STRATHCLYDE

DISCUSSION PAPERS IN ECONOMICS

EVALUATING THE USEFULNESS OF FORECASTS OF RELATIVE GROWTH

BY

GRANT ALLAN

NO. 12-14

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STRATHCLYDE
GLASGOW

Evaluating the usefulness of forecasts of relative growth

Grant J. Allan[1]

Draft version – Not for citation
May 2012

Abstract: Forecasts of differences in growth between countries serve an important role in the justification of governments' fiscal policy stances, but are not tested for their accuracy as part of the current range of forecast evaluation methods. This paper examines forecasted and outturn growth differentials between countries to identify if there is usefulness in forecasts of "relative" growth. Using OECD forecasts and outturn values for GDP growth for (combinations of) the G7 countries between 1984 and 2010, the paper finds that the OECD's success in predicting the relative growth of G7 countries during this period is good. For each two-country combination results indicate that relative growth forecasts are less useful for countries which have smaller outturn growth differentials.

Keywords: macroeconomic forecast evaluation; relative forecast accuracy; OECD.

JEL codes: E01, O47, E02.

[1] Research Fellow, Fraser of Allander Institute, Department of Economics, University of Strathclyde, Sir William Duncan Building, 130 Rottenrow, Glasgow, G4 0GE, Email: grant.j.allan@strath.ac.uk; Tel: +44 (0) 141 548 3838.

1. Introduction

The accuracy of forecasts of macroeconomic variables, such as growth, is the subject of a large academic literature. Conventional forecast accuracy tests compare forecasted and outturn values, describing differences as the "error" of the forecast. Quantitative forecast accuracy statistics have been used to reveal the accuracy of a forecaster or forecasting organisation (Artis, 1996, Granger, 1996; Melliss and Whittaker, 1998; Timmerman, 2006) or to compare the accuracy of different forecasters (Mills and Pepper, 1999; Öller and Barot, 2000; Pons, 2000; Batchelor, 2001; Loungani, 2001). A second group of forecast evaluation techniques explores forecasts *qualitative* accuracy. This includes the "directional test" (Merton, 1981) in which forecasts of the direction of change in a variable between years are compared to the actual direction of change in that variable. For example a forecast may indicate whether annual GDP growth will be above (below) a specific rate. The directional test uses contingency tables to evaluate the independence of forecasts and outturn change in a variable and also has significant usefulness for the users of forecasts. For example, Melander *et al.* (2007, p. 17) note that "this [test] may be more important than the forecast [quantitative] value". Being able to correctly predict the direction of change is the only aspect of the usefulness of forecasts discussed in Stekler (1991)'s review of forecast evaluation techniques. There is, however, another way in which forecasts could be useful to users: where forecasts for more than one country are produced at the same time and countries *relative* growth prospects are compared.

Comparisons of national growth prospects are a standard part of modern political and economic discourse. In support of the economic position adopted in the UK Coalition Budget of March 2011 (HM Treasury, 2011), the Chancellor George Osborne stated – "In recent months, many other countries have seen their ratings downgraded and their borrowing costs soar. Our

country's fiscal plans have been strongly endorsed by the IMF, by the European Commission, by the OECD, and by every reputable business body in Britain… [European Commission forecasts] show that the UK is forecast to grow more strongly in the coming year than Spain, Italy, France, the average for the Eurozone and the average for the EU". In the same month, the Financial Secretary to the Treasury (Mark Hoban MP) further directly compared countries growth forecasts, stating: "The IMF's latest forecast shows the UK economy growing this year, and growing faster than the economies of France, Germany and Italy next year" (Hansard, 2011, col. 750).

Comparisons of national growth forecasts are not limited to the UK. During a speech in January 2012, the New Zealand Prime Minister John Key, when discussing the impact of his government's fiscal plans for the new parliament on economic growth, stated: "So anyone who complains that New Zealand isn't growing at four, five or six per cent a year right now is on the wrong planet. In fact, we are doing better than most developed countries. In both 2012 and 2013, the New Zealand economy is forecast to grow more strongly than the Eurozone, the UK, Japan, the United States and Canada. So we are in relatively good shape" (Key, 2012).

