Evaluating the forecast quality of GDP components: An application to G7*

Paulo Júlio† Pedro M. Esperança‡

April 20, 2012

Abstract

We evaluate the quality of OECD's and IMF's forecasts for real GDP growth and for GDP expenditure components. We use a scaled statistic to compare the prediction models' performance across GDP components with different volatilities and decompose the GDP forecast error into the corresponding component contributions. Moreover, we use two recently proposed statistics—Mean of Total Weighted Absolute Error and Mean of Total Weighted Squared Error—to evaluate the overall accuracy of component predictions. We conclude that overpredictions in investment and net exports explain GDP overpredictions at 1–year horizons. Accurate GDP forecasts for same–year predictions are mostly explained by canceling out effects in component prediction errors—mainly in exports and imports—rather than by accurate component predictions. We also show that forecasts are in general inefficient for both GDP and its components and that the 2008 crisis had a large negative effect on the quality of forecasts being issued, but not on the predictive quality of forecast models.

JEL Classification: C52, C53, E37
Keywords: Forecast evaluation, GDP expenditure components, Mean of total weighted absolute error, Mean of total weighted squared error, G7, 2008 crisis

*We are extremely grateful to Ariane Dinis and Gustavo Almeida for excellent research assistance. The opinions expressed in this article represent the views of the authors and do not necessarily correspond to those of the Portuguese Ministry of Economy and Employment.

†*Office for Strategy and Studies,* Portuguese Ministry of Economy and Employment. Corresponding author: paulo.julio@gee.min-economia.pt

‡Currently a postgraduate student at the University of Oxford.

1 Introduction

The literature has focused primarily on the quality of forecasts for real Gross Domestic Product (GDP) and for other macroeconomic variables (*e.g.* inflation or unemployment), but has not satisfactorily addressed the forecast accuracy of the major expenditure components of GDP—private consumption (C), government consumption (G), investment (I), exports (X) and imports (M).[1] In a recent article, Júlio et al. (2011) have analyzed, for the first time, the quality of forecasts for GDP expenditure components. The authors showed that overpredictions in investment and exports explain most of Portuguese GDP overpredictions at 1–year horizons. GDP forecast bias diminishes significantly for same–year predictions, a fact that is mostly explained by canceling out effects in component predictions errors rather than by accurate component predictions. The authors have also proposed two new statistics—Mean of Total Weighted Absolute Error (MTWAE) and Mean of Total Weighted Squared Error (MTWSE)—to objectively evaluate the overall accuracy of component predictions.[2]

This article analyzes the forecast quality of GDP expenditure components for G7 countries across 3 dimensions: bias, accuracy, and efficiency. We use forecast data issued by the Organization for Economic Cooperation and Development (OECD) and by the International Monetary Fund (IMF) for the 1993–2010 period, and evaluate both 1–year ahead and same–year predictions. Our focus lies on the overall quality of institutions' forecasts, and thus we pool evaluation statistics across countries in order to obtain an aggregate overview of the main features driving these forecasts. An evaluation of the forecast quality at the country level is relegated to the appendix. In addition, we propose panel versions of two types of efficiency tests presented in the literature, and analyze the effects of the 2008 crisis on the quality of forecasts.

We conclude that forecast accuracy is lower for investment, exports and imports, but forecast models perform comparatively worse predicting private consumption and government consumption. This is explained by different volatility levels of GDP components. At 1–year horizons, GDP overpredictions are mostly explained by investment and net exports. GDP forecast bias diminishes substantially at same–year horizons, a fact that is explained by canceling out effects in component prediction errors—mainly exports and imports—rather than by accurate component predictions. The MTWAE and MTWSE statistics suggest that OECD's forecasts for GDP components are more accurate than IMF's forecasts. We also show that forecasts are in general inefficient, both for GDP and its components, but inefficiency is more acute in government consumption predictions. The

[1] In this article, we always refer to real growth rates, even if not explicitly stated.

[2] Ash et al. (1998) also evaluate the quality of forecasts for GDP expenditure components, but solely through a directional analysis approach. Timmermann (2007) explores IMF's forecasts for the current account for several world regions, but does not address the quality of forecasts for other GDP components. Frankel (2011) studies forecasts for real growth rates and budget balances and argues that over–optimism in official forecasts can help explain excessive budget deficits.

2008 crisis had a large negative effect on the quality of forecasts—mainly for investment, exports and imports—but this was due to an increase in volatility rather than to a decrease in the performance of forecast models.

The accuracy of macroeconomic forecasts originating from both public and private institutions has been subject to a close inspection since the 1970s[3] and the literature has kept growing in recent years. For instance, Fildes and Stekler (2002) have conducted a survey on the state of macroeconomic forecasting focusing their analysis on studies made for the United States (US) and the United Kingdom (UK). Öller and Barot (2000) analyzed OECD's forecasts and national institutions' forecasts for GDP growth and inflation, concluding that there are no systematic differences in predictive quality across institutions and that all tend to overestimate at longer horizons. Loungani (2001) and Melander et al. (2007) also analyze GDP growth forecasts issued by several institutions, and conclude that these display very similar degrees of accuracy. In contrast, Pons (2000) finds OECD's forecasts for GDP growth to be superior to those issued by IMF, although the author does not detect a consistent pattern of over or underestimation. The accuracy of forecasts for GDP expenditure components is never addressed in any of these studies.

This article is organized as follows. The next section introduces the statistical methodology used to evaluate forecast quality. Section 3 describes the data. Section 4 analyzes the results. Section 5 evaluates the effects of the 2008 crisis on the quality of forecasts. Section 6 summarizes the results of efficiency tests. Section 7 concludes.

2 Methodology

This section presents the methodology. Section 2.1 introduces notation, whereas Sections 2.2 and 2.3 put forward the pooled versions of some well–known statistics widely used in forecast analysis. Section 2.4 introduces a pooled scaled statistic, which allows the comparison of forecast models' accuracy across GDP expenditure components with different volatility levels. Section 2.5 details the methodology used to evaluate the contributions of GDP expenditure components to the average GDP forecast error and presents pooled versions of two recently proposed statistics that assess the overall accuracy of component predictions. The exposition follows closely Júlio et al. (2011). Finally, Section 2.6 presents two simple efficiency tests.

2.1 Notation

Let the subscript j index the country and the subscript t index the *forecast period* (the period for which the forecast was produced), $j = 1, \dots, N$ and $t = 1, \dots, T$. Define $F_{jt}(s)$ as the s–period (or s–step) ahead forecast for the target variable A_{jt}. The variable s is

[3] See for instance Stekler (1972, 1987), McNees (1976, 1978, 1986, 1988), Zarnowitz (1979, 1984), Holden and Peel (1985, 1990), Clemen and Winkler (1986), Nordhaus (1987), Joutz (1988).

known as the *forecast horizon* or *time span*: the number of periods between the production of the forecast $F_{jt}(s)$ and the actual realization A_{jt}. The forecast error—the difference between actual and forecasted values for a given variable in country j—is

$$e_{jt}(s) = A_{jt} - F_{jt}(s) \tag{1}$$

for $j, t, s \in \mathbb{N}_0$. From (1) it is clear that a positive forecast error implies an underestimation, whereas a negative error implies an overestimation, of A_{jt}. Henceforth the forecast horizon s will be suppressed for notational convenience, if not strictly needed.

We analyze 1–, 2–, 3– and 4–step ahead forecasts, so that $s \in \{1, 2, 3, 4\}$: 1–step (3–step) ahead forecasts are those issued on the Autumn of the same (previous) year, and 2–step (4–step) ahead forecasts are those issued on the Spring of the same (previous) year.

