Cross-country Evidence on the Quality of Fiscal Forecasts [*]

João Tovar Jalles[$] Iskander Karibzhanov[#] Prakash Loungani[+]

April, 2011

ABSTRACT

This paper uses mean and dispersion forecasts from Consensus Economics and finds (in general) evidence for biasedness, inefficiency and lack of accuracy of budget balance forecasts for a sample of 29 countries between 1993-2009, grouped into developed and emerging and developing countries. The distribution of projection errors appears to be slightly twisted to over-prediction of budget balance. Moreover, errors in forecasting the budget balance can be explained by wrong forecasts of GDP growth and inflation forecast errors. The extent of budget balance overprediction and the dispersion of forecasts increases during recessions. "Group think" may be another reason that may explain inefficiency of budget balance forecasts and yet another reason results from failure in accounting international linkages. Finally, there is a sense of "pessimism" amongst forecasters during recovery periods.

Keywords: forecast comparison, business cycles, recession, recovery, herding
JEL codes: C53, E27, E37, E62, D8

[*] Acknowledgments: The authors are grateful to Jair Rodriguez and Angela Espiritu for research assistance. Most of the research was conducted while Joao Tovar Jalles and Iskander Karibzhanov were visiting the IMF whose hospitality was greatly appreciated. The usual disclaimer applies. The views expressed are those of the authors and do not necessarily represent those of the IMF or IMF policy.
[$] Corresponding author: University of Cambridge, Faculty of Economics, Sidgwick Avenue, CB3 9DD, Cambridge, UK. E-mail: jodstj2@cam.ac.uk
[#] University of Minnesota, Department of Economics, 1925 S 4th street STE 4-101 Minneapolis, USA. E-mail: karib003@umn.edu
[+] International Monetary Fund, Research Department, 700 19th Street NW, Washington DC, 20431, USA. E-mail: ploungani@imf.org. Tel.: +1-202-623-7043

1. Introduction

Fiscal issues have come to command first-order importance in the discussion of today's economic policy, particularly in the context of the debt crisis faced by Euro zone countries. The accountability of governments for the use of public funds in western democracies is at the root of budgetary procedures and has, ultimately, led to the creation and continuous development of fiscal forecasting methods. Such accountability requires a clear understanding of the impact of macroeconomic developments and discretionary government action on both revenue and expenditure components. Over the years, a key issue in the design of fiscal policy rules has been the accuracy of government budget forecasts. Not surprisingly, therefore, fiscal forecasting and monitoring has always received attention from policy makers, monetary policy authorities, international economic organizations, financial market analysts, rating agencies, research institutions and the general public.

In Europe, more recently, the need to monitor whether the EU Member States comply with the Maastricht Treaty's and the Stability and Growth Pact's fiscal rules, has fuelled academic interest in fiscal forecasting (see Poplawski-Ribeiro and Rulke, 2011 for a recent analysis). Nevertheless, evaluating forecast performance and accuracy might still be a crucial element of forecasting procedures. It is relevant for surveillance and checking purposes, and allows for improving forecasts by learning from past errors. Sizeable, systematic or biased forecast errors in particular items would (ideally) allow fiscal analysts to identify weaknesses in their forecasting procedures, in terms of methods, discussions or decision-making processes.

This paper proposes to address the following questions: i) How do fiscal forecasts behave and perform statistically between countries and over time? ii) What sensitivity analysis can be made at different forecast horizons, that is, is there a marked difference between current-year and year-ahead predictions? iii) Do forecasters, on average, under or over-predict the budget balance and for how long? iv) Are fiscal forecasts accurate during recessions (and related crises) and recoveries episodes? v) What is the difference between developed and emerging and developing countries and for which type of crises? vi) What happens to forecasts in the run-up to recessions? vii) Is there evidence of herding or "group-think" behaviour in forecast revisions? viii) Can we talk about a sense of optimism or pessimism once recoveries are already in motion?

To answer these questions, we use both mean and dispersion forecasts from Consensus Economics and we rely on time series methods and regression analysis (both cross-sectional and panel) to find evidence for unbiasedness, efficiency and accuracy of budget balance forecasts for a sample of 29 countries between 1993-2009 (using monthly data), grouped into developed and emerging and developing countries.

When addressing at the literature on fiscal forecasting two broad topics emerge. First, the performance of fiscal forecasters has been extensively scrutinized. Government targets, in particular, have been criticized for being consistently biased as a result of setting unrealistic, politically motivated targets (Strauch et al., 2004; Moulin and Wierts, 2006), sometimes leading to claims that public finance projections in Europe should be produced by independent authorities[1]. Since national governments express their views about the outlook for fiscal policy in the form of annual targets and plans rather than projections or forecasts, the activities of revenue estimation and spending planning are key in the elaboration of annual budgets and the determination of (multi-annual) targets. In contrast, all other institutions involved in fiscal forecasting (EC, ECB, National Central Banks, OECD, IMF, national research institutes, financial institutions and rating agencies) aim to assess whether public finances are evolving in line with official budgetary targets, and to provide a timely warning when they are moving away from those targets[2]. Governments' fiscal policies are often under fierce scrutiny, and deviations of actual paths of key budgetary items from those initially planned may generate a great deal of debate and criticism. While the literature on forecast performance broadly supports the view that fiscal forecasts prepared by governments tend to be biased[3] and inaccurate, it seems that the causes for forecast errors have been insufficiently covered in literature Second, much effort has been made in order to identify best forecasting practices; however, no definitive conclusions have been reached.

---

[1] The topic of an independent authority in charge of fiscal (and macroeconomic) projections has recently received a great deal of attention (see, e.g., Fatas et al. 2003).

[2] Most studies on forecast track records tend to signal that projections by the EC for European countries are the most accurate within international organisations publishing fiscal forecasts (in part due to its independence) (Keereman, 1999; Artis and Marcellino, 2001; and Melander et al., 2007). Nevertheless, some recent research (Bruck and Stephan, 2006) has challenged EC projections for presenting a number of shortcomings, including the correlation of forecast errors with the political cycles of some of countries.

[3] Jonung and Larch (2006) claim that in some euro area countries biased forecasts have played an important role in generating excessive deficits in the past.

Our findings suggest that, the distribution of projection errors appears to be slightly twisted to over-prediction of budget balance, which might be evidence for the presence of bias in the pool and this seems to be particularly true for industrialized countries. Overall, forecasts are biased and inefficient as found in the literature. Moreover, errors in forecasting the budget balance-to-GDP ratio can be explained by wrong forecasts of GDP growth and inflation forecast errors. Both in advanced and emerging and developing economies, on average, the extent of budget balance overprediction increases during recessions. During recession (and recovery) episodes forecasts generally appear to be efficient, with a few exceptions depending on the particular type of crises originating them. In addition, the dispersion of forecasts rises during recession episodes. We also find that in both groups of economies, middle revisions in one direction generally are followed by final revisions in the same direction. Moreover, "group think" may be another reason that may explain inefficiency of budget balance forecasts, particularly in emerging and developing economies. Finally, although during recoveries budget balances are also difficult to predict, their forecasts during recovery episodes tend to be more correct on average than those of recessions. Additionally, there is a sense of "pessimism" going on amongst forecasters up to 4 years after the end of the recession in the case of advanced economies.

In addition to this introduction, the paper has 5 more sections organized as follows. Section 2 we briefly present a review of the current state of the literature around fiscal forecasting. The following section describes our data sources and variables' definitions. In Section 4 we simultaneously discuss our methodological approach as well as our first set of results. Section 5 addresses the behaviour and performance of fiscal forecasts around turning-points (recession and recovery episodes). The last section concludes.

2. Literature Review on the Analysis of Fiscal Forecasts

In the academic literature there is a wealth of papers evaluating forecast records, more often for the macroeconomic side of the economy than for the fiscal side. Any accuracy comparison has to be taken with caution. Firstly, the information set available when generating real-time forecasts tends to be much smaller than that available when performing ex-post comparisons[4]. Secondly, evaluating point forecasts of certain variables (such as Net Lending/Borrowing of the General Government) does not allow an exhaustive evaluation of a set of projections where all economic and fiscal variables are jointly determined.

Given the role played by both government revenue and expenditure forecasts in budgeting processes[5], almost all national fiscal policy agencies have implemented some sort of forecasting procedure based on either judgment, simple regression equations, time series methods, structural macro-econometric models, or some combination of these. The need to design a procedure flexible enough to accommodate the day-to-day requirements of fiscal policy decision-making tends to create tensions with the use of appropriate tools. This trade-off has guided practitioners and scholars in the discussion of which procedure would best fit policy requirements and formal correctness.

With these tensions in mind, we find two main themes discussed in the related literature. Firstly, some papers discuss the appropriate procedures for fiscal forecasting, in many cases by means of performance and accuracy comparisons. Secondly, conditional on a given procedure, other papers discuss the properties of the produced forecasts, in terms of (systematic) biases and violation of the rationality hypothesis. However, it is not clear which method fiscal and monetary authorities, international economic organisations, financial market analysts, rating agencies or research institutes should be adopting when preparing their forecasts. Bretschneider et al. (1989) compare the forecasting accuracy of different forecasting methods. Their results favour a combination of judgment and simple econometric equations, against time series and complex econometric models. Grizzle and Klay (1994) and Lawrence et al. (1998) also show evidence for combining judgement and simple methods against more complicated or automated techniques on the basis of transparency. Nazmi and Leuthold (1988) and Baguestani and McNown (1992) still ascertain time series techniques as

[4] For instance, GDP revisions, lags in fiscal data availability, or frequent revisions, and, most importantly, changes in announced policy actions or the appearance of non-announced policy measures, requires a careful check of the information available when making forecasts, so that the assessment is fair and informative.

[5] An overwhelming majority of the existing papers dealing with short- and medium-term fiscal projections or attempting to assess the effect of the cycle on the budget focus on government revenues, in particular on important items such as Wage Taxes, Sales Taxes or Value Added Taxes, and Social Contributions (see Lawrence et al., 1998, and Van den Noord, 2000). However, it is not unusual to see short-term forecasting models of the spending side of the budget as, for example, in Mandy (1989) for unemployment insurance funds, or Tridimas (1992), Giles and Hall (1998), Sentance et al. (1998) and Pike and Savage (1998) for an integrated view of both revenue and spending sides.

viable for tax receipts forecasting, while Litterman and Supel (1983) and Fullerton (1989) provide some evidence to support the combination of different forecasting techniques. Sentance et al. (1998), Pike and Savage (1998), Giles and Hall (1998), Cao and Robidoux (1998) and Willman et al. (2000) present the fiscal side of structural macroeconomic models. Macroeconomic models, as iteration tools for preparing the budgetary forecasts, allow for estimating the effects of fiscal policy on economic activity. Furthermore, they ensure the consistency between the macroeconomic, inflation and budget projections. Nevertheless, it is often the case that such models are too aggregated to produce sufficiently detailed government revenue and expenditure projections, which are necessary for a thorough assessment of public finances.[6] Some international organisations have chosen to follow an iterative process to overcome the shortcomings related to the use of large scale macroeconomic models in the context of the forecast exercises (ECB, 2001).

Conditional on a given procedure, a great deal of literature has analysed the potential bias the political and institutional process might have on revenue and spending forecasts (e.g. Plesko, 1988; Feenberg et al., 1989; Auerbach, 1995, 1996; Jennes and Arabackyj, 1998; Bruck and Stephan, 2006; Jonung and Larch, 2006 and references therein), and the nature and properties of forecast errors within national states (e.g. Gentry, 1989; Baguestani and McNown, 1992; Melliss, 1996; Melliss and Whittaker, 1998; Strauch et al., 2004; Mühleisen et al., 2005; Moulin and Wierts, 2006 and references therein) and international organisations (e.g. Artis and Marcellino, 1998; Pons, 2000; Golosov and King, 2002 and references therein).

The main lessons we can draw from the literature are that: i) there is some consensus on the politically and institutionally-generated bias in fical forecasts; ii) a good forecasting assessment should rely on a combination of different procedures, tools and methods Next we briefly address data related issues and then move into methodological considerations and our first set of results.

3. Data

[6] For a discussion on fiscal blocks in leading macroeconomic models see Bryant and Zhang (1996a and 1996b), European Commission's QUEST model in Roeger and in't Veld (1997), IMF's MULTIMOD in Laxton et al. (1998), and Eurosystem models in Fagan et al. (2001) and Fagan and Morgan (2005).

In the past decade there has been a huge growth in published economic analysis emanating from banks, corporations and independent consultants around the world, and a parallel growth in "consensus forecasting" services which gather together information from these disparate private sources. Each month since 1989, the Consensus Economics service has published forecasts for major economic variables prepared by panels of 10-30 private sector forecasters, initially for the G7 countries but subsequently for over 70 other economies. Below the individual forecasts for each variable, the service publishes their arithmetic average, the "consensus forecast" for that variable. Consensus forecasts are known to be hard to beat. While individual private sector forecasts may be subject to various behavioural biases (Batchelor and Dua, 1992), many of these are likely to be eliminated by pooling forecasts from several forecasters.

