be a substitute for forecasting; on the contrary, for any of the purposes listed above forecasts are needed for the evolution of the relevant indicators. Here, the relevance of the present study should become apparent. Students of the analytics of economic policy coordination (see, for example, Cooper (1985)) have long stressed the significance of agreement about propositions in positive economics to the success of international policy coordination: that agreement must embrace both the evaluation of responses of performance indicators to policy indicators and the baseline forecast evolution of the indicators. The Fund's World Economic Outlook (WEO) has long been in the business of projecting the latter, making forecasts of the development of the performance indicators subject, essentially, to starting assumptions about policies. For the successful functioning of an indicator system, the degree of forecasting accuracy must be tolerably good, given the alternatives. This setting provides more than adequate motivation for an examination of the Fund's forecasting track record as distilled from the projections published in the WEO and publications of the same kind circulated internally within the Fund for nearly a decade before regular publication began in 1980. (It is important to note at the outset that this analysis covers only the short-term forecasts of the WEO--that is, those covering the current year and one year ahead. Medium-term projections and scenarios have increasingly become an intimate part of the world economic outlook process but are not dealt with here.)

The study is organized as follows. The first section briefly reviews the Fund's forecasting methods and discusses a number of

criteria that will be used to evaluate the track record of the WEO projections. Then follows a detailed analysis of the accuracy of WEO projections for output, inflation and the balance of payments. The third section compares the WEO projections with those of the Organization for Economic Cooperation and Development (OECD) as well as with those of a number of national forecasting agencies. Section IV attempts to identify the reasons for some of the forecast errors. The conclusions of the study are contained in Section V. A number of appendices contain additional analytical material.

I. Forecasting Methods and Criteria for Evaluation

1. Nature of the forecasting exercise

This is not the place to explain the construction of the WEO forecasts in detail. (Goldstein (1986) may be consulted for such a description). But it is essential to spell out some of the principal characteristics of the forecast process, for these affect what post-mortem techniques can be used.

First, it is important to stress the conditional nature of the forecasts. They are prepared on certain assumptions about "exogenous" variables: fiscal and monetary policy, exchange rates, and oil prices are the leading variables in question. The basic assumption about policies is that "present policies" will be held unchanged during the forecast period, though "present policies" are interpreted to include any currently known announcements about future policy adaptations and may also "encompass certain policy adaptations or changes that seem likely to occur even though they have not been announced by the authorities". 1/ Exchange rates are currently projected at the real (formerly, nominal) levels prevailing at a recent base date, whilst the oil price is also usually projected (in the absence of more specific indicators otherwise) as constant in real terms.

The reasons these variables are treated in this particular way are perhaps mixed. The treatment of policies follows the customary practice of national official forecasting and its many derivatives where a prime originating purpose of the forecasting exercise is to provide a consistency check on policy itself. A similar justification applies here too and the WEO draws conclusions for desirable policy adjustment from its analysis of the future outlook. For market-based policy instruments such as interest rates, WEO projections must also be inhibited by the knowledge that a Fund forecast might move the market in a way which could force the hand of a member government, which would be

---

1/ This quotation, which could have been drawn in identical or very similar terms from other WEOs, comes from the published WEO of May 1980.

an embarrassing prospect. 1/ Somewhat similar considerations may affect the treatment of oil prices, but WEO practice here is like that of other forecasters and to this extent reflects a belief that predicting the timing and magnitude of changes in oil prices is a particularly hazardous undertaking.

The fact that WEO forecasts are conditional on assumptions about policy and oil prices suggests that, subject to measurement problems, it is important to allow for the falsification of the conditional assumptions in reviewing the track record. (See Section IV.) The position with respect to exchange rates is rather different. The typical conditional projection in the World Economic Outlook of an unchanged pattern exchange rates cannot be defended on the argument that exchange rates are a policy instrument, at least not one independent of fiscal and monetary policy, but rather because the undoubted power of the Fund to "move markets" would make it inappropriate for the publication of exchange rate forecasts.

Strictly speaking, since policy adjustment is not allowed to take the strain of supporting the pattern of exchange rates assumed and the exchange rate is not allowed to take the strain of supporting the set of policies assumed, the collection of conditional assumptions the WEO is forced to make about these variables can only be squared with theoretical considerations by invoking "portfolio shifts" of just the right type and magnitude to sustain them. In principle, nothing is more likely than that this assumption of accomodating portfolio shifts will fail. But this cannot mean that it would be right to treat deviations of exchange rates from their "forecast" paths as a reason for forecast error elsewhere. First of all, the failure of the exchange rate assumption to materialize may reflect a failure of other parts of the forecast just as much as the other way around; second, pragmatically but most importantly, despite theoretical considerations, the power of structural models to predict the exchange rate is extremely low. In practice, it is not clear that the conditional exchange rate baseline projection--which is, after all, a form of random walk prediction--can be significantly bettered. Finally, also pragmatically, exchange rate effects take a considerable time to work through on to output (although less time on to prices); they could have little impact within the typical short-term forecast period. On the other hand, it seems fair to say that failures of the exchange rate assumption may have more rapid and noticeable effects on balance of payments forecast errors, and since these turn out to be the most problematic part of the track record, some attempt to relate them to exchange rate forecast errors seems worthwhile.

1/ Thus, despite the flexibility with which "current policy" is interpreted, WEO procedures fall far short of the contemporary identification in the literature of "policy" with a rule for the adjustment of policy instruments (and so "policy change" with a change in the rule) and are closer to the traditional identification of "policy" with the instrument settings themselves (and "policy change" with a change in these settings).

A second important characteristic of the WEO forecasting exercise is that it is comparatively informal, much more so, say, than the leading forecasting models in the United Kingdom. (For a recent review of these the interested reader may consult Wallis et al., 1986.) While model-based exercises are conducted at various stages of the production of the forecast, there is no computer-based "world model" behind the forecast as a whole. This is not necessarily a drawback in itself, but the implication for post mortem analysis is that it is not possible to decompose an ex post forecast error into exogenous variable, judgmental, and model-based error in the way that would be appropriate, and feasible, for a model-based exercise (see Osborne and Teal (1979) for an original exercise of this type). Nevertheless, it should be possible, measurement problems permitting, to relate the forecast errors to exogenous variable errors, as discussed below.

A third important characteristic of the WEO forecast procedure is that it has, at its heart, a consistency check not shared by national forecasters. As described in Goldstein (1986), original country-desk-based forecasts, prepared against environmental assumptions specified by the Research Department, are aggregated to check for the consistency of their trade and balance of payments implications. Identified discrepancies are then removed by an iterative process in which the country desk forecasts are successively revised, until the check is satisfied. The opportunity, and indeed the need, to conduct this check obviously arises from the closed economy nature of world forecasting which contrasts with the open economy basis of national forecasting. It would be useful to identify a way of confirming the value of consistency checks. One approach might be to compare the ex post accuracy of the initial and final forecasts made in each round, but the records available do not allow this comparison to be made. An additional problem relates to the fact that there is a significant discrepancy in the world current account, which may reduce the value of the consistency check.

2. Criteria of forecast quality

Given the selection of variables to be examined, the principal tools used for assessing the WEO forecasts in the sections below comprise the following: inspection of forecast error summary statistics; investigation of systematic bias in the forecasts taken over a long period; comparison with alternative forecasts; and the investigation of the rationality of forecast error. These checks are supplemented by an identification of outstanding episodes in the track record and an attempt to explain these in more detail by recourse to narrative material. Some explanation of these tools is in order.

a. Summary statistics. The principal summary statistics deployed in examining the WEO track record are the average absolute error of forecast, the root mean square error, and the Theil inequality statistic. Because a forecast error series may display both positive and negative errors the simple mean may be a highly misleading indicator of accuracy, and for this reason the average absolute error is preferred.

As a basis for comparing this statistic among series, the mean absolute value of the realized series itself is also presented. It is commonplace in economic analysis to prefer a measure which penalizes a large deviation more highly than a series of smaller ones of equivalent total size; for analytical tractability a quadratic measure is often used and for this reason the root mean square error (RMSE) is a preferred statistic in studies like the present one. This statistic too needs to be normalized in some fashion to facilitate comparison between series and the study makes use of such a normalization, in the Theil Inequality Statistic which can be generally defined as the ratio of the RMSE of the forecast under consideration to the RMSE of an alternative forecast. In the main text tables displayed below, this alternative is provided by the naive "no change" forecast, where the forecast for year t of variable x is the t-1 value for x (where x may be the growth rate of real GDP or the rate of inflation). 1/ (In Appendix D, we also consider the Theil Statistics produced by the alternative naive standard that the forecast of x for year t corresponds to the ten-year moving average of x.)

b. Realization--forecast regressions. The efficiency of forecasting may be tested by performing the regression of the realizations on the forecasts themselves, as $R(t) = a + b\,F(t) + u(t)$. A perfect forecast would identify the intercept in such a regression as zero, the slope as unity and yield a correlation coefficient of 1.00. Where knowledge of the realization-forecast relationship itself can reduce the forecast error variance these conditions will not hold. It seems a natural interpretation, within the terms of this regression framework, to identify a failure of the two expectations about the intercept and slope terms with the presence of bias; but this inference is not necessarily correct. 2/ The essence of the matter is that the realization-prediction regression detects whether the pattern of forecast errors can be related to the level of the forecast, not whether the average error is significantly different from zero, which can be tested for directly by measuring the average error and asking whether it is significantly different from zero. 3/ In answering some of these questions it is useful to supplement the results that can be obtained for specific countries (areas or aggregates) by pooling the data. While this procedure permits the benefit of offsetting country error, it enhances the power of significance tests.

c. Comparisons. "Absolute" measures of forecast accuracy are useless in themselves; they need to be related, on the one hand, to the standards of accuracy required by the purpose for which they are sought and, on the other, to comparable measures generated by alternative

---

1/ In the original case considered by Theil (1966) the "no change" naive forecast referred to the levels of the series and the RMSE of such a forecast is of course simply the RMSE of the series itself.

2/ See Holden and Peel (1987).

3/ See Appendix C for a fuller discussion of these points.

forecasting techniques. In the latter category, the normalizations already noted compare the forecasts with those generated by two alternative prediction schemes. It would be possible also to generate univariate time series and multi-variate (Bayesian vector autoregression) models as a further source of alternative forecasts; ex post facto it might well prove possible to generate a model in this class which would be superior to the WEO forecasts, but the achievement would not be very interesting because the alternative model does not represent a feasible alternative forecasting technique. Even if models of this class, possessing superior forecasting qualities, could be built on a purely ex ante bias, their usefulness and plausibility would be in doubt if they did not enforce consistency and could not accomodate variation for policy or environmental change. Given these drawbacks, this type of alternative was not explored.

The alternative actual forecasts with which the WEO forecasts are compared here are those produced by the OECD and, following Llewellyn and Arai (1984), by a set of national forecasters. With the OECD the comparison is with another international agency producing forecasts of a nearly comparable scope in country coverage, assumptions and detail. However, a difficulty with both types of comparison is that it is not possible to align the forecast dates exactly and so differences in the information sets conditioning the forecasts inescapably contaminate the comparisons.

3. Explanations for forecast errors

Given the conditional nature of the forecasts, explained above, testing the rationality of the forecasts involves assessing the contributions of "innovations" (unexpected changes) in the "exogenous" variables. The principal difficulties in implementing this approach are measurement problems. While it is possible to derive a reasonably satisfactory series for the innovations in oil prices, it is less easy to do this for fiscal policy and appears not to be feasible for monetary policy. In the case of fiscal policy the problem is less conceptual than practical: series of fiscal policy anticipations and outturns exist, but are for one reason or another less than satisfactory. For monetary policy there is the substantial conceptual problem that indicators like the growth rate of the money supply reflect not only policy but the economy more generally. While measures of fiscal policy like the fiscal impulse or the structural budget balance attempt to normalize for the influence of the economy, no comparable measure exists for monetary policy. Hence, even though fiscal policy and oil prices are not the only driving variables in world economic forecasting, for practical reasons it is only the contribution of innovations in these variables to explanations of the forecast error that is assessed. In a fully "rational" forecast these variables should only appear in the form of current innovations; neither lagged innovations nor actual values should in principle explain current errors if the forecasters have fully taken on board the implications of previous changes and have a correct

model of the significance of their own current anticipations of these variables for those they are forecasting. 1/

The role of systematic analysis of the complete time series of forecasts is not to avoid the challenge of historical analysis of forecast error so much as to provide a context for it and to avoid the trap of choosing specific explanations which fit the facts in any one episode but have no overall power to improve the forecasts in general. Moreover, there is a dimension of forecasting quality which lends itself best to graphical and narrative analysis and this is the question of turning-point error. An allegedly common failing of forecasts is the failure to spot the significant cyclical turning points.

4. The data base

The data base used in this study comprises the forecasts in the published versions of the WEO and similar data from the earlier comparable unpublished documents. The nature of this data base, in terms of the forecast horizons used and regularity of the forecast exercises conducted is indicated in Table 1. This shows that while there has been some irregularity in forecast production dates, particularly up to 1982, there has nearly always been a forecast for the year in question produced in the second quarter. An earlier forecast for the year has generally been available in the fourth and often as early as the third quarter of the previous year. In the last two years, the first forward look has been taken even earlier, with forecasts for the following year appearing as early as April. 2/ Besides producing a main forecast, there have been many occasions when uncertainties about principal conditioning variables (such as oil prices and exchange rates) have been felt to be sufficiently acute as to warrant the production of variant "scenarios."

The content of the WEO is extraordinarily rich: forecasts are produced not simply for the principal variables of interest in the main countries, but in considerable detail both for these economies and as well for regional and analytical groupings embracing the entire world economy (with the exception of the U.S.S.R. and other countries of Eastern Europe that are not members of the Fund.) In order to make

---

1/ Given the conditional nature of the forecasts, there need be no necessary presumption that forecast errors should be serially uncorrelated or preserve the desirable efficiency and freedom-from-bias properties described above in the regression of realizations on the "raw" forecasts, uncorrected for the effect of "innovations". For example, because of the nature of the policy and oil price assumptions, these innovations may very well be, themselves, serially correlated. However, as will be shown, there is not a great deal of evidence of bias or inefficiency in the uncorrected forecasts in any case.

2/ For internal purposes, this practice of taking a long forward look was instituted even earlier.

Table 1. The Forecast Horizon Content of the WEO

WEO Date(s)	Forecast Horizon	WEO Date(s)	Forecast Horizon
January 12, 1971	1971	December 1, 1978	1979
May 27, 1971	1971	February 9, 13 and 15, 1979	1979
April 13, 1972	1972	June 11, 13 and 15, 1979	1979
January 31; February 22; March 1, 1973	1973	August 30, 1979	1980
June 14; August 9, 1973	1974	May 1980 2/	1980
December 21, 1973; January 4 and 31, 1974	1974	August 22, 1980	1981
March 14, 1974	1974	June 1981 2/	1981
May 22, 23 and 24; June 21, 1974	1974	August 24, 1981	1982
December 24 and 31, 1974	1975	April 1982 2/	1982
March 31, 1975	1975	August 2, 1982	1983
May 21 and 23, 1975	1975	May 1983 2/	1983
December 12 and 15-16, 1975	1976	August 19; September 16, 1983	1984
July 7 and 9; August 11, 1976	1976	April 1984 2/	1984
February 22 and 24; March 2-3, 1977	1977 1/	September 1984 2/	1985
June 29; July 5 and 11, 1977	1977	April 1985 2/	1986
December 27, 1977	1978	October 1985 2/	1986
April 3-4, and 10, 1978	1978	April 1986 2/	1987
September 6, 1978	1978 1/	October 1986 2/	1987
		April 1987 2/	1988

1/ Some figures given for the first half of the following year.
2/ Published.

progress in assessing the accuracy of the forecasts, it is necessary to make a number of decisions about variables and forecasts to exclude.

The identification of forecast error plainly requires a definition of the outturn or realization with which the forecast can be compared. Because of the incidence of revisions of economic data, there is more than one possible series of realizations that might be chosen. Investigators generally take the view that the purpose of forecasts is to be right about the true evolution of the economy and that, at any point in time, that is most nearly revealed by the latest, revised-to-date, series of data. This view, though clearly quite a persuasive one, is perhaps too "pat". The latest available set of data is not homogeneous in vintage: early data are many times revised, latest data are perhaps still preliminary or partial estimates. Rebasing economic series may make it quite inappropriate to use the latest available data as a check on the forecast: the latter will have been formulated on data with a different base, and different properties, 1/ and it may not be feasible to reconstruct the data on a consistent base. Then again, policy (and short-term forecasting post mortems) will inevitably be based on early, not subsequently revised, data. For all these reasons there is room for choice about the realization series to be used and expedient criteria may legitimately affect the decision. In the present case (as detailed in the next section), three types of realization are deployed altogether in the present study, which one is in play at any time being made clear in context. None of the more general conclusions arrived at appears to depend on the particular choice of realization series, though some conclusions are drawn from experiments involving the use of a specific series which were not, or could not be, replicated on the latest available set as was done for all the processing described in the next section.

II. Forecasting Accuracy

In this section we consider the accuracy of WEO forecasts of principal variables over the whole available period, using a selection of the standard criteria discussed in the previous section. This discussion is presented in three stages: first, a consideration of the variables selected for study; second, an analysis of the forecasts in respect of industrial countries; finally, an analysis of the forecasts for developing countries.

1/ Extreme examples arise when rebasing involves a substantially new representation for a particular activity: for example, the behavior of volume estimates of GDP may be substantially different for an economy which undergoes a resources boom dependent upon whether the relevant activity weights are pre- or post-boom.

1. Selection of Variables

As indicated in the previous section, WEO forecasts embrace a large number of variables for several individually specified countries and aggregate groupings of various kinds. A useful study necessitates the suppression of secondary detail and the selection of a primary set of variables. Recent discussion of the use of indicators in multilateral surveillance draws attention to the relationship between indicators and the transmission mechanism of economic policy. Thus, a conventional view of the latter directs attention to indicators of policy input (as, for example, the structural budget balance) at one end of the transmission mechanism and indicators of performance at the other (such as output growth, inflation, or the balance of payments); in between stand intermediate variables such as the exchange rate and perhaps interest rates. In this study, attention is directed at the indicators of economic performance, measured by real GNP/GDP growth, GNP/GDP deflator or consumer price inflation and the current account of the balance of payments. In addition, because of the special interest afforded to trade by the world context of WEO forecasts--WEO trade forecasts being often cited by national forecasters--export and import volume growth and the development of the terms of trade are also investigated.

The country coverage of the projections examined also needs to be determined. Here, the institutional importance of the Group of Seven Major Industrial Countries (G-7), their weight in world output and trade (in 1984-85, 56.9 and 53.5 percent, respectively) and the fact that the WEO has consistently provided forecasts for the G-7 members individually, dictates that the forecast record for each of these countries and for the group as a whole should be examined. At the same time, aggregates for the industrial countries as a whole and for "Europe" as a group can also be easily and usefully examined. In addition to the industrial countries, the developing countries need also to be examined; none of these is as large in combined trade and output weight as the smallest of the G-7 countries, and WEO forecasts have traditionally distinguished various groupings of the developing country bloc. The longest standing of such groupings and thus the most amenable to analysis over a reasonably long period of time are the regional groupings, where among the non-oil block, Africa, Asia, Europe, the Middle East and the Western Hemisphere are separately distinguished.

Finally, a choice has had to be made of horizon of forecast and vintage of realization or outturn data to be employed. Table 1 in Section 1 gave a brief summary of the projection content of successive WEO rounds; the variable dates of these rounds imply that whatever selection is made, no set of forecasts is homogeneous in its timing relative to the forecast horizon. However, a distinction was drawn between two groups of broadly homogeneous forecasts - "current year" (CY) forecasts, where the forecast for year t is made during the year

t itself - and "year ahead" (YA) forecasts where the forecast for year t is made in year t-1. 1/

In practice even this distinction proved an ideal rather than a rigorously enforceable practice, as the actual sourcing for the two categories of forecast shown in Table 2 illustrates. The CY forecasts are considerably more homogeneous in timing, varying only by 3 months from April to July at the maximum, compared to a maximum variation of 7 months from August to March (of the following year!) in the case of the YA forecasts. 2/ The additional variability nevertheless seemed a price worth paying to obtain a reasonably long series. In the choice of outturn data, the main analysis deploys two categories. For the CY forecasts, the outturn is identified with the "first available" estimate, the figure reported in the following year's World Economic Outlook; in the case of the YA forecasts, however, the outturn is identified with the "first settled" estimate, that available in the World Economic Outlook of the following-year-but-one (i.e. the YA forecast for 1980 is compared with the outturn data published in the forecast source in 1981). These choices of outturn data had certain specific advantages over the use of latest available estimates: first, some of the aggregates were changed in definition over the course of time, and the use of these outturn data enabled the resultant inconsistencies to be minimized or even eliminated in a way which would not have been so straightforward with latest available data. Secondly, the combination of "first settled estimates" as outturn data with the YA forecasts allowed these to be compared with OECD and national forecasts prepared on a similar basis for the paper by Llewellyn and Arai (1984) and extended in the present study. Latest available data nevertheless were used in replication of all the principal computations of the main analysis; a summary of these results appears in Appendix B.

2. Summary Statistics: Industrial Countries

Tables 3 to 6 provide evidence of the track record of WEO forecasting based on averaging over the whole period: 1971-86 for the CY forecasts, 1973-85 for the YA forecasts.

Subject to a finding (see below) of some bias when the data are pooled, the track record for output growth forecasts, in the first table, is by a small margin the best of the three. The CY forecasts show comparatively low average absolute errors compared to the mean absolute value of the series, while the Theil coefficients indicate that

---

1/ This distinction follows that in the study of forecasting in the Fund by Kenen and Schwartz, though (as described in Appendix A), the actual classification of forecasts and the outturn data employed here are different from theirs.

2/ The dates referred to are the dates of the documents used in the study, some of which are published, some unpublished. Publication lags have varied between one and two months.

Table 2. Sourcing of the Forecasts 1/

Current-Year Forecasts		Year-Ahead Forecasts	
May 27, 1971		January 12, 1971	
April 13, 1972		not available for 1972	
June 14, 1973		January 31, 1973	
May 23, 1974		December 24, 1973	
May 23, 1975		December 24 and 31, 1974	
July 7-9, 1976		December 12, 15, and 16, 1975	
June 29; July 5, 1977		March 3, 1977	
April 3-4, 1978		December 27, 1977	
June 11, 13, and 15, 1979		February 15, 1979	
May 1980	published WEOs	August 30, 1979	
June 1981		August 22, 1980	
April 1982		August 24, 1981	
May 1983		August 2, 1982	
April 1984		August 19, 1983	
April 1985		September 1984	published WEOs
April 1986		October 1985	
		October 1986	

1/ Dates refer to those of WEO documents, published where stated, otherwise unpublished. The publication lag is generally 1-2 months.

Table 3. Forecast Accuracy--Summary Statistics: Industrial Countries' Output Growth

(In percent)

	Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom	Group of Seven	Total Industrial Countries	Europe
					Current Year (1971-86)					
Mean absolute actual value	3.744	3.444	4.938	2.950	2.681	2.875	2.056	3.163	3.025	2.469
Average absolute error	0.906	0.888	1.238	0.713	1.050	1.106	0.931	0.619	0.625	0.606
RMSE	1.338	1.063	1.716	1.208	1.412	1.470	1.190	0.761	0.757	0.996
Theil's inequality statistic	0.388	0.278	0.393	0.439	0.448	0.362	0.444	0.244	0.257	0.382
Regression: intercept	-0.416	0.106	0.241	-0.321	-0.629	-0.603	-0.034	-0.027	-0.398	-0.686
	(0.66)	(0.28)	(0.26)	(0.55)	(1.09)	(0.95)	(0.10)	(0.83)	(1.18)	(1.47)
slope	1.021	0.910	0.879	1.016	0.983	1.211	0.819	1.002	1.036	1.106
	(0.14)	(0.98)	(0.74)	(0.09)	(0.10)	(0.95)	(1.47)	(0.02)	(0.39)	(0.67)
$\bar{R}^2$	0.751	0.866	0.651	0.698	0.679	0.655	0.742	0.899	0.893	0.759
					Year Ahead (1973-85)					
Mean absolute actual value	3.362	3.431	4.692	2.477	2.662	2.962	2.385	3.031	2.885	2.338
Average absolute error	1.800	1.454	1.792	1.085	1.631	1.938	1.392	1.130	1.108	1.238
RMSE	2.238	2.047	3.217	1.085	2.207	2.482	1.779	1.713	1.657	1.744
Theil's inequality statistic	0.596	0.504	0.824	0.648	0.656	0.627	0.673	0.544	0.555	0.658
Regression: intercept	-1.244	0.020	2.723	-0.671	-2.898	0.263	-0.686	-0.811	-0.943	-1.455
	(0.79)	(0.02)	(1.25)	(0.61)	(1.69)	(0.20)	(0.78)	(0.76)	(0.89)	(1.08)
slope	1.103	0.886	0.309	0.970	1.601	0.747	1.085	1.066	1.092	1.222
	(0.26)	(0.49)	(1.87)	(0.09)	(1.12)	(0.59)	(0.24)	(0.23)	(0.31)	(0.49)
$\bar{R}^2$	0.348	0.525	-0.026	0.389	0.397	0.143	0.410	0.512	0.511	0.340

Note: The definitions of Current-Year and Year-Ahead forecasts are discussed in the text and in Appendix A. Mean absolute actual is defined as $\Sigma|R_i|/n$ where R_i is the realization ("actual") in year i and n the number of years in the sample; mean absolute error is $\Sigma|F_i-R_i|/n$ where F is the forecast. RMSE is $\sqrt{\Sigma(F_i-R_i)^2/n}$ and Theil's inequality coefficient is RMSE(F)/RMSE(F,a), where F,a is a naive "no change" forecast. The regression data are for the regression of R_i on F_i and figures in parentheses are t-stats: those for the intercept test against difference from zero, those against the slope for differences from unity.

the root mean square error of the forecasts is only some 20-40 percent, in typical cases, of the error that would be incurred by a "naive" forecaster. The realization-forecast regressions provide no indication of inefficiency. 1/ As might be expected, the CY forecasts are superior by these criteria to the YA forecasts, where the RMSEs and Theil coefficients are higher, the fit of the realization-forecast regression poorer, and average absolute errors in relation to actual mean absolute values higher than they are in the CY forecasts. Even so, these results appear fairly satisfactory: the Theil statistics are all well below unity, and the average absolute errors are well below the mean absolute value of the output growth series itself.