2.2 Standard Evaluation Statistics

To evaluate the quality of forecasts, we start with the pooled versions of the standard measures of forecast evaluation. These are termed Pooled Mean Error (PME), Pooled Mean Absolute Error (PMAE), and Root of Pooled Mean Squared Error (RPMSE), and are respectively given by

$$\text{PME} := \frac{1}{NT} \sum_{j=1}^{N} \sum_{t=1}^{T} e_{jt} \tag{2}$$

$$\text{PMAE} := \frac{1}{NT} \sum_{j=1}^{N} \sum_{t=1}^{T} |e_{jt}| \tag{3}$$

$$\text{RPMSE} := \sqrt{\frac{1}{NT} \sum_{j=1}^{N} \sum_{t=1}^{T} e_{jt}^2} \tag{4}$$

PME is the average forecast error (across time and across countries), thus providing a simple measure of central tendency. A negative value means that forecasts overpredict actual values, whereas a positive value indicates an underprediction. PMAE provides a measure of the average total forecast error, regardless of the direction of the error (how much, on average, the forecasts are off-target). Hence, a lower PMAE reflects more accurate forecasts. RPMSE also provides a measure of total forecast error, but attributes disproportionally higher contributions to larger deviations from target. Thus, whereas PME measures how biased forecasts are on average, PMAE and RPMSE evaluate forecast accuracy.

2.3 Comparative Evaluation Statistics

It is also possible to construct measures that compare any two forecasting methods. A minimum benchmark that any formal forecasting model should outperform is the often–called *naïve forecast*: a 'model' predicting the same change as in the last observed period (a random walk–like behavior)

$$N_{jt}\left(s\right) = \begin{cases} A_{j,t-1} & , s \in \{1,2\} \\ A_{j,t-2} & , s \in \{3,4\} \end{cases} \tag{5}$$

for all t and j, where N_{jt} stands for the naïve forecast. At 3– and 4–period horizons we use the $t-2$ values, as the actual values of $t-1$ are not yet known. Letting η_{jt} denote the error of the naïve forecast

$$\eta_{jt} = A_{jt} - N_{jt} \tag{6}$$

we can compute the pooled version of the Theil's (1966) U2 statistic

$$\text{Pooled U2} := \sqrt{\frac{\sum_{j=1}^{N}\sum_{t=1}^{T}\left(A_{jt}-F_{jt}\right)^2}{\sum_{j=1}^{N}\sum_{t=1}^{T}\left(A_{jt}-N_{jt}\right)^2}} = \sqrt{\frac{\sum_{j=1}^{N}\sum_{t=1}^{T}e_{jt}^2}{\sum_{j=1}^{N}\sum_{t=1}^{T}\eta_{jt}^2}} \tag{7}$$

which provides a parsimonious comparative statistic of institutions' forecasts *vis-à-vis* the naïve forecasting method. A value of 1 means that naïve and institutions' forecasts have a similar forecasting ability; a value smaller than 1 implies that institutions' forecasts outperform naïve forecasts; and a value larger than 1 attests to a better forecasting accuracy of the naïve method. In this latter case, institutions' forecasts have no valuable content, on average, when compared to the naïve method.

2.4 Scaled Statistics

The previous statistics are only valid when comparing a variable's forecast coming from different institutions or forecasting methods. If one aims to compare the accuracy of institutions' forecasts across a group of variables, these statistics are inadequate, as they do not take into account the intrinsic level of volatility of each series. A more volatile series is naturally harder to predict and thus forecast errors tend to be larger; however, this does not necessarily mean that forecast models perform worse predicting that series. A comparative statistic which addresses this issue can be obtained by scaling each series' errors with the inverse of the corresponding in-sample average absolute difference between the actuals of

Table 1: *Volatility measured by V_j.*

	Canada	France	Germany	Italy	Japan	UK	US	Pooled
GDP	1.70	1.30	1.97	1.85	2.20	1.76	1.47	1.75
Priv. Cons.	1.20	0.78	0.79	1.05	1.12	1.45	0.99	1.05
Gov. Cons.	1.13	0.95	1.05	0.99	1.00	1.26	0.93	1.04
Investment	6.63	3.53	4.70	4.28	4.00	6.35	4.02	4.79
Exports	4.87	5.98	6.78	7.33	10.10	5.68	6.02	6.68
Imports	5.96	5.75	5.74	6.97	6.89	5.63	5.68	6.09

consecutive periods (a measure of volatility). Let K be the sample size. The scaled errors

$$v_{jt} = e_{jt} \left(\frac{1}{K} \sum_{k=1}^{K} |A_{jk} - A_{jk-1}| \right)^{-1} = e_{jt} V_j^{-1} \tag{8}$$

can thus be used in (3), with v_{jt} replacing e_{jt}, to obtain the Pooled Mean Absolute Scaled Error (PMASE).[4] Table 1, which presents the volatility of each series measured by V_j for G7 countries, shows that investment, exports and imports are much more volatile than the remaining series.

2.5 Contributions Analysis

It is also possible to decompose the GDP forecast error into the individual contributions of the corresponding expenditure components. This exercise enables one to identify which components contribute the most to the GDP forecast error and whether errors in forecasted expenditure components tend to add up or to cancel out. Let z_{jt} denote the effective real growth rate of variable Z in country j at year t, and z^f_{jt} the corresponding forecasted real growth rate, $Z = \{GDP, C, G, I, X, M\}$; and define $w^Z_{jt} = Z_{jt}/GDP_{jt}$—variable Z's share on GDP in country j at t. The effective real GDP growth rate in country j can therefore be decomposed into the corresponding component contributions

$$gdp_{jt} \equiv c_{jt} w^C_{j,t-1} + g_{jt} w^G_{j,t-1} + i_{jt} w^I_{j,t-1} + x_{jt} w^X_{j,t-1} - m_{jt} w^M_{j,t-1} + \epsilon_{jt} \tag{9}$$

where ϵ_{jt} is a discrepancy term which accounts for the non–additivity of component contributions resulting from chain–linked data.[5] There are two additional discrepancy sources when using forecasted data: neither the weights used by institutions nor the base year for

[4] The Mean Absolute Scaled Error was originally suggested by Hyndman and Koehler (2006). Similarly, one could also obtain the pooled version of the Root Mean Scaled Squared Error; however, we chose not to report this statistic, since it is greatly influenced by outliers.

[5] Chain–linked data is the rule followed by most statistical offices in developed countries.

those weights are known. Instead, we use effective weights, and thus[6]

$$gdp^f_{jt} \equiv c^f_{jt} w^C_{j,t-1} + g^f_{jt} w^G_{j,t-1} + i^f_{jt} w^I_{j,t-1} + x^f_{jt} w^X_{j,t-1} - m^f_{jt} w^M_{j,t-1} + \epsilon^f_{jt} \tag{10}$$

Let e^z_{jt} denote the forecast error of variable Z's growth rate, *i.e.* $e^z_{jt} = z_{jt} - z^f_{jt}$, and define the following vectors

$$\mathbf{e}_{jt} = (e^c_{jt}, e^g_{jt}, e^i_{jt}, e^x_{jt}, e^m_{jt})' \qquad \text{and} \qquad \mathbf{w}_{jt} = (w^C_{jt}, w^G_{jt}, w^I_{jt}, w^X_{jt}, w^M_{jt})'$$

Subtracting (10) from (9) and taking the average across time yields

$$\frac{1}{T}\sum_{t=1}^{T}\left(e^{gdp}_{jt} - \upsilon_{jt}\right) \equiv \frac{1}{T}\sum_{t=1}^{T}\mathbf{e}'_{jt}\mathbf{w}_{j,t-1} \tag{11}$$

where $\upsilon_{jt} = \epsilon_{jt} - \epsilon^f_{jt}$. In equation (11), $T^{-1}\sum_{t=1}^{T} e^z_{jt} w^Z_{j,t-1}$ represents the average contribution of the forecast error arising from variable Z, in percentage points, to the GDP growth forecast error in country j.