This means that, in practice, the most promising alternative to official forecasts for most users of economic forecasts is not some naive model, but a consensus of private sector forecasts. This is recognized by Artis (1996), who makes a visual comparison of IMF and Consensus Economics forecasts for real GDP and CPI inflation, and concludes that there is "little difference between WEO and Consensus errors". In a similar vein, Loungani (2001) plots real IMF and Consensus Economics GDP forecasts for over 60 developed and developing countries in the 1990s, and notes that "the evidence points to near-perfect collinearity between private and official (multilateral) forecasts …"

The data used in this paper comprises the mean and dispersion[7] of the private analysts' monthly consensus forecasts of real output growth, inflation, and budget balance-to-GDP ratios for the current and next year for the period from February 1993 to September 2009. Twenty nine countries are represented in the sample, of which 9 are advanced economies (G7[8] plus Australia and New Zealand) and 20 are emerging and developing economies. The sample is geographically diverse, covering countries in Asia, Europe and Western Hemisphere.

[7] The number of forecasters is greater than 10 for most countries and for the major industrialized countries the panelists are generally based in the countries they forecast.
[8] Some of the results are presented solely for this group of countries.

The "event" being forecasted is annual average budget balance-to-GDP ratio. Every month (or every other month in the case of some emerging and developing economies) a new forecast is made of the event. For each year, the sequence of forecasts is the 24 forecasts made between January of the previous year and December of the year in question.[9] For countries for which only bi-monthly forecasts are available, we use the preceding month forecast as values for the months for which forecast data is missing.[10]

In addition to consensus forecasts, the dataset includes actual data on real GDP growth, inflation and budget balance-to-GDP ratio from the International Monetary Fund's International Financial Statistics and a number of Central Banks and National Statistical institutions. Using quarterly GDP series, recession episodes are identified based on the classical definition of a business cycle using quarterly changes in the level of real GDP (Burns and Mitchell, 1946)[11]. Economies are classified as being in a recession in a given survey month if the respective forecasted fiscal year falls in the recession year as defined following NBER's approach for dating turning points in the business cycle. The dating of crisis follows Laeven and Valencia's (2008) database, extended by one year to 2008[12]. The sample includes 3 banking crisis in advanced economies and 12 banking crisis, 11 currency crisis and 3 debt crisis in emerging and developing economies.

In what follows, we analyse the Consensus' forecasts of budget balance-to-GDP ratios for each of the G7 countries (representing the major economic players) together with some reference groups such as the one aggregating industrialized countries and the one combining emerging and developing countries.

4. Methodology and Results

[9] The Consensus Economics provides forecasts for the public sector budget balance in absolute numbers (local currency). As most fiscal criteria and rules for budgetary discipline are formulated in terms of the balance-to-GDP ratio, we construct these by using actual (realized/observed) national GDP figures as the denominator. Alternatively, one could have followed Heppke-Falk and Hufner's (2004) approach: first, generate projected nominal GDP series (using information from Consensus' real GDP growth and inflation forecasts) and then take the respective ratio. For robustness and consistency purposes, we have checked that the correlation between our final series and one constructed via the latter method is 97 and 98 percent for the current-year and year-ahead forecasts, respectively (statistically significant at the 1% level).

[10] A similar approach is taken in Loungani et al. (2011).

[11] The NBER uses a similar definition for dating turning points in the US business cycle.

[12] For 2009, we have replicated the same information as of 2008 onto the following year.

4.1 Descriptive Statistics and Graphical Analysis

Figure 1 shows the distribution of the government's budget balance-to-GDP ratio forecast errors for both advanced and emerging and developing countries between 1993-2009, which is a presentation not found in other many other studies. The distributions of forecasts tend to converge to the distributions of actual values over time. The distributions of one-year-ahead forecasts (h=21) are centered around the sample mean, especially for advanced economies. As the forecasts horizon narrows, forecast distributions tend to widen, moving closer to the actual distributions (current-year, h=1). The differences between the forecasted and actual distributions persist mainly in the left-hand-side tail of the distribution and especially for emerging and developing economies. Alternatively, as shown in Figur 2, one can plot the Kernel density estimates[13] of the forecast errors of the budget balance for the same forecast horizons. Note that deficit is defined as a negative value, so that a positive forecast error[14] corresponds to overprediction of the budget deficit (or underprediction of the budget balance). The distribution of projection errors appears to be slightly twisted to over-prediction of budget balance, which might be evidence for the presence of bias in the pool. This seems to be particularly true for industrialized countries. In addition, as expected, there seems to be some evidence of increased accuracy as the forecast horizon shortens.

[insert Figures 1 and 2]

From a theoretical point of view (see, e.g., Auerbach, 1999) a series of forecast errors should be analysed on three fronts: accuracy, unbiasedness and efficiency. By accuracy, following Musso and Phillips (2002), we refer to two aspects of the forecast as compared to the actual outcome. The first one is how close both are in quantitative terms by means of a

---

[13] We use a kernel density estimator, which is a way of generalizing a histogram constructed with the sample data. Where a histogram results in a density that is piecewise constant, a kernel estimator results in a smooth density. The smoothing is accomplished by spreading each data point with a kernel, usually a pdf centered on the data point, and a parameter called the bandwidth. We use the Epanechnikov kernel, an inverted-U quadratic curve, which is actually the most efficient kernel and a common choice in econometrics.

[14] Defined as actual realization minus forecasted value.

number of "standard measures", while the second one refers to the capacity of the forecast to predict the direction of change in the final outcome[15].

First, define, for each country *i* during year *t*, $e_{it} = A_{it} - F_{it}$, where e represents forecast error, *F* denotes the forecast, and *A* denotes its respective realization. With regard to the "standard measures", we use three conventional error measures to assess relative performance. The first is forecast average bias (ME), defined as the average difference between the actual value and its forecasted value[16]. The second statistic used is the mean absolute error (MAE)[17]. The third statistic used is the root mean squared error (RMSE)[18]. While these measures have a number of limitations, the RMSE has invariably been used as standard for judging the quality of predictions.[19] A better approach for evaluating performance is to compare descriptive statistics with similar statistics obtained from a naive standard. Using a naive standard permits one to test whether a forecaster's errors are significantly smaller that those of the benchmark. A frequently used naive standard compares a forecaster's errors with those obtained from a no-change (random-walk) naive model. Theil's (1966) U-coefficient formalizes this comparison. If U is less than 1, the forecasts which are being evaluated have smaller errors than those of the naive model, but this result does not guarantee that the former are significantly better than the latter[20].

---

[15] Directional sign-accuracy tends to be less considered in the analysis of fiscal forecasts than quantitative measures (exceptions are Feenberg et al., 1989; Keereman, 1999; Melander et al., 2007 or Pérez, 2007).

[16] A positive value for bias indicates that on average over the whole run of forecasts for the budget balance, the actual value was under-estimated, so that the forecasts were too low. A negative bias indicates that on average the forecasts were too high.

[17] This is the average of all the differences between actual and forecast values, disregarding the sign of the error.

[18] This is computed by taking all errors, disregarding their signs as in the case of the MAE, and squaring them. The RMSE penalizes heavily forecasters who make a few large errors, relative to forecasters who make a larger number of small errors.

[19] The RMSE may be the most popular measure among statisticians partially because of its mathematical tractability. The alternative MAE, however, is, e.g., more robust to outliers. Moreover, the scale of the data often varies considerably among series and the RMSE is not unit-free. Nevertheless, both Academicians and practitioners have a stronger preference for the RMSE. More recently, researchers seem to prefer the so-called Percent Better, the Mean Absolute Percentage Error and the Relative Absolute Error. For a review of the literature, see Armstrong and Collopy, 1992; Fildes, 1992; Fair, 1980; Jenkins, 1982; Baillie et al., 1993.

[20] This naive model serves as an approximate minimum standard of comparison for variables that experience both positive and negative changes. Whatever benchmark is used in the evaluation of forecasts, the difference between the two sets of errors should be tested for statistical significance.

The ME, MAE, RMSE and Theil statistics are reported in Table 1. If we take the full sample period first, one can notice that the mean error is always smaller than one point, and the absolute error is only slightly larger (with the highest value corresponding to emerging and developing countries). Even if we take the sample's middle point and split it into two sub-periods, 1993-2001 and 2001-2009 (corresponding roughly to the sample's midpoint), we do not observe much difference in the behaviour of these measures. More interesting is to look at the year-ahead and current-year performance: as expected the latter presents larger ME, MAE and MSE vis-a-vis the former. With respect to the Theil statistics we observe that they are generally below unity, therefore suggesting that the consensus forecasts are "preferred" to the random-walk model. This looks like a good performance both country-by-country (and groupings) and over time.

[insert Table 1]

Another important point, as outlined above is whether a forecast can be of value to a user in the sense of accurately predicting the direction of change in the actual series (Joutz and Stekler, 2000). To examine the sign forecast accuracy, we construct the following two-by-two contingency table in which the actual and forecast data for each horizon are classified (i) by whether the actual change in a given variable is (+) or (-,0), and (ii) whether the forecast correctly or incorrectly predicted the sign. That is,

Contingency Table	
$n_{11} : \Delta A_{t+h}(+), \Delta F_{t+h}(+)$	$n_{12} : \Delta A_{t+h}(-,0), \Delta F_{t+h}(+)$
$n_{21} : \Delta A_{t+h}(+), \Delta F_{t+h}(-,0)$	$n_{22} : \Delta A_{t+h}(-,0), \Delta F_{t+h}(-,0)$

where the actual change is $\Delta A_{t+h} = A_{t+h} - A_t$ and the predicted change is $\Delta F_{t+h} = F_{t+h} - A_{t-1}$. Note that A_{t-1} is the most recent value known at the time of the forecast. The diagonal cells include numbers of correct sign forecasts and the off-diagonal cells include the numbers of incorrect sign forecasts. In testing the null hypothesis of no association between the actual and predicted changes we use Fisher's exact test. Our results are presented in Table 2. In 2 out of 7 cases (UK and Canada) we reject the null hypothesis in the case of the current year forecasts. Marginally rejected are the USA and France cases with p-values of 11% for the current year forecasts. With respect to the year-ahead forecasts we never reject the null hypothesis, strengthening the claim that fiscal forecasts are hard to carry out, as previously discussed in Section 2.

[insert Table 2]

The final question we address in this sub-section is whether errors in forecasting the budget balance-to-GDP ratio can be explained by wrong forecasts of other relevant macroeconomic variables (see Artis and Marcellino, 2001 for an analysis of IMF's fiscal forecasts). The first candidate is GDP growth - unexpected growth increases of the GDP level and decreases of the budget deficit (translating the effect of the automatic stabilizers); therefore, growth forecast errors are expected to be negatively correlated with deficit to GDP forecast errors (or conversely, positively correlated with the budget-balance to GDP ratio forecast errors). A negative effect of inflation on deficit forecast errors is also possible when the tax indexation system is not perfect. This interpretation is supported by the data. Table 3 reports results from a regression of budget balance-to-GDP forecast errors on growth and inflation forecast errors (eg and eπ). For both current and year-ahead forecasts, *eg* is significant and positive for all the 7 countries. As for inflation, current year forecast errors appear with negative signs twice, however these are statistically insignificant; for the year ahead forecast errors, Germany and Italy present negative and statistically significant coefficients for $e\pi$ at the 1% level. In Model II we run panel fixed effects[21] to take into account the time dimension and country heterogeneity within our sample and results confirm previous findings; that is, positive and statistically significant coefficients for both GDP growth and inflation forecast errors. Moreover, as in Model I, coefficients for advanced economies are higher in magnitude vis-a-vis the ones for emerging and developing countries.

[insert Table 3]

[21] This was selected as the preferred model relative to the random effects alternative with a Hausman-Wu type test.

4.2 Rationality Tests and Unbiasedness

The key for an assessment on "rationality" or "efficiency" lies, firstly, in the available information at the time the forecast was elaborated (data, policy measures), and secondly in the (optimal?) use of this information (methods, procedures).

A comparison of forecast with benchmark standards, as we did in the previous section, determines which one has smaller errors. It does not indicate how to improve upon the observed record. Rationality tests determine whether or not the predictions are optimal with regards to a particular information set (Wallis, 1989). The accuracy of a forecast is based on the properties of the forecast error. A forecast is considered to be accurate if it is unbiased and efficient. A forecast is unbiased if its average deviation from the outcome is zero. In this sense, we start by examining whether the forecasts are biased. As a rule, if forecasts are in line with the Rational Expectations Hypothesis (REH) formulated by Muth (1961), they should be unbiased. The REH states that market participants use all cost-efficient knowledge to forecast economic variables. The literature on fiscal forecasts tends to detect lack of rationality. This conclusion has been reached through investigating: (i) the presence of serial correlation in a time series of errors (as in Gentry, 1989; Campbell and Ghysels, 1995; Artis and Marcellino, 1998 and 2001; Keereman, 1999; Pons, 2000; Bruck and Stephan, 2006; Melander et al., 2007); (ii) the correlation of the forecast errors with information available at the time the forecast was performed (as in Feenberg et al., 1989; Cassidy et al., 1989; or Strauch et al., 2004).

A series of papers test for the presence of bias in fiscal forecast errors by adopting a purely statistical approach, and considering symmetric loss functions on the part of the agencies generating the forecasts. In particular, we use the following specification:

$$A_{t+h} - F_{t+h} = \alpha + \varepsilon_{t+h} \tag{1}$$

where the forecast horizon $h=1,...,24$. A_{t+h} is the actual value of variable x in $t+h$, and F_{t+h} is the forecast for $t+h$ made at t. F_{t+h} is unbiased if we cannot reject the null hypothesis that $\alpha = 0$ (proposed by Holden and Peel, 1990).[22]

Table 4 reports, for 4 different forecast horizons, the t-tests for α=0 for each member country of the G-7, the pooled sample and 2 major groups (aggregating industrialized countries and emerging and developing countries) and considering differentiated time spans. It also shows the Ljung-Box Q-statistic for testing up to 3 lags possible serial correlation in the time series of the forecast errors.

