NATIONAL INSTITUTE OF ECONOMIC RESEARCH

WORKING PAPER

No. 64
January, 1999

COMPARING THE ACCURACY OF EUROPEAN GDP FORECASTS

LARS-ERIK ÖLLER
BHARAT BAROT

WORKING PAPER No. 64
January, 1999

COMPARING THE ACCURACY OF EUROPEAN GDP FORECASTS

LARS-ERIK ÖLLER
BHARAT BAROT

Svensk resumé på sid 24.

Abstract. One-year-ahead forecasts by OECD and by national institutes of annual growth of GDP of 13 European countries are analysed for accuracy 1971-1995. Average errors were large: 1.9 % in *RMSE* and 1.4 % in *MAE*. Wilcoxon signed-rank tests showed that only four (five) OECD and two (four) institute forecast records were significantly better than an average growth forecast (the last observed growth) as measured by *RMSE*. Still, both OECD and institute forecasts could jointly beat the naive variants, but there was no significant difference in accuracy between OECD and the institutes. Few one-year-ahead forecasts were biased or had autocorrelated errors. Directional forecasts were also jointly informative. For some years, e.g. 1988, the direction was perceived wrongly by many forecasters. Only the OECD forecasts of Italy and Sweden improve significantly over time. OECD's two-year-ahead country forecasts were significantly correlated with the outcome only for half of the countries, and they were positively biased.

Keywords: European economic forecast accuracy, directional forecast accuracy.

1 Introduction

A forecast cannot be checked before it becomes worthless. However, if more than one forecast of the same variable is available and if there are records of previous forecasts and outcomes, one's confidence can be distributed according to the error-record. It should be the responsibility of forecasters to regularly report and analyse their forecast records to the consumers, using adequate statistical methods.

Large international organisations like IMF and OECD have their forecasts occasionally scrutinised, cf. Artis (1996) and OECD (1993), but there has been little interest in studying the forecasts of the most influential national macroeconomic forecasters. This is a modest attempt at comparing the accuracy of Gross Domestic Product (GDP) annual growth forecasts made by OECD for 13 European countries and forecasts made by an institute in the country, proper.

The first hazard one encounters is getting lost in dimensionality. There are many forecasters and each one forecasts many variables. Different periods can be studied and there is a multitude of ways of assessing accuracy. In order to keep the analysis manageable, we study just one variable (GDP), one main horizon (next year), and predominantly only one period.

The next obstacle is the very meaning of accuracy. Assessing the accuracy of a forecast *ex post* may seem to be a simple problem: one just measures the distance between the forecast and the known outcome. But if one is forecasting GDP, the outcome is not known in the sense of a static target. All countries first publish a preliminary figure, which can best be described as an informed guess. Successive revisions, sometimes with a 10 years lag, will reduce the amount of approximation in the figure, but never eliminate it completely. Hence comparing forecasters according to how close their hit came to a published figure is also a comparison of the "outcome" data. There are hardly any studies on accuracy of official statistics, an exception being the unpublished study by Gustafson (1994), where inconsistency in production data is demonstrated by studying two inventory series that should portray the same activity, but that are only weakly correlated.

Accuracy reports by organisations and national institutes are mostly descriptive studies, while external researchers, publishing in refereed journals, usually test forecasts and ask the obvious question, whether the they contain information for presumed users. In Ash et al. (1998) (containing an excellent list of references) OECD forecasts of G7 countries and 20 macroeconomic variables are studied and tested, using forecast records from the period 1967-1987, and for three forecast horizons: ½, 1 and 1 ½ years ahead. Taking ordinary differences of seasonally adjusted GDP, their main result is that what they call "quasi-forecasts", i.e. (directional) forecasts for the current quarter generally contain information. When the horizon is one year there is some indication of information content only for France and UK (and USA), and only for UK when forecasting the 1½ years. Stekler (1994) analyses three organisations forecasting quarterly GNP figures of the United States for the period 1972-1983. Directions and rough size of change are studied. The main conclusion again is that current quarter quasi-forecasts are valuable, while one quarter ahead the results are ambiguous.

Given this rather pessimistic starting point we set out to study the data at hand, presented in Section 2. In Section 3 we look at *average errors,* both root mean square and mean absolute errors *(RMSE* and *MAE)* and we test for information content, using the Wilcoxon signed rank test. We also test for accuracy improvement over time and for information in two-year-ahead forecasts. *Systematic errors,* i.e. bias and autocorrelation are studied in Section 4. In this section we also study bias in two-year-ahead forecasts. Non-parametric measures of accuracy based on direction allow for a different assessment. In Section 5 we look at acceleration-deceleration and test against the naive variant using a standard contingency table. Section 6 discusses the results.

An earlier version of this study, using final outcome, was published (in Swedish) in Öller et al. (1998).

2 Data [1]

Annual growth forecasts made by OECD have been collected from the December issue of *OECD Economic Outlook,* 1971-1996. When these forecasts are made, preliminary data for the first half of the current year is the latest data available. Table A.1 in Appendix 1 lists the European forecasting institutes that have kindly provided us with the data on their own forecasts. The dates of publication vary somewhat among institutes, but occur in the autumn, and concern the current and the next year. As seen from the last column, many institutes have forecasts only for more recent times (an unbalanced panel). This has necessitated separate studies of the subperiods 1980-1995 and 1985-1995. Generally, national institute and OECD forecasts are strongly connected; the coefficient of correlation is in fact 0.89. Graphs of the forecasts are shown in Exhibits 1. Because of this correlation the two forecasts are in most cases jointly good or bad. Except for varying volatility, data quality may be an issue here. Note also the general reluctance to forecast negative values. Consequently, GDP figures of some countries seem to be more difficult to forecast than the figures of others countries. Some report no negative values, although all economies have experienced contractions. Rates of –7 % for Finland and Switzerland were forecasted as non-negative!

Outcome (for both OECD and institutes) is defined as the figure published in the December issue of the *Economic Outlook* one year later. This is probably the figure most forecasters are aiming at. It is close to the, partly unknown, reference year of the percentage growth figure, cf. Ash et al. (1998).

Since 1990, OECD has also published forecasts of GDP of member countries *two years ahead*. These short time series are used for supplementary information in Sections 3 and 4.

3 Average errors and naive forecasts

The most common average error measure, *RMSE* is shown in Exhibit 2. Panel A contains OECD's country forecasts and Panel B those of the national institutes. Looking at the entire period 1971-1995, the mean errors are in the range between 1.4 % for OECD's forecast of the GDP growth in France as well as for IFO's forecast for Germany, and 2.7 % for both forecasts for Finland. Considering that these are errors in GDP growth rates, even the minimum, 1.4 p.c. points must be considered unacceptably large.