One may also be interested in pooling equation (11), in order to evaluate the average contributions of component forecast errors to the average GDP forecast error across all countries. Pooling in this context allows one to analyze global trends regarding the decomposition of GDP forecast errors. This is achieved by taking the average across countries on both sides of equation (11)

$$\frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\left(e^{gdp}_{jt} - \upsilon_{jt}\right) \equiv \frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}\mathbf{e}'_{jt}\mathbf{w}_{j,t-1} \tag{12}$$

A negative value in $(NT)^{-1}\sum_{j=1}^{N}\sum_{t=1}^{T} e^z_{jt} w^Z_{j,t-1}$ means that the component is overestimated in general, whereas a positive value has the opposite interpretation. It follows from (12) that, even if GDP forecast errors are small on average, this can result from large *cancelling out effects* in component contribution errors.

For this reason, we use two additional sets of statistics to assess the quality of forecasts. These statistics coherently aggregate individual measures of forecast accuracy for GDP expenditure components, thus evaluating the overall accuracy of component predictions. The unpooled versions of these statistics were originally proposed in Júlio et al. (2011), whereas the pooled versions are introduced herein. The first set evaluates the sum across components of the absolute distance between forecasted and actual contributions, thus

[6] An additional source of forecast errors may arise from positive or negative contributions of the statistical discrepancy ϵ_{jt} to growth. In practice, institutions have to deal with ϵ_{jt} explicitly in order to obtain GDP growth directly through the sum of component contributions. This can be done by distributing the statistical discrepancy's weight on GDP across components. The term ϵ_{jt} usually affects additivity of component contributions up to the second decimal place, thus having a negligible effect on conclusions.

assessing the total error in component forecasts. The unpooled version is termed Mean of Total Weighted Absolute Error (MTWAE), since it reflects the mean of the sum across components of absolute errors, weighted by the corresponding shares on GDP. The MTWAE for each country j is defined as

$$\text{MTWAE}_j := \frac{1}{T}\sum_{t=1}^{T} |\mathbf{e}_{jt}|' \, \mathbf{w}_{j,t-1} \tag{13}$$

where $|\mathbf{e}_{jt}|$ is a vector whose entries are the absolute values of the entries in $\mathbf{e}_t$. The corresponding pooled version (PMTWAE) is given by

$$\text{PMTWAE} := \frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T} |\mathbf{e}_{jt}|' \, \mathbf{w}_{j,t-1} \tag{14}$$

The weights reflect the relative importance of each component: those components with higher shares on GDP are naturally more important from the forecaster's point of view and should be weighted heavily. These statistics are computed for each institution and forecast horizon. Those institutions whose forecasts are associated with higher values in these statistics issue less accurate component predictions, even if GDP is accurately forecasted. Naturally, MTWAE and PMTWAE can be decomposed into component contributions, respectively

$$\frac{1}{T}\sum_{t=1}^{T} |e^z_{jt}| w^Z_{j,t-1} \qquad \text{and} \qquad \frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T} |e^z_{jt}| w^Z_{j,t-1}$$

The second set of statistics evaluate the sum across components of the squared errors, each weighted by the corresponding share on GDP. It is thus similar to the previous set, expect that larger errors contribute disproportionately more to the statistic. Letting

$$\mathbf{\Omega}_{jt} = \text{diag}(w^C_{jt}, w^G_{jt}, w^I_{jt}, w^X_{jt}, w^M_{jt})$$

the unpooled version of the statistic, named Mean of Total Weighted Squared Error (MTWSE), is defined as

$$\text{MTWSE} := \frac{1}{T}\sum_{t=1}^{T} \mathbf{e}'_{jt}\mathbf{\Omega}_{j,t-1}\mathbf{e}_{jt} \tag{15}$$

and the corresponding pooled version is

$$\text{PMTWSE} := \frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T} \mathbf{e}'_{jt}\mathbf{\Omega}_{j,t-1}\mathbf{e}_{jt} \tag{16}$$

Naturally, one can take the square root of MTWSE and PMTWSE to convert the measurement unit to the original scale. However, it may be advantageous to use instead the expressions in (15) and (16), since these can be easily decomposed into component contributions, respectively

$$\frac{1}{T}\sum_{t=1}^{T}(e_{jt}^{z})^2 w_{j,t-1}^{Z} \qquad \text{and} \qquad \frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T}(e_{jt}^{z})^2 w_{j,t-1}^{Z}$$

2.6 Efficiency Tests

The previous statistics do not attest whether it would be possible to improve issued forecasts. If one could issue a more accurate forecast with the information currently available, then improvements in quality would be possible. The tests which evaluate this feature are known as efficiency tests (Wallis, 1989; Fildes and Stekler, 2002).

A systematic bias signals that forecasts are either tendencially pessimistic or optimistic, and thus forecast accuracy could be permanently improved by adjusting predictions upwards or downwards, respectively. An unbiased forecast is a necessary condition for "weak informational efficiency." However, it is not sufficient, since efficiency also requires that forecast errors contain only unpredictable effects, *i.e.*, forecast errors cannot contain systematic information that could have been used to improve forecast accuracy.[7] In other words, forecast errors cannot be serially correlated.

In what follows, we propose panel versions of two types of efficiency tests. Panel tests allow for a considerable gain in power as compared to the corresponding time series versions. This is particularly important in this context due to the reduced time series dimension of forecast data. For the same reason, we chose not to present individual tests for each country.

To test for bias and serial correlation, we start by regressing the forecast errors on a constant and several lagged terms

$$e_{jt} = \gamma_0 + \sum_{l=1}^{p}\gamma_l \; e_{j,t-l} + \varepsilon_{jt} \tag{17}$$

The residuals are assumed to be serially uncorrelated after p is properly selected, but they may be heteroskedastic and contemporaneously correlated over j, $\mathrm{E}(\varepsilon_{jt}\varepsilon_{kt}) = \sigma_{jkt}^2,\ \forall j, k, t.$ Notice that forecast errors may be positively correlated across j, since unforeseen changes in GDP or any of its components in a large economy affects macroeconomic aggregates in other countries as well. We do not include individual–specific effects, as they are not supported by the Hausman test.

The model is estimated by OLS and parameter estimates are consistent as long as the underlying process is stationary. This requirement is obviously satisfied, since any distur-

[7] As pointed out by Holden and Peel (1990).

bance to the forecast error at t should not influence forecasts errors in the long run. Bias is evaluated by performing a z-test on the non–linear hypothesis that $\gamma_0/(1-\sum_{l=1}^{p}\gamma_l)=0$. Serial correlation, in turn, is evaluated through a χ^2 test on the null hypothesis that $\gamma_l=0,\forall l$.[8]

Another framework used to test for bias in institutions' forecasts dates back at least to Theil (1966), and is applied for instance in Joutz and Stekler (2000), Loungani (2001) and Vuchelen and Gutierrez (2005). The test (adapted to the panel framework) consists in evaluating whether the coefficients α and β in the following regression

$$A_{jt}=\alpha+\beta F_{jt}+u_{jt},\quad u_{jt}=\rho u_{j,t-1}+\varepsilon_{jt} \tag{18}$$

do not differ significantly from 0 and 1 respectively. We assume again that residuals, ε_{jt}, are heteroskedastic and contemporaneously correlated over j. The Hausman test does not support individual–specific effects, and hence we do not include them in (18).

A drawback of this approach is that serial correlation is not tested, but instead modeled by assuming that u_{jt} follows a common autoregressive process of order 1. Modeling autocorrelation is necessary, since autocorrelated residuals inflate the tests for bias, making any inference invalid.

The two tests for bias are conceptually different. The test resulting from equation (17) evaluates whether forecast errors have zero mean, whereas that resulting from equation (18) evaluates whether a regression line representing unbiased forecasts can fit the data. The conclusions of these tests may differ, for instance, if forecast errors have zero mean, but there is a tendency to overestimate when actual data takes high values and a tendency to underestimate when actual data takes low values. In this case, the former test may not reject the null of unbiasedness, whereas the latter might. Thus, the test resulting from (18) imposes stronger conditions, as it requires that forecast errors exhibit no specific patterns of over or underestimation.