[insert Table 4]

The year-ahead forecasts are marked by optimism which dissipates by Oct_t. But there is a marked difference across the two groups: the bias is much greater for emerging and developing countries than for industrialized economies. Additionally, in 2 out of 4 cases for the full sample case, industrialized countries show no bias (in particular in Oct_t and Apr_t. The null hypothesis of no bias cannot be rejected for the industrialized economies for the subperiod of 1993-2001. For emerging and developing countries (and for the sample of all countries), the null is rejected at selected forecast horizons for the full sample case and about 50% of the cases in the 1993-2001 subperiod. With respect to a country-by-country analysis we find the both the US and Canada present the higher number of rejections of the null of no bias, particularly during the 1993-2001 period. We also note that comparing both subperiods, the 2001-2009 shows more evidence in favour of unbiasedness of forecasts for each individual country. In the current-year forecasts, serial correlation is detected for US, Japan and France; in the year ahead forecasts serial correlation affects the errors of US and Canada (for the full sample). Once we split the sample into two subperiods we find no or little evidence or serial correlation of forecast errors for any of the selected forecast horizons.

A source of the problems of some of the Consensus' deficit forecasts could be the presence of structural breaks over the forecast period due to unmodelled changes in economic policy.

[22] Since the joint hypothesis (of $\alpha' = 0$ and $\beta = 1$) in Eq. 3 is sufficient but not a necessary condition.

This could also explain the good forecasting performance of the naive models, because of their robustness to breaks, e.g. Clements and Hendry (1997). A possible remedy in this case is the "intercept correction" method[23] which is implemented by adding the lagged forecast error to the actual forecast from the model, i.e.,

$$icf_h = f_h + e_{h-1} \tag{2}$$

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[24]. From Table 5, in the case of year-ahead forecasts, it can be seen that there is a generalized decrease in the absolute magnitude of the estimated coefficients, compared with the results shown in Table 4. With the intercept corrected forecast error one can never reject the null of unbiasedness for both current-year and year-ahead forecasts[25], regardless of the country or grouping.[26]

[insert Table 5]

4.3 Efficiency

Following Nordhaus (1997) a sequence of forecasts of the same event must follow a martingale. We now define initial revision of the forecast as the change in the forecast between October and April of the previous year, the middle revision as the change between

[23] The term "intercept correction" comes from the practice of those forecasters who use formal econometric models for forecasting, or absorbing into a correction the constant terms of the model's equations, persistent errors evident from their recent tracking behaviour.

[24] Actually, if the original forecast errors are uncorrelated, we are simply adding an MA(1) term to them because $ice_h = a_h - icf_h = e_h - e_{h-1}$.

[25] While the rejection of such rationality tests suggests that forecasts might have been improved, a caveat should be noted. Jeong and Maddala (1991) show that tests of rationality depend upon the statistical assumptions and procedures that are used. For example, the regression tests are not valid in the presence of unit roots (Pain, 1994). In addition, it is also possible that the forecaster's loss function is asymmetric and the predictions would still be rational even if the Mean Error is non-zero.

[26] Under the null hypothesis of unbiasedness the error term is the forecast error and should then be free of serial correlation (Artis and Marcellino, 2001). Moreover, as pointed out by Romer and Romer (2000), the order of serial correlation increases as the forecast horizon becomes longer. Hence, we apply the Wald Chi-squared test corrected for heteroskedasticity and autocorrelation of the residuals.

April of the current year and October of the previous year, and the final revision as the change between October of the current year and April of the current year. Our results from regressions of later revisions on earlier ones are shown in Table 6. The first set of regressions is for all countries. In all four regressions, there is evidence of a strong positive correlation among forecast revisions (at the 1% significance level). As shown in the next two sets of regressions, efficiency can be rejected for both industrialized and emerging and developing countries (even when the revisions are non-adjacent). Overall, there is a clear tendency for "forecast smoothing", that is, a tendency for a revision in one direction to be followed by further revisions in the same direction.

[insert Table 6]

A second test, in line with Gentry (1989), is performed and regressions following Eq. 3 are reported in Table 7.

$$A_{t+h} = \alpha' + \beta F_{t+h} + \varepsilon_{t+h} \tag{3}$$

where the forecast horizon $h=1,...,24$. A_{t+h} is the actual value of variable x in $t+h$, and F_{t+h} is the forecast for $t+h$ made at t. F_{t+h} is efficient if we cannot reject the null hypothesis that $\alpha' = 0$ and $\beta = 1$ individually and jointly.

In all but two cases (for industrialized economies at Oct_{t-1} and Apr_t, the joint hypothesis of a zero constant and a slope coefficient of unity is rejected, as indicated by the F-statistics and associated p-values reported in the last row of each block of regressions (smaller than the typical 5% level).

[insert Table 7]

Since the data are pooled across countries and over time, there is reason to suspect that the disturbance term, ε_{t+h}, in Eq. 3 would not be random. We attempt to control for some of the possible correlations by augmenting the regression to include year fixed effects and a fixed effect for whether the country is industrialized or emerging or developing. The idea is that some years may be harder to forecast than others for all countries, and industrialized countries may be easier to forecast in all years than emerging and developing economies; these differences are picked up by the fixed effects. Our results are reported in Table 7 - Model II (however, for reasons of parsimony the fixed effects are not reported). Once again, despite two cases to the contrary, the evidence suggests rejection of efficiency.

If we redo the efficiency test with intercept corrected forecasts[27] we keep on finding evidence against efficiency and now this is true for the pooled sample as well as for the industrialized and emerging and developing groupings.

Additionally, forecast efficiency implies that the deviation between the outcome and the projection is not related to information available at the time the projection was made (Barrionuevo, 1993). This condition is tested by measuring the statistical significance of the co-movements between the deviation of the outcome of the forecast and the forecast itself (the β-test) and the co-movement between the deviation of the outcome of the forecast in the current period and that in the previous period (the ρ-test). We estimate β using a least-squares regression of the forecast error on a constant term and the forecast. We estimate ρ with a regression of the current-period forecast error on a constant term and the previous period error. Therefore, a condition for efficiency is that both β and ρ be zero.

$$\begin{aligned} e_t &= \alpha + \beta F_t + u_t \\ e_t &= \gamma + \rho e_{t-1} + u_t \end{aligned} \tag{4}$$

where e and F denote forecast error and forecast, respectively and α and γ are constant terms (with u_t being an i.i.d. residual that under the null hypothesis of unbiasedness coincides with the forecast error - see e.g. Clements and Hendry, 1997). If ρ is different from zero and β is zero, the forecast is inefficient because the errors of the past are repeated in the present, and hence forecasts could be improved by adjusting them by ρ. If β and ρ are both different from zero, the inefficiency is partly due to the way in which new information is incorporated into projected values and partly because the present errors are

[27] Available from authors upon request.

highly correlated with past ones (Barrionuevo, 1993; Pons, 2000). If ρ is zero and β is different from zero, the inefficiency arises because the model used to derive the forecast is not the minimum variance model, and thus the projections could be improved by adjusting them by β. Table 8 presents our results. If we take the full sample panel, we see that only Germany's and the UK's forecasts are efficient in both year ahead and current year forecasts. If we take either the pooled sample or the emerging and developing economies we find that both β and ρ are statistically different from zero, hence inefficient. If we take the 1993-2001 subperiod, we find that US's, France's and Canada's forecasts are efficient, but pooling the 29 countries together still yields an inefficiency result. Finally, for the 2001-2009 subperiod Germany's and Canada's forecasts appear to be efficient. In many other cases, e.g. if ρ is significant the projection could be improved by adjusting for last period's error.

[insert Tables 8]

We also conducted a similar analysis on the time dimension.[28] Only in 1994 we find evidence in favour of efficiency in both current-year and year-ahead projections. Overall, we can conclude that budget balance-to-GDP ratio forecasts are inefficient.

Next we analyse the performance of fiscal forecasts during recession and recovery episodes.

5. Budget Balances, Recessions and Turning Points

The recent financial crisis has revived the interest not only in predicting recessions (especially those associated with crises) but also on the forecasting performance and accuracy of many macro aggregates that could help determining upward and downward movements in the business cycle and its corresponding turning points. In this section we examine the performance of budget balance forecasts in the same group of 29 advanced, emerging and developing economies during recessions, crises and recovery episodes. We recall our research questions: i) How well do fiscal forecasts incorporate news about

[28] Results are not shown for reasons of parsimony.

recessions and/or recovery periods? ii) Do we observe conjuntural accurate revisions taking place? iii) What is the difference between the two groups of countries and for which type of crises? iv) What sort of (useful?) information can fiscal forecasts provide in the runup to a recession?

The analysis covers a long period (17 years) and looks not only to differences in forecasting performance for different groups of economies, but also differences between regular and crisis-related recessions. The analysis covers both mean consensus forecasts and the dispersion of forecasts, and extends to recovery periods. Moreover, we also discuss the possibility of herding behaviour among forecasters.

We start by examining the budget balance-to-GDP forecast performance during recession and crisis-related episodes, using descriptive statistical analysis and simple regressions. Some basic properties of budget balance-to-GDP forecasts during recession years are summarized in Table 9. The first two rows show qualitative properties of the forecasts. As shown in the first row, as of April of the preceding year, the consensus budget balance forecast was for negative balance in 36 of the 48 episodes of recession (75% of the cases). By April of the year of the recession, a deterioration of the budget balance was forecasted in about 80% of the cases (the same applies for October). However, while forecasters do recognize a deterioration of the budget balance during recessions in the year in which they occur, the results in the second row show that the magnitude of the downward revision was almost always underpredicted. For instance, in October the forecasted budget balanced exceeded its actual realization in 26 out of the 48 episodes. The final row shows the average forecast error at the four forecast horizons over all recession episodes and also for the industrialized and emerging and developing countries separately. There is a significant downward bias in the year-ahead April forecasts that only slowly dissipates over time. This bias is larger for the emerging and developing country sample, but the qualitative pattern is similar across countries.

[insert Table 9]

Figure 3 presents a graphical summary of the ME, MAE and RMSE descriptive statistics by forecast horizon and country group[29]. It is immediate that all the three measures are much higher for Emerging and Developing economies than for Industrialized ones. Moreover, as we move from year-ahead towards current-year, we see a steadily decline in both the MAE and MSE for both groups (unconditional line). Moreover, in general, we have negative values for the forecast errors meaning that forecasters overestimate the budget balance, or in other words, they underestimate the budget deficit (in the case of the Industrialized economies, this pattern is reversed by mid of the current year). Looking at recessions, we observe that all the three measures behave much worse for emerging and developing countries. For the latter group, the MSE does not seem to decrease as we approach h=1.

[insert Figure 3]

In Figure 4 we have the Kernel density estimates of the budget balance forecast errors during recessions for different forecast horizons and split in the two groups of countries. Once again, but this time more pronounced, the distribution of projection errors appears to be twisted to over-prediction and this seems to be particularly true for industrialized countries.

[insert Figure 4]

Both in advanced and emerging and developing economies, forecasters tend to slightly overpredict the budget balance on average, and the extent of overprediction increases during recessions, especially those following banking crisis for industrialized countries and currency crisis for emerging and developing countries (Table 10)[30]. For advanced economies, the mean forecast errors are about 5 times larger during recession episodes than unconditional errors and 7 times larger during post-crisis recessions. For emerging and developing economies, the differences are 2.5 and 6 times, respectively. For recessions following banking and debt crisis for emerging and developing countries, forecast errors are smaller (not shown). The same

[29] Similar graphs for individual advanced countries are available from the authors upon request.
[30] Only the full set of recessions (recoveries) is presented; detailed information of the different types of crises (banking, debt and currency) is not shown for reasons of parsimony but it is available upon request.

pattern holds in general for the absolute forecast errors. Like Dovern, Fritsche and Slacelek (2009)[31], we also find that the dispersion of forecasts rises during recession episodes, especially those associated with banking crises for industrialized countries and currency crises in emerging and developing economies. The regression analysis confirms the above findings based on descriptive statistics (Table 11). In particular, dispersion increases in both recession and recovery period for both industrialized and emerging and developing economies. Moreover, recessions are positively and statistically correlated with absolute forecast errors for both groups of countries and these coefficients are higher when recessions follow a banking crisis. For emerging and developing countries however, there are two occasions (recessions following debt and currency crises) where we find statistically significant negative coefficients.

[insert Tables 10 and 11]

The analysis of the time profile of forecast errors shows that forecasters do anticipate a deterioration of the budget balance in the middle of the year preceding the recession year (Figure 5)[32]. A year prior to the recession year, forecasts are close to the unconditional average, especially for advanced economies. Around July-October of that year, forecasters start marking down forecasts, although they continue to overestimate the budget balance well into the recession year. This pattern holds both for advanced and emerging and developing economies. During recessions after banking crises in emerging and developing countries, forecasters usually start marking down forecasts of the budget balance at the beginning of the recession year, but up to that point we see a gradually increase in forecast errors.