The standard deviation *(SD)* of the outcome in the next column is larger for Finland than for the other countries, and hence that variable is more difficult to forecast. In the third column, *RMSE* has been divided by *SD*. Now the most accurate forecaster is IFO, followed by a large group of almost equally accurate forecasts. Finally, there are some forecasters with ratios close to, or even above unity.

The ratio *RMSE/SD* can be interpreted as a *Theil Index*. A value above unity means that the forecast cannot beat a "naive" average growth forecast and hence contains no information in the sense of *RMSE*. The naive forecast would be the best possible, if GDP would be close to a random walk with drift (average growth) in logarithms, and if nothing else would be known about GDP than its own history. A random walk with drift can be written:

$$\Delta GDP_t = \mu + a_t, \qquad (1)$$

where Δ is percentage growth, μ is average growth and the errors a_t are i.i.d. $(0,\sigma^2)$. We see that both forecasts for Norway have values above unity. A possible explanation is the unpredictability of the off-shore economy of Norway.

Given that the Theil Index is a random number we have tested if the values below unity are significantly so, using the Wilcoxon signed-rank test (S_3) in Diebold and Mariano (1995). An asterisk denotes (5%) significance. We have also tested against the naive forecast: next year's growth is equal to this year's, implying GDP would be integrated of order two. This is a rather strange assumption, and indeed proves easier for the forecasters to beat, as can be seen from the cross-markings in Exhibit 2, denoting significance. Note that the totals of both groups of forecasts and the grand total are significant, although only some of the individual forecasts are significant. Hence, increasing the number of observations through pooling indicates that, indeed, there is information in the one-year-ahead forecasts, if calculated as *annual* growth rates, not as successive differences of semiannual or quarterly figures, as in Ash et al. (1998) and Stekler (1994). Using the same test we also checked if there is a difference in accuracy between OECD and national institutes, but we found no significant difference.

Why is it so difficult for individual forecasters to beat naive model (1)? Could the GDP time series be regarded as generated by (1)? We tested if mean growth deviations could be regarded as white noise using a Breusch-Godfrey LM-test. A triangle after the country name in Panel A of Exhibit 2 marks where this hypothesis could be rejected (if available, we used longer data series). Only for five countries the hypothesis could not be rejected, so that (1) should not have been difficult to dominate.

For the shorter periods 1980-95 and 1985-95 no tests were considered because of too few observations, and one has to take the figures at their face value. The forecasts for Sweden confirm their positions, while IFO has to leave its high *RMSE* accuracy ranking in favour of a newcomer, ISCO's forecasts for Italy.

The figures under the totals show the *RMSE* when the forecasts are compared to the final outcome. In almost all cases forecasts are closer to the figure released in December next year than to the final value.

The absolute errors in Exhibit 3 present a picture close tho that of Exhibit 2. *MAE* puts less weight on large errors than *RMSE*. The fact that total mean errors are smaller in Exhibit 3 than in Exhibit 2 follows from the rule of thumb: $RMSE \cong 1.25 MAE$.

Has the development of forecasting methods during a quarter of a century resulted in improved forecast accuracy? Looking at aggregate *RMSE/SD* and *MAE/MAD* in the bottom rows of Exhibits 2 and 3, there seems to have been little change, although some forecasting institutes show declining average errors. Regarding the observed errors as a random sample, the absolute errors made by each forecaster were regressed against a constant and a time trend. Only the OECD forecasts for Italy and Sweden improved significantly over time.

We saw that, contrary to the results mentioned in the introduction there is joint information in one-year-ahead forecasts. Is there also information in the two-year-ahead forecasts published by OECD? The correlation between forecast and outcome was first tested for each country separately. A significant relationship was found only for UK and Finland. Note that there were just seven observations for each country. A more reliable test was constructed by testing two classes (testing all forecasts jointly was impossible because of heteroscedasticity) of forecasts, where the classification was based on *RMSE*. The group "better forecasts" contained those for Finland, France, Germany, Sweden, Switzerland and UK. For this (homoscedastic) group, the two year ahead forecasts are significantly correlated with the outcome, but not for the group of remaining country forecasts. Note that a two-year-ahead forecast may be based on a medium term scenario that ignores the business cycle, in which case comparing it to an outcome containing a cyclic component may be unfair.

4 Systematic errors

Is there a simple way of improving forecasting accuracy by avoiding systematic mistakes, such as *optimistic* or *pessimistic* tendencies? If one too often hits above the target one should lower the gun, and vice versa.

Testing for systematic errors is here performed by regressing the forecast errors on a constant and two lags. The lags measure *inertia* in adapting forecasts to new developments and can also be interpreted as measuring the tendency to make forecasts that are moving averages of past observations. Autocorrelated errors and bias are systematic in the sense that forecasts could be improved by just "forecasting" the error and then subtracting the error forecast from the GDP forecast. The MIFN forecast in Exhibit 1 could be a case in point. Including both bias and autocorrelation in the same test means that one simultaneously checks for systematic over- or underprediction, generally (bias) or locally (autocorrelation).

The results of the tests are shown in Exhibit 4. The figures in the left part of the table show that there is no biased forecaster on significance level 5 %, the closest being INSEE with 8 % When the same test was performed with just one lag, or when final, instead of preliminary outcome was used, the bias of the INSEE forecasts became significant on the 5 % level.

Looking at inertia, again there are very few cases. The assumption concerning autocorrelation in Norwegian forecasts is supported by the test. Surprisingly, the best forecaster according to average errors in the more recent periods, NIER could have improved accuracy even further just by correcting for inertia!

We also tested for bias in the two-year-ahead forecasts. In both groups there was significant positive bias, revealing optimism when looking two years ahead. Judgemental forecasts have been found to become more optimistic the longer the horizon, cf. Milburn (1978).

5 Directional forecasts

Leitch and Tanner (1995) suggest that the numerical accuracy measures *(RMSE, MAE*, etc.*)* have little relevance for users of forecasts in business enterprises, for which the *direction* indicated by the forecast is what matters most. One reason for this could be that businessmen look at professional forecasts in order to decide whether or not to invest in larger production capacity. In case the investor gets the wrong signal, the result will either be a loss of market shares or over-capacity. Ibid. presents evidence that there is joint directional information in the US GNP forecasts of 42 professionals.

Annual changes in GDP are almost always positive (and forecasts even more so). Given the short time series, the reader may get an idea of how well negative growths have been predicted just by studying Exhibits 1, cf. comments in Section 2. Here, another way of analysing directions has been chosen. At the end of the year, both OECD and the institutes publish forecasts of the next as well as of the *current* year. This means that the forecasts signal if there will be an *acceleration* or a *deceleration* of production. We study this second order direction, as is done in the case of prices in Ash et al. (1998).