The track record for inflation (Table 4), is marginally less satisfactory than that for output, though still overall highly acceptable. The superiority of the CY forecasts again stands out. These forecasts display, with the single exception of Germany, smaller average absolute errors, lower RMSEs and lower Theil statistics than the YA forecasts. The CY forecasts provide no evidence of inefficiency, yielding a good fit in the realization-forecast regressions. The YA forecasts provide a poorer fit in these regressions, and for Italy indicate inefficiency; for this same country, moreover, the Theil statistic exceeds unity. Elsewhere, however, the general run of evidence is favorable, even if the performance is not so good as in the nearer-term horizon of the CY forecast or the comparable output forecasts.

Turning to the evidence on export and import volume forecasts, Tables 5 and 6, the track record now suggests little difference between the CY and YA forecasts (though the CY statistics for imports are better than those for exports). In terms of overall quality, both appear equally good, with low Theil statistics suggesting generally that these forecasts provide a distinct improvement on the naive standard. There is no evidence of inefficiency and the overall fits of the realization-forecast regressions are on the whole not unreasonable except for the export growth statistics for Italy.

The record for balance of payments forecasts in Table 7 is considerably less reassuring than for output and inflation. The Theil statistics, especially for the YA forecasts, are notably high, showing that the forecasts are little better than a naive projection, while the average absolute errors are high in relation to the absolute mean values, and in two of the YA forecasts (for France and the Group of Seven) are actually somewhat higher. The realization-forecast regression suggests inefficiency in two cases, both for the YA (Group of

---

1/ Generally, in commenting on the realization-prediction regression results in what follows, we may take guidance from the t-statistics estimated on the coefficients. It is worth noting, though, that the individual t-statistics are not always a secure guide to the results to be obtained from the appropriate joint test of the restrictions. See Appendix C.

Table 4. Forecast Accuracy--Summary Statistics: Industrial Countries' Inflation

(In percent)

	Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom	Group of Seven	Total Industrial Countries	Europe
					Current Year (1971-86)					
Mean absolute actual value	7.369	6.063	5.112	8.781	4.594	13.913	11.163	5.650	6.831	8.238
Average absolute error	1.136	0.544	1.575	0.950	0.688	1.306	1.781	0.556	0.425	0.744
RMSE	1.567	0.724	2.418	1.291	0.862	1.971	2.236	0.698	0.610	0.971
Theil's inequality statistic	0.626	0.351	0.529	0.722	0.498	0.611	0.343	0.360	0.326	0.430
Regression: intercept	-0.148	-0.173	-0.038	1.852	-0.191	2.832	-0.280	0.185	0.305	1.183
	(0.13)	(0.31)	(0.04)	(1.81)	(0.26)	(2.20)	(0.22)	(0.34)	(0.61)	(1.41)
slope	1.076	1.025	0.930	0.827	1.036	0.837	1.076	0.973	0.970	0.904
	(0.50)	(0.30)	(0.53)	(1.49)	(0.24)	(1.79)	(0.71)	(0.35)	(0.43)	(0.93)
$\bar{R}^2$	0.765	0.905	0.762	0.766	0.758	0.847	0.867	0.913	0.927	0.834
					Year Ahead (1973-85)					
Mean absolute actual value	8.285	6.769	5.392	9.608	4.385	15.231	11.969	7.346	7.438	8.823
Average absolute error	1.938	1.292	2.938	1.638	0.515	3.031	2.323	1.215	1.100	1.162
RMSE	2.731	1.725	4.200	1.200	0.686	3.574	3.225	1.757	1.639	1.479
Theil's inequality statistic	0.927	0.741	0.831	0.836	0.429	1.105	0.494	0.794	0.773	0.663
Regression: intercept	1.478	1.285	0.396	3.388	-1.332	8.290	-2.433	0.917	1.064	0.201
	(0.50)	(0.73)	(0.18)	(1.75)	(2.07)	(2.31)	(1.14)	(0.47)	(0.55)	(0.10)
slope	0.944	0.865	0.858	0.741	1.245	0.491	1.350	0.922	0.899	1.052
	(0.14)	(0.51)	(0.46)	(1.26)	(1.81)	(2.09)	(1.89)	(0.29)	(0.38)	(0.22)
$\bar{R}^2$	0.281	0.443	0.364	0.501	0.875	0.203	0.813	0.476	0.468	0.608

Note: For definitions etc., see Note to Table 3.

Table 5. Forecast Accuracy--Summary Statistics: Industrial Countries' Export Growth

(In percent)

	Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom	Group of Seven	Total Industrial Countries
					Current Year (1972-86)				
Average absolute error	3.307	2.560	4.170	2.573	2.847	3.753	2.353	1.887	1.893
RMSE	5.071	3.433	5.378	3.285	3.453	4.648	2.972	2.493	2.469
Theil's inequality statistic	0.521	0.366	0.475	0.439	0.378	0.621	0.425	0.366	0.376
Regression: intercept	0.115	-0.456	1.443	-0.424	-2.189	1.317	-0.423	-0.668	-0.345
	(0.06)	(0.47)	(0.68)	(0.25)	(1.44)	(0.58)	(0.30)	(0.60)	(0.29)
slope	1.158	1.247	0.937	1.048	1.416	0.615	1.055	1.182	1.079
	(0.55)	(1.82)	(0.24)	(0.19)	(1.88)	(1.22)	(0.19)	(0.99)	(0.39)
$\bar{R}^2$	0.519	0.857	0.463	0.532	0.741	0.166	0.470	0.744	0.663
					Year Ahead (1972-85)				
Average absolute error	3.436	2.443	4.464	2.586	2.886	3.929	2.464	2.011	3.015
RMSE	5.233	3.372	5.567	3.339	3.521	4.799	3.068	2.588	3.917
Theil's inequality statistic	0.519	0.358	0.480	0.428	0.340	0.627	0.429	0.367	0.569
Regression: intercept	0.288	-0.739	1.741	-0.126	-1.976	1.512	-0.367	-0.688	-2.520
	(0.14)	(0.75)	(0.71)	(0.07)	(1.23)	(0.62)	(0.25)	(0.57)	(0.77)
slope	1.156	1.249	0.908	1.022	1.401	0.599	1.051	1.181	1.315
	(0.52)	(1.88)	(0.32)	(0.09)	(1.75)	(1.20)	(0.17)	(0.94)	(0.59)
$\bar{R}^2$	0.519	0.871	0.404	0.518	0.738	0.146	0.464	0.74	0.296

Note: For definitions etc., see Note to Table 3.

Table 6. Forecast Accuracy--Summary Statistics: Industrial Countries' Import Growth

(In percent)

	Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom	Group of Seven	Total Industrial Countries
					Current Year (1972-86)				
Average absolute error	5.233	4.187	3.920	2.353	2.433	4.250	2.933	3.141	2.740
RMSE	6.344	5.157	4.644	2.759	3.157	5.238	3.422	3.365	3.016
Theil's inequality statistic	0.508	0.202	0.279	0.230	0.530	0.422	0.430	0.263	0.276
Regression: intercept	-0.112	-0.150	-2.652	-0.366	0.205	-3.675	-0.826	-1.321	-1.446
	(0.05)	(0.04)	(1.72)	(0.36)	(0.14)	(1.89)	(0.63)	(0.96)	(1.22)
slope	1.301	1.274	1.259	1.162	0.829	1.423	1.229	1.265	1.314
	(0.90)	(1.46)	(1.54)	(1.16)	(0.86)	(1.54)	(0.99)	(1.34)	(1.69)
$\bar{R}^2$	0.505	0.764	0.798	0.830	0.540	0.65	0.66	0.74	0.78
					Year Ahead (1972-85)				
Average absolute error	5.250	3.720	3.750	2.414	2.536	4.371	3.007	3.049	3.954
RMSE	6.429	4.507	4.503	2.828	3.257	5.380	3.505	3.281	4.929
Theil's inequality statistic	0.519	0.302	0.300	0.249	0.541	0.439	0.428	0.352	0.566
Regression: intercept	-0.579	-0.919	-3.233	-0.409	0.209	-3.635	-1.008	-1.684	-4.306
	(0.25)	(0.65)	(2.25)	(0.39)	(0.14)	(1.81)	(0.73)	(1.25)	(1.26)
slope	1.332	1.288	1.266	1.158	0.829	1.450	1.238	1.273	1.573
	(0.97)	(1.79)	(1.74)	(1.09)	(0.83)	(1.56)	(1.00)	(1.43)	(1.01)
$\bar{R}^2$	0.521	0.828	0.839	0.828	0.537	0.650	0.667	0.769	0.358

Note: For definitions etc., see Note to Table 3.

Table 7. Forecast Accuracy--Summary Statistics: Industrial Countries' Balance of Payments on Current Account

(In billions of dollars)

	Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom	Group of Seven	Total Industrial Countries
					Current Year 1/ (1973-86)				
Mean absolute actual value	2.869	25.623	13.346	4.123	6.800	4.515	4.685	19.977	
Average absolute error	2.300	10.379	6.180	2.121	4.200	2.579	2.764	12.479	
RMSE	3.080	13.329	7.364	2.989	5.359	2.996	3.886	14.652	
Theil's inequality statistic	0.991	0.588	0.498	0.565	0.577	0.471	0.671	0.579	
Regression: intercept	-1.664	0.197	0.764	-0.664	0.845	0.526	0.377	-3.924	
	(1.77)	(0.50)	(0.39)	(0.68)	(0.52)	(0.53)	(0.34)	(0.81)	
slope	0.348	1.142	1.162	0.844	1.059	1.014	0.929	0.732	
	(2.60)	(1.80)	(2.03)	(0.81)	(0.37)	(0.06)	(0.31)	(1.58)	
$\bar{R}^2$	0.066	0.941	0.942	0.581	0.767	0.612	0.549	0.573	
					Year Ahead 2/ (1973-85)				
Mean absolute actual value	2.846	23.462	13.662	3.700	9.077	4.539	6.169	14.515	15.562
Average absolute error	1.323	11.838	7.885	3.800	5.746	4.285	4.192	17.762	21.220
RMSE	1.932	16.640	9.507	4.569	7.155	5.558	5.156	20.323	25.26
Theil's inequality statistic	0.764	0.817	0.840	0.904	1.058	0.944	0.923	0.912	0.967
Regression: intercept	-0.151	-4.902	1.028	-1.365	4.424	-1.295	1.721	-1.151	-5.628
	(0.20)	(0.97)	(0.31)	(1.03)	(1.46)	(0.86)	(1.08)	(0.22)	(0.92)
slope	0.768	1.039	1.073	0.333	0.478	0.409	1.093	0.414	0.155
	(1.06)	(0.27)	(0.34)	(1.77)	(1.47)	(1.39)	(0.28)	(2.09)	(2.63)
$\bar{R}^2$	0.486	0.807	0.669	-0.019	0.063	-0.006	0.45	0.09	0.068

1/ Includes official transfers.
2/ Excludes official transfers.

Note: For definitions etc., see Note to Table 3.

Seven and Total Industrial) and for the CY (Canada and Japan) forecasts, whilst the overall explanatory power of the forecast is rated very low, at least in the YA forecasts. There is, it is important to note, a considerable improvement in forecast accuracy when the forecast horizon is reduced--though the CY forecasts are still markedly inferior to forecasts of a corresponding term for output or inflation.

The relative weakness of the balance of payments forecasts revealed by these summary statistics is not unexpected and is in line with experience at a national level and with the OECD forecasting track record (see Section III below). The problem is evidently related to the sizable fluctuations that have been observed in the world current account discrepancy, particularly since the late 1970s. As the reasons for this discrepancy are imperfectly understood, WEO projections have to be based on an implicit assumption of relative stability in the projected path of the discrepancy. 1/It should also be noted that the current account is the difference between two large flows which each has a volume as well as a price component. Relatively small forecast errors in any of the underlying volume or price changes can thus induce relatively large errors in the absolute difference between the nominal flows. Finally, as discussed in Appendix F, exchange rate innovations may at times have contributed to the errors in current account projections.

Table 8 shows the forecasts for world trade and industrial countries' terms of trade. (Repeated in the table for convenience are the statistics pertaining to industrial countries' export and import volume forecasts.) The overall record for trade as conveyed in these figures is one in which the shortening of the forecast horizon contributes greatly to accuracy, with a marked decline in the error statistics (the RMSE and average absolute error more than halve between the YA and CY forecasts) and a sharp rise in the explanatory power of the forecasts. Turning to the terms of trade, there is again a marked improvement in quality as the forecast horizon is reduced, yet in both cases there is strong evidence of inefficiency, with the forecasts underestimating whenever the terms of trade improve significantly. Inspection of the time series of the errors shows that while there are large positive forecast errors associated with the two rounds of large oil price increases these are more than offset by persistent negative forecast errors in the subsequent periods.

The direct tests for bias in the forecast errors, in the sense of significant differences from zero, are reported in full in Appendix C where in addition to the processing of individual country data, tests are also conducted on the pooled set of errors. The results suggest a

---

1/ The world current account discrepancy has recently been the subject of an extensive analysis that concluded that the main source of the discrepancy arises in the services account, see International Monetary Fund (1987).

Table 8. Forecast Accuracy--Summary Statistics: Terms of Trade and Trade Volumes

(In percentage changes)

	All industrial countries		Total world trade	Industrial countries' terms of trade
	Imports	Exports		
		Current Year (1972-86)		
Average absolute error	2.740	1.893	1.667	1.177
RMSE	3.016	2.469	1.993	1.522
Theil's inequality statistic	0.276	0.376	0.311	0.319
Regression: intercept	-1.446	-0.349	1.353	0.427
	(1.22)	(0.29)	(1.55)	(1.64)
slope	1.314	1.079	1.208	1.345
	(1.69)	(0.39)	(1.42)	(4.693)
$\bar{R}^2$	0.780	0.663	0.840	0.968
		Year Ahead (1972-85)		
Average absolute error	3.954	3.015	3.569	2.185
RMSE	4.929	3.917	4.335	3.150
Theil's inequality statistic	0.566	0.569	0.608	0.547
Regression: intercept	-4.306	-2.520	-4.021	-0.620
	(1.26)	(0.77)	(1.06)	(1.04)
slope	1.573	1.315	1.489	3.075
	(1.01)	(0.59)	(0.77)	(3.86)
$\bar{R}^2$	0.358	0.296	0.272	0.725

Note: For definitions etc., see Note to Table 3.

degree of output optimism in WEO forecasts. Although individual country output forecast errors are not significant, they are predominantly of the same sign, so that upon pooling a significant amount of bias is suggested--on the order of 0.3 percent in the CY forecasts, somewhat higher at about 0.5 percent in the YA forecasts (if the 1974 error is excluded from these). The output optimism appears to have been most pronounced in the second half of the 1970s, undoubtedly reflecting the fact that the deceleration in growth in many countries was only gradually perceived as a break in trend, rather than as a cyclical downturn. On the other hand there appears to be no bias in the inflation forecasts, at least not if the 1974 YA error is excluded. Interestingly, the inflation and output errors are significantly negatively correlated in the pooled data set, lending support to the contention (see Kenen and Schwartz (1986) that the implicit WEO forecasts for nominal income are more robust than those for either real output growth or inflation.

3. Summary Statistics: Developing Countries

Statistics of the forecasting record in respect of developing countries are presented in Tables 9 to 14. These plainly show a much poorer track record than that for the industrial countries.

In the output growth forecasts, for example, a majority of the Theil statistics for the YA forecasts exceed unity, while half of those for the CY forecasts do so. Apparently, a naive prediction of no change in output growth would have been a better forecast in these instances than the actual WEO forecasts. 1/ The fit of the realization-forecast regressions is also problematic, violation of efficiency being particularly strongly indicated for the Asia group in the YA forecasts. Nevertheless, average absolute errors appear reasonably low in relation to the mean absolute value of the outturn series and there appears to be an improvement with the reduction in forecast horizon length upon moving from the YA to the CY forecast sets. Much the same statements can be made of the record in relation to forecasts of inflation. These are considerably poorer than the corresponding forecasts for the industrial countries, with some notably high Theil statistics in the YA forecasts (where all but one exceed unity), generally low overall explanatory power in the realization-forecast regression and evidence of inefficiency in several cases.

For export and import volume growth, regional detail is available only for the CY forecasts; here the evidence is somewhat more reassuring. In the statistics on the export volume forecasts, only one of the Theil statistics exceeds unity (for the Middle East grouping, for which the average absolute forecast error itself exceeds the mean

---

1/ Because of delays in the production of data for these countries, however, it must be pointed out that the greater part of "last year's growth" is estimated in any case.

Table 9. Forecast Accuracy--Summary Statistics: Non-Oil Developing Countries' Output Growth

(In percent)

	Africa	Asia	Europe	Middle East	Western Hemisphere	Total Non-Oil Developing Countries
			Current Year (1977-86) 1/			
Mean absolute actual value	2.220	5.320	2.467	4.080	3.440	3.370
Average absolute error	1.31	1.10	0.66	1.63	1.92	0.96
RMSE	1.355	1.426	0.814	2.197	2.605	1.218
Theil's inequality statistic	0.942	0.946	0.796	1.089	1.008	1.055
Regression: intercept	-1.903	4.948	1.136	-0.630	-0.056	0.155
	(1.48)	(1.15)	(1.36)	(0.34)	(0.03)	(0.09)
slope	1.246	0.065	0.518	0.987	0.773	0.827
	(0.65)	(1.24)	(1.56)	(0.04)	(0.51)	(0.45)
$\bar{R}^2$	0.520	-0.124	0.184	0.387	0.186	0.289
			Year Ahead (1979-85) 2/			
Mean absolute actual value	2.343	5.386	2.150	4.186	3.471	3.629
Average absolute error	1.50	1.83	1.23	2.66	2.64	1.66
RMSE	1.681	2.124	1.300	3.182	3.736	1.947
Theil's inequality statistic	1.075	1.370	0.738	1.933	1.167	1.206
Regression: intercept	0.475	25.481	2.911	-1.793	-3.566	3.404
	(0.17)	(5.10)	(1.27)	(0.35)	(0.67)	(0.84)
slope	0.597	-3.390	-0.255	0.960	1.318	0.047
	(0.56)	(5.22)	(1.66)	(0.05)	(0.27)	(1.14)
$\bar{R}^2$	-0.098	0.718	-0.215	0.047	0.039	-0.20

Note: For definitions etc., see Note to Table 3.

1/ Current Year data for Europe cover the period 1978-86.
2/ Year Ahead data for Europe cover the period 1980-85.

Table 10. Forecast Accuracy--Summary Statistics: Non-Oil Developing Countries' Inflation

(In percent)

	Africa	Asia	Europe	Middle East	Western Hemisphere	Total Non-Oil Developing Countries
			Current Year (1977-86)			
Mean absolute actual value	19.240	8.400	26.77	25.830	82.01	36.860
Average absolute error	4.31	1.81	5.09	6.51	16.13	6.31
RMSE	5.988	2.157	6.411	10.730	18.915	7.059
Theil's inequality statistic	1.109	0.654	0.940	1.004	0.752	0.898
Regression: intercept	22.212	-0.190	2.843	10.503	-5.441	-2.851
	(3.29)	(0.13)	(0.32)	(1.18)	(0.98)	(0.58)
slope	-0.173	1.269	1.104	0.576	1.327	1.300
	(3.04)	(1.25)	(0.25)	(1.41)	(4.17)	(1.89)
$\bar{R}^2$	-0.097	0.790	0.409	0.215	0.969	0.880
			Year Ahead (1979-85) 1/			
Mean absolute actual value	19.614	8.429	24.017	29.629	89.700	38.114
Average absolute error	2.800	2.314	8.667	16.057	39.990	12.629
RMSE	3.749	2.703	11.065	22.072	49.965	14.767
Theil's inequality statistic	0.584	1.034	1.348	1.785	2.541	3.042
Regression: intercept	5.425	2.608	53.381	48.196	-62.253	-24.322
	(0.89)	(0.64)	(2.09)	(4.07)	(1.89)	(0.55)
slope	0.830	0.898	-1.272	-0.557	3.057	2.450
	(0.49)	(0.17)	(1.76)	(4.69)	(3.17)	(0.83)
$\bar{R}^2$	0.437	0.159	-0.007	0.232	0.779	0.140

Note: For definitions etc., see Note to Table 3. Inflation is measured by consumer price indices.

1/ For Europe, Year Ahead data cover 1980-85.

Table 11. Forecast Accuracy--Summary Statistics: Non-Oil Developing Countries' Export Growth

(In percent)

	Current Year (1981-86) 1/						Year Ahead (1977-85)
	Africa	Asia	Europe	Middle East	Western Hemisphere	Total Non-Oil Developing Countries	Total Non-Oil Developing Countries
Mean absolute actual value	3.633	8.650	6.600	5.650	4.800	5.845	6.189
Average absolute error	3.15	5.90	4.12	8.53	5.00	2.14	3.078
RMSE	4.270	6.619	4.894	9.434	5.726	2.673	3.804
Theil's inequality statistic	0.986	0.740	0.591	1.204	0.880	0.528	0.845
Regression: intercept	-0.948	17.057	3.050	0.802	-6.865	0.795	2.621
	(0.15)	(1.42)	(0.29)	(0.13)	(2.39)	(0.40)	(0.25)
slope	0.674	-1.193	0.689	-0.265	1.344	0.853	0.146
	(0.27)	(0.12)	(0.16)	(1.00)	(0.86)	(0.47)	(0.44)
$\bar{R}^2$	-0.160	-0.107	-0.213	-0.237	0.675	0.320	-0.142

Note: For definition etc., see Note to Table 3. Regional details for Year Ahead data are not available.

1/ Current year data for Total Non-Oil Developing Countries cover 1972-86.

Table 12. Forecast Accuracy--Summary Statistics: Non-Oil Developing Countries' Import Growth

(In percent)

	Current Year (1979-86) 1/						Year Ahead (1977-85)
	Africa	Asia	Europe	Middle East	Western Hemisphere	Total Non-Oil Developing Countries	Total Non-Oil Developing Countries
Mean absolute actual value	4.625	6.350	3.600	7.538	8.738	5.520	5.050
Average absolute error	4.913	1.838	3.471	5.625	10.175	3.200	3.744
RMSE	6.249	3.022	5.703	7.074	10.711	4.134	5.548
Theil's inequality statistic	1.246	0.646	1.029	0.962	0.913	0.627	1.037
Regression: intercept	-3.520	0.687	3.500	-3.873	-6.228	0.540	2.621
	(2.00)	(0.22)	(1.40)	(1.49)	(1.54)	(0.31)	(0.25)
slope	-0.886	0.801	-0.662	1.066	1.172	0.832	0.146
	(1.25)	(0.49)	(2.46)	(0.17)	(0.26)	(0.52)	(0.44)
$\bar{R}^2$	-0.104	0.266	-0.006	0.495	0.236	0.282	-0.142

Note: For definition etc., see Note to Table 3. Regional detail for Year Ahead data not available.