Dispersion tends to rise in the run-up to recession, more visibly in emerging and developing economies. In advanced economies dispersion rises sharply prior to recessions associated with banking crisis (Figure 6).

[insert Figures 5 and 6]

[31] Who look at 6 other macro indicators from Consensus Economics: consumer-price inflation, nominal three-month interest rate, GDP growth, consumption growth, investment growth and unemployment rate.
[32] For emerging and developing countries similar graphs are available from the authors upon request for recessions after debt and currency crises.

5.1 On the efficiency properties of budget balances during recession episodes

As before, we examine efficiency by conducting the two types of tests described in Section 3. In what follows, we re-run those regressions but we now include interaction terms for banking, currency and debt crisis and for recession episodes to ascertain differences in the budget balance forecast efficiency during these episodes compared to averages[33]. The results suggest that on average forecast are inefficient, both in advanced and emerging and developing economies, but not during recession episodes (Table 12 - focus on specifications 3 and 6)[34]. In both groups of economies, middle revisions in one direction generally are followed by final revisions in the same direction (the coefficients on middle revisions are statistically significant). The hypothesis of forecast efficiency cannot be rejected for nonadjacent revisions, i.e., initial revisions for the case of industrialized economies. During recession (and recovery) episodes forecasts generally appear to be efficient, with a few exceptions depending on the particular type of crises originating them.

[insert Table 12]

Finally, following Eq. 3 in Section 4.3 the results for the joint hypothesis for the slope and intercept coefficients using the Wald tests is displayed in Table 13. The hypothesis of forecast efficiency is overwhelmingly rejected, including for recession episodes (apart from one exception for recessions after banking crises for emerging and developing countries). These findings appear stronger than those of the previous test.

[insert Table 13]

---

[33] We use Wald tests to check if coefficients on the initial and middle revisions are statistically different from zero.

[34] Specifications 1 and 4, and 2 and 5 do not include (on purpose) initial revisions (and interaction terms) and middle revisions (and interaction terms), respectively.

It is also possible that some herding behavior or "group thinking" is taking place among forecasters. If this is the case, dispersion of budget balance forecasts is likely to increase during periods of increased uncertainty, for example, those associated with recessions. A simple regression of dispersion of forecasts on monthly dummies in the run-up to recessions shows contradictory results for advanced countries (Table 14). Dispersion tends to decline in the three months prior to the recession (although these coefficients are insignificant). Between six and three months prior to the recession episode we find positive and statistically significant coefficients, hence favouring the conjecture of "group thinking". For emerging and developing economies, the results show that dispersion tends to rise in the run-up to a recession, with all but one coefficient positive and highly statistically significant. In both groups of economies dispersion tends ot rise with the horizon of the forecast as expected. The finding suggest that besides the failure to account fully for linkages with emerging economies, "group think" may be another reason that may explain inefficiency of budget balance forecasts. In emerging and developing economies, a rise in dispersion may serve as an early warning indicator of hard (fiscal) times ahead.

[insert Table 14]

5.2 What about recoveries?

The above analysis has documented failures in forecasting of recessions and discussed possible reasons that may explain them. Naturally a question arises about forecast performance during recoveries, another turning point of the business cycle. We explore this question briefly. In Figure 4, if we now look at the recoveries' line, we see that by April of the recovery year in advanced economies budget balances' forecast errors turn positive, meaning that we have some overestimation of the budget deficits and the forecasters' pessimism persist. Note that for emerging and developing economies the forecast errors switch to positive in October of the year preceding the recovery. Figure 7 presents the Kernel density estimates of the budget balance forecast errors during recoveries for different forecast horizons and split in the two groups of countries.

The main finding is that although during recoveries budget balances are also difficult to predict, their forecasts during recovery episodes tend to be more correct on average than those of recessions, possibly because knowledge that the economy is in a recession encourages forecasters to anticipate a recovery (Tables 10 and 11). In advanced economies, forecasters tend to overpredict budget balances during recoveries (Table 10). In emerging and developing economies, forecasters also tend to overpredict budget balances during recoveries. Budget balances' improvements during recoveries are also apparently difficult to predict, with most forecasters not being able to anticipate a budget balance reversal until about 3 months prior to an "official" recovery starts taking place (Figure 8). Hence, there is a sense of "pessimism" going on amongst forecasters up to 4 years after the end of the recession in the case of advanced economies; in other words, the "actuals" line only shifts upwards after 3 years but the budget balance predictions' path only improve as of 4 years after the end of the recession and up until then they overestimate the budget deficit. For emerging and developing countries "pessimism" lasts for a much shorter period with both "actuals" shifting in the "right" direction and budget balance forecasts being revised upward (and fast) as we move away from the end of the recession. Finally, the evidence on the efficiency of forecasts during recoveries appears to be similarly mixed, as in the case of recessions (Tables 12 and 13).

[insert Figures 7 and 8]

6. Concluding remarks

This paper assessed the performance of budget balance forecasts, paying special attention to business cycles' turning points and uncovering similarities and differences between advanced and emerging and developing countries. By means of the use both monthly mean and dispersion forecasts from Consensus Economics and relying, mostly, on time series methods and regression analysis we find, generally speaking, evidence for biasedness, inefficiency and lack of accuracy of budget balance forecasts for a heterogeneous sample of 29 countries between 1993-2009. Our results suggest that, the distribution of projection errors appears to be slightly twisted to over-prediction of budget balance and this seems to be particularly true for industrialized countries. In addition, as expected, there seems to be some evidence of increased accuracy as the forecast horizon shortens. Overall, forecasts are biased and inefficient as found in the literature (the bias is much greater for emerging and developing countries than for industrialized economies). Moreover, errors in forecasting the budget balance-to-GDP ratio can be explained by wrong forecasts of GDP growth and inflation forecast errors. Both in advanced and emerging and developing economies, on average, the extent of budget balance overprediction increases during recessions, especially those following banking crisis for industrialized countries and currency crisis for emerging and developing countries. During recession (and recovery) episodes forecasts generally appear to be efficient, with a few exceptions depending on the particular type of crises originating them. In addition, the dispersion of forecasts rises during recession episodes, especially those associated with banking crises for industrialized countries and currency crises in emerging and developing economies. We also find the in both groups of economies, middle revisions in one direction generally are followed by final revisions in the same direction. Moreover, "group think" may be another reason that may explain inefficiency of budget balance forecasts, particularly in emerging and developing economies, as a rise in dispersion may serve as an early warning indicator of hard (fiscal) times ahead. Finally, although during recoveries budget balances are also difficult to predict, their forecasts during recovery episodes tend to be more correct on average than those of recessions. Additionally, there is a sense of "pessimism" going on amongst forecasters up to 4 years after the end of the recession in the case of advanced economies.

References

1. Armstrong, J. S. and Collopy, F. (1992), "Error measures for generalizing about forecasting methods: Empirical comparisons," *International Journal of Forecasting*, 8, 69-80.
2. Artis, M.J. and Marcellino, M. (1998), "Fiscal Solvency and Fiscal Forecasting in Europe", CEPR Discussion Paper 1836.
3. Artis, M.J. and Marcellino, M. (2001), "Fiscal forecasting: The track record of the IMF, OECD and EC", *Econometrics Journal* 4, 20-36.
4. Artis, M.J. (1996), "How accurate are the IMF's short term forecasts? Another examination of the World Economic Outlook", IMF Research Department Working Paper.
5. Auerbach, A.J. (1995), "Tax Projections and the Budget: Lessons from the 1980's", *American Economic Review* 85, 165-169.
6. Auerbach, A.J. (1996), "Dynamic Revenue Estimation", *Journal of Economic Perspectives* 10, 141-157.
7. Auerbach, A.J. (1999), "On the Performance and Use of Government Revenue Forecasts, *National Tax Journal* 52, 765-782.
8. Baguestani, H. and McNown, R. (1992), "Forecasting the Federal Budget with Time series Models", *Journal of Forecasting* 11, 127-139.
9. Baillie R. T., Bollerslev T., and Mikkelsen H.-O. (1993), "Fractionally integrated generalized autoregressive conditional heteroskedasticity", Kellogg Graduate School of Management, Northwestern University, working paper 168, 1-24.
10. Barrionuevo, J. M. (1993), "How accurate are the World Economic Outlook projections?", IMF Staff Studies for World Economic Outlook.
11. Batchelor, R. and Dua, P. (1995), "Forecaster diversity and the benefits of combining forecasts", *Management Science*, 41, 1, 68-75.
12. Burns, A. F., and Wesley C. M. (1946), "Measuring Business Cycles", New York, New York: National Bureau of Economic Research.
13. Bretschneider, S.I., Gorr, W.L., Grizzle, G. and Klay, E. (1989), "Political and organizational influences on the accuracy of forecasting state government revenues", *International Journal of Forecasting* 5, 307-319.
14. Bruck, T., and Stephan, A. (2006), "Do Eurozone Countries Cheat with their Budget Deficit Forecasts", *Kyklos* 59, 3-15.
15. Campbell, B. and Ghysels, E. (1995), "Federal Budget Projections: a nonparametric assessment of bias and efficiency", *Review of Economics and Statistics* 77, 17-31.
16. Cao, J.-G. and Robidoux, B. (1998), "The Canadian Economic and Fiscal Model" 1996 version, Department of Finance, Canada, Working Paper 98/07.
17. Cassidy, G., Kamlet, M. S. and Nagin, D. S. (1989), "An empirical examination of bias in revenue forecasts by state governments", *International Journal of Forecasting* 5, 321-331.
18. Clements, M.P. and Hendry, D. F. (1997), "Intercept corrections and structural change", *Journal of Applied Econometrics*, 11, 475-494.
19. Dovern, J., Fritsche, U. and Slacalek, J. (2009), "Disagreement among Forecasters in G7 Countries," Macroeconomics and Finance Series 200906, Hamburg University, Department Wirtschaft und Politik.

20. European Central Bank (2001), "A Guide to Eurosystem Staff Macroeconomic Projection Exercises", European Central Bank, Frankfurt am Main, Germany, June.

21. Fagan, G, Henry, J. and Mestre, G. (2001), "An Area-Wide-Model (AWM) for the euro area", European Central Bank Working Paper 42.

22. Fagan, G. and Morgan, J. (eds.) (2005), Econometric models of the euro-area central banks, Edgar Elgar Publishing, Cheltenham, UK.

23. Fair, R. C. (1980), "Estimating the Expected Predictive Accuracy of Econometric Models," *International Economic Review* 21, 355--378.

24. Fatas, A., von Hagen, J., Hughes A. H., Strauch, R. and Sibert, A. (2003), "Stability and Growth in Europe: Towards a better Pact", Center for Economic Policy Research, London, UK.

25. Feenberg, D.R., Gentry, W., Gilroy, D. and Rosen, H.S. (1989), "Testing the rationality of State Revenue Forecasts", *Review of Economics and Statistics* 71, 300-308.

26. Fildes, R. (1992), "The evaluation of extrapolative forecasting methods," *International Journal of Forecasting*, 8, 81-98.

27. Fullerton Jr., T.M. (1989), "A composite approach to forecasting state government revenues: Case study of the Idaho sales tax", *International Journal of Forecasting* 5, 373-380.

28. Gentry, W.M. (1989), "Do State Revenue Forecasters Utilize Available Information?", *National Tax Journal* 42, 429-39.

29. Giles, C. and Hall, J. (1998), "Forecasting the PSBR Outside Government: The IFS Perspective", *Fiscal Studies* 19, 83-100.

30. Golosov, M. and King, J. (2002), "Tax Revenue Forecasts in IMF-Supported Programs", International Monetary Fund, WP/02/236.

31. Grizzle, G.A. and Klay, W. E. (1994), "Forecasting State Sales Tax Revenues: Comparing the Accuracy of Different Methods", *State and Local Government Review* 26, 142-152.

32. Heppe-Falk, K. and Hufner, F. (2004), "Expected budget deficits and interest rate swap spreads: evidence for France, Germany and Italy", Deutsche Bundesbank Discussion Paper 40/2004.

33. Holden, K. and Peel, D. A. (1990), "On testing for unbiasedness and efficiency of forecasts", *Manchester School* 63, 120--127.

34. Jenkins, Gwilym, M. (1982), "Some practical aspects of forecasting in organizations," *Journal of Forecasting*, 1,3-21.

35. Jennes, B. and Arabackyj, N. (1998), "Budget forecasting records of the federal and provincial governments", *Monthly Economic Review* XVII, 1, Canada.

36. Jeong, J. and Maddala, G.S. (1991), "Measurement Errors and Tests for Rationality", *Journal of Business and Economic Statistics* 9(4), 431-439.

37. Jonung, L. and Larch, M. (2006), "Fiscal policy in the EU: are official output forecasts biased?", *Economic Policy*, 491-534.

38. Joutz, F. and Stekler, H.O. (2000). "An Evaluation of the Predictions of the Federal Reserve," *International Journal of Forecasting*, 16, 17-38.

39. Keereman, F. (1999), "The track record of the Commission Forecasts", *Economic Papers* 137.

40. Lawrence, K., Anandarajan, A. and Kleinman, G. (1998), "Forecasting State Tax Revenues: a new approach" in *Advances in Business and Management Forecasting*, 2, 157-170.

41. Laxton, D., Isard, P., Faruqee, H., Prasad, E. and Turtelboom, B.(1998), "MULTIMOD Mark III: The core dynamic and steady-state models". Occasional Paper 164, IMF, Washington DC.