Exhibit 5 shows the years when the forecasters missed the sign in the changes in growth. OECD's forecast of UK's GDP has missed only three times (1977, 1986 and 1988) during the quarter of a century studied. IFO fares much worse in this comparison than when looking at average errors. In order to compare to a naive alternative the last column contains ratios of forecast misses to the number of misses when simply projecting the last acceleration/deceleration one year ahead. In analogy with Theil's Index, it can be said that the direction hasn't been predicted better than by a naive forecast if the ratio isn't below unity. There are three ratios above unity: the OECD forecast for the Netherlands and the Danish as well as the Swiss national institute forecasts. Note that neither were these forecasts found to contain information in the *RMSE*-based test.

For ratios below unity we have performed a simple 2×2 contingency table test. An asterisk in the last column of Exhibit 5 shows where forecasts contain information, which is in more than half of the cases, a larger ratio than for average forecast errors, with constant growth as the naive forecast. This supports the claim in Leitch and Tanner (1995) that macroeconomic forecasters may be better at direction than at numerical accuracy. When we tested both forecast groups jointly, the accuracy was significantly better than a naive projection. Again, there was no significant difference between OECD and national institutes.

Could acceleration/deceleration be particularly difficult to forecast some years? Exhibit 6 presents aggregated OECD and national institute directional forecast misses over time. The year 1988 stands out as the most difficult to forecast. It's not hard to guess why; in October 1987, too many forecasters still believed in an always positive correlation between the stock market and production. More surprising is the contrast between the years 1994 and 1995. In the former case, nobody missed the direction, whereas 1995 was almost as bad as 1988. It's hard to say why. On average, 25-40 % of the directional forecasts are wrong.

6 Discussion

Summarising the results, we have found that when studying annual GDP growth rates:

(1) Average errors in forecasts of GDP growth are large, considering the rather small variations and the importance of this variable.

(2) Macroeconomic forecasts as published by OECD and national institutes jointly contain information, both on an interval scale and as directional forecasts. However, we found no significant difference in accuracy between OECD and institute forecasts. Note that one knows for sure that all OECD forecasts are published in December, while some institutes may publish as early as in September.

(3) There seems to be more information in annual than in semiannual change forecasts.

(4) Only the following forecasters are significantly better than all naive autoprojections tested here: average growth, latest growth and latest direction: the OECD forecasts of France and UK, and among the institutes, NIER. The following forecasts were not found to contain any significant information: OECD forecasts of Ireland, Netherlands and Norway, and the institute forecasts made by: CPB (Nertherlands) and MIFN (Norway). Only full length records are tested.

(5) The only forecasts that had improved over time were OECD's for Italy and for Sweden.

(6) The better half of the OECD two-year-ahead forecasts may jointly contain some information.

(7) Very few systematic errors were found in the one year ahead forecasts, but two-year-ahead forecasts are biased upwards.

Trying to understand why errors are so large, recall what was said in the introduction on measuring forecast errors as the difference between the forecast and preliminary figure. However important GDP is as an indicator, its measurement is very shaky. In Exhibit 7 we have calculated the *root mean square revision.* This is a measure of the errors that are detected in the official statistics. On top of that there are all kinds of errors, e.g. from sampling. Assume that two thirds are detected and removed. Then the average statistical error would be as large as the total average forecast error of Exhibit 2. This simple calculation shows that the forecasters may be able to share the blame for bad forecasts in equal parts with the statistical authorities. Exhibit 7 reveals that some revisions are significantly biased.

One would have expected a general improvement in accuracy over time, due to better knowledge of the economy and improved forecasting techniques. Here, too, the reason to the failure is unidentifiable. One thing is clear. It is as important, although less rewarding academically, to develop more accurate national accounts, using available electronic transmission techniques, as in trying to find new ways of improving the forecasts.

Acknowledgements

We want to thank the national institutes mentioned in the Appendix for helping us with the data. Special thanks go to Svante Öberg and Erik Ruist. We are also grateful to colleagues at NIER for helpful comments and support.

References

Artis, M.J. (1996). How accurate are the IMF's short-term forecasts? Another Examination of the World Economic Outlook. *Staff Studies of the World Economic Outlook*, IMF.

Ash, J.C.K., D.J. Smyth and S.M. Heravi (1998): Are OECD forecasts rational and useful?: A directional analysis. *International Journal of Forecasting 14,* 381-391.

Diebold, F.X. and R.S. Mariano (1995): Comparing predictive accuracy. *Journal of Business and Economic Statsistics* 13, 253-263.

Gustafson, C.-H. (1994): On the consistency of data on production, deliveries and inventories in the Swedish manufacturing industry. National Institute of Economic Research and the Economic Council, Sweden, Working Paper 38.

Leitch, G. and Tanner, J.E. (1995). Professional economic forecasts: are they worth their costs? *Journal of Forecasting*, 14, 143-157.

Milburn, M.A. (1978): Sources of bias in predicting future events. *Organisational Behavior and Human Performance* 21, 17-26.

OECD Economic Outlook, 1991-1996, OECD, Paris.

Öller, L.-E. and B. Barot (1998). How reliable are NIER's forecasts? A European comparison. In Swedish in *Konjunkturläget,* March, 113-135.

Stekler, H.O. (1994): Are economic forecasts valuable? *Journal of Forecasting* 13, 495-505.