3 Data

Our dataset contains information on forecasts for *Gross Domestic Product*, *Private Consumption*, *Government Consumption*, *Investment* (namely gross fixed capital formation), *Exports* and *Imports* (all in volume percentage change), issued by OECD and IMF for G7 countries: Canada, France, Germany, Italy, Japan, UK, and US. Forecasts were retrieved from OECD's *Economic Outlook* and from IMF's *World Economic Outlook*. These institutions issue forecasts twice a year: OECD issues them on May/June and November/December and IMF on April/May and September/October. Although institutions' forecasts are not issued exactly in the same month, they are done in the same time hori-

[8] A similar framework is presented in Öller and Barot (2000). The authors, however, perform a cross–section analysis for each country, thus avoiding some panel data complications.

Table 2: *Forecast horizon and issue date.*

Forecast period	*Forecast horizon*	*Issue date*
t	$1-step$	Autumn t
	$2-step$	Spring t
	$3-step$	Autumn $t-1$
	$4-step$	Spring $t-1$

zon and should thus use similar information sets (Pons, 2000). We thus classify forecasts according to the season in which they are issued. For convenience, forecasts are labeled as 1–, 2–, 3– and 4–step ahead forecasts. Table 2 summarizes the terminology.

Actual values, which were also used to compute expenditure component shares on GDP, were taken from National Statistical Offices. The period scrutinized is 1993–2010 for same–year forecasts, and 1994–2010 for 1–year ahead forecasts.[9] The choice of the realization is not consensual in the literature, as one should find a compromise between the argument that forecasters do not know the nature of data revisions and the argument that the realization should reflect exact economic outcomes (Vuchelen and Gutierrez, 2005).[10] First releases do not incorporate all information about economic activity. When information is not available, Statistical Offices uses imputation and forecasting methods, and econometric models, to issue an estimate of economic outcomes. Thus, comparing forecasts with first releases or even with intermediate releases is equivalent to compare forecasts with an estimation of economic outcomes. Even though there might exist revisions between first and final releases that forecasters could not be aware of, most revisions derive from the incorporation of new and updated information, whose sources are usually known to forecasters. For this reason, we evaluate forecasts against final releases and not first or intermediate releases.

4 Results

In this section we present the results from the pooled statistics for the 1993–2010 period. Individual statistics for each country are relegated to the appendix.

4.1 Gross Domestic Product

Table 3 presents the pooled statistics for GDP growth forecasts. The PME statistic shows that OECD and IMF overestimate GDP growth at 1–year spans—a fact that is explained

[9] Prior to 1992, OECD and IMF reported GDP forecasts for some G7 countries and GNP forecasts for others. Only from 1993 onwards forecasted variables were harmonized, with GDP and the corresponding expenditure components being reported for all G7 countries.

[10] For instance, in an often cited article, Keane and Runkle (1990) argue that forecasters aim at predicting first releases, since they do not know in advance the nature of data revisions occurring after the date on which they make their forecasts. Similar arguments are used in Zarnowitz and Braun (1993) and Joutz and Stekler (2000). On the opposite direction, Ash et al. (1998) and Öller and Barot (2000) use data published 6 months and 12 months after the event, respectively, as these are neither flash estimates nor late revisions.

Table 3: *Pooled statistics: GDP.*

	OECD				IMF			
	4–step *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*	*4–step* *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*
PME	-0.49	-0.15	0.09	0.12	-0.46	-0.29	0.12	0.13
RPMSE	2.14	1.61	1.00	0.78	2.13	1.81	1.06	0.82
PMAE	1.60	1.23	0.76	0.61	1.57	1.36	0.82	0.62
PMASE	0.93	0.71	0.44	0.35	0.91	0.79	0.48	0.36
PU2	0.84	0.63	0.40	0.31	0.83	0.70	0.42	0.33

by overestimations for France, Germany, Italy and Japan. Same–year forecasts are more accurate, although characterized by minor underpredictions. Bias decreases consistently as the forecast horizon shortens (except for the 1–period horizon), suggesting that forecast performance improves as more recent information is incorporated into predictions. Moreover, OECD's and IMF's forecasts display similar bias except at the 3–period span, in which OECD takes a small lead.

The RPMSE and PMAE statistics point towards a negative relationship between forecast accuracy and the forecast horizon.[11] This fact illustrates the role that new and updated information has on the quality of forecasts. The most accurate 1–year ahead forecasts are issued for France, whereas the least accurate are issued for Germany and Japan. For same–year predictions, GDP forecasts display the highest accuracy in France and the lowest accuracy in Japan and the UK. Forecasts issued by OECD are slightly more accurate than IMF's at the 3–period span, but for other forecast horizons the difference is marginal.

Finally, institutions' forecasts outperform the naïve method at all time spans, as shown by the PU2 statistic. Naturally, the performance gap between institutions' forecasts and naïve forecasts widens as the horizon shortens.

4.2 GDP Expenditure Components

Table 4 presents an evaluation of forecasts for GDP expenditure components. The PME statistic suggests that forecasts for private consumption have the smallest biases at almost all time spans. Government consumption also displays a comparatively small bias and is underestimated at all horizons.

At 1–year spans, the largest biases occur in predictions for investment, exports and imports. Investment and exports are systematically overestimated, and the same holds for imports with OECD's forecasts. On the opposite direction, IMF's forecasts for imports have a comparatively small bias at the 4–period span and are downwardly biased at the 3–period span. For same–year predictions, the largest biases occur in exports and imports, as these components are clearly underestimated by both institutions. Investment is overestimated by IMF at the 2–period span, but in other cases bias is small.

[11] Since PMASE is a rescaling of PMAE, it draws exactly the same conclusions.

Table 4: *Pooled statistics: GDP expenditure components.*

		OECD				IMF			
		4–step Spr $t-1$	*3–step* Aut $t-1$	*2–step* Spr t	*1–step* Aut t	*4–step* Spr $t-1$	*3–step* Aut $t-1$	*2–step* Spr t	*1–step* Aut t
C	PME	-0.15	0.15	0.14	0.09	-0.19	-0.03	0.18	0.07
	RPMSE	1.53	1.30	1.05	0.80	1.67	1.41	1.09	0.85
	PMAE	1.20	1.01	0.75	0.61	1.28	1.10	0.83	0.62
	PMASE	1.17	0.96	0.67	0.55	1.23	1.07	0.75	0.56
	PU2	0.89	0.76	0.68	0.53	0.97	0.82	0.71	0.55
G	PME	0.31	0.31	0.21	0.14	0.34	0.24	0.32	0.16
	RPMSE	1.39	1.33	1.33	1.14	1.51	1.55	1.19	1.15
	PMAE	1.07	1.04	0.98	0.81	1.15	1.17	0.95	0.85
	PMASE	1.03	0.99	0.90	0.75	1.11	1.12	0.88	0.80
	PU2	0.88	0.85	0.96	0.83	0.96	0.98	0.86	0.83
I	PME	-1.32	-0.55	-0.09	0.08	-1.17	-0.88	-0.37	-0.04
	RPMSE	5.45	4.29	3.27	2.57	5.54	4.66	3.36	2.74
	PMAE	3.86	3.04	2.44	1.84	3.99	3.49	2.47	2.01
	PMASE	0.82	0.64	0.51	0.38	0.85	0.75	0.53	0.42
	PU2	0.77	0.60	0.48	0.38	0.78	0.66	0.49	0.40
X	PME	-1.56	-0.79	0.37	0.75	-0.31	-0.15	0.93	1.08
	RPMSE	7.38	5.84	3.69	2.35	7.35	6.59	4.31	2.89
	PMAE	5.42	4.35	2.83	1.73	5.63	5.00	3.39	2.23
	PMASE	0.83	0.66	0.44	0.28	0.86	0.77	0.53	0.36
	PU2	0.87	0.69	0.39	0.25	0.87	0.78	0.45	0.30
M	PME	-0.85	-0.18	0.35	0.77	0.11	0.59	0.76	0.92
	RPMSE	6.64	5.64	3.73	2.46	6.55	6.14	4.35	3.13
	PMAE	5.10	4.32	2.87	1.78	5.10	4.92	3.30	2.26
	PMASE	0.84	0.71	0.47	0.30	0.84	0.81	0.55	0.37
	PU2	0.80	0.68	0.43	0.29	0.79	0.74	0.51	0.36

The RPMSE and PMAE statistics suggest that forecasts for investment, exports and imports have the lowest accuracy. However, this does not imply that forecast models perform comparatively worse predicting these components, as they are also more volatile. By scaling the errors with the inverse of the volatility of each series, PMASE becomes more appropriate to make inferences about the predictive quality of institutions' forecast models across GDP components.