1/ Current year data for Total Non-Oil Developing Countries cover 1972-86, while current year data for Europe cover 1980-86.

Table 13. Forecast Accuracy--Summary Statistics: Non-Oil Developing Countries' Balance of Payments on Current Account

(In billions of dollars)

	Africa	Asia	Europe	Middle East	Western Hemisphere	Total Non-Oil Developing Countries 2/
			Current Year (1977-86) 1/			
Mean absolute actual value	9.120	12.130	5.743	11.860	19.290	42.467
Average absolute error	2.57	4.80	1.47	2.65	4.86	6.68
RMSE	3.695	6.411	2.113	3.547	6.639	8.028
Theil's inequality statistic	0.740	0.885	1.316	0.548	0.672	0.509
Regression: intercept	-4.422	-2.554	-2.762	-2.194	-1.644	-2.020
	(1.16)	(0.68)	(3.16)	(1.38)	(0.36)	(0.49)
slope	0.494	0.668	0.608	0.727	0.952	0.943
	(1.33)	(1.46)	(2.80)	(2.68)	(0.23)	(0.71)
$\bar{R}^2$	0.071	0.458	0.749	0.848	0.682	0.907
			Year Ahead (1978-85) 2/			
Mean absolute actual value	8.350	14.500	5.100	8.800	20.975	59.062
Average absolute error	2.538	6.588	1.660	3.438	10.763	21.825
RMSE	3.896	7.491	1.756	4.933	12.098	24.041
Theil's inequality statistic	0.893	1.222	0.911	0.850	1.146	1.181
Regression: intercept	-3.294	-8.187	-0.743	-4.692	-5.515	-21.479
	(0.40)	(1.75)	(0.60)	(0.72)	(0.46)	(0.86)
slope	0.622	0.399	0.703	0.366	0.697	0.597
	(0.47)	(2.30)	(1.66)	(1.10)	(0.62)	(1.08)
$\bar{R}^2$	-0.062	0.160	0.783	-0.093	0.124	0.184

Note: For definitions etc., see Note to Table 3.

1/ Current Year data for Non-Oil Developing Countries cover 1972-86, while current year data for Europe cover 1980-86.

2/ Year Ahead data for Europe cover the period 1981-85.

Table 14. Primary Product Price Forecasts

(In percent)

	Agricultural Raw Materials	Beverages	Food	Metals	Non-Oil Developing Countries' Exports
			Current Year (1981-86)		
Mean absolute average value	7.46	12.29	11.67	8.07	8.06
Average absolute error	5.652	8.095	7.562	8.533	5.887
RMSE	8.954	10.080	9.259	8.954	7.497
Theil's inequality statistic	0.800	0.534	0.556	1.209	0.668
Regression: intercept	-5.505	-1.103	-5.112	-8.248	-5.487
	(1.967)	(0.287)	(1.279)	(6.778)	(2.313)
slope	1.013	0.649	0.942	0.794	0.833
	(0.025)	(1.669)	(0.029)	(1.098)	(0.567)
$\bar{R}^2$	0.376	0.631	0.382	0.771	0.585

Note: For definitions etc., see Note to Table 3.

absolute value of the outturn series), though for import growth, Theil statistics exceed unity in three out of the seven cases. The overall explanatory power of the forecasts in the realization-forecast regression is generally low although indications of inefficiency are confined to European import growth forecasts.

The balance of payments forecasts, finally, provide some indication of weakness; the YA forecasts produce three instances of Theil statistics above unity, with evidence of bias in the Asia grouping and generally low explanatory power for the forecasts in the realization-forecast regression. The CY forecasts are somewhat better. The average absolute errors are generally lower, in relation to the absolute mean of the outturns, the Theil statistics (that for Europe excepted) are lower, overall explanatory power of the forecasts higher, though with more evidence of departure from the efficiency requirements on the parameters of the regression.

The results of directly testing for bias, both on individual area results and pooling the data as a whole are again reported in Appendix C. For the developing countries too, these tests suggest a tendency toward output optimism, at least in the year-ahead sample. Some individual area bias in inflation estimates also appears, though this is not significant when the data are pooled.

One reason for weakness in developing country forecasting is the extent to which such forecasts must rely upon projections of commodity prices, themselves known to be associated with large margins of uncertainty. Unfortunately, changes of definition and noncontinuities in reporting such forecasts in the WEO documents make it impossible to examine more than a small run of years of commodity price projections. Table 14 reports some summary statistics on forecasts made for four individual groups of commodities and the aggregate of interest here, non-oil developing countries' exports. The variability of these prices is notably high and it is not too surprising that the average absolute errors--ranging from 5.7 to 8.5 percent--are also rather big. Even so, the forecasts do at least compare well with the naive standard and only a proportion appear to infringe the efficiency criteria in the realization-forecast regression.

The summary statistics reviewed, based on the overall record, provide a number of general conclusions. First, industrial country forecasting appears to be much better than that for developing countries. This is perhaps not surprising: developed countries are better understood, data streams are not so thin and are more reliable. It should also be borne in mind that the quality of the data analyzed in these tables is less good for the developing countries due to frequent changes in definitions and coverage.

Second, among the industrial country forecasts, the balance of payments forecasts appear considerably worse than those for output, inflation, export or import volumes. This should not be cause for great

surprise (though it may be cause for concern): the balance of payments is the difference between two series; small changes in these can induce large changes in the difference, which as a data series may be volatile both in behavior and in its revisions. Moreover, the emergence of a relatively large and volatile discrepancy on the world's aggregate current account--which in principle should be in balance--casts some doubt on the quality of balance of payments data even for the industrial countries. Difficulty in forecasting the balance of payments is a common complaint among national forecasters. 1/

A third conclusion that can be drawn for the industrial countries is that the CY forecasts are superior to the YA forecasts. While it may not seem surprising (it may even appear obvious) that near-term forecasting is more accurate than longer-term forecasting, such results are not invariably recorded (see, for example, Burns (1986) for a contrary instance).

Fourth, the record appears comparatively free from inefficiency in the sense that country-by-country and area-by-area the parameters of the realization-forecast regression conform by and large to the requirements of efficiency, although upon pooling the data direct tests for the significance of forecast bias produce some evidence of an output optimism error (more pronounced for the YA than for the CY forecasts).

Finally, it should be noted that the relative inferiority of the balance of payments forecasts and of the YA forecasts for the industrial countries do not carry over to the developing countries. By and large, these results are similar to those arrived at by Kenen and Schwartz (1986) in their study of the Fund's forecasting and they were confirmed by additional replicating calculations using the latest available estimates of outturns (Appendix B).

III. Comparisons

It is natural to enquire how well forecasting in the Fund compares with other forecasts. Here we consider two alternatives, the OECD and national forecasting agencies. In these comparisons we are able to follow the lead set by Llewellyn and Arai (1984) who have already compared OECD and national forecasting records.

1. Comparisons with OECD

There have been a number of analyses of OECD's track record besides that of Llewellyn and Arai (such as Smythe (1983), Smythe and Ash (1975), and Holden, Peel, and Sandu (1987)). It would be nice to suppose that in choosing the OECD as a comparison one is also choosing a

---

1/ The following section shows that OECD's balance of payments forecasting is similarly weaker than its output and inflation records.

forecast which is--at least in more recent years--less "judgmental" and more model-based than that underlying the WEO forecasts, for this would give added point to the comparison. However, it is not clear how far such a contrast is realistic. 1/

The forecasts compared here are those for output growth, inflation and the balance of payments on current account of the Group of Seven countries individually and in aggregate. In Llewellyn and Arai (1984) attention was focused on the OECD's forecasts a year ahead, using issues of the OECD Economic Outlook for December of year t-1 for forecasts for year t, the realizations coming from the OECD Economic Outlook for December of the following year (t+1). An immediate problem is that WEO forecasts are not always based on the same information set as that conditioning the OECD forecasts, the historically less regular WEO round drawing on some forecasts made as early as August of the previous year and as late as March in the year in question. In order to achieve as close a match as possible, since 1981 year ahead forecasts in July issues of the OECD Economic Outlook have been compared with the WEO's August forecasts. But this is only a partial solution to the problem and the remaining differences in timing, whilst apparently not severe, are unfortunate. 2/ The fact that calendar time discrepancies between forecast dates are small is uncertain assurance that discrepancies between the corresponding information sets are in the relevant sense also small. The scope of the comparison is from 1973 to 1985.

Table 16 lists summary statistics for the OECD forecasts which can be compared with corresponding statistics for the World Economic Outlook (shown as the year ahead forecasts in tables 3, 4, and 7). Within Table 16 itself the summary statistics quoted show that the balance of payments is the worst forecast for the OECD too: the Theil statistics are generally higher than for the output and inflation forecasts, while the overall explanatory power in the realization-forecast regression is often very low. Similarly, average absolute errors are larger in relation to the absolute value of the balance of payments than they are for output growth or inflation (indeed they are sometimes bigger than the mean absolute value of the balance of payments itself). These features are identical to those found for the WEO forecasts of the same variables. The regression evidence suggests a departure from efficiency in the forecasts of inflation in Italy (just as in the WEO forecasts--see Table 2) and, in addition, for the balance of payments forecasts for that country and Canada.

Realizations do not differ greatly between the OECD and WEO forecasts and an initial comparison between the average absolute errors and RMSEs of the two seems reasonable. On this basis, the OECD output

---

1/ An account of OECD forecasting methods appears in Llewellyn et al. (1985).

2/ Full details of the basis for the comparisons are given in Table 15.

Table 15. Basis for OECD-WEO Comparison

Forecast for:	Published in: OECD Economic Outlook	Published in: Fund's World Economic Outlook	Realization Data for:
1973	Dec. 1972	Jan. 1973	n.r.
1974	Dec. 1973	Dec. 1973	n.r.
1975	Dec. 1974	Dec. 1974	1973
1976	Dec. 1975	Dec. 1975	1974
1977	Dec. 1976	Mar. 1977	1975
1978	Dec. 1977	Dec. 1977	1976
1979	Dec. 1978	Feb. 1979	1977
1980	Dec. 1979	Aug. 1979	1978
1981	Dec. 1980	Aug. 1980	1979
1982	July 1981	Aug. 1981	1980
1983	July 1982	Aug. 1982	1981
1984	July 1983	Aug. 1983	1982
1985	July 1984	Sept. 1984	1983
1986	Dec. 1985	Oct. 1985	1984
..	Dec. 1986	Oct. 1986	1985

Table 16. OECD Forecast Accuracy--Summary Statistics: Group of Seven

	Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom	Group of Seven
				Year Ahead (1973-85): Output Growth (In percent)				
Mean absolute actual value	3.377	3.400	4.654	2.531	2.685	2.954	2.315	2.992
Average absolute error	1.790	1.150	1.640	1.010	1.520	2.100	1.190	1.070
RMSE	2.52	1.68	2.81	1.54	2.20	2.48	1.64	1.50
Theil's inequality statistic	0.660	0.414	0.728	0.600	0.648	0.627	0.646	0.476
Regression: intercept	-0.385	-0.106	2.124	-0.062	-1.279	0.771	0.217	-0.493
	(0.21)	(0.15)	(1.10)	(0.07)	(0.85)	(0.79)	(0.37)	(0.57)
slope	0.883	1.038	0.458	0.852	1.160	0.651	0.796	1.178
	(0.24)	(1.18)	(1.51)	(0.52)	(0.33)	(1.07)	(0.84)	(0.64)
$\bar{R}^2$	0.166	0.666	0.049	0.394	0.278	0.198	0.447	0.583
				Year Ahead (1973-85): Inflation (In percent)				
Mean absolute actual value	8.346	6.672	5.377	9.961	4.369	15.369	11.831	7.346
Average absolute error	2.11	1.52	3.35	1.362	0.62	2.22	2.74	1.26
RMSE	2.877	4.617	0.739	1.735	1.677	3.026	3.459	1.775
Theil's inequality statistic	0.959	0.733	0.895	0.754	0.451	0.859	0.546	0.804
Regression: intercept	3.222	1.585	-0.210	3.337	-0.643	6.149	-0.934	1.178
	(1.05)	(1.00)	(0.72)	(1.96)	(0.95)	(2.69)	(0.36)	(0.60)
slope	0.641	0.756	0.968	0.685	1.196	0.673	1.165	0.847
	(0.97)	(1.10)	(0.07)	(0.09)	(1.23)	(0.04)	(0.74)	(0.59)
$\bar{R}^2$	0.142	0.469	0.235	0.485	0.831	0.581	0.690	0.445
				Year Ahead (1973-85): Balance of Payments (In billions of dollars)				
Mean absolute actual value	2.825	24.795	13.288	4.092	6.916	4.498	4.899	19.385
Average absolute error	3.18	15.28	7.51	3.34	4.35	4.72	3.79	16.04
RMSE	5.039	22.382	9.978	3.933	5.372	6.023	4.878	20.864
Theil's inequality statistic	1.782	1.044	0.883	0.767	0.724	1.043	0.840	0.913
Regression: intercept	-2.042	-7.620	1.108	-1.488	0.562	-2.239	0.945	-9.495
	(1.91)	(1.19)	(0.32)	(0.95)	(0.33)	(1.55)	(0.66)	(1.44)
slope	-0.020	1.155	1.059	0.639	1.163	0.111	1.221	0.698
	(4.98)	(0.78)	(0.25)	(0.96)	(0.34)	(2.30)	(0.51)	(0.81)
$\bar{R}^2$	-0.090	0.731	0.621	0.134	0.473	-0.083	0.361	0.180

Note: For definitions etc., see Note to Table 3.

growth forecasts emerge as slightly superior to those in the WEO in that in a majority of cases, the OECD error is less than that of the WEO. For inflation and the balance of payments, however, the evidence suggests the opposite. Thus, the WEO balance of payments record does not appear to be capable of improvement by using the unexploited information in OECD forecasts. We turn, in a moment, to consideration of a more formal test of this proposition.

a. Error triangles. An alternative way of comparing these forecasts is to take the growth, inflation and balance of payments forecasts together for a particular country. A typical social welfare or decision function would consider a weighted sum of squared prediction errors for these variables as the minimand; the quadratic form implies symmetrical penalties, as seems appropriate in this case, 1/ increasing sharply with the value of the error. Without access to the weights with which the squared prediction errors are to be combined, the approach cannot be used to quantify exhaustively the relative success of the two organizations. However, where each of the prediction errors for the three variables of interest for a particular country and period is less for one set of forecasts than for the other, the former set is unequivocally superior to ("dominates") the latter, whatever the appropriate weights may be. In diagrammatic terms (see Chart 1), a dominant forecast provides a triangle in output/inflation/balance of payments error space which lies wholly within the other (the chart displays triangles for Japan and France for illustrative purposes). The forecasts made for each of 13 years, for seven individual countries and the Group of Seven aggregate provide the basis for 104 such triangles. In addition, there are at least eight period-averages that can be considered. For the single-year comparisons, the absolute error clearly provides the same results as squared errors and Table 17 shows that on this basis there were 22 cases out of the total of 104 in which one set of forecasts was unequivocally better than the other. In just under two thirds of these, the WEO forecast turned out to be dominant.

Looking at the comparative performance over the period as a whole, it becomes essential to consult a quadratic error statistic: in fact, the Theil statistics already collected in Table 3, 4, 7 and 16 fit the bill well enough, and also have the merit that they standardize for the (small) differences in realizations as between the two sets of forecasts. A comparison of these statistics yields the result that, over the whole period, the OECD forecasts dominate those of the WEO for France and for the Federal Republic of Germany whilst the WEO forecasts for Canada dominate those of OECD.

1/ Since we are only considering deviations from forecasts, rather than from socially desirable target values, some of the objections directed towards the quadratic form when that is being used to derive a social welfare function to direct the optimal application of policy do not apply.

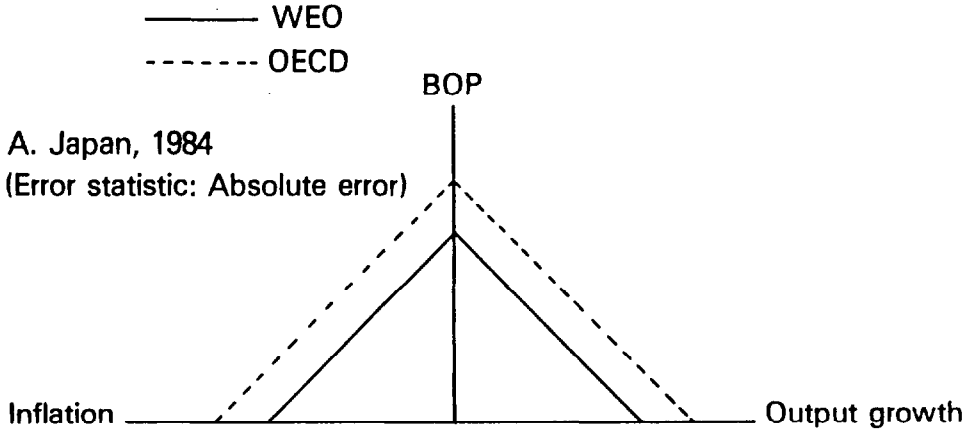
Table 17. Forecast Comparisons: International Monetary Fund and OECD

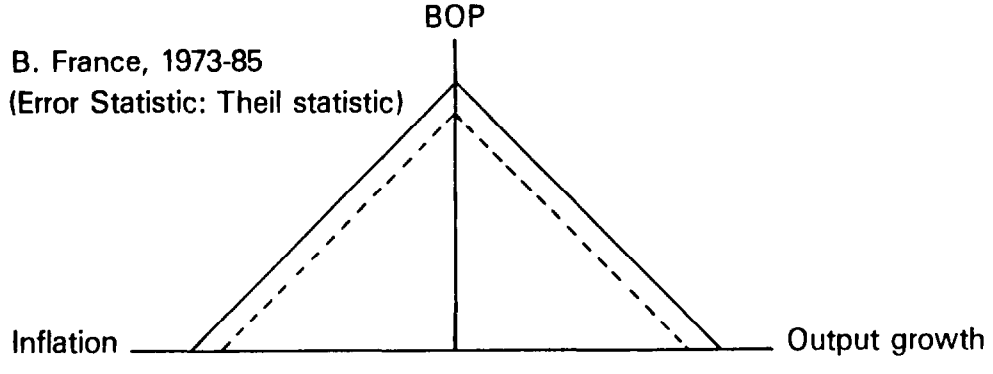
Year	Forecaster	Country	Dominant Forecast by: 1/
1973	None	None	0
1974	OECD	United States	1
1975	OECD	France	1
	Fund	Canada	1
1976	Fund	Japan, United States	2
1977	Fund	United States	1
	OECD	Group of Seven	1
1978	None	None	0
1979	Fund	Japan, Italy	2
1980	OECD	France, Group of Seven	2
1981	OECD	France	1
1982	Fund	Canada, United States United Kingdom	3
1983	Fund	Canada	1
	OECD	United States	1
1984	Fund	France, Japan, Group of Seven	3
	OECD	United Kingdom	1
1985	Fund	Canada	1
		Total	22
		Fund	14
		OECD	8

1/ That forecast with simultaneous minimum average absolute error for three variables--output, inflation, and the balance of payments.

CHART 1

ERROR TRIANGLES: ILLUSTRATIVE COMPARISONS OECD AND WEO FORECASTS





b. Error patterns. The evidence so far suggests that there is generally little to choose between the two sets of forecasts; with some slight risk of misrepresentation this might be characterized as saying that the two organizations tend to make the same errors about the same variables for the same countries at the same time. Tables 18 - 20 both confirm this and provide qualifications. These tables show the pattern of forecast errors (forecast minus realization) for the three variables of interest and the simple correlation coefficients between them, both across countries for a particular year and across years for a particular country. The latter are notably high and on the whole higher than the former. Some years of "disagreement" (1978 on output growth, 1982-85 on inflation) thus stand out, but in doing so serve to highlight the fact that the pattern of errors is usually very similar between the two organizations. In Llewellyn and Arai (1984) a similar conclusion was drawn about a comparison of the OECD and national forecasters.

c. Non-nested tests. A formal test of the proposition that the OECD forecast could improve the WEO forecast and vice-versa was nevertheless implemented. Borrowing from the methodology of non-nested tests as applied to estimated equations, it was proposed to ask whether the OECD (WEO) forecast had any power in explaining the error of the WEO (OECD) forecast. For this purpose it seemed appropriate to use, not the raw forecast error, but the residuals from the forecast-realization regression, the "structural forecast error". 1/ Thus the test conducted was, for each of the three variables (output growth, inflation, and the balance of payments) and for each of the Group of Seven countries and the aggregate to run the regressions:

$$E(WEO)(t) = a + bF(OECD)(t) + u(t)$$

$$E(OECD)(t) = c + dF(WEO)(t) + w(t)$$

where E(WEO) and E(OECD) are the respective structural errors and F(WEO), F(OECD) the corresponding raw forecast. These tests were carried out on a subset of the comparison data set (terminating in 1984, and starting in 1974 for the balance of payments). Only in three cases did the results of these regressions suggest unexploited information in the alternative forecast. Both forecasts for the balance of payments of Canada could be improved by knowledge of the other, and a reduction in the WEO error on the U.S. balance of payments could be effected by exploiting the OECD forecast. In the circumstances it seems excessively burdensome to report the findings in full.

1/ This eliminates the possibility of merging two logically distinct methods of improving a forecast--on the one hand using the knowledge to be gained from the realization-forecast regression of the forecast itself, on the other, exploiting the information content of an alternative forecast.

Table 18. Forecast Error Comparison: International Monetary Fund and OECD

Output Forecast Errors
(In percentage points)

	Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom	Group of Seven	Correlation Coefficient 1/
OECD 1973	-0.8	0.4	0.6	0.0	-0.1	-0.1	-0.3	-0.3	
Fund	-0.8	0.4	0.5	0.0	-0.3	-0.3	-1.0	0.0	0.74
1974	2.7	4.4	9.3	1.6	2.9	4.1	3.4	4.0	
	2.2	4.9	10.8	1.4	2.6	2.6	3.6	4.8	0.98
1975	2.9	-0.2	-0.1	4.2	5.7	3.5	3.6	1.2	
	1.9	0.3	1.2	4.7	5.4	4.6	4.0	2.0	0.93
1976	-0.7	-0.3	-2.1	-2.2	-2.5	-4.1	-2.1	-1.6	
	0.8	0.2	-0.7	-1.8	-2.2	-2.2	-0.7	-0.8	0.94
1977	0.8	-0.4	0.8	0.0	0.9	-2.2	-0.4	-0.4	
	1.3	0.3	1.2	0.4	2.2	-2.2	0.6	0.6	0.965
1978	0.4	-0.2	-0.6	-0.1	-0.3	-1.6	-0.3	-0.6	
	1.1	0.1	-0.8	0.1	0.0	0.4	-0.5	0.0	0.245
1979	1.2	-0.3	-1.2	0.2	-0.5	-1.5	0.8	-0.7	
	1.3	1.1	-1.1	0.0	-0.4	-0.1	1.0	0.3	0.865
1980	1.4	-1.1	0.6	0.7	0.5	-2.0	-0.2	-0.5	
	3.0	0.2	0.5	1.0	1.3	-1.2	1.4	0.6	0.865
1981	-1.8	-1.3	0.9	0.7	0.1	-0.8	0.2	-0.5	
	-2.3	-3.4	1.1	1.6	2.3	0.6	1.8	-0.6	0.83
1982	7.3	2.7	1.0	0.1	3.1	2.1	-1.0	1.9	
	5.8	2.6	2.7	0.6	3.1	2.1	-0.5	2.4	0.96
1983	-2.0	-1.2	1.0	1.8	2.0	3.7	-1.4	-0.2	
	-1.2	-1.6	-0.5	1.8	0.7	4.5	-1.6	-0.4	0.94
1984	-0.3	-2.3	-2.3	-1.1	-1.0	-0.6	-0.4	-1.7	
	-0.8	-2.5	-1.8	-0.3	-0.6	-0.2	-0.5	-1.7	0.855
1985	-1.0	-0.2	-0.8	0.4	0.3	-0.1	-1.3	-0.32	
	-0.9	1.3	-0.4	0.4	-0.1	0.2	-0.9	0.5	0.68
Correlation Coefficient	0.93	0.90	0.98	0.97	0.91	0.95	0.91		

1/ Excluding Group of Seven.