A similar sentiment was echoed in the speech by Jim Flaherty, the Canadian Minister of Finance, in March 2012, when, discussing Canada's economic prospects, he stated: "I can assure you that on the fiscal track we are on, [the net debt to GDP ratio] will return to where it was before the plan in the medium term. Also, the OECD and IMF predict that our economy will be among the leaders of the industrialised world over the next two years" (Flaherty, 2012). While the quantitative accuracy of economic forecasts is likely critical for revenue projections, and therefore spending decisions, these statements from across the world indicate that relative growth forecasts are used to justify the fiscal policy stances taken by national governments.

This paper therefore suggests a new qualitative forecast accuracy test which can be used to understand the usefulness of forecasts of relative growth between two countries produced by a single forecaster at a given point in time. This paper is concerned with developing and presenting results from a simple test for evaluating the *relative* outturn (i.e. actual) growth of nations against their forecasted relative growth. The proposed test compares the predicted sign (i.e. positive or negative) of the differential between growth forecasts and the sign of the outturn differential between outturn growth rates for those countries. This test will be useful to users of forecasts as countries forecasted relative growth rates have been used to justify fiscal plans: this test should explain whether such a use of forecasts would appear appropriate.

In Section 2 we summarise the literature on forecast evaluation techniques, showing that these omit the important "relative" dimension. Section 3 describes the calculation of statistics on relative forecast accuracy and the use of contingency tables. Section 4 outlines the data used and presents the results from assessing the accuracy of relative growth forecasts. This new qualitative forecast evaluation approach is tested using annual OECD GDP growth forecasts for the G7 countries between 1984 and 2010[2]. In aggregate, the results indicate that we can reject a null of independence between the OECD's forecasts for the relative growth of G7 countries and their outturn relative growth. Considering each two-country combination of G7 countries in turn over this period, the results are mixed. Our results suggest that countries with smaller outturn growth differentials are less likely to see their relative growth correctly predicted. Section 5 offers the conclusions from this paper and outlines some areas for further work in the evaluation of the accuracy of relative forecasts of economic growth.

---

[2] The G7 countries are the United States, Japan, Germany, United Kingdom, France, Italy and Canada.

2. Forecast evaluation approaches

2.1 Quantitative forecast accuracy tests

Conventional forecast evaluation techniques explore the *accuracy* or the *efficiency* of forecasts of a particular variable (e.g. a growth rate in year *t*, or year *t* inflation rate). Forecast *efficiency* refers to the notion that forecasts are unbiased (i.e. they are neither systematically positive nor negative predictors of the outturn value). Techniques for evaluating the *efficiency* of forecasts can involve regression analysis of forecast and outturn data. Mills and Pepper (1999, p. 250) define optimal forecasts as "both unbiased and efficient".

Here we described existing tests of forecast *accuracy*. These compare *ex ante* and *ex post* data on specific variables and calculate the "forecast error" as the difference between that forecasted and the outturn value for that variable. Figure 1 shows how forecast errors are calculated, illustrating this for a hypothetical scenario of three separate growth forecasts for country *A*'s growth in year *t*, each made in year *t*, ($F_{1,1}^{A}$, $F_{2,2}^{A}$ and $F_{3,3}^{A}$) and three outturn values for country A's growth (Y_1^{A}, Y_2^{A} and Y_3^{A})[3]. The forecast error is given by the distance between the forecast and the outturn value (e.g. the first forecast's error, $e_{1,1}^{A} = F_{1,1}^{A} - Y_1^{A}$). A positive (negative) forecast error indicates that the forecast was greater (lower) than the outturn value.

[Figure 1 here]

Empirical measures for quantitative forecast accuracy require a time series of forecast errors for a specific forecaster, and capture the quantitative accuracy of macroeconomic forecasts

---

[3] In practice, (preliminary) growth outturn values are typically available three months after the end of the final quarter of the year.

in a single variable. Such measures include Mean Error (ME), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) (some, or all, of these measures are used in, for example, Granger, 1996; Melliss and Whittaker, 1998; and Mills and Pepper, 1999; Pons, 2000; Timmerman, 2006).

