

No. 2206

**FISCAL FORECASTING:  
THE TRACK RECORD OF THE  
IMF, OECD AND EC**

Michael J Artis and Massimiliano Marcellino

***INTERNATIONAL MACROECONOMICS***



**Centre for Economic Policy Research**

# **FISCAL FORECASTING: THE TRACK RECORD OF THE IMF, OECD AND EC**

**Michael J Artis**, European University Institute and CEPR  
**Massimiliano Marcellino**, IGER, Università Bocconi

Discussion Paper No. 2206  
August 1999

Centre for Economic Policy Research  
90–98 Goswell Rd, London EC1V 7RR  
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999  
Email: [cepr@cepr.org](mailto:cepr@cepr.org), Website: <http://www.cepr.org>

This Discussion Paper is issued under the auspices of the Centre's research programme in **International Macroeconomics**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Michael J Artis and Massimiliano Marcellino

CEPR Discussion Paper No. 2206

August 1999

## ABSTRACT

### Fiscal Forecasting: The Track Record of the IMF, OECD and EC\*

We analyse the relative performance of the IMF, OECD and EC in forecasting the government deficit, as a ratio to DGP, for the G7 countries. Interesting differences across countries emerge, sometimes supporting the hypothesis of an asymmetric loss function (i.e. of a preference for underprediction or overprediction) and potential benefits from forecast pooling.

JEL Classification: C53, E17, H62

Keywords: fiscal forecasting, forecast encompassing

Michael J Artis  
European University Institute  
Badia Fiesolana  
Via dei Roccettini, 9  
I-50016 San Domenico di Fiesole  
Florence  
ITALY  
Tel: (00 39) 055 468 5777/221  
Fax: (00 39) 055 4685202  
Email: artis@datacomm.iue.it

Massimiliano Marcellino  
IGIER  
Università Bocconi  
Via Salasco, 5  
20136 Milano  
ITALY  
Tel: (00 39) 02 5836 3327  
Fax: (00 39) 02 5836 3302  
Email: massimiliano.marcellino@uni-  
bocconi.it

\*This Paper is produced as part of a CEPR research network on New Approaches to the Study of Economic Fluctuations, funded by the European Commission under the Training and Mobility of Researchers Programme (Contract No ERBFMRX-CT98-0213). We are grateful to Roberto Artoni, Clive Granger, David Hendry, to participants in the (EC)<sup>2</sup> conference and in seminars at UC Davis, Bocconi and the IMF for several critical and helpful comments. The authors are also grateful to the EUI Research Council for financial support

Submitted 1 July 1999

## NON-TECHNICAL SUMMARY

The Paper is concerned with the analysis of the quality of fiscal forecasting by the major international agencies: the IMF, the OECD and the EC, focusing particularly on the European 'Big Four' countries.

A principal reason for being interested in this subject is the fact that fiscal forecasts play a key role in the operation of the Stability and Growth Pact. Should fiscal forecasts turn out to be of poor quality, the operation of the Pact would be severely compromised. Warnings might be issued where none are deserved; equally, countries might be exonerated from blame in the conduct of their fiscal policy when in fact they should not be. Moreover, in order to aim at the position of 'surplus or close to balance' in the budget over the medium term enjoined by the Pact, governments need to know what fiscal forecasting errors they should allow for. The purpose of the medium-term target is to allow countries to feel safe in relying on the 'automatic stabilizers' whilst not producing a deficit of more than 3% of GDP which would invoke the Pact's sanctions procedure. The safety margin should include an allowance for forecasting error.

In order to investigate the quality of fiscal forecasting, the Paper takes the track records of the IMF, OECD and EC in forecasting the fiscal deficit (in ratio to GDP) and performs an analysis of the errors made in the past. For the IMF's track record, for which a somewhat longer sample of data is available, the analysis covers the G-7 countries; for the OECD and EC forecasts, however, the analysis is confined to the 'Big Four' European economies. The general findings suggest that fiscal forecasting is sufficiently accurate not to make the Stability and Growth Pact an infeasible operation. Average errors are around 1%. The presence, in the track records, of some larger errors is a warning, though, that this average experience conceals a handful of bad episodes offsetting a more favourable general run of results. Much of the error can be explained in terms of output growth forecast errors, as might be expected. When output growth forecasts prove to have been too optimistic, the deficit forecast will generally prove to have been an overestimate. The probable reason is that more-buoyant-than-expected growth results in more-buoyant-than-expected tax revenues and a consequential reduction in the realized deficit relative to the forecast.

In terms of the standard diagnostic tests administered in forecast post-mortems, fiscal forecasts prove to be average performers. There is some, but not pervasive, evidence of bias and some lack of efficiency. Bias is a tendency towards producing a significant positive or negative error; a forecast can obviously be improved by taking account of known bias. Efficiency more generally describes a situation in which a forecast cannot be improved by

incorporating additional available information; for example from any tendency of the errors to 'cycle' through time. Tests are carried out to determine whether any one agency's forecasts 'encompass' another's, in the sense of explaining that other agency's forecast errors. Between the different agencies making the forecasts, no single one emerges as clearly superior or dominant.

Of course it may be that evidence based on track records for periods when fiscal forecasts were not (most of the time, anyway) focal points in policy-making may prove misleading when it comes to periods in which, as now, with the Stability and Growth Pact in operation, the salience of such forecasts is much increased. The Paper suggests that in such periods there may be pressures on published forecast releases that would not otherwise exist. Technically, this could imply that the 'loss function' applying to the forecast would no longer be (as is routinely assumed in forecast error post-mortem analysis) symmetric, but rather asymmetric. For example, a country might issue forecasts which are 'shaded' so that the published deficit ratio foreseen for next year might be less than really appears to it most likely, in order to postpone unwelcome investigation. The Paper conducts some tests to see whether there is already much evidence of this in the sample period used in the study (ending in 1996 this includes the post-Maastricht pre-EMU entry period in which deficits have had added salience). There is some slight evidence that the forecasts might have been 'contaminated' even in this period.

## 1. Introduction

In this paper we analyse the fiscal forecast record of the major international organizations - the IMF, OECD and the EC. We submit these forecasts (forecasts for the ratio of the budget deficit to output) to a variety of tests for accuracy, efficiency and unbiasedness, using data for the G-7 countries, and, more especially for the 'European 4' subgroup of the G-7. An important motivation for this exercise comes from the recognition that deficit forecasts are playing an increasing role in macroeconomic policy decisions. This has been particularly obvious in the European context where, for example, the operating procedures of the Stability and Growth Pact involve reference to forecast values of the fiscal deficit at more than one point. In an era of fiscal consolidation, this significance of prospective as well as of actual fiscal deficits is true more generally: so soon as 'sustainability' is mentioned, the forecast path of future deficits has to be added to the historical record for evaluation. It is important to enquire, therefore, into the reliability of such forecasts. There is also a more technical motivation for the work reported here. Implicit in standard forecast evaluation practice is the assumption that forecast misses exact symmetric penalties; hence a quadratic form for the loss function is routine. Recently, however, Granger (1997) has highlighted the fact that when the loss function is not of this form, the standard properties of optimal forecasts will not hold. The political context in which fiscal deficit forecasts emerge may well be one in which the costs of forecast misses are not symmetric; and, whilst the international organizations whose forecasts we examine here are under different political pressures from those which influence national governments, they are operating in a related political environment. Fiscal forecast errors may be especially sensitive ones. In this paper we use the fiscal forecast samples to evaluate whether the predictions meet the optimality criteria of nonquadratic loss functions.

In the paper we pay somewhat more attention to the set of IMF forecasts than we do to those of the OECD and the EC, for the simple reason that the

IMF sample is somewhat larger than those we were able to collect for the other two agencies. In the next section we begin with a first evaluation of the IMF forecast record. In Section 3 we compare this with the record for OECD and EC forecasts. In Section 4 we evaluate whether there is any evidence in favour of an asymmetric loss function. Section 5 concludes.

## 2. An evaluation of the IMF forecasts

In this section we analyse the IMF forecasts of gross deficit ratios for the G7 countries.<sup>1</sup> We consider both year-ahead forecasts, which are identified with those published in October of year  $t$  for  $t+1$ , and current-year forecasts, which are those published in May of year  $t$  for year  $t$ .<sup>2</sup> They are compared with first released actual data on gross deficit ratios. Such a comparison is the most interesting from a policy perspective, and further revisions of out-turn data usually do not appear to greatly affect the results of related forecasts (e.g. Artis (1988), Gallo and Marcellino (1998)).