Table 19. Forecast Error Comparison: International Monetary Fund and OECD

Inflation Forecast Errors
(In percentage points)

		Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom	Group of Seven	Correlation Coefficient 1/
OECD	1973	-3.1	-2.1	-6.3	-1.8	-0.6	-4.0	-0.1	-2.5	
Fund		-2.8	-2.3	-6.3	-2.3	-0.4	-2.3	-0.5	-2.5	0.94
	1974	-7.1	-3.6	-11.9	-3.4	0.2	-6.9	-6.3	-4.9	
		-7.4	-4.7	-8.6	-4.4	0.2	-7.6	-4.8	-4.9	0.92
	1975	1.2	1.2	7.4	-1.3	-1.6	1.0	-7.8	1.0	
		0.5	0.5	9.0	-2.1	-1.1	-1.0	-7.8	0.8	0.975
	1976	-0.3	1.5	-0.9	0.9	0.7	-5.3	1.1	0.4	
		-0.5	0.6	-0.1	1.1	1.2	-2.8	0.0	0.3	0.92
	1977	0.6	-0.7	1.3	0.6	0.4	1.5	-1.3	0.0	
		0.3	-0.5	1.3	-0.5	0.1	2.7	-0.6	0.0	0.79
	1978	0.1	-1.3	0.5	-1.4	0.1	-0.6	1.5	-0.5	
		0.0	-1.3	-0.8	-1.2	0.1	-0.8	0.4	-0.8	0.84
	1979	-3.5	-1.6	2.5	-0.8	-0.3	-2.5	-4.6	-0.8	
		-3.4	-1.1	0.2	-1.5	-0.2	-1.8	-4.2	-1.1	0.93
	1980	-1.5	0.6	2.6	-0.5	-0.6	-3.7	-2.3	0.1	
		-3.2	0.1	4.4	-1.8	0.1	-6.2	-4.6	-0.1	0.98
	1981	-0.5	1.3	2.2	-0.4	-0.3	-0.9	3.1	1.0	
		-2.0	-1.3	1.6	-3.1	0.2	-3.8	1.2	-0.8	0.805
	1982	0.4	1.9	2.8	-1.1	-1.3	0.0	2.4	1.4	
		-0.3	1.2	1.5	0.5	-0.3	2.0	1.0	1.0	0.52
	1983	4.2	1.3	3.1	4.1	-0.5	-0.3	2.2	1.7	
		1.6	1.8	2.3	2.1	1.3	2.0	2.2	1.9	0.39
	1984	3.2	1.2	1.7	1.2	0.6	1.6	1.6	1.6	
		1.9	0.4	1.6	0.3	1.1	4.1	1.8	1.1	-0.325
	1985	1.7	1.4	0.3	0.2	0.8	-0.6	-0.8	0.5	
		1.3	1.0	-0.5	-0.4	0.4	2.3	-1.1	0.5	0.285
Correlation Coefficient		0.96	0.89	0.95	0.83	0.67	0.82	0.95		

1/ Excluding Group of Seven.

Table 20. Forecast Error Comparison: International Monetary Fund and OECD

Balance of Payments Forecast Error
(In billions of dollars)

		Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom	Group of Seven	Correlation Coefficient 1/
OECD	1973	-0.82	-5.84	5.34	1.99	-5.41	5.87	2.38	3.50	
Fund		-0.4	-6.6	6.2	1.6	-6.1	4.7	2.2	1.6	0.99
	1974	1.7	5.6	4.2	5.2	-8.6	6.45	5.80	20.35	
		1.4	6.6	3.6	4.6	-8.7	4.6	4.3	16.4	0.99
	1975	1.15	-19.2	0.7	-6.15	2.1	-5.15	-2.8	-29.35	
		0.9	-18.1	0.7	-6.4	4.8	-5.4	-3.8	-27.3	0.99
	1976	0.45	6.65	-8.2	3.6	-1.4	3.05	-0.75	3.4	
		0.4	4.4	-5.4	3.4	0.2	3.7	-2.1	4.6	0.96
	1977	0.4	12.3	-10.9	-0.45	1.3	-1.8	-1.75	-0.9	
		0.1	9.5	-8.9	-1.2	0.1	-3.4	-3.4	-7.2	0.99
	1978	0.85	-5.85	-6.5	-5.9	-5.8	-4.65	1.5	-25.85	
		0.7	-2.8	-5.8	-7.3	-6.1	-6.4	1.9	-25.8	0.92
	1979	-0.1	-7.2	20.8	0.8	7.75	-0.35	5.4	27.1	
		1.0	-6.7	19.6	0.4	9.9	-0.2	5.6	29.6	0.99
	1980	-5.9	-0.2	1.95	5.9	11.4	15.1	-7.0	21.25	
		-4.1	-5.4	9.0	7.3	14.0	14.4	-4.6	30.6	0.91
	1981	2.5	13.25	-11.45	1.25	-2.9	5.75	-11.95	-3.55	
		-0.2	12.5	-17.2	2.1	-4.5	6.5	-13.2	-14.0	0.98
	1982	-12.95	11.1	10.1	5.25	-1.55	0.5	-4.65	7.8	
		-4.7	3.8	1.4	6.4	-10.9	-1.6	-1.9	-7.5	0.63
	1983	-10.3	35.3	-0.05	-3.0	0.85	-7.25	2.65	18.2	
		-1.8	36.0	-5.2	-6.2	1.7	-2.5	-5.5	16.5	0.93
	1984	1.5	64.0	-13.25	-3.2	-4.3	1.0	0.05	45.8	
		1.3	39.0	-10.4	-1.1	-0.2	-0.2	0.8	29.2	0.96
	1985	2.65	12.70	-13.20	0.70	-3.20	4.45	-2.60	1.5	
		0.2	-2.5	-9.1	1.4	-7.5	2.1	-5.2	-20.6	0.67
Correlation Coefficient		0.87	0.94	0.91	0.96	0.89	0.95	0.85		

1/ Excluding Group of Seven.

In conclusion, it is apparent that the WEO forecasts compare well with OECD's and that there is no substantial evidence that there is unexploited information in these forecasts which the WEO could benefit from using. As the two groups of forecasters "breathe the same air," exchange information, and maintain contacts with the same national forecast agencies, the conclusion is perhaps not surprising.

2. Comparison with National Forecasts

In this section we take up the issue of how the WEO forecasts compare with national ones. In this we follow the lead set by Llewellyn and Arai (1984), using a data set similar to theirs but extended by three years to cover the period 1973-85. 1/

Llewellyn-Arai. First, recall the basis for the Llewellyn-Arai study and its principal conclusions. Llewellyn and Arai approached forecasters in a large number of OECD countries for data on output and inflation forecasts and outturns. They described the terms in which these were to be defined in the following way: "...the forecasts sought were those made in October, November or December for the following calendar year. The forecasters were asked to supply the values of the actuals against which the forecasts should be compared. This seemed the most sensible approach because only those who made the forecasts really know which series, to which base year, and so on they were trying to predict. Forecasters were asked to provide estimates of the actuals made about one year after the outcome..." 2/ The OECD forecasts against which the national data were compared were compiled on a year ahead basis matching forecasts for year t, taken from the OECD's Economic Outlook in December of year t-1, against outturns taken from the Economic Outlook in December of year t+1. The coverage of the Llewellyn-Arai paper was ambitious, in terms both of the number of countries covered and the length of time for which series were obtained, the paper deploying data for 1966-82 for all OECD countries with the principal exceptions of Australia, Canada, and Ireland.

The main results of the comparison were based on an analysis of the total forecast error distribution of OECD and the national forecasters, and analysis of that distribution in particular years. From the frequency distribution of forecast errors, it was concluded that there was a slight tendency towards skewness in the direction of "growth optimism" for output but that overall neither one of the two sets of forecasts exhibited a marked superiority over the other. Most important, perhaps, Llewellyn and Arai detected that there were certain years which were marked by positive or negative skewness, suggesting that in these periods a shock external to the OECD (such as the oil

---

1/ The author is most grateful to John Llewellyn for assisting him in the updating and to those individuals who supplied him with national forecasts and outturns.

2/ Llewellyn and Arai (1984), p. 83.

price increases) or a system-wide behavioral innovation (the decline in expenditures with rising inflation, for example) was the predominant source of error; in other years individual country errors were largely offsetting (a similar phenomenon in comparing the WEO with OECD is noted above). In the case of inflation, the error distribution was found to display a marked negative skewness, though this appeared to reflect most heavily experience in the early years of the series.

The present study sought to update the national forecast series in Llewellyn and Arai for comparison with the WEO year ahead forecast series. Those who supplied the original national forecast data were approached again and details of other forecasts were obtained from the corresponding country desk officers in the Fund. The resultant extensions, obtained for most of the series used in the original study, are shown in Appendix E. They make up series for each of the Group of Seven countries from 1973 to 1985, with the exception of Canada, and in addition cover Finland, Netherlands, Sweden, Switzerland, and Austria. 1/ As the WEO does not publish forecasts for these last countries their contribution is assessed in the context of forecasts for the aggregate "Europe."

The information conditioning the data set is less consistent than in the original Llewellyn-Arai study. Whilst the "centre of gravity" ofthe WEO year ahead forecast is only slightly earlier (October as opposed to December) than that of the national forecasts with which they are compared, it is obvious than on average a difference of two months may seriously disadvantage the WEO, particularly in situations when major disturbances occur. Given the highly significant sensitivity of forecast accuracy to the information set available to the forecaster, this mismatch in source dates implies that comparisons of forecast accuracy on this data set must be treated with care. 2/

A second source of heterogeneity in the data lies in the sector provenance of the national forecasts; some of these are "official" forecasts, whilst others emanate from the private sector. For the United States and the Federal Republic of Germany, more than one series is available: in the case of the United States as in the original study, Stephen McNees kindly provided details of a consensus forecasts. In addition, for the United States, a comparison was also made with the projections published by the Office of Management and Budget (OMB) in the Mid-Session Budget Review, the dating of which corresponds quite closely to that of the WEO year-ahead forecasts since 1980. Prior to

---

1/ Extensions of the series were also not obtained for Italy for 1983-85.

2/ The problem was smaller in the case of the OECD comparison (see Section 3.1) because this used the July issues of OECD's Economic Outlook, where appropriate, to align the source dates more closely than is possible here.

1980, the dating of the WEO corresponds more closely to the projections presented in the CEA's annual report.

Various hypotheses can be formulated about the possible biases in national forecasts. The political exposure to which official forecasts are subject might be thought to predispose them towards "errors of sentiment" - causing output growth to appear too high, or inflation too low; but it may also be argued that the normative element in official forecasts helps accuracy, for governments may be able to implement mid-term policy corrections to check deviations of variables from their target trajectories. Both private and official national forecasts might suffer from solipsistic bias because they are based on less-than-full information about international developments and the feedback between these and their own projections than about purely domestic developments. By contrast, a characteristic of forecasting by international agencies like the IMF or the OECD is that forecasting at this level requires endogenization to the forecasting process of global variables which are (mostly) treated as exogenous at the national level.

Despite the various possible sources of difference, however, the fact that so much information is passed from one forecaster to another, both domestically and internationally, is bound to narrow the scope for substantial disagreement to persist and thus to limit the extent to which forecast accuracy differs, at least over any reasonably lengthy period of time.

Tables 21 and 22 collect the forecast errors (forecast minus realization) for the different agencies covered, respectively, for output and for inflation. In Table 23, composite series for the Group of Seven and Europe, constructed as a weighted average of the national figures, are compared with the corresponding WEO forecasts. In constructing these series the missing 1984 and 1985 figures for Italy were proxied by the WEO figures for those dates, and where multiple national forecasts were available, as in the case of the United States and Germany, the one with the lowest average absolute error was selected.

Considering first the individual country forecasts for output growth, it can be seen that there is a fair measure of sympathy between the WEO and national forecast errors; but there are some exceptions and correlation coefficients in nearly every case are lower than those recorded between the OECD and WEO country forecasts, Japan affording an extreme example (though with a special explanation--see below). In the case of the Federal Republic of Germany, the WEO forecast is strongly correlated with the national forecasts; the larger positive errors of 1975, the smaller ones for the preceding year and the negative errors for 1977 are common to all four forecasts for Germany. In 1982, sizable positive errors are the common experience. In France, Italy, and the UK the patterns of error for the national forecast and the WEO are very similar. The aggregate Group of Seven and Europe output growth forecast errors are well correlated with the corresponding WEO projections.

Table 21. Output Growth Forecast Errors, Year-Ahead Forecasts 1/

(In percentage points)

	United States				Japan		France		Federal Republic of Germany				Italy		United Kingdom		
	OMB	CEA	Consensus	WEO	Official	WEO	Official	WEO	Consensus	Five Wise Men	Official	WEO	ISCO	WEO	NIESR	WEO	B 2/
1973	n.a.	0.8	0.2	0.4	4.3	0.5	0.0	0.0	-0.4	0.1	-0.9	-0.3	-0.9	-1.4	-0.8	-1.0	-0.20
1974	n.a.	3.2	2.9	4.9	2.7	10.8	2.5	1.4	2.6	2.1	0.6	2.6	3.1	2.6	0.4	3.6	+1.00
1975	n.a.	-1.0	1.0	0.3	0.9	1.2	4.2	4.7	6.1	5.6	5.6	5.4	3.7	4.6	3.5	4.0	+1.00
1976	0.2	0.0	-0.1	0.2	-0.1	-0.7	-0.2	-1.8	-1.6	-1.1	-1.1	-2.2	-3.6	-4.2	-1.3	-0.7	-0.83
1977	0.8	0.0	0.1	0.3	0.9	1.2	1.5	0.4	3.1	2.1	2.6	2.2	1.3	-2.2	-1.7	0.6	+0.67
1978	1.3	0.4	-0.1	0.1	1.3	-0.8	1.2	0.1	-0.4	0.1	0.1	0.0	-0.6	0.4	-0.5	-0.5	+0.09
1979	2.0	1.4	-0.8	1.1	0.2	-1.1	0.3	0.0	-0.4	-0.6	-0.4	-0.4	-1.0	-1.0	1.1	1.0	-0.09
1980	1.2	-0.7	-1.1	0.2	-0.2	0.5	1.1	1.0	0.7	1.0	0.7	1.3	-2.5	-1.2	1.9	1.4	+0.50
1981	-1.7	1.0	-0.7	-3.4	2.0	1.1	1.2	1.6	0.3	0.8	-0.2	2.3	-0.2	0.6	1.1	1.8	+0.33
1982	5.1	4.2	2.4	2.6	1.9	2.7	1.2	0.6	2.2	1.7	2.5	3.1	1.3	2.1	0.1	-0.5	+0.83
1983	1.0	-3.0	-1.3	-1.6	-0.3	0.5	1.1	1.8	-0.3	-0.3	-0.8	0.7	2.7	4.5	-0.8	-1.6	0.0
1984	-1.6	-1.1	-1.3	-2.5	-1.0	-1.8	-0.5	-0.3	-0.6	-0.1	-0.1	-0.6	n.a.	-0.2	-0.3	-0.5	-1.00
1985	1.6	1.5	0.7	1.3	0.3	-0.4	0.4	0.4	-0.5	0.5	0.0	-0.1	n.a.	0.2	-0.6	-0.9	+0.27
Average absolute error 1973-79	1.075 5/	0.971	0.743	1.043	1.486	2.329	1.414	1.200	2.086	1.671	1.614	1.871	2.029	2.343	1.329	1.629	
Omitting 74	1.075 5/	(0.600)	(0.383)	(0.400)	(1.283)	(0.917)	(1.233)	(1.167)	(2.000)	(1.600)	(1.783)	(1.750)	(1.850)	(2.300)	(1.483)	(1.300)	
1980-85	2.033	1.917	1.250	1.933	0.950	1.167	0.917	0.950	0.767	0.733	0.717	1.350	1.675 3/	2.100 3/	0.800	1.117	
Whole Period	1.650 5/	1.408	0.977	1.454	1.239	1.792	1.185	1.085	1.477	1.239	1.200	1.631	1.900 4/	2.255 4/	1.085	1.392	
Correlation Coefficient, WEO	0.89	0.68	0.83		0.50		0.88		0.91	0.92	0.86		0.86		0.75		

Source: Llewellyn and Arai (1984), updated and extended (Appendix D).

1/ Forecast errors are defined as forecasts minus realization values.
2/ Computed as the ratio of the difference of positive and negative errors to their combined sum and thus bounded between -1 and +1.
3/ 1980-83.
4/ 1973-85.
5/ Excluding 1973-75.

Table 22. Inflation Forecast Errors, Year-Ahead Forecasts 1/

(In percentage points)

	United States				Japan		France		Federal Republic of Germany				Italy		United Kingdom		
	OMB	CEA	Consensus	WEO	Official	WEO	Official	WEO	Consensus	Five Wise Men	Official	WEO	ISCO	WEO	NIESR	WEO	B 2/
1973	n.a.	-2.3	-2.7	-2.3	-9.6	-6.3	-2.1	-2.3	-0.6	-0.1	-0.6	-0.4	-3.0	-2.3	-2.9	-0.5	-1.00
1974	n.a.	-3.2	-3.5	-4.7	-8.1	-8.6	-4.1	-4.4	0.5	1.0	0.3	0.2	-10.2	-7.6	-2.8	-4.8	-0.67
1975	n.a.	2.3	-0.2	0.5	5.1	9.0	-3.0	-2.1	-1.3	-2.3	-1.8	-1.1	1.6	-1.0	-5.6	-7.8	-0.33
1976	2.0	0.9	0.7	0.6	0.0	-0.1	-1.9	1.1	1.4	0.9	0.9	1.2	-7.5	-2.8	-1.1	0.0	0.0
1977	0.1	-0.4	-0.3	-0.5	1.6	1.3	-0.4	-0.5	0.4	0.4	-0.1	0.1	-1.0	2.7	-1.0	-0.6	-0.33
1978	-1.1	-2.3	-1.4	-1.3	0.9	-0.8	-2.0	-1.2	0.1	-0.4	-0.4	0.1	-1.2	-0.8	-0.3	0.4	-0.50
1979	-2.2	-1.6	-1.1	-1.1	1.7	0.2	-1.5	-1.5	-0.3	-0.8	-0.3	-0.2	-3.4	-1.8	-3.2	-4.2	-0.67
1980	-0.1	-1.0	-0.2	0.1	1.8	4.4	-2.4	-1.8	-0.5	-0.5	-1.0	0.1	-5.8	-6.2	-0.2	-4.6	-0.33
1981	0.8	1.6	0.1	-1.3	1.5	1.6	-0.9	-3.1	0.4	-0.1	0.4	0.2	-3.5	-3.8	-0.4	1.2	0.00
1982	2.1	2.6	1.9	1.2	1.4	1.5	1.1	0.5	-0.3	-0.8	-0.8	-0.3	0.3	2.0	0.0	1.0	0.64
1983	2.3	1.5	1.5	1.8	1.5	2.3	-0.9	2.1	0.3	0.3	0.3	1.3	-0.1	2.0	1.7	2.2	0.67
1984	1.0	1.4	0.7	0.4	0.2	1.6	-0.5	0.3	0.6	1.1	1.1	1.1	n.a.	4.1	0.5	1.8	0.83
1985	1.3	1.1	1.0	1.0	-0.1	-0.5	-0.3	-0.4	0.3	-0.2	-0.2	0.4	n.a.	2.3	0.2	-1.1	0.0
Average absolute error (1973-79)	1.350 5/	1.857	1.414	1.571	3.857	3.757	2.143	1.871	0.657	0.843	0.629	0.471	3.986	2.714	2.414	2.614	
(Omitting 74)	1.350 5/	1.633	1.067	1.050	3.150	2.950	1.817	1.450	0.683	0.817	0.683	0.517	2.950	1.900	2.350	2.250	
1980-85	1.267	1.533	0.900	0.967	1.083	1.983	1.017	1.367	0.400	0.500	0.633	0.567	2.425 3/	3.500 3/	0.500	1.983	
Whole Period	1.300 5/	1.708	1.177	1.292	2.577	2.938	1.623	1.638	0.538	0.685	0.631	0.515	3.418 4/	3.000 4/	1.531	2.323	
Correlation Coefficient, WEO	0.809	0.80	0.94		0.92		0.64		0.86	0.80	0.86		0.83		0.80		

Source: Llewellyn and Arai (1984), updated and extended (Appendix D).

1/ Forecasts errors are defined as forecasts minus realization values for the GNP/GDP deflator (except for the United Kingdom, and NIESR where the CPI was used).
2/ Computed as the ratio of the difference of positive and negative errors to their combined sum and thus bounded between -1 and +1.
3/ 1980-83.
4/ 1973-83.
5/ Excluding 1973-75.

Table 23. Output Growth and Inflation Forecast Errors 1/

(In percentage points)

	Output				Inflation			
	National Forecasts		WEO		National Forecasts		WEO	
	Group of Seven	Europe	Group of Seven	Europe	Group of Seven	Europe	Group of Seven	Europe
1973	0.58	-0.32	0.09	-0.50	-3.38	-2.04	-2.49	-1.30
1974	2.39	1.23	4.98	2.30	-4.18	-3.22	-4.86	-3.10
1975	2.20	4.43	2.01	4.80	-0.06	-1.90	0.77	-2.00
1976	-0.09	-0.22	-0.80	-1.70	-0.31	-1.43	0.36	0.40
1977	0.64	1.15	0.58	1.00	-0.07	-0.48	0.02	0.00
1978	0.31	0.39	-0.10	0.30	-0.84	-0.55	-0.89	-0.30
1979	-0.22	0.08	0.23	0.00	-0.73	-1.52	-0.99	-1.70
1980	-0.01	1.03	0.49	1.00	-0.51	-2.00	-0.03	-2.20
1981	0.27	0.47	-0.57	1.50	0.01	-0.88	-0.69	-1.00
1982	1.95	1.27	2.19	1.60	1.20	0.07	1.06	0.10
1983	-0.62	0.14	-0.37	1.00	1.10	0.34	1.90	1.70
1984	-0.93	-0.33	-1.74	-0.40	0.51	0.29	-.98	1.10
1985	0.44	-0.08	0.61	0.00	0.50	-0.27	0.51	0.20
Average absolute error	0.82	0.86	1.14	1.24	1.03	1.15	1.20	1.16
Correlation Coefficient, WEO	0.89	0.92	--	--	0.95	0.90		

Note: The Group of Seven excludes Canada (WEO errors are assumed for national forecasts for Italy in 1984 and 1985). The aggregate for Europe is based on data for France, Italy, the Federal Republic of Germany, the United Kingdom, the Netherlands, Austria, Sweden, Switzerland, and Finland.

1/ Forecast errors are defined as forecasts minus realization values.

In the light of the mismatch of source dates discussed earlier, it may be helpful in making comparisons of forecast accuracy to divide the sample into two sub-periods and also to take into account the special features distinguishing the 1974 forecast exercise. The WEO forecasts for the years in the first half of the sample (1973-79) were prepared later in the year than those in the second sub-sample (1980-85), as indicated in Table 2. Accordingly, we might expect that the WEO forecast would reflect a relative information advantage in the first sub-sample and a corresponding disadvantage in the second sub-sample.

The 1974 forecast was prepared in December 1973 in conditions of extreme uncertainty about the supply situation in the oil market. The forecast was in fact made on a "business as usual" assumption, this phrase meaning that no physical supply shortages were assumed. As it happened, the prospect of prolonged and severe constraints of this kind was dispelled when, early in January, decisions were taken by the Organization of Arab Petroleum Exporting Countries (OAPEC) to relax the supply curbs that had been threatened but to substitute further price rises. This news prompted WEO forecasters in a January supplement to the December forecast to hazard that "All in all, it seems reasonable to conclude that current prospects for growth...are...somewhat intermediate between (a) the staff's "business as usual" projections...and (b) the much bleaker outlook that followed the imposition of OAPEC cutbacks in oil production, until their recent easing." This clearly shows that the Fund's 1974 projections had a consciously artificial element, which knowingly would produce an "optimistic" bias in output forecasts, even though the "business as usual" assumption proved to be less wide of the mark than might have seemed likely at some earlier stages. Of course, national forecasters faced similar problems in coping with the impact of the oil price increase; and some would have had an information advantage (most likely, Japan 1/), though others would not. There is a case, therefore, for omitting 1974 from the comparisons of forecast accuracy. When this is done, as can be seen from Table 21, the WEO output forecasts generally do appear more accurate (with lower absolute errors) than those of the national agencies in the first sub-sample and (with the single exception of the OMB forecast for the United States) less accurate in the second.