With the hypothetical forecast errors from Figure 1, these tests are formally:

- ME = $(e_{1,1} + e_{2,2} + e_{3,3})/n$

- MAE = $(|e_{1,1}| + |e_{2,2}| + |e_{3,3}|)/n$

- RMSE = $\sqrt{[(e_{1,1})^2 + (e_{2,2})^2 + (e_{3,3})^2]/n}$

- MAPE = $[(|e_{1,1}|/Y_1) + (|e_{2,2}|/Y_2) + (|e_{3,3}|/Y_3)]/n$

where n is the number of forecasts (i.e. three in this case) and $\bar{Y}$ is the mean outturn value over the sample. What these values report is the closeness (i.e. the accuracy) between forecasts and outturn values. Smaller values for each measure suggest that forecasts are more accurate (i.e. errors are smaller) over the period covered.

There is a significant literature examining the accuracy of forecasts made by international organisations, such as the IMF or OECD (Artis, 1996; Koutsogeorgopoulou, 2000; Pons, 2000; Timmerman, 2006,). These studies have "found that the forecast errors are by and large non-systematic, and the forecasts prepared by the IMF and OECD are generally unbiased and efficient" (Krkoska and Teksoz (2007, p. 31). In addition, many studies have compared the forecast accuracy of private and public institutions. Batchelor (2001) compares the growth forecasts of the OECD and IMF with private sector forecasts for the G7 between 1990 and 1999. He finds that private sector forecasts are less biased and have a lower mean absolute error (MAE). Similar work by Artis (1996), Mills and Pepper (1999) Öller and Barot (2000) and

Loungani (2001), found no major differences in the forecast accuracy of public or private bodies. Mills and Pepper (1999) review the accuracy of the UK's major forecasting organisations and find that there forecasts typically have high correlations and are poor predictors of major economic events.

2.2   Qualitative forecast accuracy tests

The most common qualitative test of forecast accuracy is the "directional" test (Schnader and Stekler, 1990). This examines if forecasts correctly predict the direction of change in a variable from one year to another, i.e. for a given country, is growth next year forecast to be higher or lower than growth this year? The usefulness of being able to accurately forecast the direction of change is acknowledged throughout the academic literature. Leitch and Tanner (1995) for example, argue that what "quantitative" accuracy tests are of little importance for business users of forecasts who are more concerned about the direction of the forecast (Öller and Barot, 2000). Pons (2000, p. 59) notes, this analyses is only concerned with the direction of the predicted and actual changes and that "small and large [absolute forecast] errors are treated equally in this classification scheme". Interestingly, Stekler (1991) takes the directional test as the only useful measure of forecast accuracy for decision makers.

The direction of growth from one year to the next, as forecasted by different organisations, has been studied in the recent academic literature (Öller and Barot, 2000; Pons, 2000; Ashiya, 2003; Greer, 2003; Ashiya, 2006). Studies using this measure have typically found that international organisations such as the OECD and IMF are good predictors of the direction of countries growth, particularly for closer forecast horizons (e.g. Pons, 2000).

One test for the directional accuracy of a growth forecast uses a 2x2 contingency table (Merton, 1981, Stekler, 1991). Showing forecasted increases or decreases in a variable, e.g. GDP, and outturn increases or decreases in that same variable, there are four possible domains for any forecast, which can be represented as in Table 1 (Pons, 2000). Higher directional accuracy would be consistent with the sum of the entries in the diagonal of Table 1 (n_{00}+ n_{11}) being high. Stekler (1991) follows Merton's (1981) original test and details a 2x2 contingency table similar to Table 1 in which conditional probabilities of the change in a variable being up (down) when the forecast is for it to increase (decrease) are used to test if the sign of the forecasted change is independent of the actual change.

[Table 1 here]

The most straightforward test for independence in a contingency table (Stekler, 1991) is to use the Chi-squared test. The χ^2 test statistic for Table 1, adjusted for Yates' continuity correction, can be calculated using:

$$\chi^2 = \sum_{i=0}^{1}\sum_{j=0}^{1} \frac{\left[(n_{ij} - n_{i.}n_{.j}) - 0.5)/N\right]^2}{n_{i.}n_{.j}/N} \text{ with d.f.=1.}$$

Where n_{ij} refers to the number of instances of forecasted direction *i* and outturn direction *j*, and there are N total observations. The 10%, 5% and 1% values for the test statistic of the χ^2 distribution are 3.84, 6.63 and 10.83 respectively. Our null hypothesis is that observations of forecasted change and outturn change are independent. A χ^2 value greater than the critical

values indicates that we can reject this null, and conclude that there exists a positive association between the forecast and outturn directions of growth (Greer, 2003).