The year ahead and current year forecast errors, defined as forecast minus actual values, are graphed in figure 1. Deficit is defined as a positive value, so that a positive forecast error corresponds to overprediction. There seems to be evidence of systematic overprediction for Japan, Italy and the UK, and underprediction for Canada. IMF fiscal forecasts for G7 countries typically have been technical exercises in which the short-term budget projections of the national authorities are adjusted for differences in macroeconomic assumptions. Thus, the presence of bias can reflect either an institutional bias or, more likely, a bias in the national authorities' forecasts that is transmitted to IMF forecasts via the technical adjustments made to official projections. For example, in the case of Italy the

---

<sup>1</sup>All the calculations were performed with GiveWin 9.1, see Doornik and Hendry (1997) and, e.g., Marcellino (1998a).

<sup>2</sup>These are paradigm definitions. In practice, there have been variations in the Forecast timetable. See Artis (1988) for a complete list of forecast dates corresponding to the 'year ahead' and 'current year' distinction used here.

improved performance after 1988 is probably related to the approval of Law 362, that basically made forecast values the target for economic policy. A third possible source of bias is the conditioning on actual as opposed to announced policies. The resulting forecasts can be expected to miss the impact effects of the changes in fiscal policy. Unfortunately, the original official projections used at the Fund were not available to us, so that it is difficult to distinguish among these three possible sources of bias.

The mean error (ME), mean absolute error (MAE) and root mean squared error (RMSE) are reported in Table 1. The worst performance both on MAE and on RMSE is again for Japan, Italy and UK, while the best one is for France, Germany and the US. The additional information present in the current year forecasts is useful in decreasing both the MAE and the RMSE for all the countries. We also run a Chow test for constancy of the MSE, by splitting the sample and constructing the ratio of the MSEs in the two subsamples which, under the additional hypothesis of uncorrelated normal forecast errors, is distributed as  $F(k, j)$  where  $k$  and  $j$  are the number of observations in the two subsamples. Constancy is always accepted, except for Italy, Japan and UK, when current year forecasts are used. Similar results are obtained with Hansen's (1992) test.

The mean error is usually smaller than one point, and the absolute error is only slightly larger. This looks like a good performance, but the unit of measurement matters. When the same errors are expressed as percentages of actual values, the typical range is  $\pm 50\%$ . It is not always clear what is the right scale to be used from a policy perspective. Following standard practice in the forecast evaluation literature we will continue measuring errors as differences of forecast and actual values.

We now formally analyse the unbiasedness and weak efficiency of the forecasts. It has become conventional to claim that forecasts are “unbiased” when  $\alpha_0 = 0$ ,  $\alpha_1 = 1$  in the regression

$$a_h = \alpha_0 + \alpha_1 f_h + u_h, \tag{2.1}$$



where  $a$  are the actual values,  $f$  are the forecasts, and  $u$  is an error term that under the null hypothesis of unbiasedness coincides with the forecast error (see, e.g., Clements and Hendry (1998, Ch.3)), and should then be free of serial correlation. As Holden and Peel (1990) showed, (2.1) is sufficient but not necessary for unbiasedness; rather, unbiasedness should be tested for as the condition  $\beta_0 = 0$  in the regression

$$e_h = \beta_0 + v_h, \quad (2.2)$$

where  $e$  are the forecast errors, and  $v$  the demeaned forecast errors. Weak efficiency also requires the forecast errors to be uncorrelated in time (see, e.g., Clements and Hendry (1998, Ch.3)).

Table 2 reports, for year ahead and current year forecasts, the t-tests for  $\beta_0 = 0$ ,  $\alpha_0 = 0$ ,  $\alpha_1 = 1$  ( $T0$ ,  $T1$ ,  $T2$ ), and a Lagrange Multiplier test ( $C$ ) for lack of up to second order autocorrelation in the forecast errors, which is distributed as  $F(2, H - 2)$ , where  $H$  is the number of available forecasts. Weak efficiency is accepted for all countries, except Japan, and the UK for current year forecasts. According to  $T0$ , unbiasedness is rejected only for Italy, Japan and Canada for year ahead forecasts, and for Japan and UK for current year forecasts, with borderline values for Italy and Canada. Such an outcome is coherent with the graphical and descriptive evidence provided earlier. Somewhat different results are obtained from  $T1$  and  $T2$ ; they reject unbiasedness more often. This can be a small sample issue, but it can also be due to the different null hypothesis of the tests.

The sample of observations on hand is relatively small and it is tempting to consider whether the country data sets could be pooled. The last row of Table 2 presents the results from the pooled regressions, which this time reject weak efficiency. We also checked for constancy of the parameters of the equations (2.1) and (2.2) by means of the Hansen (1992) tests. It was almost always accepted in the single country regressions, but rejected in the pooled regressions, suggesting their inappropriateness for our data.

It is now interesting to compare the IMF forecasts with those from two naive models, a random walk without drift (which implies that the optimal forecast of  $a_h$  is  $a_{h-1}$ ), and a deterministic trend model, see also Marcellino (1998b) for additional results.<sup>3</sup> The first two columns of Table 3 report the Theil statistics for the two models (*TH1* and *TH2* respectively), which in this case simply coincide with the ratio of the RMSE for the IMF forecasts to that of the naive forecasts. Hence, a value of the statistic smaller than one indicates that the IMF forecasts outperform the naive forecasts.

The results are rather surprising. For year ahead forecasts a naive model achieves a smaller RMSE than the IMF for all countries, with the exception of the US. The performance improves for current year forecasts, when only the forecasts for Italy, Germany and Japan can be beaten.

Comparing the models using a deterministic criterion such as the Theil statistic can be misleading because the differences in the chosen criterion among the models may not be significant from a statistical point of view. Therefore, Diebold and Mariano (1995) proposed to base the comparison on the statistic

$$DM = H^{-1/2} \frac{\sum_{j=1}^H d_j}{\sigma_d} \rightarrow N(0, 1), \quad (2.3)$$

where

$$d_j = g(e_{1j}) - g(e_{2j}),$$

$g$  is the loss function of interest, e.g. the quadratic loss  $g(e) = e^2$  or the absolute loss  $g(e) = |e|$ ,  $e_1$  and  $e_2$  are the errors from the two competing forecasts, and  $\sigma_d$  is the standard deviation of  $d$ . Notice that if  $DM$  is positive the loss associated with the first model is larger than that for the second one. Diebold and Mariano (1995) suggested to estimate  $\sigma_d$  with spectral based techniques but, given the

---

<sup>3</sup>Fewer forecast errors are available for the deterministic forecasts because the parameters of the model have to be estimated. We have regressed the first five actual values on a constant and a trend, and used the estimated parameters to forecast the sixth observation. Both the estimates and the forecasts are then recursively updated.

small sample available and the non correlation of  $d_j$  for almost all cases, we use the standard formula

$$\hat{\sigma}_d^2 = \frac{1}{H-1} \sum_{j=1}^H (d_j - \frac{1}{H} \sum_{i=1}^H d_i)^2.$$

The results are reported in Table 3 for both absolute loss (*DMA*), and quadratic loss (*DMS*). For year ahead forecasts, the trend model outperforms the IMF for Germany, while the random walk is better for Japan and UK, and for Italy in the case of quadratic loss. For current year forecasts the random walk is better only for Japan, while the IMF forecasts are better than the trend forecasts for France. No other loss function differentials are statistically significant from zero, even if their signs are often positive, in particular for year ahead forecasts.<sup>4</sup>

A source of the problems of some of the IMF deficit forecasts could be the presence of structural breaks over the forecast period due to unmodelled changes in economic policy. This could also explain the good forecasting performance of the naive models, because of their robustness to breaks, see e.g. Clements and Hendry (1997b). A possible remedy in this case is “intercept correction”. The term “intercept correction” comes from the practice of those forecasters who use formal econometric models for forecasting, of absorbing into a correction of the constant terms of the model’s equations, persistent errors evident from their recent tracking behaviour. Here this method is implemented by adding the lagged forecast error to the actual forecast from the model, i.e.

$$icf_h = f_h + e_{h-1}, \tag{2.4}$$

see e.g. Clements and Hendry (1997a). In practice this or other types of adjustment can be expected to be already present in forecasts from official agencies, but it seemed nevertheless worthwhile experimenting with this adjustment.<sup>5</sup>

---

<sup>4</sup>Notice that in the calculation of the Theil statistics the RMSEs are calculated over the longest available period for each model, while for the DM test the same sample period is used for both models. This explains why for Japan *DMS2* is negative even if *TH2* is larger than one.