In respect of inflation, the WEO and national forecasts are generally somewhat more highly correlated, those for France being the prominent exception; here, the weakness in the relationship between the two is especially pronounced for the 1980s (the correlation coefficient between the two forecasts for 1980-85 is only 0.40). In the United States the sign pattern of the errors in the three forecasts is highly consistent, 1982 proving--as for real output--a prominent weakness in the U.S. Mid-Session Budget Review Forecast. In Japan, also, the patterns of the forecasts are very similar overall, 1978 being a

---

1/ Japan's forecasts are on a fiscal year (April-March basis), see Appendix E.

prominent exception involving a difference of sign, though there are other years where the difference is bigger, 1975 carrying a particularly large WEO error. In Italy and the United Kingdom, a generally high degree of correspondence also includes some large error differences (1974 and 1976 in Italy, 1975 in the United Kingdom being the notable examples). By the average absolute error criterion, the WEO inflation forecasts are superior for Italy and the Federal Republic of Germany, but the margin of difference is again generally small (the principal exception being a difference of 0.8 of 1 percent in the national forecaster's favor in the case of the United Kingdom). In the aggregate series, the inflation forecast errors for the composite Group of Seven and Europe turn out to be very similar indeed. Indeed, having regard to the various qualifications concerning source-date mismatch and the status of the 1974 forecasts and to the fact that for multiple-forecast countries the forecast yielding the lowest average (output) error was used in the national forecast compilation, it could not be said that there is any significant difference between the track records of the WEO and the aggregated national forecasts of output and inflation in the Group of Seven and Europe as a whole.

Since there are two forecast variables under consideration here (a comparison with national forecasters regarding the balance of payments forecasts was not possible due to lack of comparable data) it is possible to compute on how many occasions one forecast dominates the other by counting the number of times that one forecast is better, for the pair of variables together, than the other. In such a case, whatever (positive) weighting is given to the different variable forecast misses, the forecast with the lowest pair of errors will dominate the other. In the present case there are 73 instances to consider (6 countries x 13 years minus the two years for which independent Italian forecasts are not available, minus one case in which the errors are identical). Of these 73 possible instances, 49 are decisive, 19 showing WEO as the dominant forecaster, and 30 the national forecaster. 1/ The WEO forecast was especially strong in 1978 when for five out of six major countries the WEO forecast was superior for both output and inflation. On the other hand, in both 1983 and 1985 the WEO was inferior for both output and inflation to five of the national forecasts. 2/

Turning finally to the question of the relationship across countries between the two forecast errors, the "B" statistic shown in Tables 21 and 22 shows that there are, as Llewellyn and Arai found, a

---

1/ For purposes of the count, the multiple forecasts for the United States and for the Federal Republic of Germany were reduced to one in each case by choosing that forecast with the overall lowest average absolute error in forecasting output.

2/ This apparent deterioration is consistent with the shift in sourcing dates for the WEO forecasts.

number of years when all or a majority of forecast errors are of a similar sign. 1/ In the case of output growth, 1974 and 1975 stand out as years of universally or widely shared positive error (over-optimism), as also 1982; whilst 1976 and 1984 are instances of widely shared or universal negative error. In the case of inflation, the bunching of errors occurs in 1973 and 1974 (negative errors), and again in 1982, 1983 and 1984 (positive errors).

Some part of the explanation of these episodes is explored below in Section IV. The national forecast data perform the useful function of identifying periods when an error is (or not) widely spread across countries and forecasting agencies, confirming that the explanation for a particular WEO error is (or not) specific to the World Economic Outlook. Since there is, generally, a high correlation between the episodes identified by the B statistics and major errors in WEO Group of Seven forecasts (and a correlation across countries between national and WEO forecast errors) there is confirmation that, at any rate for the most part, WEO errors tend to be shared in some fairly large degree by other agencies; they appear thus to be general products of the imprecise art of economic forecasting rather errors purely specific to the World Economic Outlook.

IV. Explanations

This section turns to explanations of the forecasting error. The approach taken is two-fold. On the one hand, an attempt is made to identify from inspection of the error patterns the most significant episodes. These are then examined in more detail, using the WEO source documents. This approach captures the uniqueness of each error-episode. The complementary approach pursued in the second part of this section is to attempt to explain systematically the forecast error, using an "innovation-accounting" approach where the error is attributed to deviations from the WEO projection of conditional assumptions about fiscal policy and oil prices.

1. Narrative Account

A convenient graphical summary of error patterns appears in Charts 2 and 3, which show, in the top panel the current year forecasts, and in the lower panels, the year-ahead forecast, together with the corresponding realization and error. Chart 2 shows output and Chart 3 inflation for the Group of Seven countries. Table 24 reports these same data, together with B statistics for the WEO country errors and the latest available estimates of realization values. A comparison of these data

---

1/ In calculating the statistic, a single national forecast was used for each country, following the same selection principle for the two cases of multiple national forecasts as in the computation of the synthetic Group of Seven and Europe aggregates.

Table 24. Group of Seven: Output and Inflation Errors in the WEO

	Output							Inflation						
	Forecasts		Outturns		Errors		Statistic 1/	Forecasts		Outturns		Errors		Statistic 1/
	Y/A	C/Y	F/A	F/S	E1	E2	B2	Y/A	C/Y	F/A	F/S	E1	E2	B2
1971	4.5	4.0	3.2	3.2	1.3	0.8	n.a.	4.2	4.9	5.5	5.5	-1.3	-0.6	-0.86
1972	n.a.	5.3	5.8	n.r.	n.a.	-0.5	n.a.	n.a.	4.3	4.2	n.r.	n.a.	0.1	n.a.
1973	6.6	7.5	6.7	6.6	0.0	0.8	-0.14	4.6	5.4	7.1	7.1	-2.5	-1.7	-1.00
1974	4.3	1.3	-0.5	-0.5	4.8	1.8	+1.00	7.3	10.9	11.9	12.2	-4.9	-1.0	-0.86
1975	0.7	-1.5	-1.6	-1.3	2.0	0.1	+1.00	11.7	10.8	10.5	10.9	0.8	0.3	-0.14
1976	4.8	6.1	5.6	5.6	-0.8	0.5	-0.71	7.6	7.5	7.1	7.3	0.3	0.4	0.0
1977	4.7	4.5	3.9	4.1	0.6	0.6	+0.71	7.0	7.0	7.1	7.0	0.0	-0.1	+0.14
1978	4.2	4.2	4.0	4.2	0.0	0.2	+0.33	6.3	6.1	7.2	7.1	-0.8	-1.1	-0.14
1979	3.8	3.5	3.5	3.5	0.3	0.0	0.00	6.7	8.0	7.6	7.8	-1.1	0.4	-0.86
1980	1.8	0.8	1.2	1.2	0.6	-0.4	+0.71	9.0	9.6	9.0	9.1	-0.1	0.6	-0.14
1981	0.6	1.6	1.4	1.2	-0.6	0.2	+0.43	8.1	8.9	8.3	8.9	-0.8	-.6	-0.14
1982	2.0	0.7	-0.4	-0.4	2.4	1.1	+0.71	7.7	7.5	6.7	6.7	1.0	0.8	-0.43
1983	2.4	1.8	2.5	2.8	-0.4	-0.7	+0.14	6.3	5.1	4.7	4.4	1.9	0.4	+1.00
1984	3.5	3.9	5.1	5.2	-1.7	-1.2	-1.00	4.7	4.2	3.6	3.6	1.1	0.6	+1.0
1985	3.5	3.2	2.7	3.0	0.5	0.5	-0.14	4.1	3.7	3.6	3.6	0.5	0.1	+0.1
1986	3.2	3.0	2.4	n.a.	n.a.	0.6	n.a.	3.5	3.1	3.1	n.a.	0.1	n.a.	n.a.

Note: Y/A = Year Ahead; C/Y = Current Year; F/A = First Available Estimate; F/S = First Settled Estimate; E1, E2 = Y/A-F/S, C/Y-F/A.

1/ Computed as (positive errors - negative errors)/(positive errors + negative errors) across Group of Seven countries (E1 errors).

CHART 2

WEO FORECAST ERRORS: GROUP OF SEVEN OUTPUT -- CURRENT YEAR AND YEAR AHEAD

(Percent)

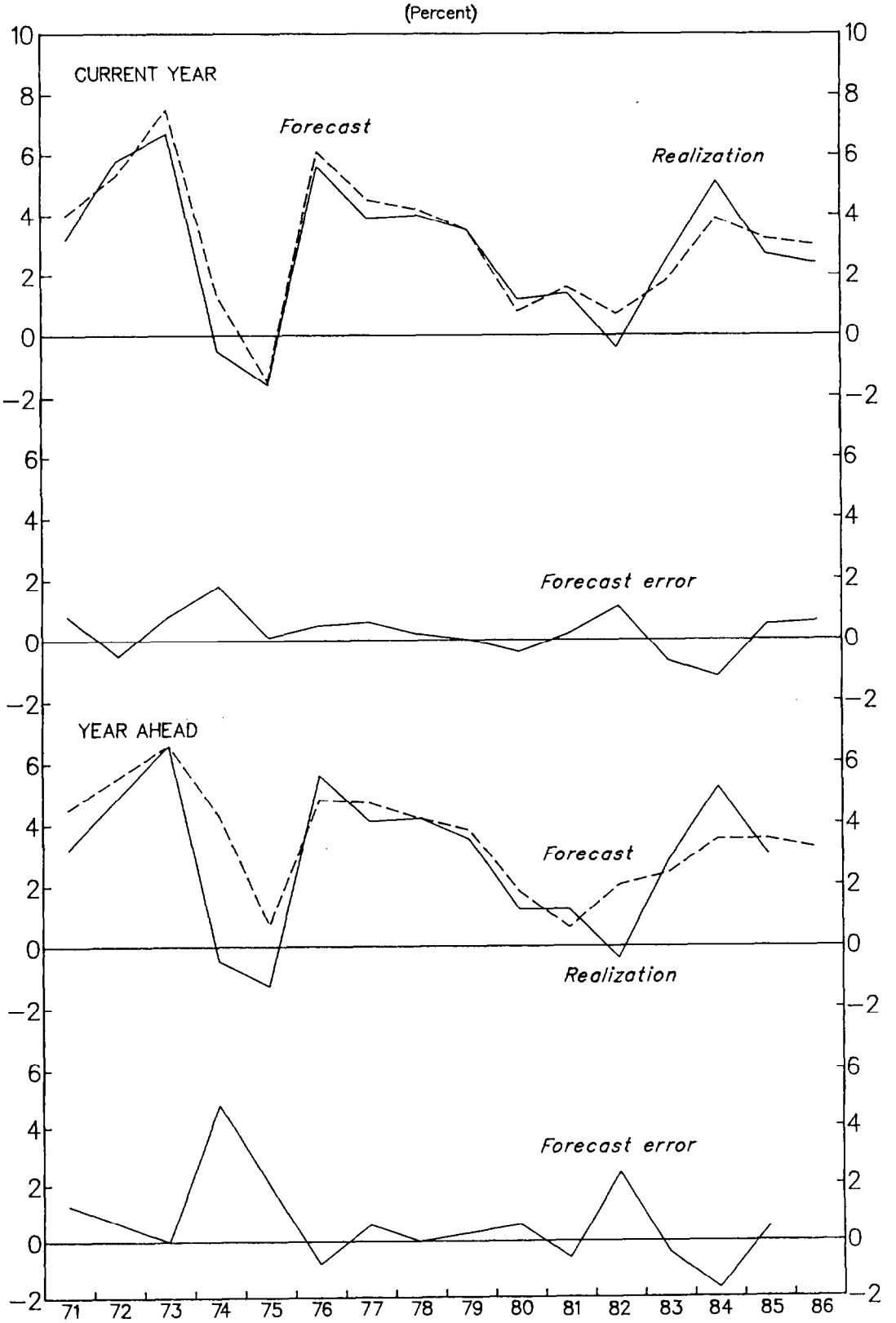
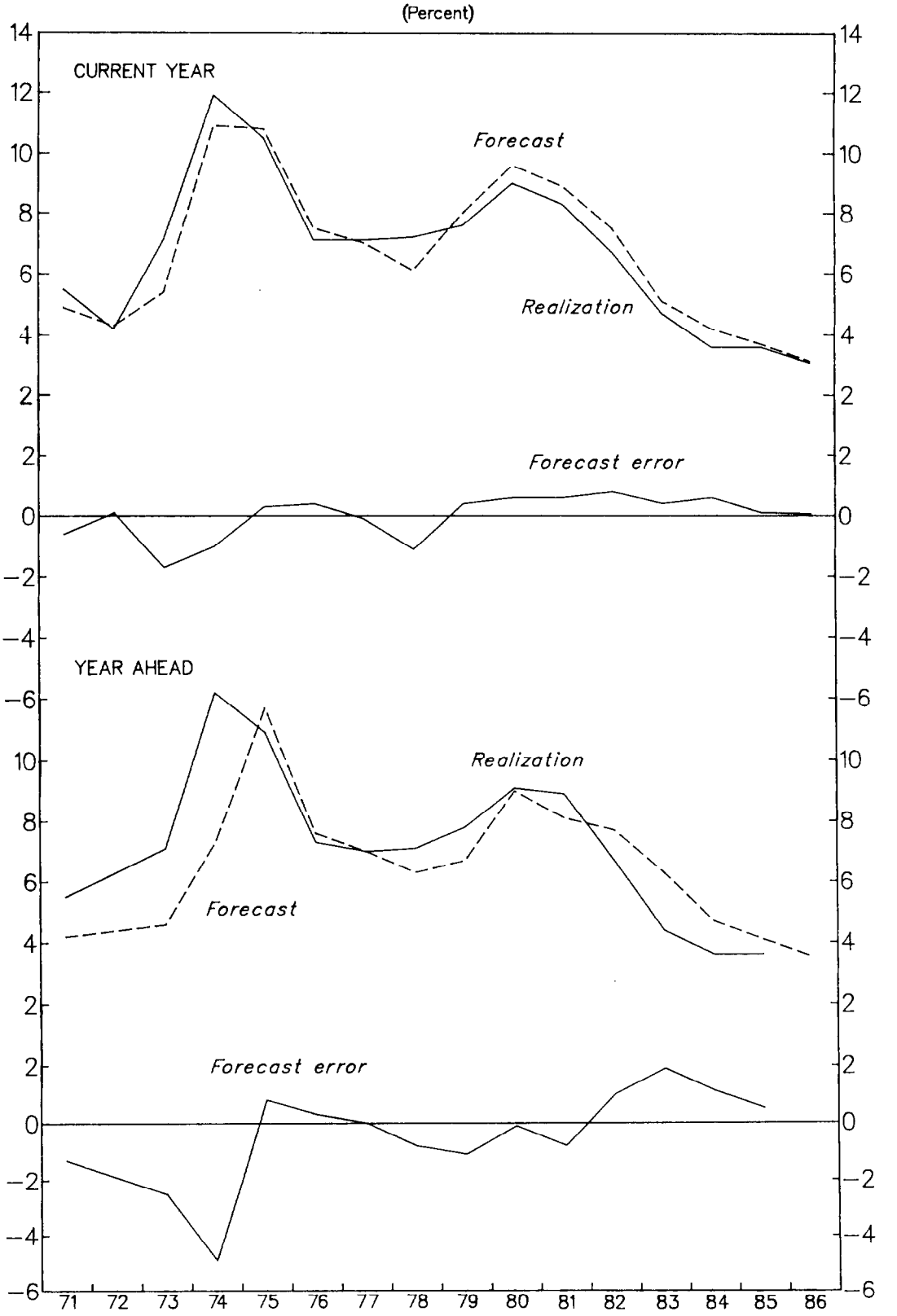


CHART 3

WEO FORECAST ERRORS: GROUP OF SEVEN INFLATION -- CURRENT YEAR AND YEAR AHEAD

(Percent)



with comparable data for the individual countries indicates, as might be expected, that the Group of Seven series are smoother and average the sometimes discrepant individual country movements. The B statistic reported is an aide-memoire in this respect (as noted earlier, by construction the statistic is bounded by -1 and +1, giving a zero reading where the country errors are offsetting in sign and a value near ± unity when they are of similar sign). 1/

a. Turning-point error. Forecasters are often accused of being able to get things right only when the outlook is unproblematically the same as the past; when things change they fail to spot the difference. Clearly there is some truth in such a charge in the present case. Chart 2 shows how the sharp downturn in output growth between 1973 and 1975 gave rise to positive errors of forecast in 1974, in particular, even on a current year basis. Similarly, the pickup in growth in 1983 and 1984 gave rise to negative errors, particularly in the latter year. On the other hand, the turning-point error occasioned by the very sharp turnaround from the 1976 recession is quite small, even on a year ahead basis while the prominent positive output error of 1982 comes at a time when output growth simply went on declining.

Turning to the inflation record, the negative error of 1974 coincided with a take-off of inflation, but its subsequent decline through 1976 and 1977 is well tracked. The renewed inflationary surge starting in 1978 before the second round of oil price increases again produces a negative error, however, whilst the subsequent disinflation seems to have been seriously unanticipated in the year ahead forecasts--though the correction in the current year forecasts is rather satisfactory.

The period as a whole can be broken up into perhaps four homogeneous sub-periods: The first round of oil price increases and its aftermath (1973 to 1975, say); the subsequent recovery (to 1978-79); the second round of oil price increases and its aftermath (1978-80 to 1982, say) followed by the "dollar shock" and the period up to the present.

Before exploring these episodes, the balance of payments forecast errors should also be recalled. By nature the B-statistics for these errors, unlike those for output and inflation, will tend towards zero, as a positive error for one member of the group will need to be accommodated by a negative error for another, except to the extent to which surpluses or deficits against the rest of the world occur, or when there are variations in the global current account discrepancy. Table 25 below thus records both the net balance of payments error for the year ahead for the Group of Seven aggregate, and the absolute error obtained by summing across the individual countries without regard to

---

1/ Because of the odd number of countries making up the Group of Group the statistic cannot be zero, except where there is a zero error for an odd number of countries.

Table 25. Group of Seven: Balance of Payments Forecast Error, Year-Ahead Forecasts

(In billions of dollars)

	Net Error 1/	Absolute Aggregate Error 2/
1973	1.6	27.8
1974	16.4	33.8
1975	-27.3	40.1
1976	4.6	19.6
1977	-7.2	26.2
1978	-25.8	31.0
1979	29.6	43.4
1980	30.6	58.8
1981	-14.0	56.2
1982	-7.5	30.7
1983	16.5	58.9
1984	29.2	53.0
1985	-20.6	28.0

1/ Absolute errors summed accross countries.
2/ Absolute errors of the Group of Seven aggregate.

sign. The latter gives an impression of the extent to which the balance of payments forecasts across countries were out of kilter. In fact the record suggests that the worst balance of payments forecast errors occurred in 1975, in 1979-80 and in 1983 and 1984 when the net error is also relatively large.

b. The first round of oil price increases and its aftermath. The first wave of oil price increases in 1973-74 contributed to raising prices steeply to double digit levels; in the subsequent adjustment the Group of Seven countries' output growth declined sharply, culminating in actual declines in 1974 and 1975. The Fund forecasters were not able to take on board the full extent of the oil price rise even in their late 1973 year ahead projections. The anticipated rise in export unit values of oil exporters for 1974 as a result very dramatically underestimated the increase to come; by May 1974 the estimated rise had been increased by a factor of four, yet still significantly underestimated the actual increase (see Appendix Table F.3).

These substantial discrepancies between forecasts and outcomes contributed powerfully, if uncontroversially, to the negative inflation forecast errors in 1973 and 1974. In no subsequent year was there an inflation error even half as large as the year ahead error for 1974. It is not surprising, either, that the same error was associated with an unforeseen recession in 1974. As discussed earlier, these forecast errors for 1974 reflected in part the uncertainties regarding the supply situation. The failure to predict the full extent of the world-wide decline in output in 1975, however, is more problematic. Oil prices rose only moderately in this year and in any case the rise was largely anticipated, whilst fiscal policies were switched to an expansionary mode in every one of the major countries except the UK (see Table F2), on a scale which was essentially unforeseen. The forecast error seems thus to have derived largely from a failure to appreciate the full extent to which the oil price rise and the reactive adjustment policies would reduce output; this mistake was widely shared by national forecasters (as indicated in the Fund's 1974 Annual Report) and the fiscal expansion was itself a response to the belated realization that the oil price increase had left real activity in the major economies in a weak condition. This was well understood by the time that the March 1975 WEO came to be issued. 1/

c. The 1976-78 recovery. The recovery from the 1973-75 downturn, which began in 1976, was foreseen as early as the May 1975 WEO, though it turned out to be understated, perhaps partly because the expansionary stance of policy was underestimated. Output anticipations in the remainder of the recovery period continued to be closely

---

1/ The appreciation of events contained in this WEO included also the point that the exchange rate assumptions used in the previous December's projections had already been shown to be substantially out of line with developments.

fulfilled and there were few aggregate disturbances until the oil price rise in 1979. Nevertheless there was quite a sizable inflation forecast error in 1978; even in the forecast made as late as April of that year, inflation was being predicted at too low a rate. In good part this seems to have been the outcome of a large error for a specific country, the United States; while the WEO forecasters correctly anticipated that inflation in Europe would accelerate in 1978, they underestimated the extent of the inflationary surge in the United States.

d. The second round of oil price increases and its aftermath. The second round of oil price increases, like the first, contributed rapidly to an acceleration in inflation and to lower output growth in the industrial world. Learning from the experience of 1974-75, and in the light of already-high inflation, countries engaged uniformly in policies of monetary restraint and, in most cases, began to apply the fiscal brake as well. Although the oil price rise itself was not predicted, its scale was of course much smaller in percentage terms than that of 1973-74 and the greater uniformity of response and the experience of the first wave of oil price increases all helped reduce the extent of consequent forecast error. In fact, aside from the initial effect of the oil price rise on the inflation forecasts for 1979, the subsequent output and inflation forecast errors appear quite small. The difficulties appear in the subsequent period.

e. The 1982 recession. This phase is characterized by quite a large error of optimism about output in 1982, when aggregate output in the Group of Seven GNP declined slightly while the forecasters, even as late as April 1982, looked for a continued rise. One of the reasons was a mini-recovery early in 1982 which was known to forecasters and was thought to presage a continued upswing. In the following two years, errors of the opposite sign were made as the world economy recovered, only to give way to renewed negative errors as the world recovery moderated. The rate of inflation declined steadily over the period persistently, somewhat faster than expected.

The forecast error for output in 1982 which became subsequently apparent in the projections prepared the year before had no obvious prompting in either oil prices or unexpected fiscal developments as the aggregate Group of Seven fiscal impulse was actually somewhat greater than had been anticipated. A continued determination to squeeze out inflation characterized the collective Group of Seven policies at this time, however, and it is pertinent to note, as extensively documented in the April 1982 WEO, that this cooperation (or rather, mutual emulation) in restraint was concentrated particularly in monetary policy. While this fact was evident at the time, the WEO forecasters do not appear to have grasped the true measure of it--or a true measure of its effects. (Nor was the error confined to the WEO--it was shared, as shown in Section III, by forecasters at OECD and by national forecasters around the world).

The suggestion that the underestimating of the strength of contractionary monetary policies and their effects underlay the forecast error can be given some support from analysis of the sectors apparently most seriously affected and the WEO's ex post comments on the "unprecedentedly" high real interest rates which emerged. Table 26 shows two indicators of monetary stringency, real interest rates (short and long) generated as the difference between nominal interest rates and either current inflation or the contemporaneous WEO inflation forecast and the growth rate of the real money supply (nominal money growth less current inflation). These data suggest that, if anything, the leap in monetary stringency was initiated in 1981 (or 1980 in the case of the real money stock) and sustained in 1982, so the weakness of the WEO forecasts may reflect both the lagged effect of the already accomplished increase in stringency and its unexpected continuation through 1982. It should also be noted that forecasters differed significantly in their assessment of the impact of monetary stringency at the time. For example, although the WEO output error for the United States was large it was much smaller than the error in the corresponding Mid-Session Budget Review and CEA forecasts for 1982 (see Section III).

f. The post-1982 recovery. Coming to terms with this mistake, the WEO forecasts for 1983 turned out to be remarkably close to the mark on output, predicting well the recovery that would ensue. Inflation, nevertheless, turned out much less strong than anticipated in the early forecasts. It is not difficult to suggest that this was related to an underestimate of the extent to which the world recession would slow home-bred inflation in the industrial countries and, most important, cause primary product prices to decline. Both in 1982 and again in 1983, against a background of declining inflation in the industrial countries, the terms of trade improved to a degree which had not been foreseen, with particularly large falls in primary product prices in 1982, following a decline in the previous year.

The upswing correctly predicted for 1983 was foreseen by WEO forecasters to continue through 1984, but not on the scale which in fact ensued. The forecasters did not catch up with this, even in their April 1984 projections. Although the sign of the output error was repeated across the set of Group of Seven countries, much the largest error and by far the greater part of the aggregate error was due to the underestimate of output growth in the United States, a mistake shared with other forecasters, including national agencies. This can be associated partly with an underestimate of the degree of fiscal expansion that policies would provide; for the rest it seems as if the rebound of private spending following the sustained reduction in inflation was larger than anticipated.