3. Relative forecast accuracy

3.1   Describing relative forecast and outturn differences

The quantitative or qualitative forecast accuracy measures examined in Section 2 do not consider how to evaluate whether *relative* forecasts of growth for different countries are accurate. What we propose is a qualitative directional test of economic forecast accuracy which can be used to evaluate the relative accuracy of forecasts, and which uses the contingency table approach of the directional test to see if forecasts of relative growth are useful for policymakers.

The relative forecast accuracy test compares the forecasted growth differential between two countries and the outturn growth differential outcomes for those two countries. For example, if country A was forecast to grow more (less) than country B, we term this a positive (negative) forecasted differential. By comparing the outturn growth data for the countries we can calculate the outturn differential, which can also be positive or negative. Put simply, if the forecasted differential and outturn differential have the same sign, then relative growth was correctly predicted. As with the directional test, we ignore the absolute size of these differentials.

Figure 2a shows how we calculate the *relative* forecast differentials for simultaneous forecasts for two countries (A,B) for a specific period such as a year, t ($fd_{t,t}^{A,B}$). Figure 2b shows the calculation of *outturn* growth differentials for those same countries in the forecasted period ($od_{t,t}^{A,B}$). As drawn, the forecasted growth differential is positive in the second forecast period (i.e. $F_{2,2}^{A} > F_{2,2}^{B}$) while the outturn growth differential is negative (i.e. $Y_{2,2}^{A} < Y_{2,2}^{B}$). For the other two

forecasts for each country the sign of the forecast and outturn growth differentials are the same, although their absolute size.

[Figure 2a here]

[Figure 2b here]

There are therefore two possible outcomes for the forecast growth differential (positive or negative) and the outturn growth differentials (positive and negative). These options give rise to four possible outcomes. Two possible outcomes relate to correct forecasts (the forecast and outturn differentials were either both positive or both negative) while two possible outcomes relate to incorrect forecasts: either the forecast was for a positive growth differential and the outturn differential negative, or *vice versa*. We can construct a 2x2 contingency table to show the outcomes of forecasted relative growth and the outturn relative growth. Replacing the forecasted and observed growth *direction* (from the directional test, described above) with forecasted and observed relative growth we can create Table 2.

[Table 2 here]

The χ^2 test statistic for relative forecast accuracy in Table 2 is again, adjusted for Yates' continuity correction, is:

$$\chi^2 = \sum_{i=0}^{1}\sum_{j=0}^{1}\frac{\left[(r_{ij} - r_{i.}r_{.j}) - 0.5)/R\right]^2}{r_{i.}r_{.j}/R} \text{ with d.f.=1}.$$

Where r_{ij} refers to the number of instances of forecasted direction *i* and outturn direction *j*, and there are R total observations. As before, the 10%, 5% and 1% values for the test statistic of the χ^2 distribution are 3.84, 6.63 and 10.83 respectively, and our null hypothesis is that all observations are independent. A χ^2 value greater than these critical values indicates that we can reject this null and say that there exists a positive association between the forecast and outturn *relative* directions of growth.

4.   Data

4.1   Sources

To demonstrate the test of relative forecast accuracy we require real GDP growth forecasts and outturn values for a sample of countries over a period of time. We use the OECD's Economic Outlook as the basis for the forecasts, taking forecasts of growth in the G7 countries from the June publication of the EO for each year between 1984 and 2010. Outturn growth for each year is taken from the latest publication of the Economic Outlook[4]. The forecasts therefore are produced at the around six months into the year to which they relate. For example, we take the forecasts for growth in 2005 for each G7 country as published in the EO from June of 2005. Relative forecast and outturn growth differentials are calculated for each two-country combination of G7 countries from the data for each country, making twenty-one two-country combinations in all (e.g. UK against USA, UK against Germany, Germany against USA, etc.). We are therefore comparing the direction of forecasted growth differentials between G7

---

[4] These were accessed in February 2012 and relate to annual growth rates of real GDP (or GNP where GDP figures do not exist, typically earlier in the sample). Figures for Germany between 1984 and 1990 are estimated by the OECD, as they relate to the period before the unification of East and West Germany. All data on forecasts and outturns growth rates are available from the author on request.

countries around the middle of each year (as published in that year) and outturn growth differentials as available now.