Footnote

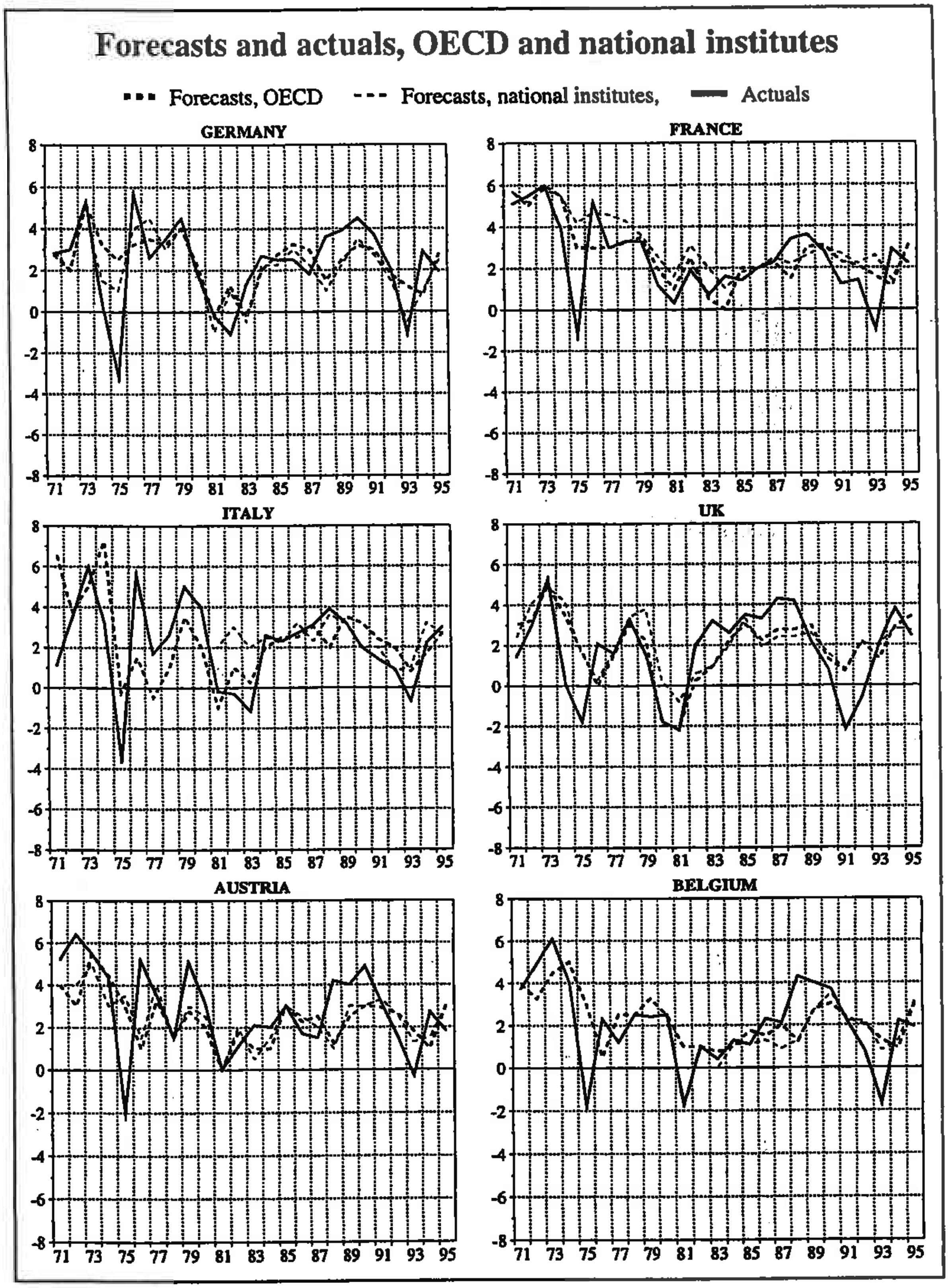
1 Data and calculations can be delivered upon request.

Appendix 1.

Table A1. National institutes

Initials	Name	Data period
IFO	Institut für Wirtschaftsforschung, Germany	1971-1995
INSEE	Direction de la prévision, France	1971-1995
ISCO	Instituto Nazionale per lo Studio della Congiuntura, Italy	1981-1995
NIESR	The National Institute of Economic and Social Research, UK	1971-1995
WIFO	Austrian Institute of Economic and Social Research, Austria	1971-1995
FPB	Federal Planning Bureau, Belgium	1983-1995
DEC	The Economic Council, Denmark	1974-1995
MIFF	Ministry of Finance, Finland	1971-1995
ESRI	The Economic and Social Research Institute, Ireland	1971-1976 1978-1995
CPB	Bureau for Economic Policy Analysis, Netherlands	1971-1995
MIFN	Ministry of Finance, Norway	1971-1995
NIER	The National Institute of Economic Research, Sweden	1971-1995
KOF	Swiss Institute for Business Cycle Research, Switzerland	1976-1995

Exhibit 1



Forecasts and actuals, OECD and national institutes
Forecasts, OECD
Forecasts, national institutes,
Actuals
GERMANY
FRANCE
ITALY
UK
AUSTRIA
BELGIUM
8
6
4
2
0
-2
-4
-6
-8
71 73 75 77 79 81 83 85 87 89 91 93 95