PMASE demonstrates that forecast models perform comparatively worse when predicting private consumption and government consumption at all horizons. Possibly, institutions' forecast models cannot accurately predict revisions in the consumption bundle carried out by households when new macroeconomic information becomes available. Government consumption is a policy–making tool, often facing unexpected increases, particularly in election years. The performance of forecast models is similar across the remaining GDP components—investment, exports and imports.

The PU2 statistic shows that forecast models perform generally better than the *naïve forecast*. Notice however that IMF's forecasts for private consumption and government consumption are only marginally better than the last observed period extrapolation for

Table 5: *Pooled contributions of expenditure components (in percentage points) to the average GDP forecast error.*

	OECD				**IMF**			
	4–step *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*	*4–step* *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*
C	-0.08	0.10	0.08	0.06	-0.10	-0.01	0.11	0.04
G	0.05	0.05	0.03	0.02	0.06	0.04	0.06	0.03
I	-0.27	-0.11	-0.02	0.03	-0.25	-0.18	-0.07	0.00
X	-0.45	-0.26	0.06	0.17	-0.17	-0.04	0.20	0.22
M	-0.30	-0.13	0.07	0.18	-0.04	0.08	0.18	0.20
υ	-0.04	-0.06	0.00	0.02	-0.04	-0.02	0.02	0.04
GDP	-0.49	-0.15	0.09	0.12	-0.46	-0.29	0.12	0.13

Notes: (i) **GDP** in the table represents the mean (across time and across countries) of e_{jt}^{gdp}, and equals the Pooled Mean Error statistic in Table 3, by definition; (ii) The sum of the contributions in the table differs from **GDP** by an error υ whose source is explained in Section 2.5.

some forecast horizons.

4.3 Decomposing GDP Forecast Errors into Component Contributions

Pooled contributions of expenditure components to the average GDP forecast error are detailed in Table 5. Notice that average discrepancies, υ, originating from the non–additivity of component contributions and from the difference between actual component shares on GDP and the shares used by institutions in forecast models, are small.

Institutions overestimate private consumption (with the exception of OECD's 3–step ahead forecasts) and underestimate government consumption at 1–year spans. Nevertheless, the contributions of these components to the GDP forecast error are small relative to other components. The remaining components are overestimated at 1–year spans (with the exception of IMF's 3–step ahead forecasts for imports). At the 4–period span, around 85% of the GDP forecast error is explained by investment and net exports. However, despite net export's similar contribution (-0.15 for OECD and -0.13 for IMF) to the GDP forecast error, OECD overestimates exports and imports by a larger magnitude on average. At the 3–period span, component contributions are smaller, leading to a less biased GDP forecast relative to the 4–period span. Overestimations in investment and net exports still explain the largest fraction of GDP overpredictions in this case. For OECD's forecasts, exports present the largest contribution to the GDP forecast error, whereas for IMF's forecasts the largest contribution comes from investment. Specific conclusions vary across countries: whereas component contribution errors for France, Germany, Italy and Japan tend to add up, contributing to larger biases in forecasted GDP, for Canada, the UK and the US they tend to cancel out, resulting in smaller biases.

Contributions of private consumption, government consumption and investment to the GDP forecast error are small at same–year spans, when compared with other time spans or components. The largest contributions are displayed by exports and imports, both under-

Table 6: *Pooled* MTWAE *statistic and its decomposition (in percentage points).*

	OECD				IMF			
	4–step *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*	*4–step* *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*
PMTWAE	4.46	3.69	2.61	1.89	4.60	4.18	2.93	2.14
				Component contributions to PMTWAE				
C	0.73	0.61	0.45	0.36	0.78	0.67	0.50	0.37
G	0.21	0.20	0.19	0.16	0.23	0.23	0.19	0.17
I	0.97	0.81	0.62	0.53	1.01	0.90	0.64	0.55
X	1.34	1.05	0.68	0.41	1.37	1.21	0.81	0.52
M	1.21	1.01	0.67	0.42	1.21	1.17	0.78	0.52

estimated. However, the effects of these components tend to cancel out, as imports enter the GDP equation with a negative sign. This results into relatively accurate GDP predictions. The *cancel out effect* is stronger for Canada, Germany and Japan. For France, Italy and the US, GDP forecasts are obtained with relatively unbiased component predictions. On the opposite direction, UK's component forecast errors tend to add up, originating a significant GDP overestimation as compared with other countries.

The pooled MTWAE and MTWSE statistics, presented in Tables 6 and 7, summarize the overall accuracy of component predictions. A lower value means that GDP forecasts are assembled with more accurate component predictions, whereas a higher value has the opposite interpretation. At 1–year horizons, PMTWAE fluctuates between 3.7 (OECD's forecasts at the 3–period span) and 4.6 (IMF's forecasts at the 4–period span) percentage points. The components which most significantly contribute to this outcome (*i.e.*, whose predictions, weighted by the component's share on GDP, are least accurate) are, by descending order, exports, imports and investment. On the opposite direction, the contribution of government consumption to the statistic is marginal. This ordering is highly correlated with the volatility of the variables. For same–year spans, the overall accuracy of component predictions increases, explaining more accurate GDP forecasts. The decomposition of PMTWAE and PMTWSE leads to similar conclusions as for 1–year spans. The highest overall accuracy in component predictions is achieved for the US and Japan, whereas the poorest performance is attained for Canada and Germany. This diversity is explained by the different accuracy levels of exports and imports across countries.

It is also evident from Tables 6 and 7 that OECD's component forecasts are more accurate than IMF's component forecasts, even though the latter displays smaller biases for some components. This is explained by cancel out effects that occur across years, which are not captured by the first moment: positive errors in some years are canceling out negative errors in other years, leading to lower biases, but also to lower accuracy. These effects are more acute for IMF's forecasts.

Table 7: *Pooled* MTWSE *statistic and its decomposition.*

	OECD				IMF			
	4–step Spr $t-1$	*3–step* Aut $t-1$	*2–step* Spr t	*1–step* Aut t	*4–step* Spr $t-1$	*3–step* Aut $t-1$	*2–step* Spr t	*1–step* Aut t
PMTWSE	9.79	6.39	2.83	1.46	9.74	7.92	3.76	1.88
	Component contributions to PMTWSE							
C	0.88	0.63	0.40	0.24	1.05	0.75	0.44	0.26
G	0.07	0.07	0.07	0.05	0.09	0.09	0.05	0.05
I	1.75	1.19	0.61	0.47	1.81	1.41	0.67	0.49
X	3.88	2.33	0.88	0.33	3.75	2.97	1.27	0.49
M	3.20	2.17	0.86	0.37	3.04	2.70	1.33	0.59

5 The Effects of the 2008 Crisis on the Quality of Forecasts

To evaluate the effects of the 2008 crisis on the quality of forecasts, we compute the same measures of forecast quality until 2007, and analyze how these have changed relatively to the complete time period. Results are displayed in Table 8.[12]

In general, the crisis contributed to increase bias (evaluated by PME), particularly at 1–year spans. However, this conclusion does not hold for all variables. Prior to 2008, 4–period ahead forecasts for GDP were downwardly biased, and the crisis strengthened this bias. OECD's 3–step ahead forecasts for GDP were upwardly biased, and IMF's 3–step ahead forecasts nearly unbiased. In the former case the crisis changed the sign of the bias, but presented no relevant effect on its absolute magnitude, whereas in the latter case the crisis originated a bias of around -0.3 percentage points. Private consumption became slightly biased (downwards) at the 4–step span, but bias decreased at the 3–period span as a result of the crisis. Government consumption remained overestimated as before the crisis. Different conclusions hold for the remaining GDP components, with PME changing between -0.8 and -0.5 percentage points at 1–year spans. Investment was overestimated before 2008, and the crisis contributed to foster this tendency. Exports and imports were underestimated by IMF and overestimated by OECD at 1–year horizons prior to 2008 (with the exception of OECD's 3–period ahead forecasts for imports). For these components, the crisis led to a substantial increase in bias for OECD's forecasts (*i.e.* overestimation increased), but to a decrease in bias for IMF's forecasts (*i.e.* underestimation decreased). The exception is OECD's 3–step ahead forecasts for imports, for which bias was reduced.