42. Laeven L. and Valencia, F. (2008), "Systemic banking crises: a new database", IMF Working Paper, WP/08/224.

43. Litterman, R.B., T.M. Supel (1983), "Using Vector Autoregressions to Measure the Uncertainty in Minnesota's Revenue Forecasts", Federal Reserve Bank of Minneapolis Quarterly Review, spring, 10-22.

44. 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.

45. Loungani, P., Stekler, H. and Tamirisa, N. (2011), "Cross-country evidence on forecasting turning-points: consensus and disagreement", IMF mimeo

46. Mandy, D.M. (1989), "Forecasting Unemployment Insurance Trust Funds: The case of Tennessee", *International Journal of Forecasting* 5, 381-391.

47. Melander, A., Sismanidis, G. and Grenouilleau, D. (2007), "The track record of the Commission's forecasts - an update", European Economy, Economic Papers 291, DG ECFIN, European Commission, Brussels.

48. Melliss, C. (1996), "The Treasury forecast record: an evaluation, ESCR Macroeconomic Modeling Bureau Discussion Paper 47.

49. Melliss, C. and Whittaker, R. (1998), "The Treasury forecast record: some new results", National Institute Economic Review, April, 65-79.

50. Moulin, L. and Wierts, P. (2006), "How Credible are Multiannual Budgetary Plans in the EU?", in Fiscal Indicators, 983-1005, Banca d'Italia.

51. Mühleisen, M, Danninger, S., Hauner, D., Krajnyáck, K. and Sutton, B. (2005), "How do Canadian Budget Forecasts Compare with Those of Other Industrial Countries?", IMF Working Paper WP/05/66.

52. Musso, A. and Phillips, S. (2002), "Comparing Projections and Outcomes of IMF-Supported Programs", *IMF Staff Papers*, 49(1).

53. Muth, J. (1961), "Rational Expectations and the Theory of Price Movements," *Econometrica* 29, 315--335.

54. Nazmi, N. and Leuthold, J.H. (1988), "Forecasting Economic Time Series that Require a Power Transformation: Case of State Tax Receipts", *Journal of Forecasting* 7, 173-184.

55. Nordhaus, W.D. (1997), "Forecasting efficiency: concepts and applications", *Review of Economics and Statistics* 69, 667--674.

56. Pain, N. (1994), "Monetary Policy Changes and Unit Root Statistics", *Bulletin of Economic Research*, 46(2), 139-45.

57. Pérez, J. J. (2007), "Leading Indicators for Euro Area Government deficits", *International Journal of Forecasting* 23, 259-275

58. Pike, T. and Savage, D. (1998), "Forecasting the Public Finances in the Treasury", *Fiscal Studies* 19, 49-62.

59. Plesko, G.A. (1988), "The accuracy of government forecasts and budget projections", *National Tax Journal* 41, 483-501.

60. Pons, J. (2000), "The accuracy of IMF and OECD forecasts for G-7 countries", *Journal of Forecasting* 19, 53-63.

61. Poplawski-Ribeiro, M. and Rulke, J-C. (2011), "Fiscal expectations under the Stability and Growth Pact: Evidence from Survey Data", IMF Working Paper wp/11/48.

62. Roeger W., J. in't Veld (1997), "QUEST: A multi-country business cycle and growth model". European Commission paper II/505-97-EN

63. Romer, C. D. and Romer, D. H. (2000), "Federal Reserve Information and the Behavior of Interest Rates", *American Economic Review*, 429-457.

64. Sentance, A., Hall, S. and O'Sullivan, J. (1998), "Modeling and Forecasting UK Public Finances", *Fiscal Studies* 19, 63-81.

65. Strauch, R., Hallerberg, M. and Von Hagen, J. (2004), "Budgetary Forecasts in Europe -- The Track Record of Stability and Covergence Programmes", European Central Bank WP 307.

66. Theil, H. (1966), "Applied economic forecasting", North-Holland, Amsterdam (1966).

67. Tridimas, G. (1992), "Budgetary deficits and government expenditure growth: toward a more accurate empirical specification", *Public Finance Quarterly* 20, 275-297.

68. Van den Noord, P. (2000), "The size and role of automatic fiscal stabilizers in the 1990s and beyond", OECD Economics Department Working Paper 230.

69. Wallis, K. F. (1989), "Macroeconomic forecasting: a survey", *Economic Journal*, 28-61.

70. Willman, A., Kortelainen, M., Männistö, H-L. and Tujula, M. (2000), "The BOF5 macroeconomic model of Finland, structure and dynamic microfoundations", *Economic Modeling* 17, 275-303.

Table 1. Descriptive Statistics

Stat.	USA	Japan	Germany	France	Italy	UK	Canada	All-pooled	Industrial	Emerg+Devel.
					Full Sample					
ME	-0.18	-0.47	0.20	-0.52	-0.37	-0.62	0.27	-0.41	-0.15	-0.62
MAE	0.97	1.73	1.01	0.84	1.01	1.16	0.74	1.61	1.01	2.08
RMSE	2.50	5.02	1.79	1.61	2.09	3.00	1.03	7.34	2.19	11.29
					1993-2001					
ME	0.49	-1.20	0.72	-0.28	0.19	-0.29	0.72	-0.45	0.06	-1.06
MAE	0.63	1.71	1.05	0.70	0.73	1.10	0.92	1.49	0.89	2.22
RMSE	0.69	4.08	1.89	0.83	0.92	1.75	1.25	7.91	1.40	15.77
					2001-2009					
ME	-0.73	0.13	-0.23	-0.71	-0.39	-0.88	-0.10	-0.40	-0.31	-0.44
MAE	1.25	1.75	0.98	0.95	0.97	1.20	0.59	1.68	1.10	2.02
RMSE	3.98	5.77	1.71	2.24	1.52	4.02	0.86	7.03	2.81	9.54
					Full Sample – Year Ahead					
ME	-0.50	-0.65	0.02	-0.71	-0.17	-0.74	0.24	-0.54	-0.30	-0.73
MAE	1.66	2.02	1.27	1.09	1.11	1.38	0.84	1.93	1.33	2.38
RMSE	5.46	6.70	2.57	2.51	1.97	4.13	1.34	8.54	3.46	12.42
Theil	0.45	0.45	0.42	0.34	0.32	0.41	0.40	-	-	-
					Full Sample – Current Year					
ME	0.01	-0.16	0.37	-0.33	-0.08	-0.41	0.30	-0.28	0.01	-0.50
MAE	0.56	1.24	0.75	0.59	0.62	0.79	0.57	1.28	0.68	1.75
RMSE	0.72	2.22	1.02	0.71	0.53	1.13	0.52	6.09	0.88	10.11
Theil	0.21	0.32	0.36	0.23	0.20	0.34	0.37	-	-	-

Source: Authors' estimates.

Notes: This table presents some descriptive statistics for 7 Developed countries, the entire sample of N=29 countries and two sub-samples containing either industrialized or emerging and developing countries. For information on the composition of each sub-sample refer to the main text. ME, MAE, RMSE and Theil stand for the mean forecast error, the absolute forecast error, the mean square forecast error and U Theil statistic, respectively (for details see the main text).

Table 2. Contingency Table – Sign Forecast Accuracy

H	Country	Correct		Incorrect		p-value
		$Y_{t+f}(+), \widehat{Y}_{t+f}(+), n_{11}$	$Y_{t+f}(-,0), \widehat{Y}_{t+f}(-,0), n_{22}$	$Y_{t+f}(-,0), \widehat{Y}_{t+f}(+), n_{12}$	$Y_{t+f}(+), \widehat{Y}_{t+f}(-,0), n_{21}$	
Current Y.	**USA**	1	3	9	3	0.11
Y. Ahead		0	5	9	2	0.18
Current Y.	**Japan**	4	5	3	4	0.34
Y. Ahead		5	6	3	2	0.14
Current Y.	**Germany**	2	3	7	4	0.16
Y. Ahead		3	3	5	5	0.24
Current Y.	**France**	5	6	4	1	0.11
Y. Ahead		3	6	5	2	0.36
Current Y.	**Italy**	6	3	3	4	0.37
Y. Ahead		6	4	3	3	0.26
Current Y.	**UK**	8	5	1	2	**0.02**
Y. Ahead		5	5	3	3	0.24
Current Y.	**Canada**	3	1	5	7	**0.06**
Y. Ahead		3	2	5	6	0.14

Source: Authors' estimates.

Notes: The first column indicates whether the forecasts refer to the current year of year ahead. Each entry from column 3-6 presents the number of counts. $\Delta A_{t+h} = A_{t+h} - A_{t-1}$ is the actual change, and $\Delta F_{t+h} = F_{t+h} - A_{t-1}$ is the predicted change in each variable. A_{t-1} is the most recent (quarterly) rate known at the time of the forecast. (+) indicates that ΔA_{t+h} $(or\ \Delta F_{t+h})$ is positive; (-,0) indicates that ΔA_{t+h} $(or\ \Delta F_{t+h})$ is negative or zero. P-value (for the Fisher's exact test) is for testing the null hypothesis of no association between the direction of change in the actual and forecast series. A Chi-square Pearson statistic would give equivalent results. A bold indicates a P-value smaller than 10%.

Table 3. The role of growth and inflation forecast errors

	egC	$e\pi C$	egY	$e\pi Y$	egC	$e\pi C$	egY	$e\pi Y$
Model	I				II			
Spec.	(1)	(2)	(3)	(4)				
USA	0.52***	0.55***	1.05***	0.72***	-	-	-	-
	(0.08)	(0.19)	(0.04)	(0.07)				
Japan	0.33**	-0.08	0.60***	1.14***	-	-	-	-
	(0.13)	(0.64)	(0.04)	(0.15)				
Germany	0.99***	0.37**	0.71***	-0.35***	-	-	-	-
	(0.05)	(0.15)	(0.03)	(0.09)				
France	0.36***	0.47***	0.82***	0.23***	-	-	-	-
	(0.09)	(0.17)	(0.04)	(0.07)				
Italy	0.55***	0.013	0.59***	-0.37***	-	-	-	-
	(0.05)	(0.14)	(0.04)	(0.09)				
UK	0.66***	0.29***	1.29***	0.47***	-	-	-	-
	(0.17)	(0.10)	(0.02)	(0.07)				
Canada	0.39***	0.10	0.39***	0.35***	-	-	-	-
	(0.09)	(0.15)	(0.04)	(0.11)				
Industrial	0.39***	0.15***	0.71***	0.25***	0.487***	0.005	0.618***	0.348**
	(0.03)	(0.05)	(0.02)	(0.04)	(0.125)	(0.197)	(0.077)	(0.147)
Emerg+Devel.	0.27***	-0.0003	0.33***	0.01***	0.129***	0.01***	0.245***	0.015***
	(0.02)	(0.0003)	(0.01)	(0.001)	(0.043)	(0.0002)	(0.055)	(0.0008)
Pooled	0.224**	0.00002	0.305***	0.0025*	0.18***	0.009***	0.299***	0.015***
	(0.089)	(0.004)	(0.043)	(0.0014)	(0.05)	(0.001)	(0.063)	(0.001)

Source: Authors' estimates.

Notes: This table presents the estimates of a regression of the budget-balance-to-GDP ratio forecast errors on the forecast errors of both GDP and inflation for 7 developed countries plus two sub-samples containing either industrialized or emerging and developing countries. Model I estimates by OLS whereas Model II estimates with panel fixed effects (this was selected as the preferred model relative to the random effects alternative with a Hausman-Wu type test). For information on the composition of each sub-sample refer to the main text. Heteroskedastic-consistent robust standard errors are reported in parenthesis. For reasons of parsimony the constant is not reported. *, **, *** indicate significance at 10, 5 and 1% levels, respectively. The suffices Y and C refer to year-ahead and current-year forecasts.