Exhibit 1 continuing

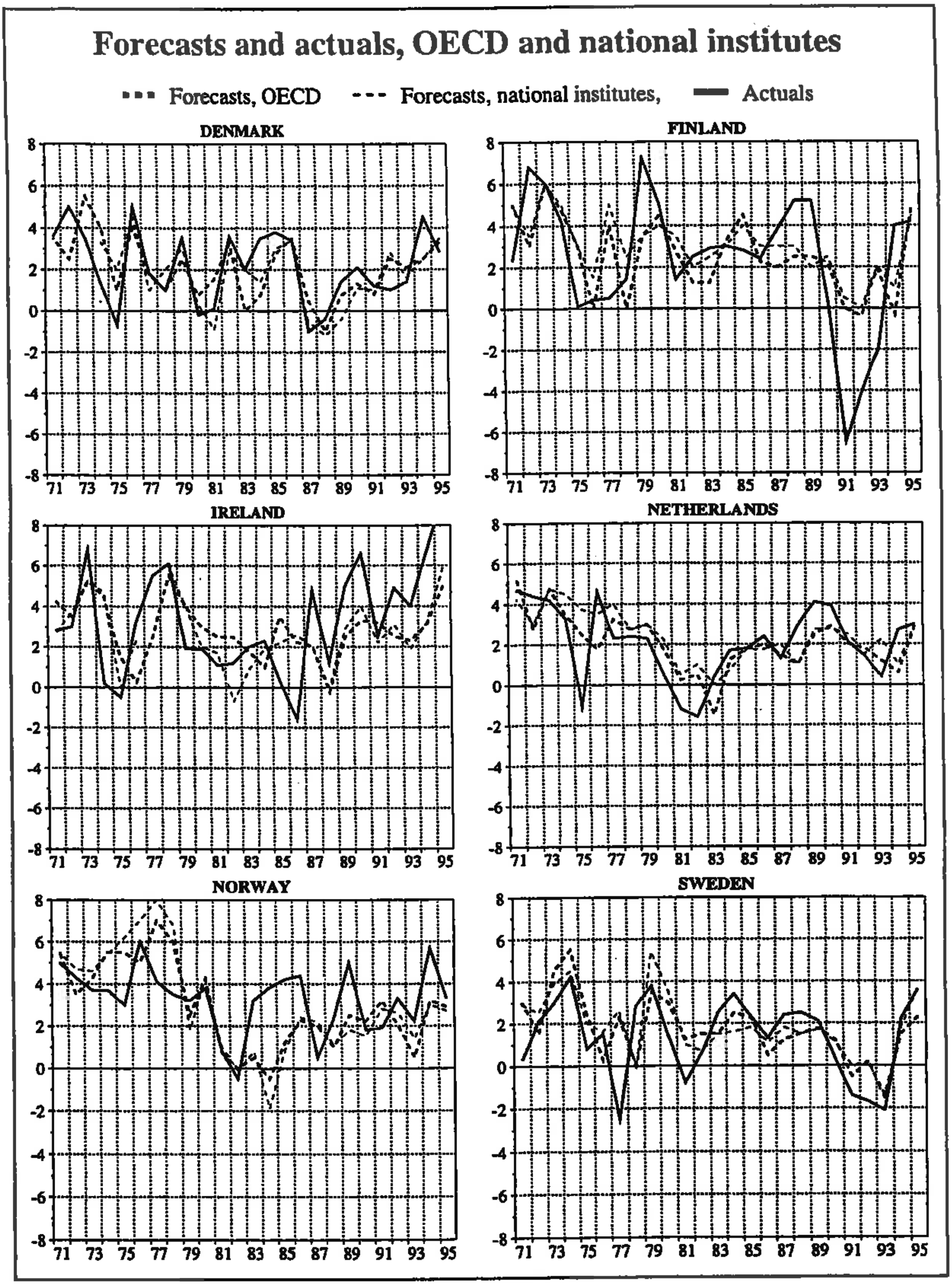


Exhibit 1 continuing

Forecasts and actuals, OECD and national institutes

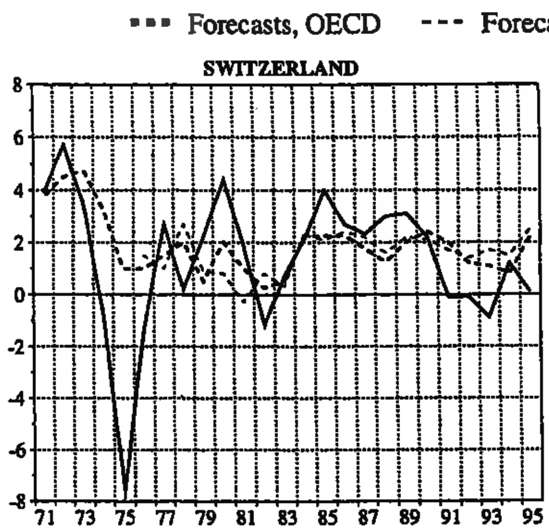


Exhibit 2. Forecasting accuracy measured by the root mean square error *(RMSE)*, and divided by the standard deviation *(SD)*

	1971 - 95			1980 -95			1985 - 95		
PANEL A	*RMSE*	*SD*	*RMSE/SD*	*RMSE*	*SD*	*RMSE/SD*	*RMSE*	*SD*	*RMSE/SD*
GERMANY▵	1.77	2.04	0.87 †	1.37	1.61	0.85	1.39	1.44	0.97
FRANCE▵	1.37	1.82	0.75 *†	1.15	1.14	1.01	1.24	1.22	1.02
ITALY	2.06	2.16	0.96	1.08	1.60	0.68	0.95	1.22	0.78
UK ▵	1.62	2.08	0.78*†	1.44	2.19	0.66	1.52	1.98	0.77
AUSTRIA	1.91	2.00	0.96†	1.36	1.39	0.98	1.53	1.43	1.07
BELGIUM	1.65	1.94	0.85	1.44	1.66	0.87	1.53	1.60	0.96
DENMARK	1.40	1.74	0.81*	1.29	1.63	0.79	1.04	1.63	0.64
FINLAND▵	2.67	3.15	0.85	2.73	3.28	0.83	3.21	3.77	0.85
IRELAND ▵	2.52	2.71	0.93	2.62	2.88	0.91	3.06	3.21	0.95
NETHERLANDS▵	1.42	1.76	0.81	1.24	1.54	0.81	1.07	1.05	1.02
NORWAY	1.89	1.53	1.24	1.98	1.64	1.21	1.82	1.50	1.21
SWEDEN▵	1.64	1.86	0.88†	1.13	1.76	0.64	1.00	1.88	0.53
SWITZERLAND▵	2.35	2.61	0.90*	1.40	1.66	0.84	1.43	1.57	0.91
Aver., OECD	1.91 (2.00)	2.15 (2.18)	0.90 (0.92)	1.64 (1.80)	1.93 (2.01)	0.66 (0.91)	1.75 (1.86)	1.96 (2.08)	0.92 (0.92)
PANEL B									
IFO	1.44	2.04	0.71*†	1.25	1.61	0.78	1.23	1.44	0.85
INSEE	1.61	1.82	0.89	1.34	1.14	1.18	1.40	1.22	1.15
ISCO	NC	2.16	NC	NC	1.60	NC	0.70	1.22	0.57
NIESR	1.78	2.08	0.86†	1.58	2.19	0.72	1.57	1.98	0.79
WIFO	1.91	2.00	0.96†	1.28	1.39	0.92	1.42	1.43	0.99
FPB	NC	1.94	NC	NC	1.66	NC	1.47	1.60	0.92
DEC	NC	1.74	NC	1.11	1.63	0.68	1.11	1.63	0.68
MIFF	2.69	3.15	0.85	2.71	3.28	0.83	3.20	3.77	0.85
ESRI	NC	2.71	NC	2.33	2.88	0.81	2.72	3.21	0.85
CPB	1.56	1.76	0.89	1.32	1.54	0.86	1.13	1.05	1.08
MIFN	2.20	1.53	1.44	2.20	1.64	1.34	1.89	1.50	1.26
NIER	1.63	1.86	0.88*†	1.14	1.76	0.65	0.88	1.88	0.47
KOF	NC	2.61	NC	1.71	1.66	1.03	1.50	1.57	0.96
Aver., institutes	1.89 (1.91)	2.15 (2.18)	0.96 (0.91)	1.71 (1.88)	1.93 (2.01)	0.91 (0.90)	1.69 (1.80)	1.96 (2.08)	0.90 (0.88)
Average, total	1.90 (1.97)	2.15 (2.17)	0.92 (0.92)	1.68 (1.84)	1.93 (2.06)	0.86 (0.91)	1.72 (1.83)	1.96 (1.95)	0.91 (0.90)

Notes: Denoting the forecast by *P* and the actual by *A*, *RMSE* and *SD* are calculated according to the formulas:

$$RMSE = \sqrt{\frac{1}{n}\sum_t (P_t - A_t)^2}\ , \qquad SD = \sqrt{\frac{\sum_t (\bar{A} - A_t)^2}{n}}\ ;$$

where $\bar{A}$ is the average growth and n is the number of observations. In *SD* the sum of squares is divided by n as in *RMSE*, regarding average growth as given. Forecasts that are significantly better than naive are marked *, if compared to average growth and by †, when naive is the last outcome. Significant deviations in GDP growth from random walk with drift is denoted by ▵ . NC means that data is not available.