The crisis also led to a substantial decrease in the accuracy of 1–year ahead predictions, but not of same–year predictions. The decrease in accuracy (measured by PMAE) was higher for more volatile components—investment, exports and imports, and affected OECD's and IMF's forecasts. The accuracy of forecasts for government consumption and private consumption were only marginally affected. Somewhat surprisingly, PMASE shows that the quality of institutions' prediction models increased in recent years, albeit

[12]For brevity, we only present the major statistics.

Table 8: *Effects of the 2008 financial crisis on pooled statistics: GDP and GDP expenditure components.*

		OECD				IMF			
		4–step Spr $t-1$	*3–step* Aut $t-1$	*2–step* Spr t	*1–step* Aut t	*4–step* Spr $t-1$	*3–step* Aut $t-1$	*2–step* Spr t	*1–step* Aut t
GDP	PME	-0.16	0.15	0.18	0.22	-0.17	0.01	0.20	0.23
	ΔPME	-0.33	-0.30	-0.09	-0.10	-0.29	-0.30	-0.08	-0.10
	ΔPMAE	0.41	0.22	0.03	0.01	0.42	0.31	0.01	0.01
	ΔPMASE	-0.14	-0.19	-0.19	-0.17	-0.12	-0.16	-0.22	-0.16
C	PME	0.04	0.32	0.25	0.19	-0.03	0.13	0.24	0.14
	ΔPME	-0.19	-0.17	-0.11	-0.10	-0.16	-0.16	-0.06	-0.08
	ΔPMAE	0.19	0.08	0.05	0.03	0.18	0.11	0.03	0.03
	ΔPMASE	-0.08	-0.18	-0.11	-0.09	-0.11	-0.14	-0.15	-0.10
G	PME	0.35	0.36	0.25	0.17	0.36	0.28	0.40	0.23
	ΔPME	-0.04	-0.05	-0.04	-0.03	-0.01	-0.05	-0.08	-0.07
	ΔPMAE	-0.04	-0.02	0.00	-0.02	0.01	0.00	-0.03	-0.02
	ΔPMASE	-0.05	-0.04	-0.01	-0.03	0.00	-0.01	-0.04	-0.03
I	PME	-0.58	-0.03	-0.09	0.17	-0.53	-0.35	-0.36	-0.02
	ΔPME	-0.73	-0.52	0.00	-0.08	-0.64	-0.52	-0.01	-0.02
	ΔPMAE	0.87	0.59	0.09	0.04	1.00	0.73	0.17	0.11
	ΔPMASE	-0.09	-0.10	-0.19	-0.16	-0.07	-0.10	-0.16	-0.15
X	PME	-0.98	-0.08	0.21	0.94	0.37	0.83	0.91	1.26
	ΔPME	-0.59	-0.71	0.15	-0.19	-0.68	-0.68	0.02	-0.18
	ΔPMAE	1.40	0.83	0.17	-0.03	1.22	0.97	0.11	0.00
	ΔPMASE	-0.03	-0.08	-0.14	-0.11	-0.08	-0.10	-0.19	-0.13
M	PME	-0.22	0.56	0.31	0.90	0.85	1.20	0.71	0.90
	ΔPME	-0.63	-0.74	0.05	-0.14	-0.74	-0.61	0.05	0.02
	ΔPMAE	1.20	0.74	0.00	-0.06	1.17	0.88	0.06	0.00
	ΔPMASE	-0.02	-0.08	-0.15	-0.10	-0.04	-0.09	-0.17	-0.11

Notes: PME is the value of the Mean Error for the 1993–2007 period. Δ corresponds to the difference in the value of the statistic between the 1993–2010 period and the 1993–2007 period. Thus, a positive (negative) value means that the financial crisis originated a positive (negative) change in the value of the statistic.

marginally. The opposing signs displayed by the changes in PMAE and PMASE are due to large increases in volatility after 2008. Thus, the fall in accuracy after the triggering of the crisis is explained by an increase in uncertainty, rather than by a decline in the quality of forecast models.

Table 9 displays the contributions of expenditure components to the average GDP forecast error prior to the crisis. When compared with the complete time period, contributions are substantially smaller, mainly those from investment, exports and imports. The crisis also led to substantial decreases in the overall accuracy of component predictions, as shown by the change in PMTWAE in Table 10, particularly at 1–year spans. Forecasts for investment, exports and imports were the prime sources of this result.

Table 9: *Effects of the 2008 financial crisis: pooled contributions of expenditure components to the average GDP forecast error.*

	OECD				IMF			
	4–step *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*	*4–step* *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*
	Average contributions until 2007							
C	0.04	0.20	0.15	0.11	0.00	0.10	0.15	0.09
G	0.06	0.06	0.04	0.03	0.06	0.05	0.07	0.04
I	-0.09	0.02	0.01	0.07	-0.08	-0.05	-0.05	0.03
X	-0.26	-0.05	0.02	0.22	0.04	0.15	0.19	0.25
M	-0.12	0.07	0.05	0.20	0.14	0.23	0.15	0.19
υ	-0.03	-0.01	0.00	0.00	-0.04	0.00	0.00	0.01
GDP	-0.16	0.15	0.18	0.22	-0.17	0.01	0.20	0.23
	Changes in contributions relative to the complete period							
$\Delta\boldsymbol{C}$	-0.12	-0.10	-0.07	-0.05	-0.10	-0.11	-0.04	-0.05
$\Delta\boldsymbol{G}$	-0.01	-0.01	-0.01	-0.01	0.00	-0.01	-0.01	-0.01
$\Delta\boldsymbol{I}$	-0.18	-0.13	-0.03	-0.04	-0.17	-0.13	-0.02	-0.03
$\Delta\boldsymbol{X}$	-0.19	-0.21	0.04	-0.05	-0.21	-0.19	0.01	-0.03
$\Delta\boldsymbol{M}$	-0.18	-0.20	0.02	-0.02	-0.18	-0.15	0.03	0.01
$\Delta\upsilon$	-0.01	-0.05	0.00	0.02	0.00	-0.02	0.02	0.03
$\Delta\boldsymbol{GDP}$	-0.33	-0.30	-0.09	-0.10	-0.29	-0.30	-0.08	-0.10

Notes: Same as in Table 5.