Table 4. Rationality, Unbiasedness and Serial Correlation

	USA	Japan	Germany	France	Italy	UK	Canada	All-pooled	Industrial	Emerg+Devel.
Spec.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Full Sample									
	Apr.(t-1)									
α	-0.68	-0.43	-0.10	-0.78**	-0.22	-0.88	0.24	-0.65***	-0.41**	-0.82***
	0.66	0.70	0.42	0.37	0.38	0.54	0.30	0.16	0.17	0.25
Q(1)	0.02	0.06	0.17	0.06	0.11	0.12	0.01	-	-	-
Q(2)	0.06	0.16	0.29	0.10	0.27	0.24	0.03	-	-	-
Q(3)	0.08	0.28	0.11	0.04	0.45	0.32	0.06	-	-	-
	Oct.(t-1)									
α	-0.46	-0.70	0.13	-0.52	-0.11	-0.75	0.22	-.43***	-0.16	-0.64***
	0.53	0.58	0.36	0.31	0.29	0.45	0.29	0.15	0.12	0.24
Q(1)	0.06	0.02	0.32	0.37	0.71	0.57	0.10	-	-	-
Q(2)	0.16	0.06	0.32	0.65	0.84	0.66	0.25	-	-	-
Q(3)	0.22	0.11	0.18	0.24	0.87	0.69	0.43	-	-	-
	Apr.(t)									
α	-0.11	-0.01	0.30	-0.34*	-0.17	-0.43*	0.35*	-0.28**	-0.05	-0.46**
	0.25	0.39	0.24	0.19	0.20	0.24	0.18	0.13	0.08	0.22
Q(1)	0.02	0.04	0.82	0.33	0.56	0.09	0.10	-	-	-
Q(2)	0.06	0.06	0.91	0.57	0.85	0.20	0.25	-	-	-
Q(3)	0.04	0.13	0.80	0.25	0.81	0.28	0.36	-	-	-
	Oct.(t)									
α	0.18*	-0.17	0.42*	-0.23	0.003	-0.42	0.33*	-0.33**	0.13**	-0.60**
	0.10	0.32	0.20	0.14	0.13	0.26	0.17	0.16	0.06	0.24
Q(1)	0.10	0.32	0.20	0.14	0.14	0.26	0.17	-	-	-
Q(2)	0.90	0.08	0.94	0.03	0.30	0.59	0.57	-	-	-
Q(3)	0.99	0.09	0.57	0.08	0.39	0.83	0.55	-	-	-

(cont.)	USA	Japan	Germany	France	Italy	UK	Canada	All-pooled	Industrial	Emerg+Devel.
Spec.	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
					1993-2001					
					Apr.(t-1)					
α	1.11**	-1.61	0.53	-0.16	0.47	0.24	0.70*	-0.65**	0.15	-1.54***
	0.31	0.71	0.59	0.35	0.54	0.59	0.35	0.29	0.20	0.53
Q(1)	0.94	0.66	0.20	0.17	0.67	0.16	0.86	-	-	-
Q(2)	1.00	0.39	0.44	0.36	0.91	0.37	0.38	-	-	-
Q(3)	0.34	0.59	0.64	0.36	0.91	0.39	0.52	-	-	-
					Oct.(t-1)					
α	0.70**	-1.13	0.57	-0.22	0.06	-0.18	1.10***	-.50**	0.13	-1.22**
	0.29	0.64	0.48	0.30	0.40	0.47	0.26	0.25	0.14	0.50
Q(1)	0.78	0.24	0.19	0.29	0.16	0.07	0.61	-	-	-
Q(2)	0.95	0.47	0.39	0.56	0.32	0.18	0.56	-	-	-
Q(3)	0.64	0.54	0.60	0.65	0.51	0.26	0.76	-	-	-
					Apr.(t)					
α	0.61***	-0.66	0.78*	-0.24	0.11	-0.40	0.82**	-0.37	0.05	-0.93*
	0.13	0.45	0.39	0.36	0.22	0.51	0.28	0.25	0.13	0.56
Q(1)	0.37	0.18	0.65	0.45	0.03	0.11	0.88	-	-	-
Q(2)	0.20	0.40	0.86	0.73	0.02	0.27	0.04	-	-	-
Q(3)	0.28	0.42	0.95	0.88	0.04	0.41	0.09	-	-	-
					Oct.(t)					
α	0.24***	-0.63	0.79*	-0.36	-0.03	-0.78*	0.50	-0.35	0.14	-0.74
	0.06	0.46	0.35	0.29	0.16	0.37	0.33	0.31	0.11	0.54
Q(1)	0.51	0.38	1.00	0.07	0.17	0.80	0.72	-	-	-
Q(2)	0.19	0.42	0.66	0.20	0.38	0.87	0.66	-	-	-
Q(3)	0.30	0.56	0.84	0.33	0.30	0.96	0.59	-	-	-
					2001-2009					
					Apr.(t-1)					
α	-1.70	0.22	-0.34	-1.09*	-0.12	-1.61*	-0.40	-0.51**	-0.73***	-0.39
	0.98	1.14	0.57	0.59	0.49	0.83	0.43	0.20	0.27	0.28
Q(1)	0.25	0.20	0.24	0.33	0.68	0.66	0.24	-	-	-
Q(2)	0.27	0.44	0.28	0.28	0.26	0.32	0.46	-	-	-
Q(3)	0.22	0.36	0.11	0.20	0.27	0.39	0.48	-	-	-
					Oct.(t-1)					
α	-1.24	-0.32	-0.07	-0.73	-0.59	-1.21	-0.30	-0.28	-0.34*	-0.25
	0.92	1.08	0.55	0.58	0.52	0.82	0.46	0.18	0.20	0.27
Q(1)	0.45	0.07	0.35	0.79	0.49	0.60	0.39	-	-	-
Q(2)	0.57	0.20	0.32	0.96	0.15	0.57	0.66	-	-	-
Q(3)	0.58	0.20	0.15	0.52	0.18	0.71	0.66	-	-	-
					Apr.(t)					
α	-0.52	0.16	0.09	-0.40	-0.23	-0.50	0.03	-0.16	-0.11	-0.18
	0.37	0.54	0.27	0.25	0.31	0.24	0.19	0.16	0.12	0.23
Q(1)	0.09	0.04	0.07	0.58	0.28	0.97	0.09	-	-	-
Q(2)	0.20	0.12	0.18	0.85	0.45	0.09	0.22	-	-	-
Q(3)	0.11	0.14	0.20	0.24	0.51	0.19	0.21	-	-	-
					Oct.(t)					
α	0.24	0.17	0.20	-0.12	0.14	-0.12	0.19	-0.26	0.18**	-0.45*
	0.16	0.46	0.17	0.12	0.22	0.38	0.15	0.17	0.08	0.25
Q(1)	0.53	0.21	0.25	0.23	0.58	0.27	0.81	-	-	-
Q(2)	0.79	0.44	0.49	0.38	0.83	0.53	0.46	-	-	-
Q(3)	0.89	0.45	0.70	0.13	0.91	0.66	0.54	-	-	-

Source: Authors' estimates.

Note: The dependent variable is Consensus forecast error. Each cell reports the results of a regression of forecast errors on a constant for 7 Developed countries, the entire sample of N=29 countries and two sub-samples containing either industrialized or emerging and developing countries. For information on the composition of each sub-sample refer to the main text. We consider the full time span and then also the 1993-2001 and 2001-2009 periods in separate. Heteroskedastic-consistent robust standard errors are reported in parenthesis. The Ljung-Box Q statistic is used to measure serial correlation and the Q statistic up to M lags may be expressed as $Q(M) = \frac{T(T+2)\sum_j^M \rho_j^2}{T-j}$. Under the null hypothesis of no serial correlation, Q is asymptotically distributed as a χ^2. *, **, *** indicate significance at 10%, 5% and 1% levels, respectively.

Table 5. Rationality and Unbiasedness: Intercept Corrected Forecasts

	USA	Japan	Germany	France	Italy	UK	Canada	All-pooled	Industrial	Emerg+Devel.
Spec.	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
					Full Sample					
					Current Year					
α	-0.10	0.16	-0.15	0.07	0.01	0.16	0.07	0.02	0.01	0.03
	0.25	0.30	0.33	0.25	0.28	0.36	0.22	0.11	0.10	0.19
					Year Ahead					
α	-0.07	0.16	-0.08	0.11	-0.04	0.15	0.03	0.10	0.02	0.16
	0.68	0.71	0.52	0.41	0.46	0.74	0.34	0.15	0.20	0.22

Note: The dependent variable is "intercept corrected" Consensus forecast error as in Eq. (5) and (6) in the main text. Each cell reports the results of a regression of forecast errors on a constant for 7 Developed countries, the entire sample of N=29 countries and two sub-samples containing either industrialized or emerging and developing countries. For information on the composition of each sub-sample refer to the main text. Heteroskedastic-consistent robust standard errors are reported in parenthesis. *, **, *** indicate significance at 10%, 5% and 1% levels, respectively.

Table 6. Test of Efficiency based on forecast revisions

Country group	Dependent Variable	Independent Variables			R-squared
		Middle revision	Initial revision	Constant	
Spec.		(1)	(2)	(3)	
All	Final Revision	0.36***	-	0.07***	0.16
		(0.02)		(0.01)	
All	Final Revision	-	0.18***	0.03*	0.02
			(0.02)	(0.01)	
All	Final Revision	0.34***	0.07***	0.09***	0.17
		(0.02)	(0.02)	(0.01)	
All	Middle Revision	-	0.27***	-0.15***	0.04
			(0.03)	(0.01)	
Industrialized	Final Revision	0.33***	-	-0.10***	0.24
		(0.02)		(0.01)	
Industrialized	Final Revision	-	0.19***	-0.13***	0.05
			(0.02)	(0.01)	
Industrialized	Final Revision	0.31***	0.05**	-0.12***	0.25
		(0.02)	(0.020	(0.01)	
Industrialized	Middle Revision	-	0.38***	-0.03*	0.09
			(0.02)	(0.01)	
Emerging and Developing	Final Revision	0.36***	-	0.13***	0.16
		(0.02)		(0.01)	
Emerging and Developing	Final Revision	-	0.17***	0.08***	0.02
			(0.02)	(0.02)	
Emerging and Developing	Final Revision	0.35***	0.07***	0.16***	0.17
		(0.02)	(0.02)	(0.02)	
Emerging and Developing	Middle Revision	-	0.27***	-0.18***	0.04
			(0.03)	(0.02)	

Source: Authors' estimates.

Note: The dependent variable is identified in column 2 as either Final Revision or Middle Revision (for definition on Revisions refer to the main text). As for the independent variables included in each regression, these are identified in columns 3 and 4 together with a constant term (column 5). Heteroskedastic-consistent robust standard errors are reported in parenthesis. *, **, *** indicate significance at 10%, 5% and 1% levels, respectively.

Table 7. Test of Efficiency

Country group	Independent Variables	Dependent Variable: "actual" budget balance (%GDP)							
Model		I				II			
(fixed effects)		(no)				(yes)			
		Apr.(t-1)	Oct.(t-1)	Apr.(t)	Oct.(t)	Apr.(t-1)	Oct.(t-1)	Apr.(t)	Oct.(t)
Spec.		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All	Constant	-1.16***	-0.86***	0.81***	-0.81***	-2391.74**	234.91	-2481.96***	176.36*
		(0.25)	(0.23)	(0.07)	(0.26)	(1106.07)	(448.53)	(769.79)	(104.39)
	Forecast	0.74***	0.78***	-0.68***	0.71***	0.74***	0.79***	0.74***	0.68***
		(0.07)	(0.07)	(0.23)	(0.08)	(0.07)	(0.06)	(0.07)	(0.08)
	F-statistic	10.73	7.01	4.45	6.19	5.81	6.06	6.24	7.58
	p-value	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Industrialized	Constant	-0.59**	-0.18	-0.05	0.17**	-2461.35***	-145.72	-1135.92**	48.48
		(0.24)	(0.15)	(0.11)	(0.07)	(910.81)	(457.93)	(477.03)	(92.04)
	Forecast	0.92***	0.99***	1.00***	1.04***	0.84***	0.94***	0.90***	1.01***
		(0.06)	(0.04)	(0.03)	(0.02)	(0.06)	(0.03)	(0.04)	(0.02)
	F-statistic	3.23	0.84	0.17	3.00	4.97	1.06	3.39	0.22
	p-value	0.04	0.43	0.84	0.05	0.00	0.35	0.03	0.80
Emerging and Developing	Constant	-1.49***	-1.29***	-1.07***	-1.19***	-25.93	275.40	192.82***	337.54**
		(0.36)	(0.35)	(0.36)	(0.34)	(171.15)	(721.14)	(67.85)	(131.62)
	Forecast	0.62***	0.64***	0.70***	0.62***	0.65***	0.64***	0.64***	0.56***
		(0.12)	(0.10)	(0.10)	(0.10)	(0.11)	(0.11)	(0.10)	(0.10)
	F-statistic	8.58	7.12	4.87	6.98	5.13	6.34	6.65	8.84
	p-value	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00

Source: Authors' estimates.

Note: The regression is expressed as $A_t = \beta_0 + \beta_1 F_t + u_t$, where A is the actual realization and F is the forecast. The F-statistic and associated p-value are for the test of the null hypothesis that $\beta_0 = 0$ and $\beta_1 = 1$. Country fixed effects are included in each regression of Model II but not reported for reasons of parsimony. Heteroskedastic-consistent robust standard errors are reported in parenthesis. *, **, *** indicate significance at 10%, 5% and 1% levels, respectively.