<sup>5</sup>In (2.4) we are adding the forecast error to the forecast with a weight of one. As an

From Table 4, in the case of year ahead forecasts, it is seen that there is a marked deterioration in the weak efficiency property of the intercept corrected forecast error. Actually, if the original forecast errors are uncorrelated, we are simply adding an MA(1) term to them because

$$ice_h = a_h - icf_h = e_h - e_{h-1}. \quad (2.5)$$

There also no major changes in the unbiasedness properties of the forecasts, which remain biased for Canada, Italy and Japan. The results are better for current year forecasts, when all the forecast can be considered as unbiased after intercept correction according to the  $T0$  test, even if some correlation is introduced in the forecast errors for Canada and France.

The final question that we address in this section is whether errors in forecasting the deficit to gdp ratio can be explained by wrong forecasts of other relevant macroeconomic variables. A natural candidate is gdp growth. Unexpected growth increases the level of gdp and decreases that of the deficit (reflecting the operation of the automatic stabilizers); hence, growth forecast errors should be negatively correlated with deficit to gdp forecast errors. A negative effect of inflation on deficit forecast errors is also possible when the tax indexation system is not perfect.

This interpretation appears to be supported by the data. Table 5 reports results from a regression of deficit to gdp forecast errors on growth and inflation forecast errors ( $eg$  and  $e\pi$ ). For year ahead forecasts,  $eg$  is significant and negative for Canada, France, Japan and UK. A significant and negative coefficient for  $e\pi$  is found only for Canada and the UK. A similar pattern emerges also with current year forecasts, but the coefficients are significant only for Canada and the UK.

In summary, the IMF forecasts are weakly efficient for the G7 countries with the exception of Japan, but they seem to be slightly upward biased for Italy, Japan

---

alternative, the weight can be determined optimally (in the sense of minimizing the MSE) as the coefficient of  $e_{t-1}$  when  $e_{t-1}$  is included as a regressor in (2.1). Optimal determination of the weight does not affect the subsequent analysis; these results are available upon request.

and UK, and downward biased for Canada. Intercept corrections can improve the performance on the bias criterion of current year forecasts. In a comparison with simple random walk and trend forecasts, the IMF forecasts often lead to higher quadratic and absolute loss, and the difference is statistically significant for a few countries in the case of year ahead forecasts.

### **3. OECD and EC forecasts**

In this section we analyse the OECD and EC forecasts for gross deficit ratios, and compare them with those from the IMF. We focus on the four European countries in the G7, namely, France, Germany, Italy and UK, both to reduce the volume of results to be presented and because these are the most interesting countries to analyse in the light of the deficit requirements of the Maastricht Treaty and of the Growth and Stability Pact. OECD deficit forecasts for all EU countries are analyzed in Artis and Marcellino (1998).

For the OECD, year-ahead forecasts are identified with those published in December of year  $t$  for  $t+1$ , and current-year forecasts are those published in June of year  $t$  for year  $t$ . For the EC, year-ahead forecasts are those published in the Autumn (October) of year  $t$  for  $t+1$ , and current-year forecasts those released in the Spring (April) of year  $t$  for year  $t$ . It should also be recalled that there are some minor differences across the agencies in the definition of the deficit, so that in the construction of the forecast errors, defined as forecast minus actual values, we use the actual (first released) values from the proper agency. The year ahead and current year forecast errors are graphed in figures 2 and 3 for IMF, OECD and EC.

From Table 6, the mean forecast error is rather low, in practice always smaller than 0.5 points, for both the OECD and the EC. The MAE and RMSE are smaller than those from IMF forecasts for Italy and the UK, with the EC doing better than the OECD, but over a different sample period. The IMF performs better for France and Germany, with the exception of current year forecasts for Germany

for which the EC achieves slightly smaller values. As for the IMF, the larger information set exploited in current year forecasts is useful in decreasing both the MAE and the RMSE.

The results from this first comparison should be interpreted with care because the sample sizes are different, in particular that for the EC is rather short, and the differences in the loss function may not be statistically different from zero. To address these two issues, we now apply the Diebold and Mariano (1995) test, with an absolute and quadratic loss function, using for each comparison the common longest available sample period. We recall that if the statistic is positive the loss associated with the first set of forecasts is larger than that for the second one. The results are summarised in Table 7.

For year ahead forecasts, the IMF does significantly better than the OECD and the EC for France, while for Germany the difference is not significantly different from zero, even if the loss from the IMF forecasts is still the lowest. The EC performs best for the UK and Italy, but only in the former case is the loss differential with the OECD and IMF statistically different from zero. For current year forecasts, the improvement in the performance of all the three agencies makes it harder to distinguish among them; actually just two loss differentials are significantly different from zero out of 24. But the aforementioned pattern is overall still satisfied, with the IMF yielding a smaller loss for France, and the EC for Italy, while the performance of OECD and EC for the UK is rather similar.

As far as the weak efficiency and unbiasedness of the forecasts are concerned, the former hypothesis is always accepted for both OECD and EC forecasts; the tests are reported in Table 8. Such an outcome is similar to what we found for the IMF, when weak efficiency was only rejected for year ahead forecasts for the UK. Unbiasedness is also always accepted for year ahead forecasts when the  $T0$  test is used, see Table 8, while it is rejected by the  $T1$  and  $T2$  tests for Germany, and also for Italy in the case of OECD forecasts. For current year forecasts, unbiasedness is rejected for Germany (and Italy using  $T1$  and  $T2$ ) in the case of

OECD, and for the UK (and Italy using  $T1$  and  $T2$ ) in the case of the EC. For the IMF, relying on the  $T0$  test, we found that year ahead forecasts for Italy and current year forecasts for the UK were biased.

The results so far seem to indicate that different agencies can do better for different countries, the IMF for France and Germany, the OECD and the EC for Italy and the UK. We now consider this issue in further detail by analysing whether the forecast errors from one agency can be explained by the forecasts from another, i.e. we look for (MSFE) forecast encompassing, see e.g. Chong and Hendry (1986), Lu and Mizon (1991), Ericsson (1992), and Marcellino (1998). The basic forecast encompassing regression is

$$e_{i,h} = \gamma_0 + \gamma_1 f_{j,h} + u_{i,h}, \quad (3.1)$$

where  $i = IMF, OECD, EC$ ,  $i \neq j$ ,  $h = 1, \dots, H$ , and the statistic of interest is the t-test for  $\gamma_1 = 0$ . Yet, with (3.1) we can only make bivariate comparisons. In order to evaluate whether the forecasts from two agencies are useful for explaining the forecast errors of the third one, we can extend (3.1) to

$$e_{i,h} = \delta_0 + \delta_1 f_{j,h} + \delta_2 f_{k,h} + u_{i,h}, \quad (3.2)$$

where  $k \neq i, j$ , and test for  $\delta_1 = 0$ ,  $\delta_2 = 0$  either jointly with an F-test or separately with two t-tests.

The regression equation (3.2) can be rewritten as

$$a_{i,h} = \delta_0 + f_{i,h} + \delta_1 f_{j,h} + \delta_2 f_{k,h} + u_{i,h}. \quad (3.3)$$

If we remove the hypothesis that the coefficient of  $f_{i,h}$  is equal to one, we get a third version of the forecast encompassing regression

$$a_{i,h} = \zeta_0 + \zeta_1 f_{i,h} + \zeta_2 f_{j,h} + \zeta_3 f_{k,h} + u_{i,h}, \quad (3.4)$$

or

$$e_{i,h} = \zeta_0 + (\zeta_1 - 1) f_{i,h} + \zeta_2 f_{j,h} + \zeta_3 f_{k,h} + u_{i,h}. \quad (3.5)$$

The hypothesis of interest in (3.5) is  $\zeta_2 = 0$ ,  $\zeta_3 = 0$ , which can again be tested by either two t-tests or an F-test.