Table 10: *Effects of the 2008 financial crisis: pooled* MTWAE *statistic and its decomposition.*

	OECD				IMF			
	4–step *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*	*4–step* *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*
	PMTWAE and component contributions until 2007							
PMTWAE	3.36	3.00	2.43	1.81	3.54	3.37	2.74	2.02
C	0.61	0.56	0.42	0.34	0.66	0.60	0.49	0.35
G	0.22	0.21	0.19	0.17	0.23	0.23	0.19	0.17
I	0.75	0.65	0.55	0.48	0.77	0.71	0.57	0.48
X	0.94	0.81	0.63	0.42	1.02	0.94	0.77	0.51
M	0.84	0.77	0.64	0.41	0.86	0.89	0.73	0.51
	Changes relative to the complete period							
ΔPMTWAE	1.10	0.69	0.18	0.08	1.06	0.81	0.19	0.12
ΔC	0.12	0.05	0.03	0.02	0.12	0.07	0.01	0.02
ΔG	-0.01	-0.01	0.00	-0.01	0.00	0.00	0.00	0.00
ΔI	0.22	0.16	0.07	0.05	0.24	0.19	0.07	0.07
ΔX	0.40	0.24	0.05	-0.01	0.35	0.27	0.04	0.01
ΔM	0.37	0.24	0.03	0.01	0.35	0.28	0.05	0.01

6 Testing for "Weak Informational Efficiency"

As shown in Section 5, the 2008 crisis led to a significant decrease in the quality of forecasts. The large forecast errors for the 2008–2010 period constitute atypical (influential) observations, which greatly affect OLS estimates in regressions (17) and (18). As such, we restrict the tests for "weak informational efficiency" to the subsample period 1993–2007.[13]

Table 11 presents the efficiency tests for the model in equation (17). We included only one lagged term, since there was no evidence of higher order serial correlation. Recall that forecasts are efficient in this context if and only if they are unbiased and serially uncorrelated. The former requires that $\gamma_0/(1-\gamma_1)=0$, whereas the latter imposes $\gamma_1 = 0$. Evidence suggests that forecasts are, in general, unbiased. The main exceptions are forecasts for government consumption at several time spans, and 1–step ahead forecasts for imports and exports. Serial correlation characterizes forecasts for private consumption and government consumption at all horizons, and forecasts for GDP at most time spans. Investment forecasts are also serially correlated in some cases.

An alternative and more robust test for bias, using model (18), is displayed in Table 12. The χ^2 test on the joint hypothesis $\alpha = 0$ and $\beta = 1$ indicates that forecasts are in general inefficient. In particular, the null hypothesis of efficiency is rejected for government consumption at all time spans (with 1 exception), and for GDP at all but the 2–period span. One also rejects the null hypothesis of efficiency at 1–year spans for private consumption, exports and imports (with 1 exception). Results for GDP are consistent, for instance, with those in Loungani (2001), who finds that Consensus Forecasts are biased for 1–year ahead predictions.

There are several reasons which explain inefficiency in forecasts for government consumption. First, this variable is often used by policy–makers to manipulate the economic cycle and to boost GDP. These changes in policy are often unexpected and difficult to predict, even in the short run. Second, the supply of public goods is often chosen by bureaucrats, whose decisions may be driven by self–interests, such as power or reputation, rather than by an optimal allocation rule (*i.e.* Samuelson rule). This may originate an unpredictable increase in government expenditure. Third, legislators and officials may grant numerous favors to interest groups in response to rent–seeking efforts. Several of these favors may lead to an higher government size than previously anticipated. Finally, policy–makers may deliberately increase public expenditures in election years beyond the budgeted, in order to impress voters with an increase in the provision of public goods or publicly–financed goods.

[13] In what follows, we use a 5% significance level unless otherwise stated.

Table 11: *Tests for "weak informational efficiency" (1993–2007 period): Model in equation (17).*

		OECD				IMF			
		4–step *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*	*4–step* *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*
GDP	$\gamma_0/(1-\gamma_1)$	-0.21	0.08	0.20	0.23***	-0.18	-0.08	0.26	0.25**
		(0.40)	(0.27)	(0.19)	(0.09)	(0.43)	(0.35)	(0.20)	(0.12)
	γ_1	0.35**	0.18	0.32**	0.28**	0.43***	0.35**	0.20	0.33***
		(0.14)	(0.15)	(0.13)	(0.13)	(0.13)	(0.14)	(0.13)	(0.12)
C	$\gamma_0/(1-\gamma_1)$	-0.02	0.23	0.29*	0.18	0.03	0.05	0.28	0.15
		(0.37)	(0.22)	(0.15)	(0.11)	(0.40)	(0.29)	(0.20)	(0.12)
	γ_1	0.57***	0.30**	0.28**	0.35***	0.66***	0.51***	0.27**	0.33***
		(0.12)	(0.13)	(0.13)	(0.12)	(0.11)	(0.12)	(0.13)	(0.13)
G	$\gamma_0/(1-\gamma_1)$	0.46**	0.41**	0.27	0.18	0.46**	0.32	0.45***	0.21**
		(0.23)	(0.18)	(0.19)	(0.15)	(0.19)	(0.22)	(0.14)	(0.10)
	γ_1	0.42***	0.42***	0.39***	0.30***	0.28**	0.33***	0.29***	0.28**
		(0.10)	(0.11)	(0.10)	(0.11)	(0.11)	(0.12)	(0.11)	(0.11)
I	$\gamma_0/(1-\gamma_1)$	-0.52	-0.03	0.15	0.37	-0.52	-0.41	-0.14	0.22
		(0.73)	(0.53)	(0.41)	(0.33)	(0.65)	(0.63)	(0.32)	(0.34)
	γ_1	0.29**	0.15	0.06	0.15	0.30**	0.28**	0.04	0.18
		(0.13)	(0.14)	(0.13)	(0.13)	(0.13)	(0.14)	(0.12)	(0.11)
X	$\gamma_0/(1-\gamma_1)$	-1.34	-0.34	0.23	0.86***	-0.04	0.44	1.00	1.20**
		(1.16)	(0.75)	(0.70)	(0.32)	(1.27)	(1.1)	(0.67)	(0.49)
	γ_1	0.05	-0.19	0.05	0.22	0.15	0.10	-0.11	0.24*
		(0.18)	(0.16)	(0.15)	(0.15)	(0.17)	(0.17)	(0.15)	(0.13)
M	$\gamma_0/(1-\gamma_1)$	-0.60	0.32	0.53	0.90***	0.49	0.87	1.00	1.15***
		(1.15)	(0.80)	(0.57)	(0.33)	(1.23)	(1.09)	(0.65)	(0.44)
	γ_1	0.15	-0.06	-0.12	0.27**	0.23	0.12	-0.12	0.16
		(0.17)	(0.16)	(0.13)	(0.14)	(0.17)	(0.17)	(0.14)	(0.11)
Observations		91	91	98	98	91	91	98	98

Notes: (i) Efficiency tests for the model $e_{jt} = \gamma_0 + \gamma_1\, e_{j,t-1} + \varepsilon_{jt}$; (ii) Bias is evaluated by testing the null hypothesis $\gamma_0/(1-\gamma_1) = 0$, whereas serial correlation is evaluated by testing $\gamma_1 = 0$; (iii) Higher orders of serial dependence were insignificant and thus not included in the final specification; (iv) Panel–corrected standard errors in parenthesis; (v) *, ** and *** represent rejections at 10, 5 and 1 percent significance levels, respectively.

7 Concluding Remarks

This article analyzed the quality of forecasts for real GDP growth and for the corresponding expenditure components. We showed that GDP is in general overestimated at 1–year spans, due to overpredictions for investment and net exports. For same–year spans, GDP forecasts are accurate, but this is often achieved through an overestimation of exports and imports, whose effects cancel out. Nevertheless, the overall accuracy of component predictions—evaluated through two recently proposed statistics, the Mean of Total Weighted Absolute Errors and the Mean of Total Weighted Squared Errors—is substantially low, meaning that GDP forecasts are assembled with rather inaccurate component predictions. Investment, exports and imports are the major contributors to this outcome. The 2008 crisis has

Table 12: *Tests for "weak informational efficiency" (1993–2007 period): Model in equation (18).*

		OECD				IMF			
		4–step *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*	*4–step* *Spr t−1*	*3–step* *Aut t−1*	*2–step* *Spr t*	*1–step* *Aut t*
GDP	χ^2 test	24.97	7.58	3.18	11.54	17.84	7.63	2.74	7.97
	p-value	0.00	0.02	0.20	0.00	0.00	0.02	0.25	0.02
C	χ^2 test	12.66	10.95	3.84	4.90	21.85	10.58	2.53	2.05
	p-value	0.00	0.00	0.15	0.09	0.00	0.01	0.28	0.36
G	χ^2 test	16.13	12.04	12.14	5.78	34.49	41.81	13.99	15.86
	p-value	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.00
I	χ^2 test	12.28	3.14	0.06	6.14	7.98	1.29	0.97	2.85
	p-value	0.00	0.21	0.97	0.05	0.02	0.53	0.61	0.24
X	χ^2 test	11.12	1.80	0.13	15.99	21.09	6.34	2.11	9.58
	p-value	0.00	0.41	0.94	0.00	0.00	0.04	0.35	0.01
M	χ^2 test	22.85	10.99	0.39	12.63	16.07	19.59	1.75	3.64
	p-value	0.00	0.00	0.82	0.00	0.00	0.00	0.42	0.16
Observations		98	98	105	105	98	98	105	105

Notes: (i) Efficiency tests for the model $A_{jt} = \alpha + \beta F_{jt} + u_{jt},\ u_{jt} = \rho u_{j,t-1} + \varepsilon_{jt}$; (ii) The χ^2 test evaluates the null hypothesis that $\alpha = 0$ and $\beta = 1$ in the equation above.

worsened significantly the quality of forecasts being issued. We have also demonstrated that forecasts are in general inefficient. This result is stronger for government consumption forecasts, which are both biased and serially correlated.