Table 8. Rho and Beta efficiency regressions (country dimension)

	USA	Japan	Germany	France	Italy	UK	Canada	All-pooled	Industrial	Emerg+Devel.
Spec.	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
Full Sample										
Current Year										
β	0.10	-0.36***	0.12	0.08	-0.08	-0.10	0.05	-0.18***	0.00	-0.29***
	(0.09)	(0.10)	(0.21)	(0.12)	(0.07)	(0.10)	(0.10)	(0.07)	(0.03)	(0.10)
ρ	0.57**	0.50***	0.06	0.22	0.13	-0.11	0.15	0.67***	0.33***	0.70***
	(0.20)	(0.14)	(0.29)	(0.25)	(0.26)	(0.37)	(0.19)	(0.08)	(0.08)	(0.09)
Year Ahead										
β	0.02	-0.63***	-0.35	-0.08	-0.40**	-0.07	-0.11	-0.25***	-0.08	-0.37***
	(0.23)	(0.15)	(0.34)	(0.19)	(0.14)	(0.15)	(0.18)	(0.07)	(0.06)	(0.12)
ρ	0.52***	0.42**	0.31	0.42***	0.35**	0.08	0.43*	0.58***	0.40***	0.62***
	(0.12)	(0.18)	(0.26)	(0.12)	(0.16)	(0.36)	(0.22)	(0.06)	(0.09)	(0.07)
1993-2001										
Current Year										
β	0.08	-0.45***	0.15	0.28	0.04	0.12	0.13	-0.35**	0.03	-0.69***
	(0.05)	(0.05)	(0.58)	(0.33)	(0.06)	(0.11)	(0.14)	(0.15)	(0.04)	(0.23)
ρ	0.30	0.45***	-0.35	0.24	-0.71**	0.25	0.14	0.83***	0.38***	0.86***
	(0.26)	(0.07)	(0.39)	(0.50)	(0.25)	(0.32)	(0.24)	(0.17)	(0.11)	(0.19)
Year Ahead										
β	-0.12	-0.56***	-0.38	0.28	-0.23	0.34*	0.04	-0.33**	-0.02	-0.51*
	(0.27)	(0.08)	(0.76)	(0.45)	(0.20)	(0.17)	(0.19)	(0.17)	(0.07)	(0.30)
ρ	-0.15	0.59	-0.62*	0.42	-0.45	0.50	0.14	0.78***	0.42***	0.79***
	(0.50)	(0.33)	(0.29)	(0.29)	(0.35)	(0.36)	(0.27)	(0.12)	(0.14)	(0.13)
2001-2009										
Current Year										
β	0.11	-0.13	0.21	0.06	-0.29	-.25***	0.08	-0.12**	0.00	-0.18**
	(0.10)	(0.17)	(0.10)	(0.15)	(0.22)	(0.02)	(0.05)	(0.06)	(0.05)	(0.09)
ρ	0.54*	0.59	0.63***	0.18	0.39	-1.39	0.09	0.58***	0.37***	0.60***
	(0.30)	(0.36)	(0.15)	(0.23)	(0.29)	(0.55)	(0.32)	(0.09)	(0.12)	(0.10)
Year Ahead										
β	-0.14	-0.79	-0.20	-0.15	-0.83*	-0.29***	0.08	-0.22***	-0.09	-0.32***
	(0.32)	(0.54)	(0.40)	(0.25)	(0.37)	(0.07)	(0.21)	(0.08)	(0.09)	(0.12)
ρ	0.35*	0.39	0.42	0.34*	0.27	-0.25	0.28	0.48***	0.35***	0.51***
	(0.16)	(0.26)	(0.26)	(0.15)	(0.23)	(0.31)	(0.23)	(0.06)	(0.11)	(0.06)

Source: Authors' estimates.

Note: β is the estimated coefficient from a least-squares regression of the forecast error on the forecast, as specified in Eq. 8 in the main text. ρ is the estimated coefficient from a least-squares regression of the current period forecast error in the forecast error of the previous period, as specified in Eq. 8 in the main text. Estimations are carried out for 7 Developed countries, the entire sample of N=29 countries and two sub-samples containing either industrialized or emerging and developing countries. For information on the composition of each sub-sample refer to the main text. Heteroskedastic-consistent robust standard errors are reported in parenthesis. *, **, *** indicate significance at 10%, 5% and 1% levels, respectively.

Table 9. Forecast performance of Consensus Budget Balance-to-GDP ratio during recession episodes

	Apr.(t-1)	Oct.(t-1)	Apr.(t)	Oct.(t)
Number of episodes during recessions where a forecast was negative (Forecast <0)	36	39	38	40
Number of episodes where forecast was too optimistic (Forecast > Actual)	33	37	35	26
Average forecast error (all episodes)	-1.66	-1.24	-1.09	-0.62
Industrialized countries	-1.39	-0.67	-0.86	-0.18
Emerging and Developing countries	-2.06	-1.98	-1.39	-1.06

Source: Authors' estimates.

Note: Refer to the main text for details.

Table 10. Descriptive Statistics: Consensus Forecasts of Budget Balance-to-GDP ratio

	Forecast Errors	Absolute Forecast Errors	Dispersion of consensus forecasts	Forecast Errors	Absolute Forecast Errors	Dispersion of consensus forecasts
Spec.	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Unconditional					
	Advanced Economies			**Emerging and Developing Economies**		
Obs.	3652	3652	3652	4766	4766	4766
Mean	-0.15	1.01	0.47	-0.62	2.08	0.81
S.D.	1.47	1.08	0.47	3.30	2.64	0.64
Min.	-8.08	0.00	0.00	-19.78	0.00	0.00
Max.	3.91	8.08	4.04	7.67	19.78	7.98
	Recessions					
	Advanced Economies			**Emerging and Developing Economies**		
Obs.	704	704	704	525	525	525
Mean	-0.76	1.12	0.50	-1.54	2.66	1.03
S.D.	1.25	0.94	0.53	4.62	4.08	0.82
Min.	-5.56	0.00	0.00	-19.60	0.00	0.00
Max.	3.52	5.56	4.04	7.24	19.60	4.16
	Recoveries					
	Advanced Economies			**Emerging and Developing Economies**		
Obs.	320	320	320	384	384	384
Mean	-0.15	0.97	0.64	-0.27	2.28	1.17
S.D.	1.34	0.94	0.58	3.34	2.45	0.84
Min.	-5.37	0.00	0.00	-19.60	0.00	0.04
Max.	-5.37	5.37	3.08	5.97	19.60	4.26

Source: Authors' estimates.

Table 11. Regression Analysis: Consensus Forecasts of Budget Balance-to-GDP ratio under Recessions

	Advanced Economies			
	Absolute forecast errors		**Dispersion of consensus forecasts**	
	Coefficient	**Standard errors**	**Coefficient**	**Standard errors**
Spec.	**(1)**	**(2)**	**(3)**	**(4)**
Recessions	0.28***	0.03	0.08***	0.01
Recoveries	0.17***	0.04	0.26***	0.02
Horizon	0.03***	0.001	0.01***	0.001
Constant	0.31***	0.02	0.23***	0.01
Obs.		3238		3238
R-squared		0.17		0.11
	Emerging and Developing Economies			
	Absolute forecast errors		**Dispersion of consensus forecasts**	
	Coefficient	**Standard errors**	**Coefficient**	**Standard errors**
Recessions	0.84***	0.14	0.27***	0.03
Recoveries	-0.39**	0.18	0.39***	0.04
Horizon	0.04***	0.006	0.007***	0.001
Constant	1.36***	0.095	0.58***	0.02
Obs.		3706		3706
R-squared		0.04		0.10

Source: Authors' estimates.

Note: The dependent variable is absolute forecast errors or the dispersion of consensus forecasts. *, **, *** indicate significance at 10%, 5% and 1% levels, respectively.

Table 12. Test of Efficiency: Consensus Forecasts' Revisions of Budget Balance-to-GDP ratio

	Advanced Economies						Emerging and developing Economies					
	Coeff.	**St. error**	**Coeff.**	**St. error**	**Coeff.**	**St. error**	**Coeff.**	**St. error**	**Coeff.**	**St. error**	**Coeff.**	**St. error**
Spec.	**(1)**		**(2)**		**(3)**		**(4)**		**(5)**		**(6)**	
Constant			0.07	0.10	-0.06	0.13	0.29**	0.11	0.30**	0.13	0.25**	0.12
Dummy for recession	-0.39**	0.16	-0.48***	0.13	-0.34**	0.15	-0.62	0.60	-0.14	0.70	0.09	0.72
Dummy for recovery	0.12	0.21	-0.05	0.21	0.03	0.36	-0.21	0.33	-0.16	0.50	0.11	0.39
Middle revision	0.48**	0.24			0.63**	0.25	0.59***	0.19			0.50**	0.22
Middle revision after recessions	-0.37	0.27			-0.54*	0.27	-0.40	0.39			-0.26	0.45
Middle revision recoveries	-0.11	0.55			-0.14	1.56	0.20	0.45			0.42	0.63
Initial revision			-0.24*	0.15	-0.42***	0.14			0.43***	0.16	0.26	0.18
Initial revision after recessions			0.37**	0.16	0.54***	0.16			-0.34	0.64	0.06	0.65
Initial revision recoveries			0.32	0.25	0.33	0.38			-0.36**	0.18	-0.10	0.21
Obs.		42.00		39.00		39.00		119.00		108.00		108.00
R-squared		0.34		0.29		0.41		0.31		0.15		0.37
<u>P-values for Wald Tests</u>												
Middle revisions after recessions		0.38				0.38		0.57				0.55
Middle revision recoveries		0.45				0.75		0.06				0.12
Initial revision after recessions				0.05		0.08				0.88		0.61
Initial revision recoveries				0.73		0.79				0.42		0.10

Source: Authors' estimates.

Note: The dependent variable is the final revisions. Regressions include dummy variables for recession and recovery episodes. The hypothesis for the Wald test is that the value of the respective total effect is not statistically different from zero. *, **, *** indicate significance at 10%, 5% and 1% levels, respectively.

Table 13. Test of Bias: Consensus Forecasts of Budget Balance-to-GDP ratio under Recessions and Recoveries

	Advanced Economies		Emerging and developing economies	
	Coeff.	**St. error**	**Coeff.**	**St. error**
Constant	-0.49***	0.15	0.28	0.31
Dummy for recession	-1.31***	0.08	-0.68**	0.30
Dummy for recovery	-1.21***	0.10	0.19	0.31
Forecast	0.94***	0.01	0.66***	0.02
Forecast after recession	-0.15***	0.02	-0.41***	0.09
Forecast after recovery	-0.19***	0.02	-0.07	0.08
Obs.		3022		3258
R-squared		0.90		0.47
<u>P-values for Wald Tests</u>				
Forecast		0.00		0.00
Forecast after recession		0.00		0.00
Forecast after recovery		0.00		0.00

Source: Authors' estimates.

Note: The dependent variable is actual values. The joint Wald test focuses on the hypothesis that the coefficient on the respective consensus forecast is equal to 1 and the respective constant is equal to 0. *, **, *** indicate significance at 10%, 5% and 1% levels, respectively.

Table 14. Dispersion of Consensus Forecasts of Budget Balance-to-GDP ratio in the Run-up to Recession

	Advanced Economies		Emerging and developing economies	
	Coeff.	**St. error**	**Coeff.**	**St. error**
12 months before the onset of recession	0.12	0.09	0.24*	0.14
11 months before the onset of recession	0.12	0.09	0.46***	0.14
10 months before the onset of recession	0.14*	0.09	0.46***	0.13
9 months before the onset of recession	0.17**	0.09	0.57***	0.14
8 months before the onset of recession	0.10	0.08	0.41***	0.13
7 months before the onset of recession	0.13*	0.08	0.30**	0.13
6 months before the onset of recession	0.09	0.08	0.30**	0.13
5 months before the onset of recession	0.18**	0.08	0.22*	0.13
4 months before the onset of recession	0.14*	0.08	0.27**	0.13
3 months before the onset of recession	-0.04	0.09	0.29**	0.13
2 months before the onset of recession	-0.03	0.09	0.23*	0.13
1 months before the onset of recession	-0.04	0.09	0.10	0.13
Horizon	0.01***	0.00	0.01***	0.00
Constant	0.30***	0.02	0.64***	0.02
Obs.		3652		4766
R-squared		0.04		0.03

Source: Authors' estimates.

Note: The dependent variable is the dispersion forecasts. The regression includes dummy variables for the 12 months preceding the onset of a recession, a variable measuring the horizon of the forecast and a constant. *, **, *** indicate significance at 10%, 5% and 1% levels, respectively.