If a set of forecasts encompasses all the others, in the sense of explaining their related forecast errors, without being encompassed, it is a suitable candidate as the preferred forecast. Yet, this seldom happens, and a more common situation is that of mutual encompassing or lack of encompassing. In the former case, the forecasts can be evaluated on the basis of other criteria, such as the MSFE or MAE comparisons that we considered before. In the latter case, there is evidence that the specification of the underlying models should be somewhat improved. Yet, this is seldom feasible in the case of large macromodels, and a more usual procedure is to combine the forecasts themselves into one that has better properties than each of them separately. Actually, (3.4) can be also viewed as a standard forecast pooling regression, see e.g. Granger and Newbold (1986, Ch. 9). The estimated values of  $\zeta_0$ ,  $\zeta_1$ ,  $\zeta_2$  and  $\zeta_3$  represent the optimal weights for the pooled forecast

$$pf_{i,h} = w_0 + w_1 f_{i,h} + w_2 f_{j,h} + w_3 f_{k,h},$$

which is the one that minimizes the MSFE.

Tables 9 and 10 contain the t-tests for  $\gamma_1 = 0$  in (3.1), and  $\delta_1 = 0$ ,  $\delta_2 = 0$  in (3.2). Starting with year ahead forecasts, it turns out that the IMF forecast errors can be explained by the OECD and EC forecasts for Germany, Italy and UK; the OECD forecast errors can instead be explained by either IMF forecasts or EC forecasts or both for all the four countries; the IMF and OECD forecasts are instead statistically significant for explaining the EC forecast errors only for Germany. A similar pattern emerges also for current year forecasts, the major change being that the forecast errors for the UK by any one agency cannot be explained by the forecasts from the other ones. These results indicate a comparative advantage of the EC forecasts, and suggest the potential usefulness of forecast pooling, the final issue that we analyse.

When the hypothesis that the coefficient of  $f_{i,h}$  in (3.2) is equal to zero is



relaxed and (3.5) is used as the regression equation, fewer forecasts are significant in explaining the forecast errors from one agency, see Table 11. This can be due either to having relaxed an improper assumption, or to the fact that the sample size is rather short and the regressors collinear, which can inflate the standard error of the estimators and bias the t-tests towards accepting their null hypothesis. The former possibility seems to be the most plausible, because when we excluded the EC forecasts from the comparison in order to increase the sample size and decrease the problem of collinearity, the results did not change substantially. From Table 11, it appears that the major potential gains from forecast combination are now for France in the case of year ahead forecasts, and Italy for the OECD when using current year forecasts.

Summarizing, so far in this section we have analysed the OECD and EC deficit forecasts and compared them with the IMF forecasts using several criteria. On the basis of standard MSFE and MAE comparisons, the IMF seems to perform better for France and Germany, and the EC and the OECD for Italy and UK. Yet, often the difference in the loss functions is not statistically different from zero. The performance in terms of unbiasedness and efficiency is rather similar across the three statistical agencies. When using standard forecast encompassing regressions there seems to be an advantage of the EC forecasts, whose related forecast errors cannot be explained by other forecasts. But when the less restrictive forecast pooling version of the encompassing regression is adopted, this advantage disappears, as well as more generally the scope for forecast pooling, even if there still seem to be cases where it could be useful, e.g. for France.

In appraising the comparative performance of these forecasts, it must be borne in mind that they are not made at precisely the same times, and that the information set available to the forecasters is not identical. This also opens up the possibility of “herding behaviour”, i.e., the first published forecasts can exert a direct influence on later released competing forecasts, see e.g. Trueman (1994). Such behaviour can be rational in a highly uncertain environment, when the goal

is a good performance in terms of the public's assessment of the organization's forecasting ability. If this were the case, we would expect the forecast errors of the IMF, OECD and EC to be highly correlated, and more so for the year ahead forecasts (when uncertainty is higher). Actually, this pattern appears to emerge in Table 12, in particular for France and UK.

Of course, other explanations for these results are possible, e.g. a different interpretation during the current year of the conditionality of the forecasts. The forecasts we consider here are all issued with the statement that they are "based on present policies", a phrase which means that credible statements, e.g. of government expenditure plans already authorized by the respective legislatures will certainly be reflected in the forecasts whilst statements of ambitious targets, even from influential politicians, will be discounted (in effect) as wishful thinking. In a period when the achievement of fiscal criteria has been so strongly enjoined by inter-governmental treaty and peer group pressure, distinctions of this type may become more controversial. A vivid illustration was provided by the EC's November 1996 forecast for Germany's deficit ratio which reflected Germany's announced *target*; EC forecasters argued that the target itself represented "present policies" and in the political circumstances of the time was not to be treated as falling on the "wishful thinking" side of the dividing line.

#### **4. On the loss function**

An assumption that we have maintained so far is that the loss function of the statistical agencies is quadratic. If this is not the case but the loss is a generic function,  $c(e)$ , most of the conventional properties of the forecast errors from optimal forecasts are no longer valid. In particular, they are no longer necessarily unbiased and uncorrelated in time, and  $\alpha_0$  and  $\alpha_1$  in (2.1) can be different from zero and one, see e.g. Granger (1997). We now study whether a non linear loss function can be the cause of the rejection of some of these hypotheses for the IMF forecasts (see Table 2).

Granger (1997) shows that the aforementioned properties will hold for the first derivative of the loss function,  $c'(e)$ , namely:

- i)  $c'(e)$  is unbiased;
- ii)  $c'(e)$  is uncorrelated in time;
- iii)  $\delta_0 = 0$  and  $\delta_1 = 0$  in the regression

$$c'(e_h) = \delta_0 + \delta_1 f_h + u_h. \quad (4.1)$$

Most of the literature on non quadratic loss functions focused on the derivation of the optimal forecasts for particular choices of  $c(e)$ , see e.g. Christoffersen and Diebold (1994). We follow a different route, we assume that the forecasts from the IMF are optimal for (the expected value of)  $c(e)$ , and check whether the properties i), ii) and iii) are satisfied. We consider two rather standard choices for  $c(e)$ , the asymmetric quadratic function

$$c(e) = \begin{cases} ae^2 & e \leq 0 \\ be^2 & e > 0 \end{cases}, \quad (4.2)$$

and the linex function

$$c(e) = (\exp ce) - ce - 1. \quad (4.3)$$

When  $b/a > 1$  or  $c > 0$  ( $b/a < 1$  or  $c < 0$ ) there is a higher loss from positive (negative) forecast errors, i.e., from overprediction (underprediction). For  $b/a$  close to one or  $c$  close to zero the loss functions can be well approximated by a quadratic function.

We now have to choose the value of the parameters of  $c(e)$ . We select  $b/a$  and  $c$  so that the empirical counterpart of condition i) is satisfied, i.e. we adopt the values of  $b/a$  and  $c$  such that the sample mean of  $c'(e)$  is zero. They are reported in Table 13. The largest values of  $b/a$  are for Canada and France (underprediction is preferred), the smallest ones for Italy and Japan, and UK for current year

forecasts (overprediction is preferred). This is a consequence of the values of the mean forecast error which is negative for the first set of countries and positive for the second one, see Table 1.

We can now verify whether conditions ii) and iii) hold or not. From Table 13, the non correlation of  $c'(e)$  is always accepted for current year forecasts and it is rejected for year ahead forecasts only for Japan and UK. Yet, this represents an improvement with respect to the quadratic loss results in Table 2 only for current year forecasts for Japan.

The hypothesis that  $\delta_0 = 0$  and  $\delta_1 = 0$  (condition iii) is instead accepted only for Canada, France, and UK for year ahead forecasts, which though represents an improvement because it was rejected before for the first two countries, compare Table 13 and 2. For current year forecasts the hypothesis of interest is only accepted for Canada and UK, as for the quadratic function case.

In summary, the assumption that the IMF forecasts are optimal for a non quadratic loss function appears reasonable only for Canada and France, where overprediction seems to be more problematic than underprediction, even if such a conclusion can depend on our hypotheses on the loss function and its parameters.

## 5. Conclusions

In this paper we set out to review the accuracy of short term forecasts of budget deficit ratios by the three major international agencies making such forecasts - the IMF, OECD and EC. Each agency forecasts twice a year and we compared a short term and a slightly longer term forecast for each of these agencies. A principal motivation for doing so was the evidence that such forecasts have come to play a more central role, especially in Europe, in macroeconomic policy adjustment than in earlier decades.