Exhibit 3. Forecasting accuracy measured by mean absolute error (*MAE*), and divided by the mean absolute deviation (*MAD*)

	1971-95			1980-95			1985-95		
PANEL A	*MAE*	*MAD*	*MAE/MAD*	*MAE*	*MAD*	*MAE/MAD*	*MAE*	*MAD*	*MAE/MAD*
GERMANY	1.29	1.54	0.84†	1.15	1.24	0.93	1.21	1.04	1.16
FRANCE	0.95	1.46	0.65*†	0.90	0.89	1.01	0.96	0.93	1.03
ITALY	1.53	1.64	0.93	0.90	1.37	0.66	0.75	0.93	0.81
UK	1.25	1.66	0.75†	1.18	1.82	0.65	1.29	1.57	0.82
AUSTRIA	1.38	1.65	0.84*†	1.13	1.13	1.00	1.29	1.20	1.08
BELGIUM	1.20	1.45	0.83*†	1.04	1.29	0.81	1.21	1.13	1.07
DENMARK	1.13	1.50	0.75 *†	1.08	1.39	0.78	0.89	1.35	0.66
FINLAND	2.14	2.37	0.90	2.17	2.52	0.86	2.70	3.22	0.84
IRELAND	2.19	2.22	0.99	2.27	2.38	0.95	2.81	2.54	1.11
NETHERLANDS	1.09	1.41	0.77	0.99	1.26	0.79	0.79	0.88	0.90
NORWAY	1.57	1.15	1.37	1.63	1.37	1.19	1.64	1.28	1.28
SWEDEN	1.32	1.52	0.87 †	1.01	1.51	0.67	0.89	1.65	0.54
SWITZERLAND	1.64	1.98	0.83*	1.18	1.43	0.83	1.25	1.41	0.89
Aver., OECD	1.44 (1.51)	1.66 (1.64)	0.87 (0.92)	1.28 (1.37)	1.51 (1.49)	0.86 (0.92)	1.36 (1.44)	1.47 (1.51)	0.94 (0.98)
PANEL B									
IFO	1.11	1.54	0.72*†	1.03	1.24	0.83	1.01	1.04	0.97
INSEE	1.14	1.46	0.78	1.07	0.89	1.20	1.03	0.93	1.11
ISCO	NC	1.64	NC	NC	1.37	NC	0.50	0.93	0.54
NIESR	1.41	1.66	0.85 †	1.32	1.82	0.73	1.23	1.57	0.78
WIFO	1.38	1.65	0.84*†	1.01	1.13	0.89	1.15	1.20	0.96
FPB	NC	1,45	NC	NC	1.29	NC	1.16	1.13	1.03
DEC	NC	1.50	NC	0.89	1.39	0.64	0.89	1.35	0.66
MIFF	2.10	2.37	0.89*	2.03	2.52	0.81	2.70	3.22	0.84
ESRI	NC	2.27	NC	1.94	2.38	0.82	2.45	2.54	0.96
CPB	1.22	1.41	0.87*	1.08	1.26	0.86	0.90	0.88	1.02
MIFN	1.71	1.15	1.49	1.66	1.37	1.21	1.58	1.28	1.23
NIER	1.18	1.52	0.78*†	0.91	1.51	0.60	0.70	1.65	0.42
KOF	NC	1.98	NC	1.36	1.43	0.95	1.21	1.41	0.86
Aver., institutes	1.41 (1.41)	1.66 (1.64)	0.90 (0.90)	1.30 (1.40)	1.51 (1.49)	0.87 (0.92)	1.27 (1.36)	1.47 (1.51)	0.88 (0.92)
Average, total	1.43 (1.47)	1.66 (1.64)	0.89 (0.91)	1.29 (1.38)	1.51 (1.49)	0.86 (0.92)	1.32 (1.40)	1.51 (1.51)	0.90 (0.95)

Notes: Denoting the forecast by *P* and the actual by *A*, *MAE* and *MAD* are calculated according to the formulas :

$$MAE = \frac{1}{n}\sum_t | P_t - A_t | \quad , \qquad MAD = \frac{\sum_t | \bar{A} - A_t |}{n}$$

where Ā is the mean of the actuals and n is the number of observations. NC means that data is not available.

Exhibit 4. Testing for systematic errors: bias and autocorrelation

The figures are the probability values of Student's t and Fisher's F under the null hypotheses:

$$H_{01}: C = 0 \wedge H_{02}: \alpha = \beta = 0,$$

respectively, in

$$e_t = C + \alpha e_{t-1} + \beta e_{t-2} + \varepsilon_t$$

	1973 -1995	1982-1995	1985-1995	1973-1995	1982-1995	1985-1995
PANEL A	P-value:C Bias			P-value: (α, β) Autocorrelation		
GERMANY	0.61	0.65	0.57	0.88	0.71	0.84
FRANCE	0.33	0.70	0.74	0.77	0.87	0.80
ITALY	0.96	0.76	0.61	0.19	0.27	0.70
UK	0.73	0.80	0.79	0.23	0.15	0.30
AUSTRIA	0.30	0.46	0.73	0.52	0.88	0.99
BELGIUM	0.49	0.90	0.90	0.93	0.55	0.75
DENMARK	0.88	0.48	0.64	0.09	0.14	0.55
FINLAND	0.59	0.64	0.60	0.44	0.35	0.54
IRELAND	0.41	0.32	0.33	0.22	0.25	0.65
NETHERLANDS	0.93	0.70	0.24	0.63	0.85	0.90
NORWAY	0.55	0.12	0.14	0.17	0.39	0.48
SWEDEN	0.44	0.85	0.75	0.77	**0.00***	**0.05***
SWITZERLAND	0.32	0.80	0.73	0.09	0.31	0.39
PANEL B						
IFO	0.64	0.15	0.19	0.61	0.66	0.94
INSEE	0.08	0.35	0.60	0.94	0.67	0.77
ISCO	NC	NC	0.36	NC	NC	0.17
NIESR	0.44	0.98	0.93	0.08	**0.01 ***	0.07
WIFO	0.33	0.55	0.78	0.25	0.98	0.98
FPB	NC	NC	0.17	NC	NC	0.10
DEC	NC	0.09	0.81	NC	0.27	0.80
MIFF	0.18	0.33	0.42	0.13	0.18	0.33
ESRI	NC	0.18	0.27	NC	0.52	0.49
CPB	0.34	0.91	0.54	0.58	0.53	0.54
MIFN	0.71	0.13	0.13	**0.04 ***	0.37	0.39
NIER	0.36	0.94	0.55	0.72	**0.00 ***	**0.04 ***
KOF	NC	NC	0.58	NC	NC	0.90