In 1985 and again in 1986 forecasting error was small. Against a background in which the sharp expansion of 1984 gave way to lower growth rates, WEO forecasters tended to overestimate (although only slightly) the extent of output growth in their year-ahead forecasts; at this time, and especially in outlining the prospects for 1986, the WEO forecasters

Table 26. Group of Seven: Indicators of Monetary Stringency 1/

(In percent)

	Real Interest Rates (Using actual inflation)		Real Interest Rates (Using WEO inflation forecast)		Real Money Supply Growth
	Short term	Long term	Short term	Long term	
1971	-3.02	--	-1.29	--	3.57
1972	-1.83	--	-2.56	--	5.29
1973	4.35	-0.26	4.37	2.16	2.65
1974	-1.57	-2.73	-2.53	-1.97	-5.03
1975	-16.63	-1.95	-12.67	-1.82	-2.41
1976	-3.45	1.00	-4.15	1.23	0.84
1977	-5.78	0.70	-4.89	1.25	1.22
1978	-2.00	1.12	-0.34	2.29	3.71
1979	1.76	1.29	1.92	1.45	1.68
1980	3.39	1.99	3.08	1.67	-3.66
1981	5.54	4.52	5.30	4.28	-2.23
1982	4.86	5.54	4.21	4.90	-0.17
1983	4.62	6.23	4.05	5.66	5.45
1984	5.77	7.15	5.45	6.84	3.07
1985	4.90	6.13	4.75	5.99	5.04
1986	3.80	4.41	3.73	4.34	8.25

1/ Nominal output, current exchange rate weights.

came to lay particular stress on the effect of the decline in oil prices as a factor significantly improving the short-run growth prospects of the industrial countries. In retrospect (see the April 1987 WEO which examines the point in some detail), the extent to which the decline in oil prices would have this effect was over-estimated; albeit the output error in the predictions for 1986 was not large.

This pursuit of episodic detail clearly suggests that the major forecasting error can be associated with the first round of oil price increases. The second wave of oil price increases also created some obvious problems for forecasters. But some large errors have occurred which cannot be explained in this way. The effect of collective policies of restraint, especially on the monetary side as in 1982, appears to have been underestimated, for example, and private sector responses first to the pressure of high inflation, then to its decline have been similarly incompletely understood. Forecaster's "caution," a tendency to miss some turning points, particularly those provoked by novel types of disturbance, or to understate the strength of the turn can also be discerned. These weaknesses should be put in perspective: the errors described appear to have been widely shared by other forecasting agencies, national and international, as analyzed in Section III, and as this evidence suggests, given the state of knowledge, were probably inescapable. In any case, most of the errors seem small and generally were quickly corrected. Just how small is "small" in this context depends of course on the purpose to which the forecasts are put.

2. Effects of Unanticipated Changes in Fiscal Policy and Oil Prices

This sub-section reports a formal attempt to ascertain the extent to which the WEO forecast error can be attributed to a falsification of the environmental assumptions on which the forecasts are conditioned. As noted above, in making a forecast, WEO forecasters make assumptions about external environmental factors, of which the most important appear to be oil prices and fiscal policy. Whilst the forecasts, as already noted, also make conditional projections of other variables like exchange rates and interest rates, it is arguable that these are to a considerable extent driven by fiscal policy and oil prices, together with variables which are endogenous to the forecast in any case. Of course, monetary policy may also have an exogenous component which may affect economic developments independently of variations in fiscal policy and oil prices. However, even though this factor clearly played an important role in the early 1980s, the identification and quantification of exogenous monetary policy changes is extremely hazardous.

With this caveat, the task here is to relate the forecasting error to unexpected changes in fiscal policy and oil prices. Two alternative approaches to the one adopted here should be mentioned. One, adopted by Kenen and Schwartz (1986) uses the forecast-realization framework to relate forecast misses to presumptive policy change; their approach, though, is really devised to test the stronger proposition that where

the forecast itself projects an "unacceptable" scenario (inflation too high, output growth too low, for example), policies will be revised and will change the outcomes to the extent that it looks as though the forecast was wrong. A second approach, appropriate in the case of forecasts produced from a formal model, would be to estimate equations of the type $R(t) = a + bP(t) + cF(t) + v(t)$, where P is the unexpected change in conditioning factors already expressed through the model in terms of the realized forecast variable. The investigator would then look for $a = 0$, $b = 1$, $c = 1$. 1/ This approach is not open to us here, since no estimates of oil price and fiscal policy multipliers that are consistent with World Economic Outlook estimates are available. Consequently, the formal approach adopted is to identify the unexpected changes and then to estimate their contribution to the forecast error through simple regressions of the type:

$$E(t) = a + bI(t) + v(t)$$

where I is a vector of the unexpected changes in conditioning variables. Some related hypotheses can be explored at the same time. For example Llewellyn and Arai (1984) concluded from their study of OECD and national forecasting errors that OECD forecasters were prone to overlook the lagged effects of large policy change; this hypothesis can be explored by relating forecast error to realized policy (and oil price) variables as well as to the innovations in them.

a. Measuring the unexpected changes. Because of the presumption that both policy and oil price changes take time to work through the economy to affect variables of interest, exploring the current year forecasts for the effects of changes in exogenous variables seemed less likely to yield interesting results than an investigation of the year-ahead forecasts; the latter were consequently employed.

The first requirement then is to obtain a measure of the unexpected changes themselves that are relevant to the year ahead horizons. Fortunately, the WEO forecasters have always given an explicit projection for oil prices in the form of a forecast of major oil exporters' export unit values, so that it is a straightforward matter to compute the unexpected oil price change. This series is shown in Table 27. (There is a discussion above of the context in which some of these errors arose, whilst the forecast and realization series themselves are shown in Appendix Table F.1).

The measurement of unexpected fiscal policy changes is somewhat less straightforward and is prone to significant observation error. However, WEO forecasts have in later years generally provided a quantified forecast of the "fiscal impulse"; the series can only be extended backwards by "translating" the qualitative comment then available to comparable terms. At the same time, estimates of the

---

1/ See, for example, Savage (1983).

realized fiscal stimulus have varied and the latest available consistent set of data does not appear to extend back beyond 1977, though at the risk of generating a spurious break these can again be extended backwards by using some earlier WEO estimates of the fiscal impulse in these years. More of the forecast figures are available for the narrowly defined central government fiscal impulse than for the broader definition of general government fiscal impulse, though the qualitative comment might sometimes apply equally to either and it is not obvious what assumption to make about the forecasters' implicit treatment of the "non-central" part of the fiscal impulse when comment is only about the central government part. Data on what the central government fiscal impulse actually was are not available before 1977. However, since the fiscal impulse is (with sign reversed) closely related to the first difference of the structural budget balance, an alternative set of realizations can be obtained from OECD estimates of this concept and in this way, two alternative series on unexpected fiscal policy changes were generated, as shown in Table 27. (A fuller account of the data base generation for these series is given in Appendix F.)

b. The Results. Given the measurement of the unexpected changes, the next step was to formulate the regressions. For this purpose the year ahead forecast errors for output growth, inflation and the balance of payments were related, for each country, to the oil price change, and, separately, to the "own" fiscal policy change and to the "foreign" fiscal policy change. In order to generate suitable estimates of the latter, unexpected changes in individual countries' fiscal policy were aggregated, using GDP or import weights. In addition, since the series of unexpected oil price changes is dominated by the two rounds of oil price increases, an alternative specification using separate dummy variables for the two events was also used. There was considerable agreement between the various specifications: overwhelmingly, oil price changes (especially those of 1973-74) contributed to inflation forecast errors (those for the Federal Republic of Germany being the only exception); while unexpected oil price changes also accounted for a number of the output and balance of payments errors. By contrast, fiscal policy changes proved relatively unimportant in explaining forecast error and, where they were significant, did not always carry the theoretically expected sign.

Table 28, by way of example, tabulates the results arrived at when treating the oil price changes as a continuous variable. Appendix F tabulates additionally the results of treating the innovations as a discontinuous series (dummy variables). While the significance of the oil price effect is expected, the relative unimportance of unexpected fiscal policy changes (and the prevalence of "wrong signs") is somewhat disappointing. In addition to problems with the quality of the fiscal policy data series per se, some other explanations also suggest themselves. To begin with, impulse measures are best regarded as measures of "policy stance"; they are imprecise about timing and weight expenditures and taxation revenues equally, so for these reasons the measure may not be a good one. It is also true that some fiscal actions

Table 27. Unexpected Changes in Driving Variables, 1973-85

	Fiscal Policy							
	Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom	Oil Prices
	(In percent of GNP)							(In percent)
FP1 1/ 1973	0.4	0.5	-1.2	0.6	0.3	-1.5	-2.5	
FP2 2/	0.4	-0.2	-0.3	0.3	1.1	0.4	-3.1	-34.8
1974	0.6	0.8	0.3	0.1	-0.4	1.1	0.2	
	0.9	0.9	0.2	0.3	-1.6	0.5	-0.1	-192.8
1975	-2.25	-3.2	-2.2	-2.6	-1.2	-1.5	-2.5	
	-3.25	-1.6	-2.4	-1.1	-2.4	-2.0	0.8	1.2
1976	0.4	0.6	0.3	1.6	-0.2	0.3	1.6	
	0.4	0.8	-1.0	0.5	0.7	0.7	-1.2	-1.0
1977	-2.0	0.7	0.1	-1.2	0.4	1.5	0.3	
	-0.4	0.9	-0.8	-1.9	1.2	1.8	0.4	-0.9
1978	-0.9	0.2	0.3	-0.9	0.4	-3.5	-1.5	
	-0.2	0.6	-2.3	-1.5	0.1	-1.6	-1.6	-0.4
1979	-0.1	0.4	0.0	-0.9	0.2	0.9	0.1	
	0.8	-0.1	1.7	0.1	-0.3	-1.4	-0.1	-35.8
1980	-0.2	-0.7	-0.3	0.5	0.3	-0.4	1.4	
	-0.8	-0.8	-0.1	1.1	-0.5	0.8	0.9	-43.6
1981	0.8	-0.2	0.2	-1.3	0.2	-0.9	1.2	
	0.6	-0.7	0.4	-1.3	-0.2	-3.7	2.7	0.7
1982	-1.9	-0.4	-0.2	1.0	0.6	-1.1	0.2	
	-0.9	-0.9	-0.2	0.4	0.7	0.1	1.0	8.8
1983	-1.2	-0.6	0.2	-0.1	-0.4	0.3	-0.7	
	-2.2	0.3	0.3	-0.1	0.9	2.5	-1.3	13.9
1984	-1.5	-1.0	-0.1	-0.3	-0.8	1.0	-0.6	
	-1.2	-1.1	0.5	0.0	0.0	-1.7	-1.1	1.0
1985	-1.2	-0.3	0.1	0.1	0.2	-1.0	-0.1	
	-1.3	-0.8	0.0	0.4	0.3	-1.4	0.1	4.3

1/ Figures of forecast fiscal impulse less realized central government fiscal impulse.

2/ Figures of forecast fiscal impulse less the (negative) of the change in the structural budget balance.

Table 28. Explaining Forecast Error: Unexpected Changes in Driving Variables
(Year Ahead Forecasts 1973-85)

	Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom
Output growth							
Constant	0.441	-0.354	-0.119	0.150	1.118	-0.263	-0.080
	(0.52)	(0.61)	(0.20)	(0.34)	(1.46)	(0.32)	(0.26)
Own fiscal policy	-0.123	-0.395	-0.486	-0.702	-1.165	0.043	0.600*
	(0.15)	(0.62)	(0.97)	(1.63)	(1.47)	(0.10)	(3.18)
External fiscal policy	-0.752	-0.137	0.425	-1.576*	-0.243	2.453	-1.325*
	(0.44)	(0.09)	(0.39)	(2.36)	(0.16)	(1.70)	(2.74)
Oil price	-0.012	-0.028*	-0.050*	-0.01	0.004	-0.012	-0.023*
	(0.83)	(2.81)	(4.90)	(1.29)	(0.24)	(0.84)	(4.38)
$\bar{R}^2$	-0.202	0.293	0.661	0.277	0.136	0.013	0.723
Inflation							
Constant	-0.718	0.157	0.836	-0.147	0.363	-0.060	-0.066
	(1.79)	(0.49)	(0.77)	(0.34)	(1.34)	(0.07)	(0.09)
Own fiscal policy	-0.812**	-0.577	-0.292	0.749	0.141	0.414	-0.326
	(2.11)	(1.65)	(0.33)	(1.78)	(0.50)	(0.84)	(0.70)
External fiscal policy	0.376	0.958	-3.145	0.693	0.557	-1.066	2.950*
	(0.47)	(1.15)	(1.63)	(1.06)	(1.04)	(0.68)	(2.47)
Oil price	0.035*	0.024*	0.042*	0.028*	0.003	0.045*	0.037*
	(5.20)	(4.55)	(2.31)	(3.87)	(0.48)	(2.91)	(2.87)
$\bar{R}^2$	0.821	0.692	0.426	0.507	0.073	0.395	0.434
Balance of payments on current account							
Constant	-0.407	9.313	-3.241	1.257	2.126	1.047	-3.602*
	(0.54)	(1.65)	(1.02)	(1.00)	(0.69)	(0.56)	(1.20)
Own fiscal policy	0.630	3.750	2.197	3.249*	-5.034	-0.292	-2.141*
	(0.86)	(0.60)	(0.84)	(2.64)	(1.58)	(0.29)	(2.93)
External fiscal policy	-1.500	14.473	-5.054	2.445	6.933	3.342	-1.373
	(0.98)	(1.05)	(0.90)	(1.28)	(1.13)	(1.03)	(0.73)
Oil price	-0.005	0.074	-0.075	-0.011	0.088	-0.029	-0.047*
	(0.43)	(0.77)	(1.39)	(0.51)	(1.40)	(0.92)	(2.32)
$\bar{R}^2$	-0.134	-0.092	0.021	0.400	-0.030	-0.031	0.475

Figures in parentheses are t-ratios.
* Significant at the 5 percent level.
** Significant at the 10 percent level.

have long lags and these may be longer than the forecast horizon. Finally, as we know that fiscal policy responds to both the realized and (conditionally) forecast path of economic variables, it may also respond to the WEO error if that error is in some sense a measure of the disappointment of desirable targets experienced as the forecast period unfolds--thus for example an ex-post over-optimistic forecast for output growth may become widely anticipated during the year and result in expansionary revisions of fiscal policy. 1/ This would lead to an observed association of positive forecast error with a negative impact from (own) fiscal impulse changes, the "wrong" sign. Clearly this two-way causation issue is not important for oil prices and neither can it be invoked to explain the "wrong" sign on the external fiscal policy variable. There is always, of course, the possibility that fiscal crowding out is more common than usually supposed. 2/

c. Actual policy effects. Even when unexpected policy and oil price changes have been taken fully on board, are their effects fully comprehended? Llewellyn and Arai indicated that they may not have been, and it would be consistent with the suggestion above that coping with a change of a novel and unexpected kind or magnitude should cause forecast error both before the change is fully known and immediately afterwards, before its implications have been taken on board. In the first place, errors arise from unanticipated changes in the multiplicand; then as a result of error in the multiplier. The sign of the latter error is in principle ambiguous.

To test this proposition formally we extended regression equations of the type reported in Table 28 by adding actual oil price changes and actual fiscal policy variables, again using the alternative measures of fiscal policy and alternative weighting schemes. With the exception that lagged actual oil price change often proved significant among the additional variables, none of the variants afforded results offering substantially reconditioned conclusions. Thus the results for oil price changes reported in Table 28 remained substantially unchanged; actual fiscal policy (lagged) turned out to be significant only in the balance of payments results for the U.S., with a (positive) sign, suggesting that WEO forecasters may have overestimated the fiscal impact on the current account. Lagged oil price realizations appeared significant in a number of cases, apparently contributing to positive output forecast errors in Germany and France, a positive inflation error in Japan (but a negative one in the United Kingdom) and a positive balance of payments

---

1/ We know that there is not a great deal of difference between WEO and national country forecasts and it may make a good deal of sense to treat the latter (at least where they emanate from a government source) as a picture (given the constraints on doing better) of a desired evolution of the economy--otherwise actual and promised policy would be changed to make it so.

2/ The estimating equations do not provide for any conditioning on monetary policy to help cope with this point.

forecast error in the Federal Republic of Germany. On this basis it would appear possible to overdo the extent to which large unexpected change yields a residue of "model" error in the immediately succeeding period.

V. Conclusions

The scope of this study has been restricted to the examination of the WEO forecasting record for the principal performance indicators for the major industrial countries and corresponding aggregates and for area-groups of non-oil developing countries. Necessarily set on one side has been the mass of detail presented in the forecasts pertaining to the components of national expenditure and the medium-term "scenarios" formulated in recent years.

Several criteria were used in evaluating the forecasts: the computation and evaluation of various summary statistics of forecast accuracy, bias, and efficiency; comparisons with alternative forecasts--naive forecasts and forecasts produced by the OECD and by national forecasting agencies; the examination of turning-point errors and forecast performance in defined episodes; and, finally, some attempt to explain forecast error in terms of unanticipated developments in policy variables and oil prices.

In judging the WEO forecast performance, a number of points must be kept in mind. Most important, it has to be recognized that the period since the inception of the WEO as a regular forecasting exercise has been extraordinarily rich in economic upheavals, which have made the odds against forecasting formidable. It should also be recalled that the objective of the WEO is not to forecast the most likely outcome but rather to provide conditional estimates of economic developments under the assumption of unchanged policies and exchange rates. Indeed, the quintessential purpose of the WEO exercise is to assist the Fund in carrying out bilateral as well as multilateral surveillance by helping to identify tensions in the projections that may call for policy adjustments or may result in exchange rate changes. Finally, it must remain true that the standard of accuracy required in a forecast is relative to the task for which the forecasts are required. Adapting the results recorded here to this criterion must remain an exercise for the interested reader.

Against this background the forecast performance appears to have been reasonably accurate, particularly for output and inflation, with the industrial country forecasts generally more accurate than those for the developing countries. Although the average absolute error for year-ahead forecasts of growth in the major seven industrial countries as a group over the 1973-85 period has been slightly above 1 percentage point, this result is strongly influenced by a few outliers attributable to some of the large disturbances experienced over the period, notably the first major round of oil price increases in 1973-74 and the tight

monetary conditions in the early 1980s. Excluding 1974 and 1982, the mean absolute error for year-ahead output growth in the major seven industrial countries has been only 0.7 percentage point. While the biggest errors can be related to the large fluctuations in oil prices, the significance of unforeseen fiscal policy developments in explaining forecast error seems strictly limited. The results also show that forecast accuracy is quite sensitive to forecast lead time, so that the errors typically diminish as more information becomes available, particularly for the industrial countries.

The forecasts for output growth, both for industrial and for developing countries, appear to have suffered from a degree of "optimism bias" in the sense that output forecasts have been mostly on the high side in relation to realized values. This bias in an optimistic direction is concentrated in the 1971-80 period. It undoubtedly reflects the fact that the slowdown in growth in this period was only gradually perceived to be a break in trend growth rather than primarily a cyclical phenomenon. Since 1980, output forecast errors have been more evenly distributed. Moreover, there is little evidence of inefficiency in the WEO forecasts, in the sense that the forecast errors cannot systematically be explained by the level of the forecast itself and are not obviously statistically biased. The WEO forecasts also appear to be efficient in that they are generally incapable of being improved by adding information from the available forecasts produced by the OECD or by national forecasters.

As between the performance indicators considered, forecasts for the current account of the balance of payments are inferior to those for output and inflation, at least for the industrial countries. This result, which might appear surprising for forecasts that are prepared on an internationally consistent basis, must be qualified in two important respects. First, the current account is the balance of very large gross flows of exports and imports (of goods, services and transfers); even small errors in the growth estimates of exports or imports may thus show up as large errors in the current account. The second qualification concerns the questionable quality of balance of payments statistics as evidenced by the large and highly volatile discrepancy on the world current account, particularly since the late 1970s. As argued in a recent report on this problem (IMF (1987)), large fluctuations in the current account discrepancy have undoubtedly been an important source of error, not only for balance of payments forecasts but also for world growth projections.

With respect to the track record of the WEO relative to national forecasting agencies, comparisons are hampered by differences in source dates of available forecasts, making generalization hazardous. However, it would appear that the WEO forecasts do not generally provide any distinct improvement over those of national agencies in forecasting national output growth and inflation. Indeed, one outcome of the comparisons with other forecasters is the finding that there is a high degree of common sharing in the principal forecasting errors. The

largest of these are traceable to the two large oil price rises, especially to the first (1973-74), though there are turning-point errors outside of these episodes which also appear to be widely shared by national and international forecasters. Although this may seem disappointing, it would be a mistake to conclude that international forecasts such as those of the WEO are therefore redundant. Indeed, in most cases, the national forecasts are prepared on the basis of assumptions about each country's international environment that typically originate from forecast exercises like those of the Fund or the OECD. Projections prepared on an internationally consistent basis are also necessary as an input into Fund multilateral surveillance activities and into any attempt to coordinate economic policies among countries. Such projections are also required as a basis for monitoring the situation of the indebted developing countries.

The question remains whether the WEO's forecast accuracy can be significantly improved. It would probably not be helpful to be overly ambitious in this respect. Nevertheless, the results of this study do suggest that there may be scope for improvement in several areas. In particular, it seems clear that a reduction in the magnitude and, more especially, the volatility of the world current account discrepancy could only enhance the quality of an internationally consistent exercise such as the WEO. An early implementation of the recommendations contained in the recent report on the discrepancy may well be a necessary condition for significant improvements in the accuracy of the projections. Another conclusion that has emerged from this report is the sensitivity of forecast accuracy to lead time, which underlines the importance of being able promptly to take into account any new information that becomes available. This raises the question of whether the accuracy of the WEO could be improved by a more widespread use of formal, model-based methods which would reduce processing time and would allow more frequent ad hoc updates of the forecasts. The ready availability of such methods would permit a given baseline projection based on the judgment of individual desk officers to be adjusted incrementally at short notice for changes in the main exogenous assumptions underlying the projections and would ease the task of providing scenario analyses. Whilst a move in this direction should not be expected to yield early dividends it is probably also true that a more formal methodology, simply by being more explicit, more easily allows constructive post mortem analysis of forecast error which should help to improve forecast performance over time.

APPENDIX A

Data Definitions

This appendix explains various problems of definition, and the differences between the basic definitions and data used here and those used by Kenen and Schwartz in their 1986 study. 1/

Changing Definitions

The basis for the principal choices of variables to investigate is indicated in the text. This note discusses some problems arising out of the changing definitions of various of the aggregates investigated and the solutions adopted.

1. Industrial countries

To obtain lengthy series for the Group of Seven and the "Europe" aggregates, not provided in early WEO documents, figures were reconstructed from the country detail and "total industrial" category. Group of Seven figures for output growth and inflation were reconstructed from individual country data using preceding year GDP/GNP (current price, current exchange rates) weights. For consistency, later Group of Seven aggregates from the World Economic Outlook, which after 1981 were based on the average of three preceding years' current price output weights, were also recomputed onto the same basis. 2/ (In the later OECD comparison, OECD Group of Seven aggregates were regenerated in the same way.) For the early years, figures for the "Europe" subaggregate were not presented in WEO documents, but were reconstructed as the residual from available data on the "Total Industrial" and the "Non-European" countries separately shown (at the time the latter category comprised only Canada, the U.S. and Japan).

In 1980, the definition of the total industrial group was expanded (and that of Europe correspondingly) to include an additional six countries--Australia, Finland, Iceland, Ireland, New Zealand, and Spain. Thus, in the May 1980 WEO, both the forecast and the outturn for the previous year are for this group; and this necessitated the reconstruction of a current year realization for 1979 using the WEO figures for the major countries supplemented by contemporary evidence of the outturn for the other industrial countries. For this purpose figures from OECD Country Surveys and OECD Economic Outlook (July 1980) were used. A similar problem arose for the realization of the year-ahead forecasts, both for 1979 and for 1980 where outcome figures for

---

1/ The basic data used in the study can be obtained from the author or from the Current Studies Division of the Fund's Research Department.

2/ Current price GDP/GNP data (in billions of dollars at current exchange rates) were conveniently obtained from OECD's 1986 edition of its National Accounts: 1960-1984, and updated from OECD's Main Economic Indicators.

the major individual countries available in the August 1980 and August 1981 WEOs were supplemented by data from the OECD's July 1980 and 1981 Economic Outlook.