It is common in forecasting post-mortems to encounter the finding that 'balance' variables - the current account of the balance of payments or the budget deficit - are by far the least well forecast values in the set of leading macro vari-

ables of interest. In our case, however, we focussed on the forecast of deficits expressed as a ratio to GDP. This evidently takes care of the worst of the problem of forecasting the actual balance itself. Mean errors are usually below 0.5 points. And, whilst naive predictors seem to perform well relative to the forecasts under examination, the differences are not in general statistically significant. The relative accuracy of the forecasts from the different agencies was also examined; no single agency is 'best' for all countries, but there seems to be some 'specialization' - the EC seemed to perform particularly well for Italy, for example. We noted that these differences might be partially explained by reference to differences in the timing of forecasts (hence, available information sets) and also by differences in the interpretation of the conditionality ("present policies") of the forecasts. As the deficit forecasts have been especially sensitive politically, we thought it possible that the symmetric loss function normally applied in forecast post-mortems might be inappropriate; but on careful examination we found it difficult to sustain this proposition generally.

## References

- [1] Artis, M. (1988), "How accurate is the World Economic Outlook? A post-mortem on short term forecasting at the International Monetary Fund", *IMF Staff Studies Behind the World Economic Outlook*, 1-49.
- [2] Artis, M. (1997), "How accurate are the IMF's short term forecasts? Another examination of the World Economic Outlook", *IMF Staff Studies Behind the World Economic Outlook*.
- [3] Artis, M. and Marcellino, M., (1998), "Fiscal Solvency and Fiscal Forecasting in Europe", C.E.P.R. W.P. 1836.
- [4] Chong, Y.Y. and Hendry, D.F. (1986), "Econometric evaluation of linear macro-economic models", *Review of Economic Studies*, 53, 671-690.

- [5] Christoffersen, P.F. and Diebold, F.X. (1994), "Optimal prediction under asymmetric loss", N.B.E.R. Technical W.P. 167.
- [6] Clements, M.P. and Hendry, D.F. (1997a), "Intercept corrections and structural change", *Journal of Applied Econometrics*, 11, 475-494.
- [7] Clements, M.P. and Hendry, D.F. (1997b), *The Zeuthen Lectures on Forecasting*, MIT Press, *forthcoming*.
- [8] Clements, M.P. and Hendry, D.F. (1998), *Forecasting Economic Time Series*, Cambridge: Cambridge University Press.
- [9] Diebold, F.X. and Mariano, R.S. (1995), "Comparing predictive accuracy", *Journal of Business and Economic Statistics*, 13, 253-263.
- [10] Doornik, J.A. and Hendry, D.F. (1997), *GiveWin - An interface to empirical modelling*, London: International Thomson Business Press.
- [11] Ericsson, N.R. (1992), "Parameter constancy, mean square forecast errors, and measuring forecast performance: An exposition, extensions, and illustration", *Journal of Policy Modeling*, 14, 465-495.
- [12] Gallo, G.M. and Marcellino, M. (1998), "Ex-post and ex-ante forecasting with provisional data", *Journal of Forecasting*, *forthcoming*.
- [13] Granger, C.W.J. (1997), "Outline of forecast theory using generalized cost functions", mimeo, UCSD.
- [14] Granger, C.W.J. and Newbold, P. (1986), *Forecasting economic time series*, 2nd edn., New York: Academic Press.
- [15] Hansen, B.E. (1992), "Testing for parameter instability in linear models", *Journal of Policy Modeling*, 14, 517-533.
- [16] Holden, K. and Peel, D.A. (1990), "On testing for unbiasedness and efficiency of forecasts", *The Manchester School*, 58(2), 120-127.

- [17] Lu, M. and Mizon, G.E. (1991), "Forecast encompassing and model evaluation", in Hackl, P. and Westlund, A.H. (eds.), *Economic structural change, Analysis and Forecasting*, 123-138, Berlin: Springer-Verlag.
- [18] Marcellino, M. (1998a), "PcGive Professional 9.1 for Windows: A Review", *International Journal of Forecasting*, forthcoming.
- [19] Marcellino, M. (1998b), "Further results on MSFE encompassing tests", I.G.I.E.R. WP 143.
- [20] Trueman, B. (1994), "Analyst forecasts and herding behavior", *Review of Financial Studies*, 7, 97-124.

## Figures

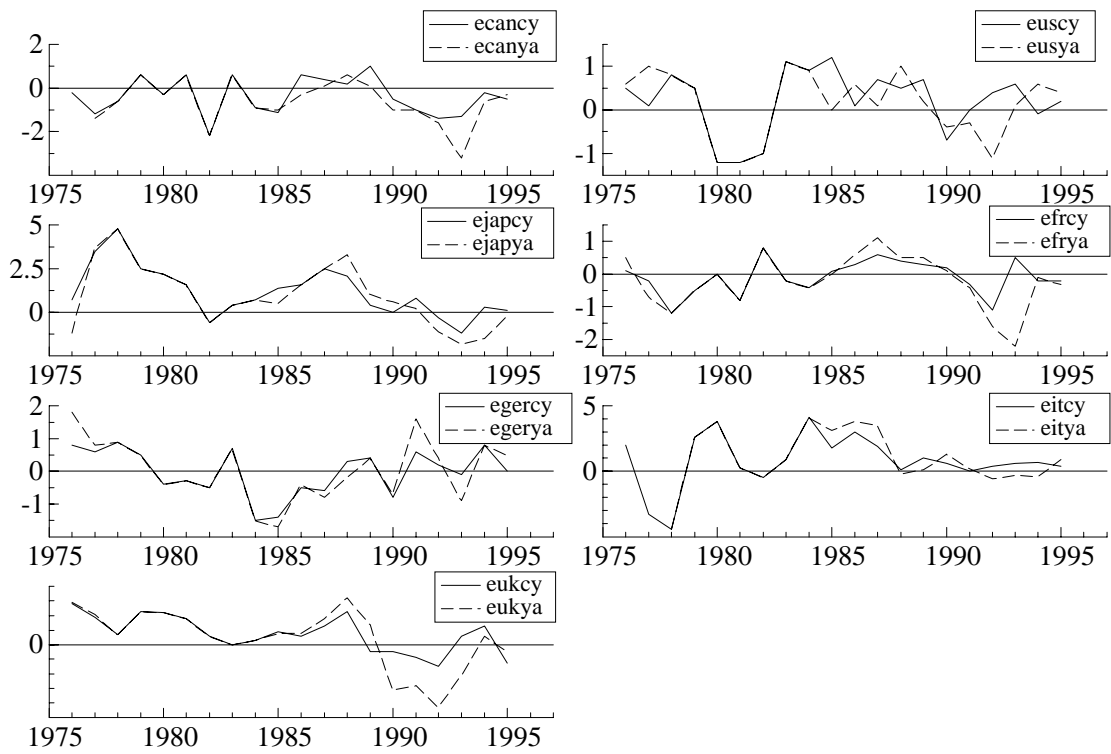


Figure 1. IMF - Year ahead and current year forecast errors.