Figure 1: Distributions of Actual and Forecasted Budget Balances (% GDP): Consensus 1993-2009

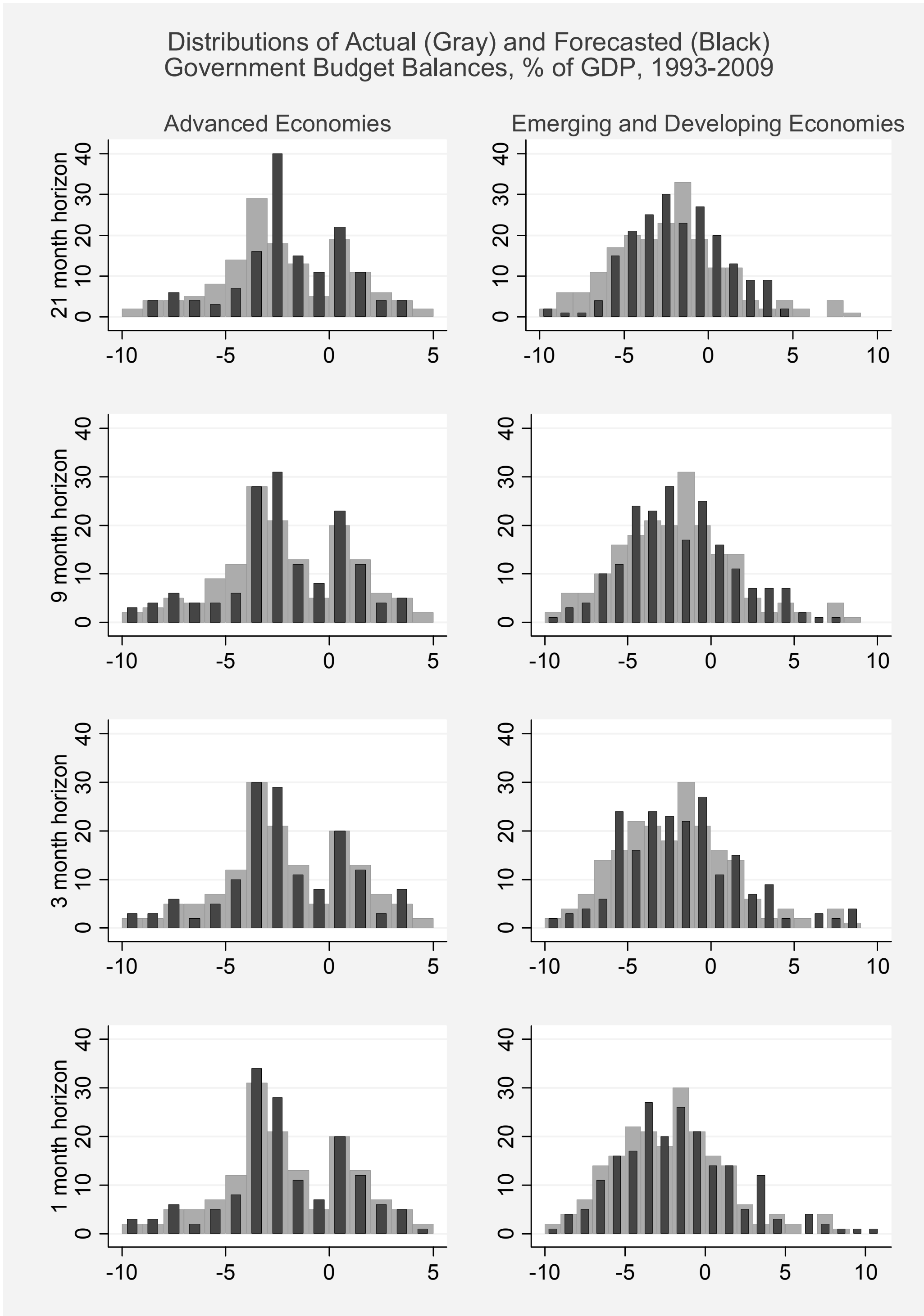

Figure 2: Kernel Density Estimates of Budget Balances Forecast Errors (% GDP): Consensus 1993-2009

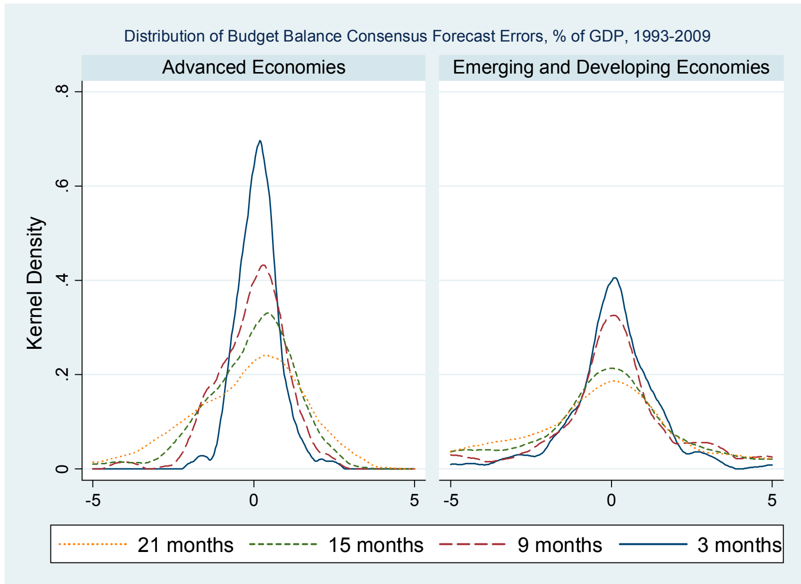

Figure 3: Descriptive Statistics of Budget Balances Forecast Errors (% GDP): Consensus 1993-2009

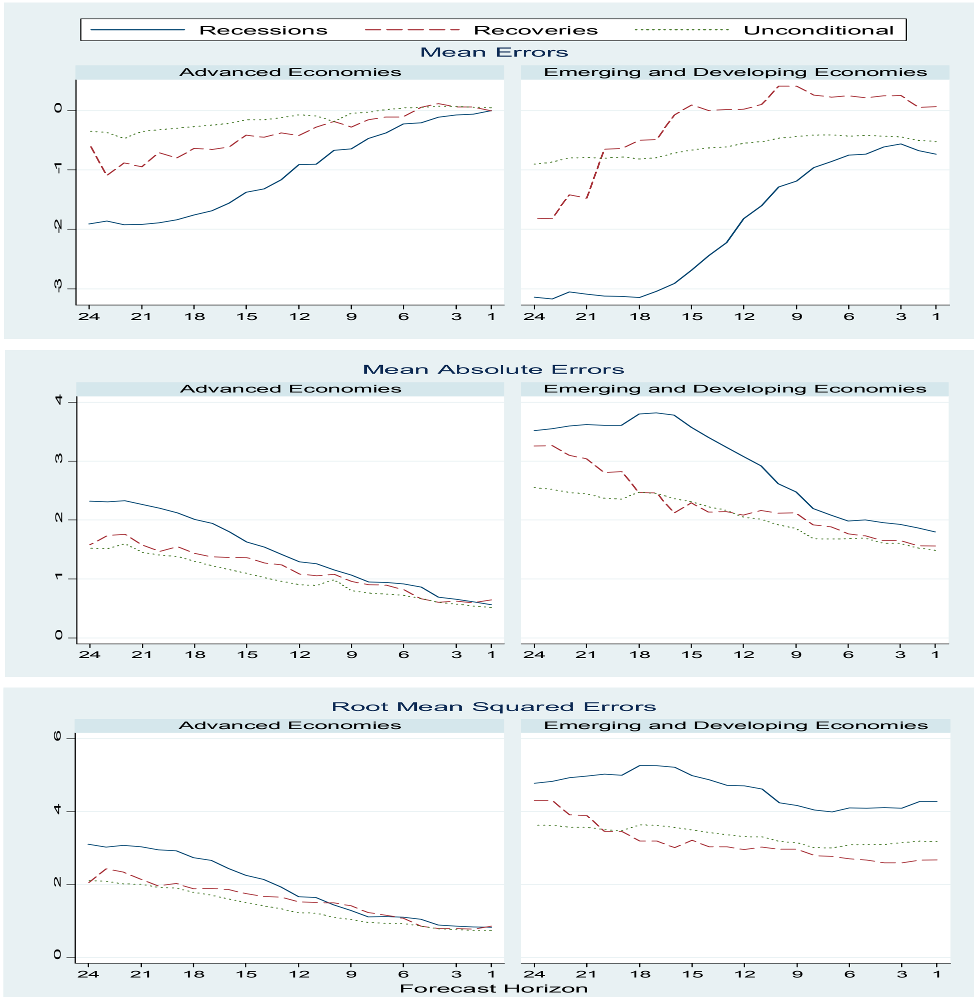

Figure 4: Kernel Density Estimates of Budget Balances Forecast Errors (% GDP) during Recessions: Consensus 1993-2009

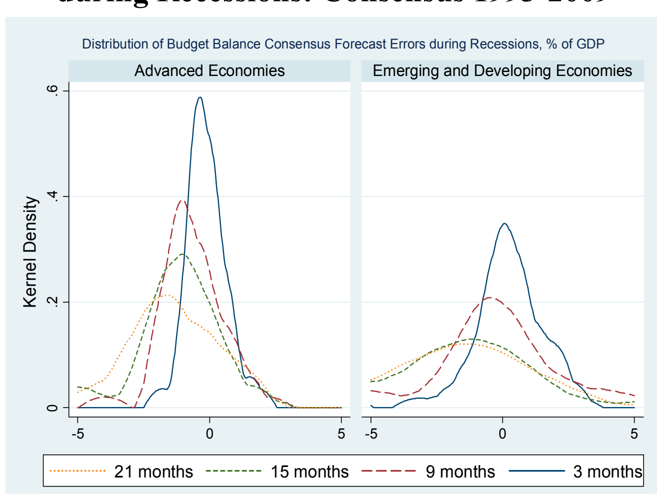

Figure 5: Actual and Forecasted Budget Balances (% GDP) during Recessions: Consensus 1993-2009

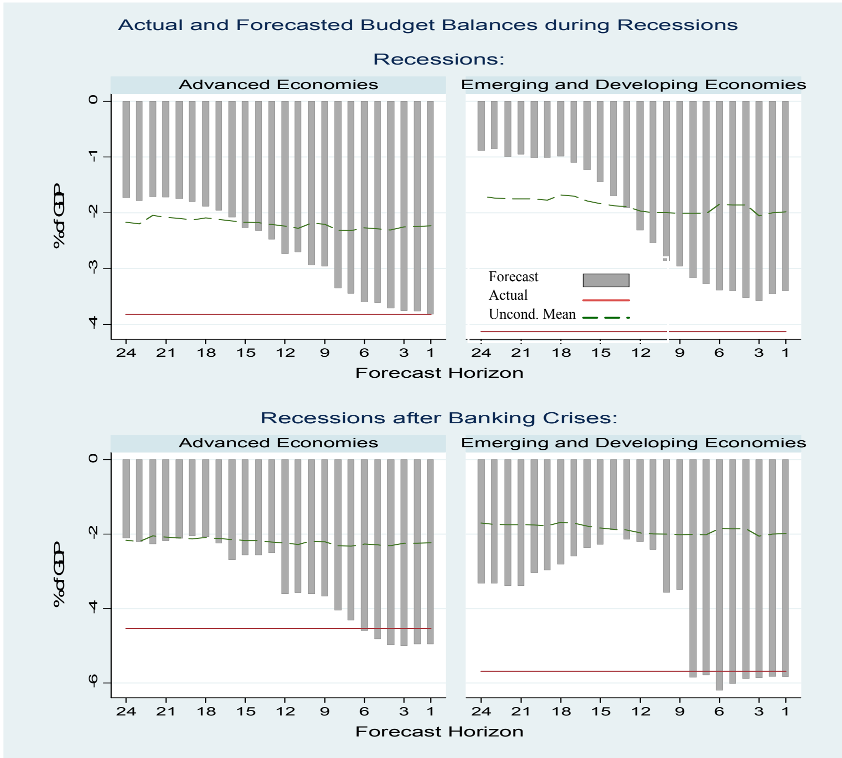

Figure 6: Dispersion of Budget Balances Forecasts (% GDP) during Recessions: Consensus 1993-2009

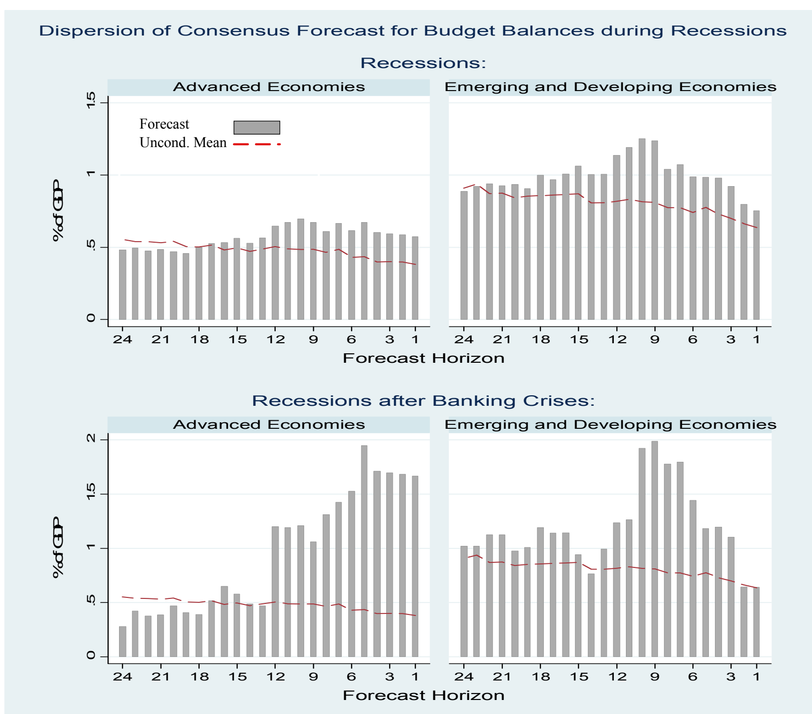

Figure 7: Kernel Density Estimates of Budget Balances Forecast Errors (% GDP) during Recoveries: Consensus 1993-2009

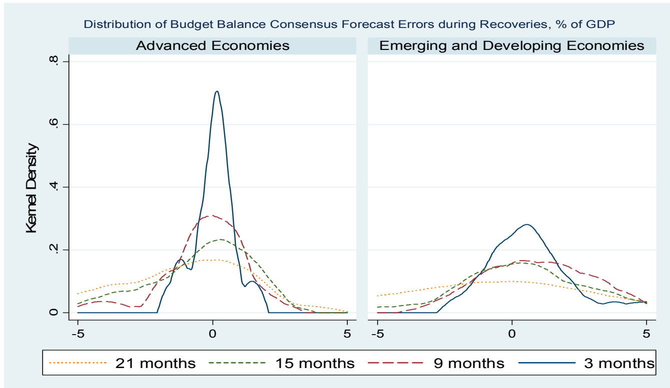

Figure 8: Actual and Forecasted Budget Balances (% GDP) during Recoveries (up to 4 years after recession ended): Consensus 1993-2009