2. Developing countries

A number of classification changes affect the aggregates and regional data for the developing countries. First, the position of China was not recognized in the data until April 1980 and even then could not be included in the May WEO of that year. Then, there was a major change of classification in 1980 in which the previous mutually exclusive groupings of "industrial," "more developed primary producers" and "developing countries" were dissolved into two major groups--"industrial" and "developing"; finally, South Africa has been treated differently from time to time, with resultant breaks in the series and non-comparable forecast and realization data in some years. In the case of China, some series were shown both including and excluding China and facilitated a reduction in the extent of non-comparabilities.

a. Reclassification. With effect from the May 1980 WEO the previous classification was abandoned in favor of a new one. Before the change, three mutually exclusive groupings were identified:

(1) Industrial (14 countries);

(2) primary producers in more developed areas (13, including South Africa). This category was also called "other developed" or "more developed primary producing countries"; and

(3) primary producing countries in the developed areas (all other countries not in (1) or (2)). Also called "less developed countries" or "developing countries."

This category was in turn split into:

(a) major oil exporters; and

(b) other developing countries.

In the reclassification, category (2) was dissolved, seven member countries joining the industrial country bloc (including Australia and New Zealand) and six joining the developing country bloc (including South Africa). As before, the latter category was split into "oil" and "non-oil" sub-categories. As a result, forecast and realization data for the total and for the regional aggregates for Africa and Europe are not fully comparable for 1979.

b. China. Data for China were introduced into the Fund's statistics from April 1980, but the May 1980 WEO was already prepared by then and did not reflect the change. The 1981 WEO data on output growth and inflation gave figures excluding as well as including China and no non-comparability of forecast and realization data arises. In export

and import growth the data give rise to a non-comparability in 1981 where actuals (totals) include, but forecasts exclude, China. Subsequent WEOs give data for the totals (forecast and realization) which include China and for the Asia regional category provide data which exclude China. From the 1985 WEO, however, both the total and the regional data include China. This gives rise to another non-comparability in 1984.

Differences from Kenen-Schwartz

While the principal distinction between current year and year ahead forecasts made here follows Kenen and Schwartz (1986), the actual classification of sources is a little different, and different outturn data are used in the two studies.

The most important differences in the sources are that, for the current year forecasts for 1977 this study uses the June-July documents rather than the March 1977 document used by Kenen and Schwartz; and for the 1979 forecasts it uses the June documents rather than the February ones. For the year ahead forecasts, unlike Kenen and Schwartz this study uses a January 1973 source for 1973 forecasts and the March 1977 documents to give forecasts for 1977 where they have none; for the sake of a more complete record the study also uses the February 1979 documents for 1979 which give more information than the December 1978 document used by Kenen and Schwartz. Finally, for 1985 and 1986, we use the published WEOs dated September 1984 and October 1985 rather than the internal August-dated documents apparently used by Kenen and Schwartz.

The net effect is a somewhat more complete year-ahead record at the cost of a greater dispersion of source dates. For the current year forecasts the alterations make little difference in either respect and essentially accomodate the change in year ahead sources, while ensuring that no single source is used for both kinds of forecast. 1/

The outturn data used here are, as explained in the main text, "first available" estimates for current year forecasts and "first settled" estimates for the year ahead forecasts. Kenen and Schwartz mainly use latest available data (with a supplementation of earlier series in an attempt to cope with definition changes).

1/ An exception is the detailed forecasts of export and import growth, for which the current year forecast series can only be made complete by using the February 1979 document.

Replications with Latest Available Outturns

All the principal statistical procedures carried out for the current year and year ahead forecasts against their respective first available and first settled outturns were replicated using the latest available data set derived from the data tape underlying the April 1987 WEO. Because of the change in definitions, data for the total industrial and Europe aggregates were not reprocessed; nor were the developing country forecast errors reprocessed. Thus the replications pertain to the seven industrial countries for output growth, inflation, exports, imports and the balance of payments for the two forecast horizons. It would be excessively tedious to report the results in full, and what follows is a table summarizing the number of countries (out of seven) for which the error statistics deteriorated upon use of the latest available data set, that is, for which the average absolute errors, root mean square errors and Theil statistics increased in value and for which the multiple correlation coefficient in the realization-forecast regression declined in value. By subtraction, the number of cases of improvement can be derived. The tabulation conveys the general impression of some deterioration in the forecast when judged against the latest available estimates of outturn. However, the deterioration is generally marginal and the increase in the error statistics is generally rather small. It is interesting to find that the current year balance of payments forecasts are the chief exception to the general rule, though it still remains true that the balance of payments are on the whole the poorest forecasts among the five variables considered.

Table B.1. Deterioration in Summary Error Statistics when Latest Available Outturn Data are Used

		AAE Increases	RMSE Increases	$\bar{R}^2$ Falls	Theil Increases
Year ahead:	Output	4	3	3	4
	Inflation	4	4	4	7
	Exports	5	6	3	5
	Imports	5	6	5	6
	Balance of payments	3	5	3	2
Current year:	Output	3	3	3	5
	Inflation	5	5	5	7
	Exports	5	5	2	4
	Imports	4	7	4	7
	Balance of payments	2	3	3	7

Note: This table covers the seven countries for which the statistics show a deterioration when the latest available data are used.

APPENDIX C

Bias and Efficiency in the Forecasts

In the text tables and discussion, attention is drawn to the fit of the realization-forecast regression, $R(t) = a + bF(t) + u(t)$ (1)

A perfect forecast would have $a = 0$, $b = 1$, and deviations from these values indicate an inefficiency in the forecast in the sense that the forecasts could be improved by knowledge of these parameters. Subtracting the forecast from each side of (1) gives

$$E(t) = R(t) - F(t) = a - (1-b)F(t) + u(t) \quad (2)$$

This makes it clear that estimates of (1) or (2) are tests of whether the error can be explained by the forecast itself, and also suggests an economical way of examining whether the forecasts are on average biassed or not. This is simply to estimate

$$E(t) = R(t) - F(t) = c + w(t) \quad (3)$$

that is, to regress the error on a constant term. The value of the constant term is then the average error itself and the t-statistic generated on this parameter tests whether the error is significantly different from zero.

Put this way, it is clear that a set of forecasts might fail to pass the efficiency restrictions on the parameters of (1) and (2), yet produce no evidence of average bias. (For example, a highly inefficient set of forecasts might produce offsetting errors.) Holden and Peel (1987) produce examples of this. We would not expect the opposite to hold, however. If the efficiency restrictions hold in (1) and (2), then bias should not be evident in (3). That this is so can be readily seen by noting that (3) can be considered as a restriction of (2) (hence of (1)) in which b is set to unity. But if the joint restrictions are satisfied, the restriction of b to unity should not significantly disturb the estimate of the constant term from zero. It follows that cases where bias is detected yet the t-statistics on the individual coefficients in the realization-forecast regression do not suggest inefficiency are cases in which evaluating the individual coefficients is an insecure means of inferring the results of a test of the validity of the joint restriction. There are a number of examples of this in the developing country estimates in this study.

In conducting the tests of (1) and (3), there seems to be no a priori reason why the often rather small country data sets should not be pooled in the interests of obtaining more precise error bands. (Thus the pooled current year data sample for output and inflation disposes of 7 x 16 = 112 observations as compared with 16 in the individual-country cases whilst the pooled year ahead sample comprises 91 observations against 15 for individual countries.) In any case the validity of the pooling can be tested for by including country shift and slope dummies.

Industrial Countries

Table C1 summarizes the results of estimating equation (3) for each country, the aggregates and for the pooled sample. In order to test for the validity of the pooling the equation was first estimated with country dummies and the constant term suppressed, as

$$E(t) = bD(t)i + w(t) \qquad (4)$$

$$i = 1...7$$

From the results obtained a further specialization of the dummies was determined and the equations were re-run with a constant and selected dummy variables. In the event, only in the current year balance of payments equation was a country dummy (for the United States) significant. In the light of this result, the fully pooled data estimates are displayed alongside the country and aggregate estimates in Table C1. The results in this table reveal bias, among the individual country estimates, only in a handful of cases: the current year forecasts of output for the Federal Republic of Germany tend toward a negative bias (over-forecast) as do the year-ahead forecasts for this country and for France. No bias is revealed among the individual country estimates for inflation except for France in the year-ahead forecasts, and no bias at all is suggested in the balance of payments forecasts, either on a current year on year-ahead basis.

Nevertheless, as can be seen, for each of the individual countries with the sole exception of Italy, the output bias--though not individually significant is consistently of the same (negative) sign. It is not too surprising, then, that the pooling reveals this bias to be significant. In a similar way, there is a predominance of positive signs in the year ahead country inflation errors, significant only for France, but upon pooling significant generally. The extent of the bias in the case of current year output forecasts is quite small (0.3 of a percentage point), but it is more than twice as large in the year ahead forecasts and 0.6 of a percentage point in the inflation forecasts on this basis. However, truncating the sample of year ahead forecasts so as to omit 1974, for which a case can be made (see the text discussion of the comparison of WEO and national forecasts), has the effect of removing completely the significance of the average inflation error, and reduces the size of the output bias (to 0.5 of a percentage point) without removing its significance.

The data pooling thus suggests that there is a tendency to output optimism in the Fund's forecasting, with the output forecasts being pitched too high. Less strongly, there is a suggestion that the forecasts for inflation are on average pitched too low. Interestingly, the signs of these errors are offsetting and a regression of inflation forecast error on output error using the pooled data set confirmed a significant negative relationship, a result which might have one of a

Table C.1. Industrial Countries: Forecast Average Errors and Significance Levels

	Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom	Group of Seven	Total Industrial Countries	Europe	Pooled Data
Current year											
Output	-0.363 (-1.090)	-0.150 (-0.552)	-0.363 (-0.839)	-0.263 (-0.865)	-0.681 (-2.123)	-0.119 (-0.314)	-0.331 (-1.127)	-0.267 (-1.449)	-0.425 (-2.847)	-0.431 (-1.858)	-0.324 (2.593)
Inflation	0.381 (0.393)	-0.019 (-0.100)	-0.431 (-0.703)	0.444 (1.391)	-0.025 (-0.112)	0.675 (1.408)	0.538 (0.959)	0.028 (0.160)	0.100 (0.643)	0.438 (1.945)	0.221 (1.381)
Balance of payments	-0.071 (-0.084)	-3.579 (-1.005)	2.743 (1.444)	-0.293 (-0.355)	1.057 (0.726)	0.493 (0.601)	0.336 (0.313)	0.693 (0.171)	...	...	0.010 (0.148)
Year ahead											
Output	-0.877 (-1.475)	-0.300 (-0.513)	-1.054 (-1.201)	-0.800 (-1.909)	-1.000 (-1.921)	0.262 (0.202)	-0.515 (-1.049)	-0.598 (-1.290)	-0.646 -1.467)	-0.838 (-1.900)	-0.708 (3.096)
Inflation	1.077 (1.487)	0.431 (0.893)	-0.431 (-0.357)	1.023 (2.063)	-0.193 (-1.101)	1.015 (1.027)	1.308 (1.537)	0.366 (0.737)	0.346 (0.748)	0.392 (1.194)	0.602 (2.073)
Balance of payments	0.400 (0.733)	-5.361 (-1.179)	1.654 (0.612)	-0.538 (-0.388)	0.950 (0.502)	-1.254 (-0.802)	1.915 (1.386)	...	-2.923 (-0.398)	6.100 (1.254)	-0.287 (-0.322)

Note: Data show the average error (actual minus forecast) and, in parentheses, the t-statistic obtained from a regression of the error series on a constant term.

number of interpretations. A plausible candidate explanation is that it reflects the prevalence of supply side shocks over the period. An implication is that WEO forecasts of nominal income are superior to those for its components (as Kenen and Schwartz (1986) also argue).

The pooling procedure can be applied equally to the realization-forecast regression (1); in this case the validity of the pooling can be tested for by proceeding from an initial specification in which shift and slope dummies are included for every country (the constant again suppressed). Again, the results suggested little need for any "country effects," and the results of the fully pooled data regression are accordingly shown in Table C2.

Developing countries

Table C3 tabulates average forecast errors (again, actual minus forecast) for the developing country areas, derived from regressing the errors on a constant. The figures in parentheses are t-ratios. Because of the comparative lack of degrees of freedom, the critical values of the t-ratio are somewhat higher than customary: significant estimates are shown by footnote 2. With a smaller data set for individual areas than for the individual Group of Seven countries, pooling is an attractive option and the acceptability of this procedure was assessed by initially regressing the errors on a set of area dummies. Inspection for significant differences between the coefficients attracted by these dummies led to the retention of a separate dummy only for the Western Hemisphere in respect of the inflation forecast errors. 1/

As can be seen, the pooled data would support the contention that, as in the case of the developed countries, WEO output forecasts tend toward optimism, though here the bias is only significant for the year ahead forecasts. Whilst not many of the individual area average output growth forecast errors are statistically significant they are all of the same sign. The pooling does not suggest a bias in other forecasts, even those for inflation, where most, but not all, forecast errors are positive in sign and most, but not all, are individually significant. Finally, the balance of payments forecasts, neither individually nor pooled appear to be biassed.

Table C4 tabulates the results of restricted pooled data estimates for the realization-forecast regression, already reported for the individual areas and the aggregate in the text tables. Again, the validity of the pooling was examined by first regressing the outturns

---

1/ Accordingly the figures reported for these cases are the values of the constant term in a regression which also included a dummy variable for Western Hemisphere observations.

Table C.2. Industrial Countries: Realization-Forecast Regressions on Pooled Data

		Constant Term	Slope	$\bar{R}^2$
Current year:	Output	-0.178 (-0.906)	0.951 (0.968)	0.764
	Inflation	0.117 (0.370)	1.013 (-0.380)	0.887
	Balance of payments	0.445 (0.741)	1.141 (-5.043)	0.945
Year ahead:	Output	-0.140 (-0.329)	0.823 (1.584)	0.371
	Inflation	0.259 (0.408)	1.041 (-0.602)	0.721
	Balance of payments	-0.302 (-0.339)	1.071 (-1.203)	0.787

Note: Terms in parentheses are t-statistics; those for the slope coefficient test for difference from unity.

Table C.3. Developing Countries: Average Forecast Errors and Significance Levels

	Total Non-Oil Developing Countries	Africa	Asia	Europe	Middle East	Western Hemisphere	Pooled Data
Current year							
Output growth	-0.600 (-1.698)	-1.090 (-4.66) 1/	-0.400 (-0.877)	-0.100 (-0.350)	-0.690 (-0.993)	-0.860 (-1.049)	-0.639 (-2.611)
Inflation	6.310 (5.982) 1/	2.050 (1.093)	1.630 (3.461) 1/	5.090 (3.918) 1/	-0.770 (-0.216)	16.130 (4.899) 1/	2.00 2/ (1.661)
Balance of payments	0.453 (0.212)	0.390 (0.318)	2.200 (1.096)	-0.738 (-1.064)	1.430 (1.322)	-0.760 (-0.346)	0.815 (0.976)
Year ahead							
Output growth	-1.171 (-1.845)	-1.071 (-2.025)	-0.543 (-0.648)	-0.500 (-1.096)	-2.029 (-2.027)	-2.186 (-1.767)	-1.288 (-3.297) 1/
Inflation	12.629 (4.042) 1/	2.514 (2.215)	1.943 (2.532) 1/	8.667 (2.817) 1/	-3.714 (-0.418)	39.986 (3.976) 1/	2.119 2/ (0.679)
Balance of payments	15.225 (1.342)	0.538 (0.369)	1.313 (0.471)	-0.900 (-0.433)	2.213 (1.328)	1.213 (0.267)	0.968 (0.805)

Note: See note a to Table C.1.

1/ Significant at 0.05 level.
2/ Excluding Western Hemisphere.

Table C.4. Developing Countries: Realization-Forecast Regressions on Pooled Data

	Constant	Slope	Dummy Variables (Western Hemisphere) shift	slope	$\bar{R}^2$
Current year					
Output growth	-0.288 (-0.470)	0.911 (0.624)	-	-	0.454
Inflation	5.566 (2.508)	0.803 (1.572)	-11.007 (-1.761)	0.525 (3.840)	0.954
Balance of payments	-1.291 (-0.842)	0.849 (1.624)	-	-	0.677
Year ahead					
Output growth	-0.552 (-0.438)	0.838 (0.616)	-	-	0.218
Inflation	13.965 (3.259)	0.381 (3.252)	-76.218 (-3.262)	2.676 (5.442)	0.877
Balance of payments	-2.951 (-1.381)	0.704 (0.365)	-	-	0.410

Note: Figures in parentheses are t-statistics; those on the constant and on both dummy variable terms test for difference from zero, whilst those on the slope coefficient test for a difference from unity.

on a full set of area shift and slope dummies, leading to a more specialized set of dummies on the basis of the consideration of the significance of differences between the dummy variable coefficients. Further runs then showed that no dummy variables were significant except for those pertaining to inflation in respect of the Western Hemisphere. These regressions suggest that the output and balance of payments forecasts are efficient in the sense that the constant terms are not significantly different from zero, the slope terms not different from unity. The inflation forecasts are another story, however, even when the Western Hemisphere observations are treated separately. The current year errors suggest that when inflation is above about 7 percent, it is likely to be underforecast, whilst the year ahead errors suggest that inflation is overforecast at levels below about 37 percent.

An Alternative Naive Forecast

The Theil statistics reported in the main tables compare the RMSE of the WEO forecasts with the RMSE of an alternative naive "no change" forecast, where the forecast for year t is set equal to the realization for t-1.

A plausible alternative naive forecast is one that projects the established trend. Table D1 shows the Theil statistics computed for this alternative naive standard, 1/ where the trend values were identified with a moving average of the past ten years' output growth or inflation. (It was not felt appropriate to treat the balance of payments in this way.) Because the computed "trend" is backward-looking, this alternative standard might be expected to represent a harder standard to beat for a variable with a comparatively short cycle, like output, than the conventional "no change" naive standard underlying the Theil statistics reported in the text tables. For a variable with an evolving trend or longer cycle, like inflation, however, it is likely to prove an easier standard.

Industrial countries

This expectation is born out for the industrial countries where the Theil Statistics quoted for output growth in the Table are uniformly higher than those reported in the text tables (though not, in general, much higher) whilst those given for inflation are uniformly lower. The relatively poorer standard of accuracy of the year ahead forecasts is repeated, but there is no case where the trend alternative would have outperformed the WEO forecasts.

Developing countries

The table also tabulates alternative Theil statistics, generated from a "forecast-trend" assumption for the developing countries; in the same way as for the developed countries, a ten-year moving average of past output growth (inflation) was used as the naive forecast. As can be seen, there are a handful of instances where such a naive forecast would have produced better forecasts (by a minimum RMSE criterion) than the procedures actually used; but a similar statement was already available with respect to "no change" forecasts and in fact in nearly every case the "no change" forecast is a harder standard to beat than the "trend" alternative--that is, the Theil statistics tabulated in Table D2 are nearly always smaller than those reported for the conventional "no change" forecast in the text tables. There are only three exceptions to this, all of them arising in the year ahead output growth forecasts. Here, for Africa and the Middle East the "trend"

1/ As the ratio of the RMSE of the WEO forecast to the RMSE of the "trend as forecast" alternative.

naive forecast is better than both the "no change" naive forecast and the forecasts actually made; with respect to Europe, whilst the "trend" naive forecast is a harder standard to beat than the "no change" forecast, the forecasts actually made outperform both.

Table D.1. Alternative Theil Statistics

	Output Growth Current Year	Inflation Current Year	Output Growth Year Ahead	Inflation Year Ahead
Industrial countries				
Canada	0.431	0.412	0.659	0.631
United States	0.344	0.258	0.618	0.536
Japan	0.436	0.435	0.868	0.659
France	0.504	0.402	0.683	0.493
Germany, Fed. Rep. of	0.546	0.450	0.791	0.386
Italy	0.456	0.297	0.783	0.520
United Kingdom	0.498	0.270	0.740	0.401
Group of Seven	0.284	0.224	0.623	0.488
Total industrial	0.406	0.199	0.645	0.471
Europe	0.439	0.270	0.730	0.431
Non-Oil Developing countries				
Africa	0.770	0.681	1.239	0.570
Asia	0.911	0.572	1.170	0.756
Europe	0.232	0.529	0.520	1.038
Middle East	1.055	0.625	2.004	1.128
Western Hemisphere	0.714	0.433	0.861	0.976
Total	0.624	0.461	0.938	0.942

Note: Using ten-year moving averages as the "naive" forecast.

Basic Data Extensions, National Forecasts

As indicated in the main text the comparison with national forecasts was carried out by extending the base already assembled by Llewellyn and Arai in their OECD study (Llewellyn and Arai (1984). Data were sought from the same institutions on the same basis as in that study, or from corresponding country desk officers in the Fund. The paradigm return would be of forecasts for year t prepared in the last quarter of t-1, with outturns taken from data available in t+1; however, there were numerous variations on the paradigm. Practice regarding the vintage of data used to describe the outturn clearly varied; more important, perhaps, the date of preparation of the forecasts is not constant across forecasters relative to the forecast horizon. Table E1 tabulates the dates recorded for the Summit Seven forecasts discussed in the text. Assuming that forecasts of fourth quarter origin were in fact made available in December and treating the two month lead time of the Japanese financial year forecasts as equivalent to that of a November forecast for the calendar year the unweighted "centre for gravity" of these forecasts is December. The WEO year ahead forecasts with which the national forecasts are compared in the main text have a considerable dispersion over the period in their lead time on the forecast; the average of these lead times is two months, giving a centre of gravity for these forecasts of October. 1/

Table E2 provides the data obtained as updates and additions to the series published by Llewellyn and Arai. Additional detail is included for the U.S. consensus forecasts because of a change in definition of the "actuals" supplied for these forecasts. For the United States, data are also included showing the track record of the projections (or "economic assumptions") contained in the Mid-Session Budget Reviews, published by the Office of Management and Budget in July-August. Since 1980, the dating of these forecasts thus corresponds quite closely to that of the WEO, whereas the dating of the CEA's projections contained in its Annual Report (published in February) corresponds more closely to that of the available WEO projections prior to 1980.

1/ No account in any of this is taken of publication lags which may vary between the forecasts.

Table E.1. Six Major Industrial Countries:
Dates of National Forecasts

	Forecast	Date 1/
United States	OMB (Mid-Session Budget Review)	July-August, t-1
	CEA (Annual Report)	February, t
	Consensus	Fourth Quarter, t-1
Japan	Official	January, t-1 for FYt
France	Official	September, t-1
Germany, Fed. Rep. of	Official	January, t
	Five Wise Men	November, t-1
	Official	October, t-1
Italy	ISCO	September-October, t-1
United Kingdom	NIESR	February, t

1/ Data given for forecasts of output growth and inflation in year t. Data refers to publication, when there is one, rather than to preparation per se. In some cases, the date shown is approximate.