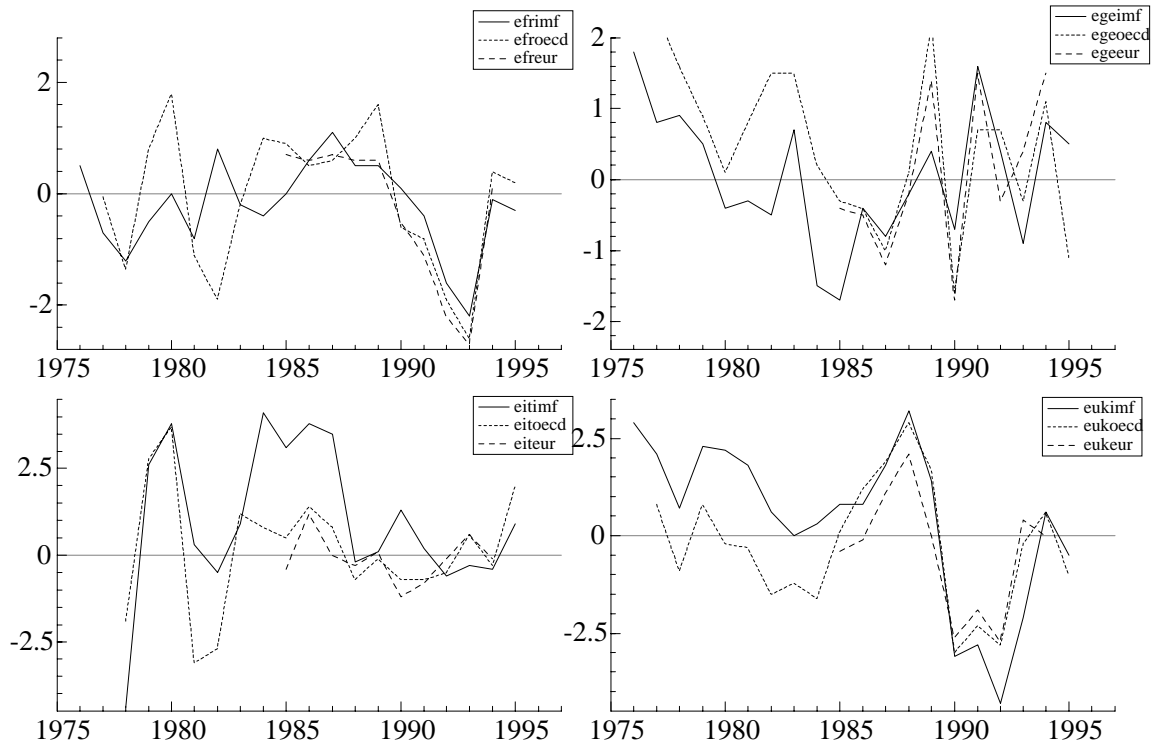


Figure 2. Year ahead forecast errors - IMF, OECD, EC.

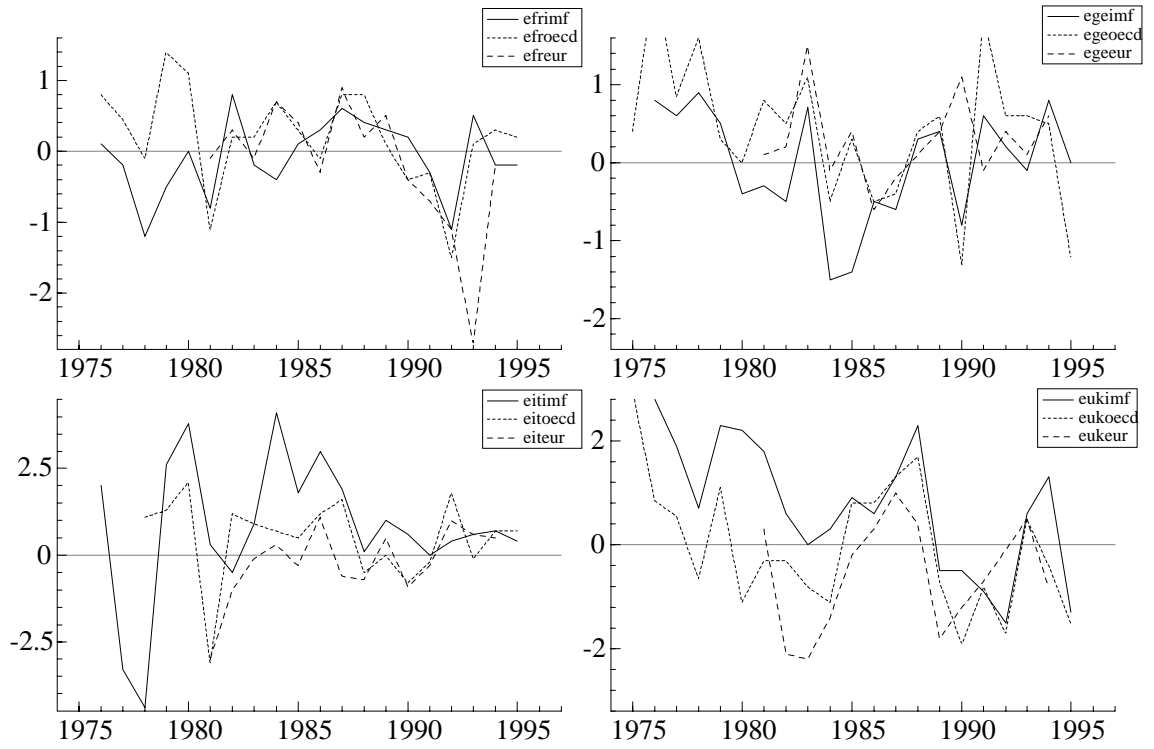


Figure 3. Current year forecast errors - IMF, OECD, EC

## Tables

Table 1: IMF Forecasts - Descriptive Statistics

Country	<i>MEY</i>	<i>MAEY</i>	<i>RMSEY</i>	Sample	<i>MEC</i>	<i>MAEC</i>	<i>RMSEC</i>	Sample
Canada	-0.62	0.89	1.17	77-95	-0.37	0.77	0.91	76-95
France	-0.22	0.63	0.83	76-95	-0.09	0.42	0.53	76-95
Germany	0.05	0.79	0.83	76-95	-0.02	0.60	0.70	76-95
Italy	0.99	1.67	2.28	78-95	0.80	1.62	2.14	76-95
Japan	0.96	1.60	2.01	76-95	1.18	1.39	1.84	76-95
U.K.	0.44	1.72	2.06	76-95	0.75	1.22	1.44	76-95
U.S.	0.14	0.66	0.76	76-95	0.21	0.63	0.73	76-95

The suffices Y and C refer, respectively, to year ahead and current year forecasts.

ME, MAE, and RMSE are the mean, mean absolute, and root mean square errors.

Table 2: IMF Forecasts - Weak Efficiency and Unbiasedness Tests

Country	$T0Y$	$CY$	$T1Y$	$T2Y$	$T0C$	$CC$	$T1C$	$T2C$
Canada	-2.66**	0.11	-2.78**	-1.96*	-1.93	0.13	-1.65	-1.11
France	-1.17	1.83	-2.26*	-1.90	-0.75	0.01	-2.15*	-2.01*
Germany	0.23	0.65	-5.83**	-6.71**	-0.09	0.60	-4.31**	-4.57**
Italy	1.98*	0.46	-4.22**	-5.02**	1.75	3.26	-4.27**	-4.99**
Japan	2.36*	4.40*	-1.76	-4.01**	3.60**	5.73*	-1.19	-4.27**
U.K.	0.94	7.50**	-0.93	-1.44	2.64**	1.82	0.11	-1.68
U.S.	0.78	2.60	-2.20*	-2.70**	1.26	0.24	-2.90**	-3.82**
Pool	1.86	23.8**	-3.75**	-6.70**	3.47**	14.01**	-2.61*	-6.54**

\* and \*\* indicate significance at 5% and 1% levels.

The suffices Y and C refer to year ahead and current year forecasts.

T0 is the (t-) test for zero mean forecast errors ( $\beta_0 = 0$  in (2.2)).

C is the (LM) test for uncorrelated forecast errors ( $v_h$  in (2.2)).

T1 and T2 are the (t-) tests in the weak efficiency regression ( $\alpha_0 = 0$  and  $\alpha_1 = 1$  in (2.1)).

Table 3: IMF Forecasts - Comparison with naive predictors

Country	Year Ahead					
	<i>TH1</i>	<i>TH2</i>	<i>DMA1</i>	<i>DMS1</i>	<i>DMA2</i>	<i>DMS2</i>
Canada	0.98	1.13	-0.10	-0.07	0.16	0.48
France	1.24	0.96	1.31	1.28	-0.69	-0.10
Germany	1.46	2.28	1.81	1.46	3.01**	2.42*
Italy	1.51	1.96	1.65	2.07*	0.74	1.49
Japan	2.00	1.63	3.05**	2.49*	0.86	0.42
U.K.	1.29	0.92	2.85**	2.23*	-0.05	-0.60
U.S.	0.83	0.77	-0.88	-0.98	-1.32	-1.77

Country	Current Year					
	<i>TH1</i>	<i>TH2</i>	<i>DMA1</i>	<i>DMS1</i>	<i>DMA2</i>	<i>DMS2</i>
Canada	0.77	0.88	-0.30	-0.89	-0.25	-0.41
France	0.79	0.62	-0.75	-1.07	-2.22*	-2.24*
Germany	1.11	1.74	0.77	0.56	1.97*	1.84
Italy	1.42	1.84	1.56	1.85	0.16	0.81
Japan	1.83	1.49	2.25*	2.05*	0.10	-0.24
U.K.	0.90	0.64	0.15	-0.63	-1.56	-1.62
U.S.	0.80	0.75	-0.98	-1.07	-1.06	-1.55

\* and \*\* indicate significance at 5% and 1% levels.