Using the G7 countries provides us with a manageable sub-grouping of OECD economies. It is also typical that the growth prospects of each country within the G7 be compared against other G7 countries, rather than developing countries, which may be on quite different growth trajectories. All of the example of statements by finance ministers seen in the introduction, for example, compare against other development economies. Many of the quantitative and qualitative tests of forecast accuracy have used the forecasts made by international organisations (such as the OECD) as they are timely, publicly available and produced for a range of countries at one instance.

4.2   Calculation of relative forecast accuracy measures

To illustrate the calculations, we show the relative growth forecasts and outturn data for the UK compared to Germany between 1984 and 2010. This demonstrates many of the issues with regard to the calculation of this measure and the test statistic. Figure 3 shows the forecasted and outturn growth differentials between the UK and Germany. The year to which the forecast and outturn growth differential relates is shown on the horizontal axis. The forecast and outturn growth differentials are compared to the UK, therefore a positive value indicates that for that year forecast (or outturn) growth was higher (i.e. more positive) than Germany.

[Figure 3 here]

Figure 3 shows that in fourteen years[5] of our sample the forecasted and outturn growth differential was positive, i.e. the UK was forecast to grow by more than Germany in the year and did so, while in eight cases[6] the forecast differential and outturn differential was negative. In all then, in twenty-two years of the twenty-seven years, the growth differential was correctly forecast. Other years saw incorrect forecasts. By grouping forecast and outturn growth differentials by their signs, as described in Section 3, we can construct 2x2 contingency tables for each two-country comparison, as well as for all two-country comparisons.

5.   Results

5.1   Testing relative forecast accuracy using contingency tables and χ^2 - all observations

Table 3 shows the forecast and outturn relative growth contingency table for the twenty-one combinations of two-country comparisons in each of the twenty-seven years of the sample period (1984 to 2010). This gives a total of 567 relative growth forecasts and outturn observations, which must lie in one of the four quadrants of Table 3[7].

[Table 3 here]

Looking at Table 3, we can see that in 455 out of 567 observations, the forecasted relative growth differential was the same as the outturn differential. This means that considering the aggregate forecasts, the OECD's forecasts of the relative growth of G7 countries was correct on

---

[5] Specifically, 1985, 1987, 1988, 1993, 1994, 1995, 1996, 1997, 2001, 2002, 2003, 2004, 2005 and 2009.
[6] Specifically, 1984, 1989, 1990, 1991, 1992, 2007, 2008 and 2010.
[7] The construction of the two-country combinations of G7 member countries means that we have compared each combination on one occasion only. For example, we have compared the UK to the USA, and not also the other way round. This means that each year of positive forecasted relative growth would not be matched by a year of negative relative growth – we have not compared the USA to the UK *as well as* the UK and USA. Therefore in Table 1 we do not have equal numbers of positive and negative forecasts or observed outturn growth observations.

over 80% of occasions. The χ^2 statistic for Table 3 is 207.5 – significantly above the 1% critical value of 10.83 – and indicating that we can reject the null hypothesis that the forecasts and observations are independent. There is therefore a statistically significant link between forecasts of relative growth and the outturn growth differential. This indicates overwhelming evidence that the OECD's forecast of the relative annual growth of G7 countries over the sample period can be useful to users of these forecasts seeking to understand countries' relative growth prospects over the horizon we have used.

We might expect that forecasts of relative growth would become less useful as the size of the forecast differential reduces, i.e. relative growth forecasts may become less useful for years in which nations are forecast to have "broadly similar" growth prospects. We can examine if this is true by only considering those relative growth forecasts which lie close to zero. Specifically, we have recreated Table 3 for all the forecast differentials between -1 and 1. In total, over the sample, there were 267 such forecasts. We calculate an χ^2 statistic for a new 2x2 contingency table for these (small) forecast differentials and whether the outturn differential was positive or negative. For forecast differentials between -1 and 1, the χ^2 statistic is 30.1, which is again significant at the 1% level. Indeed, only for forecast differentials between -0.2 and 0.2 are relative forecasts of growth unable to reject a null of independence at the 1% level for our complete sample of forecast and outturn growth differentials (although the χ^2 statistic is still significant at the 5% level).

5.2 Testing relative forecast accuracy using contingency tables and χ^2 - country comparisons

Table 4 shows the results from examining each of the twenty-one possible two-country combination of G7 countries within our sample and time period. This shows that we also find significant χ^2 statistics for many of the two-country comparisons, indicating that forecasts and outturn growth differentials are not independent. Rows in the table are sorted by the size of the χ^2 statistic, indicating that more significant relationships (less independent forecasts and outturn differentials) are at the top of this table.