Table E2. Updates to the Llewellyn-Arai Data Base

	GNP			GNP Deflator		
	Forecast	Actual	Error	Forecast	Actual	Error
Country: United States						
Source: Consensus 1/						
1973	6.1	5.9	0.2	3.3	6.0	-2.7
1974	1.1	-1.8	2.9	5.9	9.4	-3.5
1975	-0.8	-1.8	1.0	9.1	9.3	-0.2
1976	5.9	6.0	-0.1	6.0	5.3	0.7
1977	5.0	4.9	0.1	5.5	5.8	-0.3
1978	4.3	4.4	-0.1	5.9	7.3	-1.4
1979	2.4	3.2	-0.8	7.4	8.5	-1.1
1980	-1.3	-0.2	-1.1	8.8	9.0	-0.2
1981	1.2	1.9	-0.7	9.5	9.4	0.1
1982	0.5	-1.9	2.4	7.9	6.0	1.9
1983	2.4	3.7	-1.3	5.3	3.8	1.5
1984	5.2	6.5	-1.3	4.8	4.1	0.7
1985	3.4	2.7	0.7	4.3	3.3	1.0
Source: Council for Economic Advisers						
1983	3.1	6.1	-3.0	5.6	4.1	1.5
1984	4.5	5.6	-1.1	5.0	3.5	1.5
1985	4.0	2.5	1.5	4.3	3.2	1.1
Source: Office of Management and Budget (Mid-Session Budget Review)						
1976	6.3	6.1	0.2	7.1	5.1	2.0
1977	5.7	4.9	0.8	6.0	5.9	0.1
1978	5.3	4.0	1.3	6.3	7.4	-1.1
1979	4.3	2.3	2.0	6.6	8.8	-2.2
1980	1.0	-0.2	1.2	8.9	9.0	-0.1
1981	0.3	2.0	-1.7	10.0	9.2	0.8
1982	3.4	-1.7	5.1	8.0	5.9	2.1
1983	4.4	3.7	1.0	6.5	4.2	2.3
1984	5.2	6.8	-1.6	4.8	3.8	1.0
1985	4.3	2.7	1.6	4.7	3.4	1.3
Country: Japan						
Source: Official 2/						
1983	3.4	3.7	-0.3	2.1	0.6	1.5
1984	4.1	5.1	-1.0	1.7	1.5	0.2
1985	4.6	4.3	0.3	1.4	1.5	-0.1
Country: Germany, Federal Republic of						
Source: Consensus 3/						
1983	0.1	1.3	-0.3	3.5	3.2	0.3
1984	2.0	2.6	-0.6	2.5	1.9	0.6
1985	2.0	2.5	-0.5	2.5	2.2	0.3
Source: Official						
1983	0.5	1.3	-0.8	3.5	3.2	0.3
1984	2.5	1.6	-0.1	3.0	1.9	1.1
1985	2.5	2.5	0.0	2.0	2.2	-0.2
Source: Five Wise Men						
1983	1.0	1.3	-0.3	3.5	3.2	0.3
1984	2.5	2.6	-0.1	3.0	1.9	1.1
1985	3.0	2.5	0.5	2.0	2.2	-0.2
Country: France						
Source: Official 4/ 5/						
1984	1.0	1.5	-0.5	6.7	7.2	-0.5
1985	1.8	1.4	0.4	5.5	5.8	-0.3
Country: United Kingdom						
Source: National Institute of Economic and Social Research 5/ 6/						
1983	1.4	2.2	-0.8	7.1	5.4	1.7
1984	2.2	2.5	-0.3	5.6	5.1	0.5
1985	2.8	3.4	-0.6	5.6	5.4	0.2
Country: Austria						
Source: Institut für Wirtschaftsforschung 5/ 6/						
1983	0.5	1.9	-1.4	4.3	3.3	1.0
1984	1.5	2.2	-0.7	5.3	5.6	-0.3
1985	3.0	2.9	0.1	4.0	3.2	0.8
Country: Finland						
Source: Ministry of Finance 4/ 5/ 6/						
1984	3.0	2.8	0.2	8.0	7.1	0.9
1985	3.5	2.8	0.7	6.0	5.9	0.1
Country: Netherlands						
Source: Central Planbureau 6/						
1983	0.1	0.8	-0.7	4.4	2.9	1.5
1984	0.9	1.8	-0.9	3.0	2.5	0.5
1985	1.8	2.1	-0.3	1.5	2.6	-1.1
Country: Sweden						
Source: National Institute of Economic and Social Research 4/ 6/						
1984	2.1	4.0	-1.9	7.4	8.0	-0.6
1985	2.1	2.2	-0.1	4.8	7.4	-2.6
Country: Switzerland						
Source: Arbeitsgruppe für Wirtschaftsprognose						
1983	-1.4	1.0	-2.4	4.0	2.7	1.3
1984	1.5	2.6	-1.1	3.0	3.3	-0.3
1985	1.7	4.0	-2.3	2.5	3.6	-1.1

1/ Forecasts taken from fourth quarter reports of the ASA and the National Bureau of Economic Research.
2/ Fiscal year (April-March) forecasts.
3/ Joint forecasts of the five leading institutes.
4/ 1983 data in Llewellyn-Arai (1984).
5/ GDP, not GNP.
6/: Consumer price index, not GNP deflator.

Fiscal Policy, Oil Prices and the Exchange Rate

In this appendix we describe further the basis for the fiscal policy assumptions from which the information on unexpected changes described in the text are obtained, the expected signs on variables in the associated regressions and a summary of some of the additional results obtained for fiscal policy and oil price innovations together with a note on the effect of exchange rate projection error.

Fiscal policy assumptions

The WEO forecasters reckon to take into account existing policies including announced and probable changes during the course of the forecast period. The quantification of this policy assumption--in the form of an estimate of the "fiscal impulse"--is more recent, however. 1/ To fill the gap, the sequence of source documents for the forecasts classified as "year ahead" was examined for qualitative comment that might be translated into quantitative terms. Relevant extracts are noted below, together with a Table (Appendix Table F1) showing the resultant quantified forecasts and realizations. The latter, drawn from the WEO 1987 data tape, extend back to 1977. Figures for earlier years are derived from an earlier WEO source but do not appear to be fully homogeneous with the later series. In the text, reference is made to an alternative set of realizations, identified as the (negative) of the first difference of estimates of the structural budget balance. For these data the paper by Price and Mueller (1984), supplemented by later figures from issues of the OECD's Economic Outlook, was used. These data are shown in Table F2. The two alternative realization series (Tables F1 and F2) lead to two corresponding alternative policy innovation measures, FP1 and FP2, respectively.

WEO extracts on fiscal policy

January 1973 WEO (for 1973). No basis for any assumption is given for Canada, Japan, or the U.S. But for France the "stance of fiscal policy... is expected to be more expansionary than in 1972", for the Federal Republic of Germany "...some reduction in the stimulative effect of fiscal policy..." is noted, whilst for Italy "the expansionary impact of fiscal policy is expected to increase in 1973..." For the U.K., comment suggests possibly some contraction.

December 1973 WEO (for 1974). The outlook was dominated by consideration of the effect of the oil crisis. For Canada, U.S., France, and the U.K., there is no comment. In Japan, fiscal policy is

---

1/ The fiscal impulse measure is described in Heller et al. (1986) where it is compared with other measures of fiscal stance, including the structural (cyclically-corrected) budget balance. The first difference of the latter (with sign reversed) is a close correlate in principle of the fiscal impulse.

Table F.1. Forecasts and Realizations of the Fiscal Impulse
(In percent of GNP)

	Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom
				Forecasts 1/			
1973	0.0	0.0	0.0	0.25	-0.25	0.25	-0.25
1974	0.0	0.0	-0.25	0.0	0.25	0.25	0.0
1975	0.25	0.0	0.25	0.0	0.50	0.0	0.25
1976	0.0	-0.5	0.0	-0.1	-0.5	-1.0	-1.0
1977	-0.2	0.7	0.3	-1.5	0.1	0.7	-1.3
1978	0.5	0.25	0.5	0.0	0.5	0.25	0.5
1979	-0.5	-0.4	1.1	-0.8	0.3	-0.8	-0.7
1980	-0.25	-0.25	-0.25	-0.25	-0.25	-0.25	-0.25
1981	-0.25	-0.25	-0.25	-0.25	-0.25	-0.25	-0.25
1982	-0.5	0.0	-0.5	1.00	-0.5	-0.5	-0.5
1983	-0.6	1.0	-0.25	0.0	-0.25	0.25	0.0
1984	0.0	-0.6	-0.5	-0.4	-0.2	0.0	-0.3
1985	-0.7	0.0	-0.5	-0.2	-0.3	-0.2	-0.4
				Realizations 2/			
1973	-0.4	-0.5	1.2	-0.3	-0.5	1.7	2.2
1974	-0.6	-0.8	-0.5	-0.1	0.6	-0.8	-0.2
1975	2.5	3.2	2.4	2.6	1.7	1.5	2.7
1976	-0.4	-1.1	-0.3	-1.7	-0.3	-1.3	-2.6
1977	1.8	0.0	0.2	-0.3	-0.3	-0.8	-1.6
1978	1.4	0.1	0.2	0.9	0.1	3.8	2.0
1979	-0.4	-0.8	1.1	0.1	0.1	-1.7	-0.8
1980	-0.1	0.4	0.1	-0.8	-0.6	0.1	-1.7
1981	-1.1	0.0	-0.5	1.1	-0.5	0.7	-1.4
1982	1.4	0.4	-0.3	0.0	-1.1	0.6	-0.7
1983	0.6	1.6	-0.5	0.1	0.2	0.0	0.7
1984	1.5	0.4	-0.4	-0.1	0.6	-1.0	0.3
1985	0.5	0.3	-0.6	-0.3	-0.5	0.8	-0.1

1/ See text for sources. Note that fiscal impulse is akin to the first difference of a structural budget balance measure times -1; that is, a plus indicates expansion.

2/ From 1977, data are from WEO, March 1987 (unpublished) and the WEO data base of this date. Data for 1973-76 are from WEO, September 1979 (unpublished). From comment and available overlapping data, it is clear the series suffers a discountinuity in 1976-77.

Table F.2. Fiscal Policy: Changes in the Structural Budget Balance 1/

(In percent of potential GNP)

	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985
Canada	+0.4	+0.9	-3.5	+0.4	-0.2	-0.7	+1.3	-0.5	+0.9	-0.4	-1.6	-1.2	-0.6
United States	-0.2	+0.9	-1.6	+1.3	+0.2	+0.3	+0.3	-0.5	+0.9	-0.9	-0.7	-0.5	-0.8
Japan	-0.3	+0.4	-2.6	-1.0	-1.1	-2.8	+0.6	+0.2	+0.6	+0.3	+0.5	+1.0	+0.5
France	0.0	+0.3	-1.1	+0.6	-0.4	-1.5	+0.9	+1.3	-1.0	-0.6	-0.1	+0.4	+0.6
Germany, Fed. Rep. of	+1.3	-1.8	-2.9	+1.2	+1.1	-0.4	-0.6	-0.2	+0.1	+1.2	+1.1	+0.2	+0.6
Italy	+0.1	+0.2	-2.0	+1.7	+1.1	-1.8	-0.6	+1.0	-3.4	+0.6	+2.2	-1.7	-1.2
United Kingdom	-2.8	-0.1	+0.5	-0.2	+1.7	-2.1	+0.6	+1.1	+2.9	+1.5	-1.3	-0.8	+0.5

Sources: For 1973-79, data are from Price and Mueller (1984), Table 2; data for later years are from OECD Economic Outlook (December 1984; June 1985; May 1986; December 1986). There may be some minor inconsistencies in the data.

1/ The sign convention is that a + sign is contractionary (a rise in surplus or a fall in deficit), a - sign is expansionary (increase in surplus, or increase in deficit).

"likely to be more conservative" while in Germany a slight relaxation is hinted at. The WEO is sceptical of efforts in July to restrain fiscal policy.

December 1974 WEO (for 1975). For Canada some stimulus is indicated; for the U.S. the "...broad maintenance of the present fiscal stance". In Japan "...relaxation of fiscal policy is noted, with "stimulative action" in the Federal Republic of Germany and "no change" in Italy. In the U.K., the "November budget would ease the cash position of industry."

December 1975 WEO (for 1976). A full set of forecast impulse figures is provided (except for Canada). They are derived from central government fiscal balance data.

March 1977 WEO (for 1977). A full set of forecast fiscal impact figures is provided, now including Canada, again for the central government.

December 1977 WEO (for 1978). No estimates are given, but there is a comment that "For 1978, a perceptible shift in the orientation of fiscal policies appears to be in the making". There is also a comment on the stimulatory measures applied in 1977 QIV by the U.K., the Federal Republic of Germany, Japan, Canada, and the likely stimulus to come in U.S. and Italy. Only France has indicated no relaxation though the WEO thinks this "may be modified" in the course of the year.

February 1979 WEO (for 1979). A fully quantified set of measures is available, this time with figures on a general government basis also available for three countries.

August 1979 WEO (for 1980). No quantification, just a general reference to "caution" in fiscal policy.

August 1980 WEO (for 1981). No overall quantification, but a comment that no tax cut has been assumed for the US and a general reference to the "restrictive" stance of fiscal policy (with no comment on any likely relaxation).

August 1981 WEO (for 1982). Not many figures are quoted, but for 1982, "...estimates indicate a withdrawal of stimulus for all countries except the United States and France. For the United States, no impulse either positive or negative, is projected; for France, an expansionary impulse equivalent to 1 percent of GDP".

August 1982 WEO (for 1983). There is a general reference to the policy analysis of the March 1982 WEO as "still valid." This noted, inter alia, that "the general trend towards less expansionary fiscal policies observed since 1979...is estimated to have continued in 1981" and went on ..."It is expected to intensify in 1982 but to moderate in

1983." For the U.S., there is a sceptical discussion of administration policy claims, which allow one to suggest a forecast of expansion, while for the Federal Republic of Germany, the discussion indicates "restraint" for Japan. Perhaps a minor withdrawal of stimulus. There is a specific forecast for Canada that the cyclically adjusted budget deficit would decline by less than 1 percent of GNP in 1983, but very general indications for the remaining countries--perhaps no change in France, an unwanted stimulus in Italy.

August 1983 WEO (for 1984). A full set of quantified estimates for Central Government fiscal impulse.

September 1984 WEO (for 1985). A full set of quantified estimates for both the central and general government. For 1985, the differences are very small.

Oil price forecasts

The oil price series employed is the export unit value of major oil exporters for which WEO forecasts are available from the source documents. The full series of forecasts, realizations and innovations is shown in Table F3. Because the innovations are dominated by the two rounds of oil price increases, dummies for these two shocks were used in the regressions as an alternative to the continuous series.

Regression evidence

In the regressions the forecast error is regressed on the oil price and fiscal policy innovations. The latter are split into two components--"own" fiscal policies and "external" (other countries') fiscal policies. The weighting for the latter is supplied either by GNP or by import weights.

As the fiscal impulse measure is signed positively for output-expansionary fiscal policy (assuming a broadly "Keynesian" mechanism), positive innovations (forecast minus realization) indicate that policy was less expansionary than predicted, and might be expected to be associated with a positive (forecast minus realization) output error. Thus the expected sign on the fiscal policy innovations in the output error equations is positive, unless there is complete crowding out. 1/ The signs to be "expected" in the inflation error equations are less obvious to the extent that the inflation response to fiscal policy depends on the response of the exchange rate, for which a priori indications are themselves ambiguous. Finally, in the balance of payments equations, "own country" and "external" policy would on the

---

1/ And assuming the absence of any significant reverse feedback from output error to fiscal innovation.

Table F.3. Oil Prices: Forecast and Realization
(In percentage change)

	1973	1974	1975	1976	1977	1978	1979
Forecast	5.2	33.0	6.3	5.3	8.5	0.0	10.1
Realization	40.0	225.8	5.1	6.3	9.4	0.4	45.9
Error	-34.8	-192.8	1.2	-1.0	-0.9	-0.4	-35.8

	1980	1981	1982	1983	1984	1985
Forecast	20.0	10.5	4.5	2.5	-1.5	0.0
Realization	63.6	9.8	-4.3	-11.4	-2.5	-4.3
Error	-43.6	0.7	8.8	13.9	1.0	4.3

same broadly Keynesian grounds, be expected to have opposing effects. The pattern of "expected" signs indicated by these considerations is summarized in Table F4.

The regression results are summarized in Tables F5A and F5B for the measure of fiscal policy FP2. Use of the alternative measure, FP1, does not greatly recondition the results. There is a discussion in Section IV of what happens when the regressors are extended to include actual oil price and fiscal policy variables.

Exchange rate effects

As noted in the text, WEO forecasts dispose of a "working assumption" about exchange rates, latterly that real exchange rates remain constant over the forecast horizon at the levels reached in a recent period, formerly that nominal exchange rates will remain at recent levels. The precise nature of these assumptions, for the set of Year Ahead forecasts defined in this study is set down in Table F6. The associated ex post error (defined as the forecast minus actual in percent of actual) is shown in Table F7.

The exchange rate assumption is not on quite the same level as the assumptions made about fiscal policy and oil prices and, as argued in the main text, it does not seem appropriate to treat the ex post exchange rate errors in quite the same fashion as the fiscal policy and oil price innovations. In any case, it would not be difficult to argue that, the lags in the process being as long as they are, exchange rate errors would not be likely to impact strongly on output growth or inflation over the typical forecast horizon. It might be supposed that the balance of payments forecasts, on the other hand, could be materially affected by these errors. It should be noted that, as the balance of payments data are defined in dollar terms, they will not, except for the United States, reflect the effects of changes in the value of national currency numeraires. A particular reason for examining the contribution of the exchange rate assumption error to the balance of payments forecast errors is of course simply that the balance of payments appears to be least well forecast among the principal magnitudes (though, as noted in the text this is by no means a WEO-specific phenomenon).

In light of these considerations the contribution of exchange rate assumption error to balance of payments errors was examined by regressing the latter on the former. In view of the likely lags it was felt appropriate to perform this regression only for the year ahead forecasts. 1/

1/ In principle, because of the importance of lags in the export and import responses to exchange rate changes, the within-year profile of the exchange rate error might be examined as well. However the regressions described were fitted to annual average errors in both variables.

Table F.4. "Expected" Signs in the Regression of Forecast Error on Unexpected Changes

	Own Fiscal Policy	Other Fiscal Policy	Oil Prices	Oil Dummies
Output	+	+	-	+
Inflation	+	+	+	-
Balance of payments	-	+	-	+

Table F.5. The Effects of Unexpected Changes in Fiscal Policy and Oil Prices on Forecast Error: Significant Results

	Fiscal Policy and Oil Prices		Fiscal Policy and Oil Price Dummies	
	GNP weights	Import weights	GNP weights	Import weights
Output:	At 5%, U.S.: oil Japan: oil France: FP (other) 1/ U.K.: FP (own) FP (other) 1/ Oil	At 5%, U.S.: oil Japan: oil France: FP (other) 1/ U.K.: FP (own) FP (other) 1/ Oil	At 5%, U.S.: D1 Japan: D1 France: FP(other) 1/ U.K.: FP (other) 1/ FP (own) D1	At 5%, U.S.: D1 Japan: D1 U.K.: D1 FP (own) FP (other) 1/
	At 10% additionally, None	At 10% additionally, Italy: FP (other) 1/	At 10% additionally, None	At 10% additionally, Germany: FP (own) 1/ France: FP (other) 1/
Inflation:	At 5%, Canada: oil U.S.: oil Japan: oil France: oil U.K.: oil FP (other) Italy: oil	At 5%, Canada: oil U.S.: oil Japan: oil France: oil U.K.: oil FP (other) Italy: oil	At 5%, Canada: D1, D2 U.S.: D1 Japan: D1 France: D1 U.K.: D1, D2 FP (other) Italy: D1	At 5%, Canada: D1, D2 FP (own) 1/ U.S.: D1 Japan: D1 France: D1 U.K.: D1, D2 FP (other) Italy: D1
	At 10% additionally, Canada: FP (own) 1/	At 10% additionally, None	At 10% additionally, France: D2 FP (own) Italy: D2	At 10% additionally, France: D2 Italy: D2
Balance of payments	At 5%, U.K.: oil U.K.: FP (own) France: FP (own) 1/	At 5%, U.K.: oil FP (own) France: FP (own) 1/	At 5%, Japan: D2 U.K.: FP (own) Germany: D2	At 5%, Japan: D2 U.K.: FP (own) Germany: D2
	At 10% additionally, None	At 10% additionally, Canada: FP (other) 1/	At 10% additionally, U.K.: D1, D2 France: FP (own) 1/	At 10% additionally, U.K.: D1, D2 France: FP (own) 1/

1/ "Incorrect" sign, see Table E.4.

Table F.6. Exchange Rate Forecast Assumptions: "Year-Ahead" Forecasts

WEO Source	Year Forecast	Constancy Assumption in WEO
Jan. 1971	1971	(Presumably at recent "fixed" levels)
Jan. 1973	1973	(Presumably at recent "fixed" levels) 1/
Dec. 1973	1974	"Present" levels 2/
Dec. 1974	1975	Values of Oct. 21-25, 1974 3/
Dec. 1975	1976	Average October 1975
Mar. 1977	1977	Average December 1976
Dec. 1977	1978	Average November 1977
Feb. 1979	1979	Average December 1978
Aug. 1979	1980	Average July 1979
Aug. 1980	1981	Average July 1980
Aug. 1981	1982	Average May-July 1981
Aug. 1982	1983	Exchange rates "in effect just after the realignment of rates among the European Monetary System countries in mid-June 1982" 4/
Aug. 1983	1984	Average May 1983
Sept. 1984	1985	Average June 1984
Oct. 1985	1986	Rates on July 22, 1985 5/
Oct. 1986	1987	Real levels of 2 weeks prior to September 5, 1986.

1/ For purposes of Table E7 assumed to be December 1972.
2/ For purposes of Table E7 assumed to be November 1973.
3/ For purposes of Table E7 assumed to be average October 1974.
4/ For purposes of Table E7 assumed to be July 1982.
5/ For purposes of Table E7 assumed to be July 1985.

Table F.7. Exchange Rate Assumption Error

(In percent)

	Canada	United States	Japan	France	Federal Republic of Germany	Italy	United Kingdom
1973	3.18	9.46	-6.01	-4.90	-10.22	10.66	3.95
1974	-3.01	-2.61	4.04	7.20	-2.29	8.01	0.15
1975	4.69	3.70	-2.89	-5.19	-1.37	-2.66	8.54
1976	-5.11	-1.77	-3.69	4.97	-6.43	20.66	11.49
1977	5.19	1.58	-8.94	-0.41	-1.64	3.63	-2.42
1978	6.03	5.46	-11.50	-0.55	-4.14	4.52	1.88
1979	-0.98	-0.92	12.38	-1.22	-2.69	-0.92	-6.83
1980	0.06	-1.28	4.23	-1.68	-1.39	3.87	-3.75
1981	-2.39	-13.57	-11.00	13.69	9.40	15.43	0.65
1982	-1.35	-8.06	5.43	8.60	-6.04	6.99	5.61
1983	-5.22	-3.43	-10.70	6.25	-2.05	4.81	11.03
1984	1.42	-9.03	-5.55	6.62	4.23	8.31	7.97
1985	1.74	-6.37	-2.11	0.34	1.59	7.09	1.49
1986	11.74	21.81	-22.93	-5.36	-9.78	-4.50	14.37

The results of the regressions were to produce insignificant coefficients on the exchange rate error term except in two cases, those of Japan and the United States. In the former case a negative coefficient was found, in the latter a positive one. The negative sign is what would be predicted on a "partial equilibrium" basis where the unexpected exchange rate movement can be thought of as an autonomous factor and volume effects dominate valuation effects. Accordingly, the opposing result found for the United States suggests the significance of J-curve effects. More particularly, the results probably reflect the comparatively large fluctuation in the exchange rate error assumption (Table F7) for these two countries (especially in 1986).

References

Ash, J.C.K. and D.J. Smythe, "The Accuracy of OECD Forecasts of the International Economy," paper presented at the International Symposium on Forecasting (London, 1984).

Burns, Sir Terence, "The Interpretation and Use of Economic Predictions," Proceedings of the Royal Society, Series A (London, 1986).

Cooper, R.N., "Economic Interdependence and Coordination of Economic Policies" in Handbook in International Economics, Vol. 2, ed. by R.W. Jones and P.B. Kenen, (Amsterdam: Elsevier, 1985).

Crockett, Andrew and Morris Goldstein, "Strengthening the International Monetary System: Exchange Rates, Surveillance and Objective Indicators," Occasional Paper, No. 50 (Washington: International Monetary Fund, 1987).

Goldstein, Morris, "Global Effects of Fund-Supported Adjustment Programmes," Occasional Paper, No. 46 (Washington: International Monetary Fund, 1986).

Heller, P.S., R.D. Haas, and N.H. Mansur, "A Review of the Fiscal Impulse Measure," Occasional Papers, No. 44, (Washington: International Monetary Fund, 1986).

Holden, K. and D.A. Peel, "On Testing for Unbiassedness and Efficiency of Forecasts," University College of Aberystwyth Discussion Papers (Aberystwyth, 1987).

Holden, K., D.A. Peel, and B. Sandhu, "The Accuracy of OECD Forecasts," (mimeographed; University of Liverpool, 1987.

International Monetary Fund, Report on the World Current Account Discrepancy," (Washington: International Monetary Fund, 1987).

Kenen, P.B. and S.B. Schwartz, "An Assessment of Macroeconomic Forecasts in the International Monetary Fund's World Economic Outlook," Working Papers in International Economics, (Princeton University: International Finance Section, 1986).

Llewellyn, J. and H. Arai, "International Aspects of Forecasting Accuracy," OECD Economic Studies, (Autumn 1984), pp. 73-117.

Llewellyn, J., S. Potter, and L. Samuelson, Economic Forecasting and Policy (London: Routledge and Kegan Paul, 1985).

Osborne, D.R. and F. Teal, "An Assessment and Comparison of Two NIESR Econometric Forecasts," National Institute Economic Review, No. 88 (1979), pp. 50-62.

Price, R.W.R. and Muller, P. (1984), "Structural Budget Indicators and the Interpretation of Fiscal Policy Stance in OECD Economies", OECD Economic Studies, Autumn, pp. 28-71.

Savage, D., "The Assessment of the National Institute's Forecasts of GDP, 1959-82," National Institute Economic Review (August, 1983).

Smythe, D.J., "Short- Run Macroeconomic Forecasting: The OECD Performance", Journal of Forecasting, Vol. 2, (1983), pp. 37-49.

Smythe, D.J. and J.C.K. Ash, "Forecasting Gross National Product, the Rate of Inflation and the Balance of Trade: The OECD Performance", Economic Journal, Vol. 85, (1975), pp. 361-64.

Theil, H. Applied Economic Forecasting (North Holland: Amsterdam, 1966).

Wallis, K.F. (eds.) et al. Models of the UK Economy 3, (London: Oxford University Press, 1986.