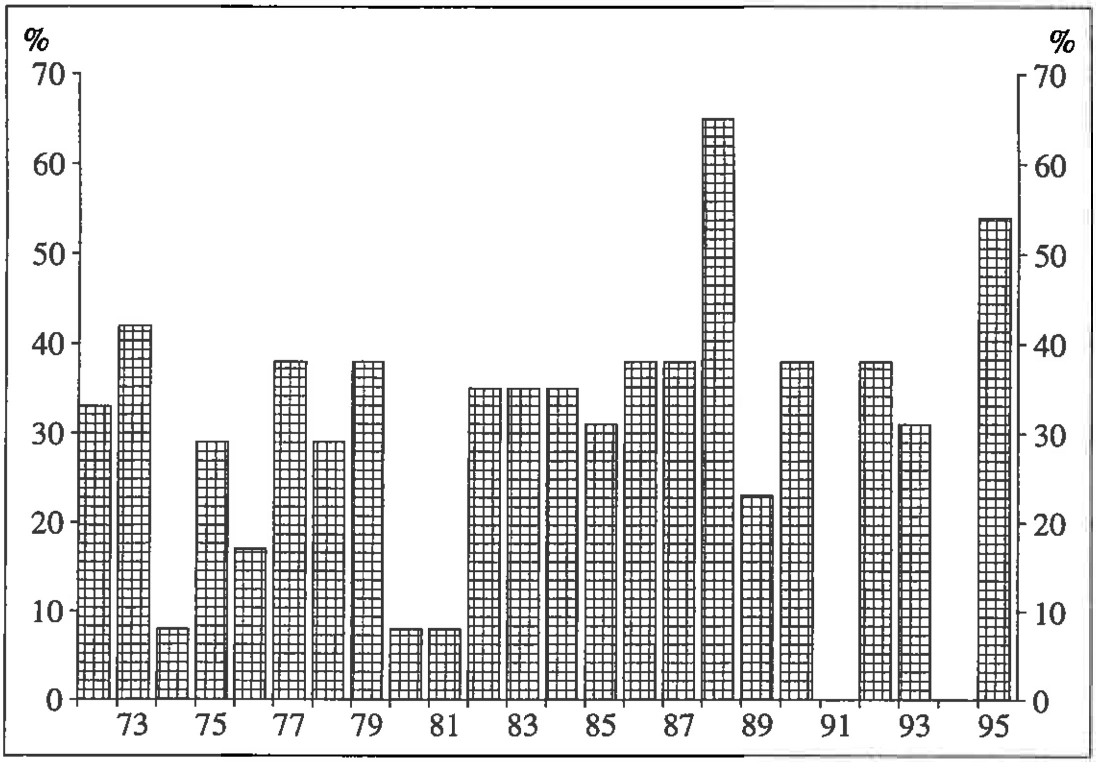
Notes: If the estimated intercept $\hat{C}$ is significantly different from zero we say that the forecast is biased. The F test of $\hat{\alpha}$ and $\hat{\beta}$ signals if there is autocorrelation in the errors. Bias means optimism/pessimism, while autocorrelated errors reveal tendency towards inertia. Significance at the 5% level is denoted by *. NC means that data is not available.

Exhibit 5. Years when acceleration/deceleration, was missed 1971 - 1995 and the ratio between these and naive forecast misses (relative numbers in parentheses)

PANEL A	1971 - 1995 Years missed	Relative numbers and 2 × 2 contingency table test results.
GERMANY	72, 75, 82, 85, 86, 87, 88, 89, 90, 95	10/11 (0.91)
FRANCE	72,79, 84, 85, 88, 89, 95	7/13 (0.54) *
ITALY	74, 76, 78, 82, 88, 92	6/10 (0.60) *
UK	75, 77, 81	3/13 (0.23) *
AUSTRIA	72, 73, 87, 95	4/14 (0.29) *
BELGIUM	72, 73, 78, 79, 80, 83, 86, 88 , 92, 93, 94, 95	12/17 (0.71)
DENMARK	73, 78, 84, 85, 86, 88, 90, 92	8/13 (0.62)
FINLAND	73, 82, 85, 88, 89	5/8 (0.63) *
IRELAND	77, 81, 85, 86, 90, 95	6/12 (0:50) *
NETHERLANDS	73, 75, 79, 81, 82, 83, 84, 86, 87, 88, 89 94	12/10 (1.20)
NORWAY	73, 74, 75, 77, 78, 84, 85, 86, 88, 92	10/11 (0.91)
SWEDEN	77, 87, 88, 92, 95	5/11 (0.45) *
SWITZERLAND	72, 73, 79, 85, 88, 89, 92, 93 , 95	9/11 (0.82)
PANEL B		
IFO	72, 75, 82, 86, 88, 89, 90, 95	8/11 (0.73)
INSEE	72, 83, 85, 89, 93, 95	6/13 (0.46) *
ISCO (85 -95)	87	1/2 (0.50)
NIESR	75, 77, 79, 81, 87	5/13 (0.38) *
WIFO	72, 83, 88, 90, 95	5/14 (0.36) *
FPB (85 - 95)	86, 88, 92, 95	4/6 (0.66)
DEC (80 - 95)	81, 83, 84, 85, 86, 88, 89 ,90, 91, 92, 93	11/8 (1.38)
MIFF	73, 82, 88, 89	4/8 (0.50) *
ESRI (80 - 95)	80, 81, 82, 86, 90	5/10 (0.50)
CPB	73, 75, 77, 79, 82, 87, 88, 93, 95	9/10 (0.90)
MIFN	73, 74, 75, 77, 78, 84, 85, 86, 88, 92	10/11 (0.91)
NIER	77, 84, 88, 92, 93	5/11 (0.45) *
KOF (80 - 95)	85, 88, 89, 92, 93, 95	6/5 (1.20)

Notes: Border cases where the forecast expresses no change in growth are classified as correct if the change in the outcome is smaller than one half percentage point. Analogously, a forecast is considered correct if it predicts an acceleration/deceleration of less than half a percentage point in cases where the outcome is no change in growth. Remaining cases are registered as wrong. A * signifies 5% confidence. NC means that data is not available.

Exhibit 6. Aggregated (OECD and institutes) error frequencies for directional forecasts



Note that this histogram is based on fewer forecasters in the beginning than in the end of the period 1972 - 1995.

Exhibit 7. *RMS* and bias of revisions, 1971 - 1995

	RMS	*Bias*
GERMANY	0.64	0.08
FRANCE	0.57	0.11
ITALY	2.25	0.31
UK	1.64	**0.43 ***
AUSTRIA	0.60	-0.12
BELGIUM	1.46	0.35
DENMARK	0.82	0.03
FINLAND	0.98	**0.33 ***
IRELAND	1.54	**1.21 ***
NETHERLANDS	1.80	**0.36 ***
NORWAY	1.28	0.49
SWEDEN	1.46	**0.49 ***
SWITZERLAND	0.69	0.17

Notes: Average RMS: 1.31. Revision = final - preliminary outcome. Bias and autocorrelation (not significant) tested as in Exhibit 6. A * signifies 5% significance.