The suffices 1 and 2 refer to random walk and trend forecast comparisons.

Th is the Theil (RMSE) ratio

DMA and DMS are the Diebold-Mariano tests in (2.3) for absolute and quadratic loss.

Table 4: IMF Forecasts - Intercept Corrections

Country	Year Ahead				Current Year			
	<i>T0FF</i>	<i>CFE</i>	<i>T1FF</i>	<i>T2FF</i>	<i>T0FF</i>	<i>CFE</i>	<i>T1FF</i>	<i>T2FF</i>
Canada	-3.41*	5.83**	-5.07**	-3.51**	-0.05	8.57**	-3.65**	-3.88**
France	-1.44	11.9**	-4.15**	-3.7**	-0.09	5.19*	-2.86**	-3.22**
Germany	-0.05	3.73*	-8.33**	-10.9**	-0.21	4.36	-5.56**	-5.92**
Italy	3.04**	2.26	-6.21**	-8.57**	-0.14	0.01	-6.34**	-6.53**
Japan	2.82**	16.0**	-3.01**	-8.20**	-0.10	0.01	-2.49*	-2.89**
U.K.	0.90	17.9**	-3.27**	-5.13**	-0.75	0.41	-2.72**	-3.00**
U.S.	0.77	17.0**	-2.99**	-3.65**	-0.08	0.62	-4.21**	-4.64**

\* and \*\* indicate significance at 5% and 1% levels.

The suffix FF refers to statistics calculated with the intercept corrected forecasts (see (2.4)).

See the notes to Table 2 for a description of the tests.

Table 5: The role of growth and inflation forecast errors

Country	$egY$	$e\pi Y$	$egC$	$e\pi C$
Canada	-0.32**	-0.33**	-0.43**	-0.19
France	-0.39**	0.13	0.08	0.22
Germany	-0.10	-0.18	-0.25	-0.26
Italy	-0.41	0.05	-0.53	0.31
Japan	-0.58*	-0.02	-0.21	-0.01
U.K.	-0.95*	-0.52*	-0.71**	-0.53**
U.S.	-0.11	0.04	-0.13	-0.44

\* and \*\* indicate significance at 5% and 1% levels.

The suffices Y and C refer to year ahead and current year forecasts.

Table 6: Descriptive Statistics

OECD								
Country	<i>MEY</i>	<i>MAEY</i>	<i>RMSEY</i>	Sample	<i>MEC</i>	<i>MAEC</i>	<i>RMSEC</i>	Sample
France	-0.09	1.02	1.22	77-95	0.20	0.55	0.70	76-95
Germany	0.48	0.97	1.18	77-95	0.42	0.80	0.98	75-95
Italy	0.17	1.36	1.72	78-95	0.49	0.98	1.24	78-95
U.K.	-0.26	1.32	1.60	77-95	-0.03	1.04	1.22	75-95

EC								
Country	<i>MEY</i>	<i>MAEY</i>	<i>RMSEY</i>	Sample	<i>MEC</i>	<i>MAEC</i>	<i>RMSEC</i>	Sample
France	-0.32	0.98	1.25	85-94	-0.14	0.64	0.90	81-94
Germany	0.05	0.91	1.07	85-94	0.24	0.55	0.73	81-94
Italy	-0.10	0.48	0.64	85-94	-0.19	0.76	1.02	81-94
U.K.	-0.41	1.13	1.54	85-94	-0.56	0.97	1.20	81-94

The suffices Y and C refer, respectively, to year ahead and current year forecasts.

ME, MAE, and RMSE are the mean, mean absolute, and root mean square errors.



Table 7: Forecast Comparison - Diebold Mariano Tests

Country	Year Ahead						"Winner"
	<i>DMAIO</i>	<i>DMSIO</i>	<i>DMAIE</i>	<i>DMSIE</i>	<i>DMAOE</i>	<i>DMSOE</i>	
France	-2.92**	-3.06**	-2.34*	-1.85	0.88	0.29	IMF
Germany	-1.18	-1.44	-0.54	-0.63	-0.52	-0.14	IMF
Italy	0.89	1.43	1.87	1.92	1.37	0.62	EC
U.K.	1.42	2.01*	6.50**	3.24**	2.83**	3.50**	EC

Country	Current Year						"Winner"
	<i>DMAIO</i>	<i>DMSIO</i>	<i>DMAIE</i>	<i>DMSIE</i>	<i>DMAOE</i>	<i>DMSOE</i>	
France	-1.15	-1.43	-0.97	-1.03	-0.56	-0.70	IMF
Germany	-1.45	-1.49	1.38	0.79	2.12*	1.28	EC
Italy	1.32	1.68	0.89	1.01	1.40	2.03*	EC
U.K.	1.43	1.82	0.02	-0.12	0.03	-0.32	EC/OECD

\* and \*\* indicate significance at 5% and 1% levels.

The suffices IO, IE and OE refer to IMF-OECD, IMF-EC and OECD-EC comparisons.

DMA and DMS are the Diebold-Mariano tests in (2.3) for absolute and quadratic loss.

Table 8: Weak Efficiency and Unbiasedness Tests

OECD								
Country	<i>T0Y</i>	<i>CY</i>	<i>T1Y</i>	<i>T2Y</i>	<i>T0C</i>	<i>CC</i>	<i>T1C</i>	<i>T2C</i>
France	-0.31	2.52	-1.57	-1.64	1.28	0.42	0.18	-0.47
Germany	1.88	0.01	-2.94**	-4.51**	2.14*	0.30	-2.59*	-4.44**
Italy	0.41	3.58	-4.00**	-4.15**	1.78	0.80	-2.43*	-2.75**
U.K.	-0.37	2.96	-1.52	-1.36	-0.11	0.67	-1.61	-1.88

EC								
Country	<i>T0Y</i>	<i>CY</i>	<i>T1Y</i>	<i>T2Y</i>	<i>T0C</i>	<i>CC</i>	<i>T1C</i>	<i>T2C</i>
France	-0.79	4.25	-0.93	-0.66	-0.76	1.43	-1.35	1.85
Germany	0.14	0.67	-2.44*	-3.15**	1.95	0.16	0.40	-0.31
Italy	-0.47	0.13	-1.67	-1.63	-0.74	0.19	-2.59**	-2.52*
U.K.	-0.83	1.34	-1.22	-0.91	-2.05*	2.48	-3.09**	-2.079*

\* and \*\* indicate significance at 5% and 1% levels.

The suffices Y and C refer to year ahead and current year forecasts.

T0 is the (t-) test for zero mean forecast errors ( $\beta_0 = 0$  in (2.2)).

C is the (LM) test for uncorrelated forecast errors ( $v_h$  in (2.2)).

T1 and T2 are the (t-) tests in the weak efficiency regression ( $\alpha_0 = 0$  and  $\alpha_1 = 1$  in (2.1)).

Table 9: Forecast Comparison - Encompassing Tests, Year Ahead

France						
	$F_{IMF}$	$F_{OECD}$	$F_{EC}$	$F_{IMF}, F_{OECD}$	$F_{IMF}, F_{EC}$	$F_{OECD}, F_{EC}$
$e_{IMF}$	—	-0.52	-0.04	—	—	-0.71, 0.68
$e_{OECD}$	-0.65	—	-0.54	—	-2.64**, 2.15*	—
$e_{EC}$	-1.77	-0.94	—	-1.84, 1.14	—	—

Germany						
	$F_{IMF}$	$F_{OECD}$	$F_{EC}$	$F_{IMF}, F_{OECD}$	$F_{IMF}, F_{EC}$	$F_{OECD}, F_{EC}$
$e_{IMF}$	—	-2.94**	-2.48**	—	—	-1.37, 0.37
$e_{OECD}$	-3.09**	—	-1.38	—	-0.46, -0.40	—
$e_{EC}$	-2.12*	-3.28**	—	-0.47, -1.58	—	—

Italy						
	$F_{IMF}$	$F_{OECD}$	$F_{EC}$	$F_{IMF}, F_{OECD}$	$F_{IMF}, F_{EC}$	$F_{OECD}, F_{EC}$
$e_{IMF}$	—	-6.05**	-4.67**	—	—	-3.64**, 1.76
$e_{OECD}$	-1.40	—	-3.74**	—	0.52, -1.86	—
$e_{EC}$	-0.66	-1.07	—	1.61, -1.84	—	—

U.K.						
	$F_{IMF}$	$F_{OECD}$	$F_{EC}$	$F_{IMF}, F_{OECD}$	$F_{IMF}, F_{EC}$	$F_{OECD}, F_{EC}$
$e_{IMF}$	—	-0.50	0.08	—	—	-2.48*, 2.44*
$e_{OECD}$	-1.23	—	-0.43	—	-2.20*, 1.89	—
$e_{EC}$	-1.41	-1.19	—	-1.07, 0.82	—	—

\* and \*\* indicate significance at 5% and 1% levels.

t-tests for  $\gamma_1 = 0$  in (3.1), and  $\delta_1 = 0, \delta_2 = 0$  in (3.2).