[Table 4 here]

Two of the two-country comparisons of forecast and outturn relative growth comparisons (Japan and France, and UK and Japan) are significant at the 1% level. The Japan-France comparison has the highest χ^2 statistic, as well as the highest accuracy of relative forecasts. The relative growth forecast between these countries was correct 25 of 27 times, while between the UK and Japan the sign of forecast growth differential was correct 24 times. As with Figure 3, we can show the quantitative scale and direction of forecast and outturn relative growth differentials for these two significant forecasts. Figure 4 shows the case of Japan and France, while Figure 5 shows the comparison between the UK and Japan.

[Figure 4 here]

[Figure 5 here]

Looking firstly at Figure 4, we can see the high level of agreement between the forecasted and outturn growth for Japan and France. Breaking this period down, in the first part of the sample – from 1984 to 1991 – Japan was forecast to grow more than France, and did so, giving positive forecast and outturn differentials. From 1997 to 2002, the reverse was true, with Japanese growth forecast and outturn growth negative compared to France. From 2003 to 2007 forecast and outturn growth differentials were all small in absolute terms (not exceeding 1 percentage point in any year) with Japan particularly badly hit by the financial crises of 2008 and 2009, but (as forecasted) responding with faster growth than France in 2010.

Figure 5 shows for the UK and Japan a similarly high level of similarity between forecast and outturn growth differentials. Rather than three periods where forecast and outturn differentials varied in sign – like for Japan and France – this comparison has two distinct periods, i.e. pre-1992 (inclusive) and post-1992.Before that year, UK growth was typically (8 out of 9 times) expected to be lower than Japanese growth, and for these years it was only 1986 where that was not a correct prediction. Between 1993 and 2009, however the OECD's prediction for the relative growth of the UK compared to Japan was positive, i.e. the UK was expected to grow faster, and for each of these years (with the exception of 2006) the UK did grow faster than Japan.

Table 4 also shows us that six of the twenty-one two-country comparisons do not have significant χ^2 statistics, indicating that in these cases we cannot reject our null of independence between forecast and outturn relative growth. These cases include the lowest relative forecast accuracy we find, which is between the US and Canada. The OECD's forecasts during the sample period correctly predicted the relative growth between these countries on 59% of the years covered (getting the relative growth incorrect on 11 of the 27 years included).

The comparison of forecasted and outturn growth for the US and Canada is worthy of closer examination. Figure 6 shows the size (and signs) of the forecast and outturn growth differentials over the sample period for this two-country comparison. As before, this Figure is drawn with the same *y*-axis scale as Figures 1, 4 and 5.

[Figure 6 here]

Figure 6 shows that there are many years where the forecast and outturn growth differential between the US and Canada do not have the same sign. There are – as Table 4 indicates – more years where the forecasted differential is negative, but the outturn positive, i.e. where USA growth was forecasted to be weaker than Canada but was not, than the opposite. One interesting feature of Figure 6 is that the absolute magnitudes of the differentials are very small. Many of the years of incorrect relative forecasts we have reported here, are where small negative (positive) forecasted growth differentials have seen small positive (negative) outturn growth differentials. There is no year when the forecasted growth differential was greater than 1.9 percentage points, while the largest outturn growth differential was in 1992, when the USA saw a GDP increase of 2.5 percentage points more than Canada.

We have seen that the US and Canada - countries with a high degree of interdependence between each other – display smaller outturn growth differentials (and indeed also small forecasted growth differentials). This will impact upon the ability of forecasted growth differentials to match with outturn growth differentials. To explore this, Figure 7 shows the correlation between the size of average (mean) absolute outturn growth differentials for each two-country comparison of the G7 countries and the percentage of years that the relative growth

differential matched the outturn growth differential for that comparison (i.e. the final column in Table 4).

[Figure 7 here]

A simple line of best fit between the points in Figure 7 gives an upward sloping line, with an R2 of 37%. This indicates that there is some relationship between the size of outturn growth differentials and the likelihood that relative growth would be correctly forecast. The range around the line – with some significant outliers – suggests that this relationship is more complex than this simple test could capture. For instance, it appears that the relative growth of USA and Canada is less accurately predicted than the relative growth between France and Italy (with relative growth correctly predicted 74% of the time) – despite the mean outturn growth differentials over the sample being (marginally) smaller between France and Italy.