Sammandrag på svenska

En användare av europeiska tillväxtprognoser kan inte skaffa sig en uppfattning om deras tillförlitlighet om det inte är känt hur bra prognosmakaren lyckats förut. Detta är ett försök till en sådan jämförande analys av BNP-prognoserna för 13 europeiska OECD-länder, gjorda dels av OECD, dels av prognosinstitut i respektive land (en preliminär rapport på svenska med delvis andra data ingick i *Konjunkturläget*, mars 1998). Vi undersöker träffsäkerheten under perioden 1971-1995. Man kan göra detta på ett otal sätt som kan ge mycket olika resultat. Här har ett brett spektrum valts för att ge en så mångsidig bild som möjligt. Prognosen görs i slutet av året och gäller nästa års tillväxt, då annat inte anges.

Genomsnittfelen är stora: 1,9 procentenheter mätt med roten av medelkvadratfelet och 1,4 om man mäter med absolut medelfel. Ett Wilcoxon-test visar att bara fyra (fem) OECD-prognoser och två (fyra) av åtta testade nationella instituts prognoser var signifikant bättre än naiva prognoser som bara anger den genomsnittliga (den senast observerade) tillväxten. Men betraktas OECD-prognoserna som aggregat och likaså institutens, är dessa signifikant bättre än båda de naiva alternativen. Däremot hittades ingen signifikant skillnad i träffsäkerhet mellan OECD och de nationella instituten. Det fanns nästan ingen bias eller autokorrelation i prognosfelen. Bara OECD:s prognoser för Italien och Sverige uppvisar en signifikant förbättring över tid.

Även prognosriktningen undersöktes. Eftersom prognosen i slutet av året även innehåller en siffra för innevarande år kan man undersöka om riktningen, i form av acceleration, respektive retardation, träffat rätt. Även här innehåller prognoserna information, sedda som aggregat, men som enskilda prognoser bara i ungefär hälften av fallen. Riktningen uppgavs fel av ett stort antal prognosmakare 1987 på hösten.

Vi studerade även de två-års prognoser OECD publicerat sedan 1990. De visade sig vara signifikant korrelerade med utfallet bara för hälften av länderna, betraktade som ett aggregat. Dessa prognosfel uppvisade signifikant positiv bias.

En jämförelse med statistiska revideringar visar att prognosmissar till ungefär lika delar kan skyllas på genuina prognosfel som på fel i statistiken.

Titles in the Working Paper Series

Number	Author	Title	Date
1	Warne, Anders and Anders Vredin	Current Account and Business Cycles: Stylized Facts for Sweden	12/89
2	Östblom, Göran	Change in Technical Structure of the Swedish Economy	12/89
3	Söderlind, Paul	Mamtax. A Dynamic CGE Model for Tax Reform Simulations	12/89
4	Kanis, Alfred and Aleksander Markowski	The Supply Side of the Econometric Model of the NIER	11/90
5	Berg, Lennart	The Financial Sector in the SNEPQ Model	2/91
6	Ågren, Anders and Bo Jonsson	Consumer Attitudes, Buying Intentions and Consumption Expenditures. An Analysis of The Swedish Household Survey Data	4/91
7	Berg, Lennart and Reinhold Bergström	A Quarterly Consumption Function for Sweden 1979-1989	10/91
8	Öller, Lars-Erik	Good Business Cycle Forecasts - A Must for Stabilization Policies	2/92
9	Jonsson, Bo and Anders Ågren	Forecasting Car Expenditures Using Household Survey Data	2/92
10	Löfgren, Karl-Gustaf, Bo Ranneby and Sara Sjöstedt	Forecasting the Business Cycle Not Using Minimum Autocorrelation Factors	2/92
11	Gerlach, Stefan	Current Quarter Forecasts of Swedish GNP Using Monthly Variables	2/92
12	Bergström, Reinhold	The Relationship Between Manufacturing Production and Different Business Survey Series in Sweden	2/92
13	Edlund, Per-Olov and Sune Karlsson	Forecasting the Swedish Unemployment Rate: VAR vs. Transfer Function Modelling	3/92
14	Rahiala, Markku and Timo Teresvirta	Business Survey Data in Forecasting the Output of Swedish and Finnish Metal and Engineering Industries: A Kalman Filter Approach	3/92

Number	Author	Title	Date
15	Christofferson, Anders, Roland Roberts and Ulla Eriksson	The Relationship Between Manufacturing and Various BTS Series in Sweden Illuminated by Frequency and Complex Demodulate Methods	4/92
16	Jonsson, Bo	Sample Based Proportions as Values on an Independent Variable in a Regression Model	10/92
17	Öller, Lars-Erik	Eliciting Turning Point Warnings from Business Surveys	11/92
18	Forster, Margaret M	Volatility, Trading Mechanisms and International Cross-Listing	12/92
19	Jonsson, Bo	Prediction with a Linear Regression Model and Errors in a Regressor	12/92
20	Gorton, Gary and Richard Rosen	Corporate Control, Portfolio Choice, and the Decline Banking	1/93
21	Gustafsson, Claes-Håkan and Åke Holmén	The Index of Industrial Production - A Formal Description of the Process Behind it	2/93
22	Karlsson, Tohmas	A General Equilibrium Analysis of the Swedish Tax Reforms 1989-1991	2/93
23	Jonsson, Bo	Forecasting Car Expenditures Using Household Survey Data - A Comparison of Different Predictors	2/93
24	Gennotte, Gerard and Hayne Leland	Low Margins, Derivative Securitites, and Volatility	2/93
25	Boot, Arnoud W.A. and Stuart I. Greenbaum	Discretion in the Regulation o U.S. Banking	4/93
26	Spiegel, Matthew and Deane J. Seppi	Does Round-the-Clock Trading Result in Pareto Improvements?	4/93
27	Seppi, Deane J.	How Important are Block Trades in the Price Discovery Process?	4/93
28	Glosten, Lawrence R.	Equilibrium in an Electronic Open Limit Order Book	4/93
29	Boot, Arnoud W.A., Stuart I Greenbaum and Anjan V. Thakor	Reputation and Discretion in Financial Contracting	4/93

NATIONAL INSTITUTE OF ECONOMIC RESEARCH
Box 3116, SE-103 62 Stockholm, Sweden,
TELEPHONE: +46 8 453 59 00, TELEFAX: +46 8 453 59 80, E-MAIL: ki@konj.se, HOMEPAGE: www.konj.se

ISSN 1100-7818