Table 10: Forecast Comparison - Encompassing Tests, Current Year

France						
	$F_{IMF}$	$F_{OECD}$	$F_{EC}$	$F_{IMF, OECD}$	$F_{IMF, EC}$	$F_{OECD, EC}$
$e_{IMF}$	—	-0.81	1.11	—	—	-0.93, 1.26
$e_{OECD}$	0.13	—	-1.37	—	-1.82, 1.71	—
$e_{EC}$	1.02	1.33	—	0.44, -0.92	—	—

Germany						
	$F_{IMF}$	$F_{OECD}$	$F_{EC}$	$F_{IMF, OECD}$	$F_{IMF, EC}$	$F_{OECD, EC}$
$e_{IMF}$	—	-2.82**	-0.18	—	—	-1.77, 0.08
$e_{OECD}$	-5.63**	—	-0.10	—	-3.82**, 0.49	—
$e_{EC}$	-1.40	-0.18	—	-2.43*, 1.88	—	—

Italy						
	$F_{IMF}$	$F_{OECD}$	$F_{EC}$	$F_{IMF, OECD}$	$F_{IMF, EC}$	$F_{OECD, EC}$
$e_{IMF}$	—	-1.05	-3.80**	—	—	0.51, -2.53**
$e_{OECD}$	-0.99	—	-2.49*	—	-0.62, -1.98*	—
$e_{EC}$	-0.65	-1.23	—	0.44, -1.08	—	—

U.K.						
	$F_{IMF}$	$F_{OECD}$	$F_{EC}$	$F_{IMF, OECD}$	$F_{IMF, EC}$	$F_{OECD, EC}$
$e_{IMF}$	—	-0.91	-0.22	—	—	-0.79, 0.65
$e_{OECD}$	-0.62	—	-0.13	—	-0.11, 0.06	—
$e_{EC}$	-1.13	-1.41	—	0.60, -0.98	—	—

\* and \*\* indicate significance at 5% and 1% levels.

t-tests for  $\gamma_1 = 0$  in (3.1), and  $\delta_1 = 0, \delta_2 = 0$  in (3.2).

Table 11: Forecast Pooling - Coefficients

France								
	Year ahead				Current year			
	<i>const.</i>	$F_{IMF}$	$F_{OECD}$	$F_{EC}$	<i>const.</i>	$F_{IMF}$	$F_{OECD}$	$F_{EC}$
$e_{IMF}$	-1.67**	-2.19**	-0.95*	2.41**	-0.25	-0.44	0.20	0.14
$e_{OECD}$	-2.56**	-2.33**	-2.54**	3.87**	-0.06	0.03	-0.24	0.26
$e_{EC}$	-2.74**	-2.54**	-1.65**	3.14**	-0.96	-0.17	0.28	-0.31

Germany								
	Year ahead				Current year			
	<i>const.</i>	$F_{IMF}$	$F_{OECD}$	$F_{EC}$	<i>const.</i>	$F_{IMF}$	$F_{OECD}$	$F_{EC}$
$e_{IMF}$	-1.91**	-1.02**	-0.59	0.51	-1.44*	-0.94	0.08	0.05
$e_{OECD}$	-1.40*	-0.19	-1.33*	0.85	-0.97	-0.43	-0.26	0.04
$e_{EC}$	-1.39*	-0.19	-0.40	-0.13	-0.47	-0.91*	0.28	0.02

Italy								
	Year ahead				Current year			
	<i>const.</i>	$F_{IMF}$	$F_{OECD}$	$F_{EC}$	<i>const.</i>	$F_{IMF}$	$F_{OECD}$	$F_{EC}$
$e_{IMF}$	-5.38*	-0.40	-1.02	0.88	-5.92**	-0.41**	0.40	-0.60
$e_{OECD}$	-6.03**	0.29	-0.75	-0.15	-7.61**	0.34*	-0.88**	-0.22
$e_{EC}$	-6.27**	0.29	0.26	-1.20*	-7.54**	0.31	0.11	-1.20**

U.K.								
	Year ahead				Current year			
	<i>const.</i>	$F_{IMF}$	$F_{OECD}$	$F_{EC}$	<i>const.</i>	$F_{IMF}$	$F_{OECD}$	$F_{EC}$
$e_{IMF}$	-0.17	-1.71	-0.52	2.10	0.08	-0.45	0.04	0.36
$e_{OECD}$	-0.39	-0.61	-1.41	1.87	-0.22	0.45	-0.74	0.21
$e_{EC}$	-0.83	-0.65	-0.23	0.66	-0.74*	0.58	-0.10	-0.66*

\* and \*\* indicate that the associated t-tests reject non significance at 5% and 1% levels.

The samples are 1985-1994 (year ahead) and 1981-1994 (current year).

The optimal pooling coefficient on the forecasts which are from the same agency as the dependent variable in the regression is one plus the reported value.

Table 12: Correlation of Forecast Errors, 1985-1994

Year Ahead					
France			Germany		
	IMF	OECD		IMF	OECD
OECD	0.89	1	OECD	0.65	1
EC	0.94	0.97	EC	0.70	0.86
Italy			UK		
	IMF	OECD		IMF	OECD
OECD	0.73	1	OECD	0.96	1
EC	0.25	0.78	EC	0.88	0.95

Current Year					
France			Germany		
	IMF	OECD		IMF	OECD
OECD	0.82	1	OECD	0.63	1
EC	0.22	0.46	EC	-0.03	-0.23
Italy			UK		
	IMF	OECD		IMF	OECD
OECD	0.52	1	OECD	0.86	1
EC	0.27	0.52	EC	0.51	0.71

Table 13: IMF Forecasts - Asymmetric Loss Function  
Weak Efficiency Tests

Country	Year Ahead				$\hat{c}$	$CL$	$T1L$	$T2L$
	$\widehat{b/a}$	$CQ$	$T1Q$	$T2Q$				
Canada	5.53	0.69	-1.59	-1.68	2.05	0.74	-1.16	-1.23
France	2.05	1.90	-1.61	-1.82	0.77	1.57	-1.69	-1.91
Germany	0.88	0.44	-5.75**	-6.41**	0.00	—	—	—
Italy	0.26	0.16	-2.88**	-2.95**	-0.40	0.09	2.32*	2.38*
Japan	0.25	3.89*	-2.62**	-3.07**	0.00	—	—	—
U.K.	0.59	7.79**	-1.56	-1.60	-0.19	6.57**	1.72	2.09*
U.S.	0.65	2.87	-2.48*	-2.65**	-0.46	2.92	2.43*	2.61**

Country	Current Year				$\hat{c}$	$CL$	$T1L$	$T2L$
	$\widehat{b/a}$	$CQ$	$T1Q$	$T2Q$				
Canada	2.85	0.35	-1.30	-1.37	1.22	0.21	-1.24	-1.29
France	1.54	0.02	-1.78	-2.02*	0.71	0.01	-1.73	-1.96*
Germany	1.05	0.59	-4.31**	-4.62**	0.06	—	—	—
Italy	0.34	3.16	-3.74**	-3.83**	0.34	2.32	3.28**	3.36**
Japan	0.08	0.58	-2.10*	-2.42*	0.00	—	—	—
U.K.	0.24	1.63	-1.02	-1.22	0.00	—	—	—
U.S.	0.51	2.29	-3.24**	-3.50**	0.00	—	—	—

\* and \*\* indicate significance at 5% and 1% levels.

The suffices Q and L refer to the asymmetric quadratic and linex loss functions.

C is the (LM) test for non correlation of  $c'(e)$ .

T1 and T2 are the (t-) tests for  $\delta_0 = 0$  and  $\delta_1 = 0$  in (4.1).