5.   Conclusions

Conventional forecast evaluation approaches compare the forecasted value of an economic variable to the outturn value of that variable. This is useful for understanding the accuracy and efficiency of forecasts, with a large academic literature evaluating the forecasts of different international, private and public forecasters. By looking at the forecasted *relative* growth of national economies, this paper explicitly captures an important dimension of forecasts: The relative forecasts of national economic growth have been used to justify the fiscal stances adopted by governments' across the world. By explicitly considering the direction of the relative

forecast differential between two countries and the outturn growth differential, a standard contingency table test can be used to explore if relative growth forecasts contain information of use to users of forecasts.

We have tested the accuracy of relative forecast of real GDP and outturn data for the G7 countries between 1984 and 2010, produced by the OECD. The results show that in aggregate, forecasts of relative growth are good predictors of actual relative growth. Only for forecast differentials of less than 0.2 in absolute terms can we not reject a null of independence between forecast and outturn growth differentials. Examining country-by-country comparisons, we see that the accuracy of relative growth forecasts has been very good for some country combinations, but less accurate for some countries. There is also shown to be a relatively weak relationship between the size of outturn growth differentials and the accuracy of annual relative growth forecasts during this period.

As the first paper exploring the accuracy of forecasts of relative growth, there is scope for further research into relative forecasts. For instance, we have only considered the forecasts by one international organisation (the OECD) over a relatively short forecasting horizon (i.e. within the year to which the forecast relates). The simultaneous release of forecasts for different countries by other international organisations (such as the IMF) could explore whether different organisations have the same level of success in forecasting relative growth. These organisations typically forecasts for two years into the future, which would allow the same test to be used on longer forecast horizons. Secondly, this paper has examined the G7 countries. Further work could explore if relative growth forecasts offer the same usefulness for developing or other non-G7 countries. Finally, it might be possible to explore and explain some of the differences

between forecasted and outturn relative growth differentials by exposure to country-specific or international shocks.

References

Artis MJ. 1996. How accurate are the IMF's short-term forecasts? Another examination of the world economic outlook. *International Monetary Fund Working Paper*: WP/96/89.

Ashiya M. 2003. The directional accuracy of 15-months-ahead forecasts made by the IMF. *Applied Economics Letters* **10**: 331-333.

Ashiya M. 2006. Are 16-month-ahead forecasts useful? A directional analysis of Japanese GDP forecasts. *Journal of Forecasting* **25**: 201-207.

Batchelor R. 2001. Bias in macroeconomic forecasts. *International Journal of Forecasting*. **23**: 189-203.

Clarke GM, Cooke D. 1998. *A basic course in statistics*. 4th edition. Arnold, London.

Diebold FX, Lopez JA. 1996. Forecast evaluation and combination. *National Bureau of Economic Research (NBER) Technical Working Paper Series, Technical Working Paper No. 192*, March 1996.

Flaherty J. 2012. Speech by the honourable Jim Flaherty, Minister of Finance, at a lunch hosted by the Canadian Club of Toronto. 30th March 2012, available online at http://www.fin.gc.ca/n12/12-035_1-eng.asp.

Granger CWJ. 1996. Can we improve the perceived quality of economic forecasts?. *Journal of Applied Econometrics* **11**: 455-473.

Greer M. 2003. Directional accuracy tests of long-term interest rate forecasts. *International Journal of Forecasting* **19**: 291-298.

Hansard 2011. House of Commons, Tuesday 1st November 2011. Column 739-778.

Key J. 2012. Speech to Waitakere Business Club, 26th January 2012, available online at http://beehive.govt.nz/speech/pm-speech-waitakere-business-club.

Koutsogeorgopoulou V. 2000. A Post-Mortem on Economic Outlook Projections. *OECD Economics Department Working Papers*, No. 274, OECD Publishing. http://dx.doi.org/10.1787/351822846618.

Krkoska L, Taksoz U. 2007. Accuracy of GDP growth forecasts for transition countries: Ten years of forecasting assessed. *International Journal of Forecasting* **23**: 29-45.

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

Loungani P. 2001. How accurate are private sector forecasts? Cross-country evidence from consensus forecasts of output growth. *International Journal of Forecasting* **17**: 419-432.

Melander A, Sismanidis G, Grenouilleau D. 2007. The track record of the Commission's forecasts – an update. *European Economy Economic Papers* 291: October 2007, available online at http://ec.europa.eu/economy_finance/publications/publication9291_en.pdf.

Mellis C, Whittaker R. 1998. The Treasury forecasting record: some new results. *National Institute Economic Review* **164**: 65-79.

Merton R.C. 1981. On market timing and investment performance 1: An equilibrium theory of value for market forecasts. *Journal of Business* **54**: 363-406.

Mills TC, Pepper GT. 1999. Assessing the forecasters: an analysis of the forecasting ,records of the Treasury, the London Business School and the National Institute. *International Journal of Forecasting* **15**: 247-257.

OECD. Various years. *Economic Outlook*, OECD, Paris.

Öller L-E, Barot B. 2000. The accuracy of European growth and inflation forecasts. *International Journal of Forecasting* **16**: 293-315.

Pons J. 2000. The accuracy of IMF and OECD Forecasts for G7 Countries. *Journal of Forecasting* **9**: 53-63.

Schnader MH, Stekler HO. 1990. Evaluating predictions of change. *The Journal of Busines*s **63**: 99-107.

Stekler HO. 1991. Macroeconomic forecast evaluation techniques. *International Journal of Forecasting* **7**: 375-384.

Timmerman A. 2006. An evaluation of the World Economic Outlook Forecasts. IMF working paper, WP/06/59, International Monetary Fund Working Paper, March 2006, available online at http://www.imf.org/external/pubs/ft/wp/2006/wp0659.pdf.

Table 1: Directional accuracy contingency table

		Predicted change		
		Positive	Negative	Total
Actual change	Positive	n_{00}	n_{01}	$n_{0.}$
	Negative	n_{10}	n_{11}	$n_{1.}$
	Total	$n_{.0}$	$n_{.1}$	N

Table 2: Relative forecast accuracy contingency table

		Forecasted relative growth		
		$fd_t^{ab} > 0$	$fd_t^{ab} \leq 0$	Total
Outturn relative growth	$od_t^{ab} > 0$	r_{00}	r_{01}	$r_{0.}$
	$od_t^{ab} \leq 0$	r_{10}	r_{11}	$r_{1.}$
	Total	$r_{.0}$	$r_{.1}$	R

Table 3: Relative forecast accuracy contingency table

		Forecasted relative growth		
		$fd_t^{ab} > 0$	$fd_t^{ab} \leq 0$	Total
Outturn relative growth	$od_t^{ab} > 0$	236	70	306
	$od_t^{ab} \leq 0$	42	219	261
	Total	278	289	567

Table 4: Results from country-by-country comparisons

		Correct		Incorrect		
	X^2 statistic	FD and OD both positive	FD and OD both negative (or same)	FD negative or same, OD positive	FD positive, OD negative or same	% correct
Japan vs. France	16.217***	14	11	1	1	93%
UK vs. Japan	13.232***	16	8	3	0	89%
UK vs. France	10.632**	16	7	4	0	85%
USA vs. France	9.674**	19	5	2	1	89%
USA vs. Japan	8.344**	17	6	2	2	85%
USA vs. Germany	8.344**	17	6	2	2	85%
UK vs. Germany	8.269**	14	8	4	1	81%
USA vs. Italy	8.104**	20	4	1	2	89%
Germany vs. Canada	7.542**	4	19	4	0	85%
Japan vs. Italy	6.006*	13	8	2	4	78%
Germany vs. France	5.542*	7	14	4	2	78%
UK vs. USA	4.768*	5	16	5	1	78%
Japan vs. Canada	4.688*	6	15	3	3	78%
Italy vs. Canada	4.407*	3	20	3	1	85%
UK vs. Italy	4.219*	18	4	3	2	81%
Germany vs. Italy	3.546	9	10	2	6	70%
France vs. Italy	3.546	13	7	5	2	74%
France vs. Canada	3.271	3	19	4	1	81%
Japan vs. Germany	2.984	11	8	3	5	70%
UK vs. Canada	1.500	4	15	6	2	70%
USA vs. Canada	0.390	7	9	7	4	59%

Notes: *** = significant at the 1% level, ** = significant at the 5% level, * = significant at the 10% level.