In this article, we pooled all evaluation statistics across G7 countries, in order to obtain an overview of the overall quality of institutions' forecasts. Naturally, forecasts for different countries might present distinct behaviors. In the appendix, we illustrate how forecast quality changes across countries. Therein, we conclude that forecasts for the US are the most accurate, whereas those Canada and Germany are the least accurate.

References

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

Clemen, R. T. and R. L. Winkler (1986): "Combining economic forecasts," *Journal of Business and Economic Statistics*, 4, 39–46.

Fildes, R. and H. O. Stekler (2002): "The state of macroeconomic forecasting," *Journal of Macroeconomics*, 24, 435–468.

Frankel, J. (2011): "Over-optimism in forecasts by official budget agencies and its implications," Working Paper No 17239, National Bureau of Economic Research.

Holden, K. and D. A. Peel (1985): "An evaluation of quarterly National Institute forecasts," *Journal of Forecasting*, 4, 227–234.

——— (1990): "On testing for unbiasedness and efficiency of forecasts," *The Manchester School*, 58, 120–127.

Hyndman, R. J. and A. B. Koehler (2006): "Another look at measures of forecast accuracy," *International Journal of Forecasting*, 22, 679–688.

Joutz, F. and H. O. Stekler (2000): "An evaluation of the predictions of the Federal Reserve," *International Journal of Forecasting*, 16, 17–38.

Joutz, F. L. (1988): "Informational efficiency tests of quarterly macroeconometric GNP forecasts from 1976 to 1985," *Managerial and Decision Economics*, 9, 311–330.

Júlio, P., P. M. Esperança, and J. C. Fonseca (2011): "Evaluating the forecast quality of GDP components," GEE Papers No. 41, Office for Strategy and Studies, Portuguese Ministry of Economy and Employment.

Keane, M. and D. Runkle (1990): "Testing the rationality of price forecasts: new evidence from panel data," *American Economic Review*, 80, 714–735.

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.

McNees, S. K. (1976): "An evaluation of economic forecasts: extension and update," *New England Economic Review*, 20–44.

——— (1978): "The 'rationality' of economic forecasts," *American Economic Review*, 68, 301–305.

——— (1986): "Forecasting accuracy of alternative techniques: a comparison of US macroeconomic forecasts," *Journal of Business and Economic Statistics*, 4, 5–15.

——— (1988): "How accurate are macroeconomic forecasts?" *New England Economic Review*, 15–36.

Melander, A., G. Sismanidis, and D. Grenouilleau (2007): "The track record of the Commission's forecasts—an update," European Economy—Economic Papers 291, Directorate General Economic and Monetary Affairs, European Commission.

Nordhaus, W. D. (1987): "Forecasting efficiency: concepts and applications," *Review of Economics and Statistics*, 69, 667–674.

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

Pons, J. (2000): "The accuracy of IMF and OECD forecasts for G7 countries," *Journal of Forecasting*, 19, 53–63.

Stekler, H. O. (1972): "An analysis of turning point forecasts," *American Economic Review*, 62, 724–729.

——— (1987): "Who forecasts better?" *Journal of Business and Economic Statistics*, 5, 155–158.

Theil, H. (1966): *Applied economic forecasting*, Amsterdam: North-Holland Pub. Co.

Timmermann, A. (2007): "An evaluation of the World Economic Outlook forecasts," *IMF Staff Papers*, 54, 1–33.

Vuchelen, J. and M. I. Gutierrez (2005): "A direct test of the information content of the OECD growth forecasts," *International Journal of Forecasting*, 21, 103–117.

Wallis, K. F. (1989): "Macroeconomic forecasting: a survey," *Economic Journal*, 99, 28–61.

Zarnowitz, V. (1979): "An analysis of annual and multiperiod quarterly forecasts of aggregate income, output, and the price level," *Journal of Business*, 52, 1–33.

——— (1984): "The accuracy of individual and group forecasts from business outlook surveys," *Journal of Forecasting*, 3, 11–26.

Zarnowitz, V. and P. Braun (1993): "Twenty-two years of the NBER-ASA Quarterly Economic Outlook Surveys: aspects and comparisons of forecasting performance," in *Business Cycles, Indicators and Forecasting*, National Bureau of Economic Research, NBER Chapters, 11–94.

Appendices

A Gross Domestic Product

Table A.1 presents individual statistics for GDP growth forecasts. The ME statistic shows that OECD and IMF overestimate GDP growth in France, Germany, Italy, and Japan at 1–year spans. For these countries, bias decreases consistently until the 2–period span. For Canada there is a marginal overestimation at the 4–period span. UK's GDP is clearly underestimated, whereas for the US one cannot identify a clear pattern of over or underestimation in 1–year ahead predictions. Same–year forecasts do not display a significant bias except for the UK, where GDP is clearly underestimated.

The RMSE, MAE and MASE statistics show that the most accurate 1–year ahead forecasts are issued for France, whereas the least accurate 1–year ahead forecasts are issued for Germany and Japan. Same–year predictions display the highest accuracy in France and the lowest accuracy in Japan and the UK. These statistics also reveal that OECD's forecasts have a higher accuracy than IMF's forecasts at the 3–period span (namely in France, Germany, Italy, and Japan). At the remaining horizons, the forecast accuracy is similar, with OECD's forecasts outperforming IMF's for some countries and *vice-versa*.

The U2 statistic shows that institutions' forecasts outperform the naïve method at all time spans. However, at the 4–period span, institutions' forecasts performance is only marginally superior than the last observed period extrapolation as U2 is close to 1 for most countries.

B GDP Expenditure Components

Individual statistics for GDP expenditure components are presented in Table B.1 for OECD's forecasts and in Table B.2 for IMF's forecasts. According to ME, private consumption is overestimated in some countries and underestimated in others at 1–year spans. Same–year forecasts for this variable are nearly unbiased relative to other components, except for Canada and the UK. Government consumption is underestimated in Canada, Germany, Italy, and Japan, at all time spans. For the remaining countries, bias is negligible. Investment is overestimated in all countries except Canada and the UK at 1–year horizons. For same–year predictions, IMF also underestimates investment (although to a lesser extent), except for France and the UK, whereas for OECD results are mixed. Exports are also tendencially overestimated at 1–year spans, with the exceptions of Germany and, in the case of IMF's forecasts, Japan and the UK. For same–year predictions the reverse situation occurs, as exports are underestimated by both institutions in general, with few exceptions. As for imports, OECD's forecasts are upwardly biased at 1–year horizons, except for Germany, the UK, and the US. For same–year predictions, there is no clear pattern of over or underestimation. IMF underestimates imports at most horizons and for most countries.

As emphasized in the main text, RMSE and MAE suggest that forecasts for investment, exports, and imports are the least accurate. However, these components are also

Table A.1: *Individual statistics: GDP*

		OECD				IMF			
		4–step Spr $t-1$	*3–step* Aut $t-1$	*2–step* Spr t	*1–step* Aut t	*4–step* Spr $t-1$	*3–step* Aut $t-1$	*2–step* Spr t	*1–step* Aut t
Canada	ME	-0.4	0.0	0.1	0.3	-0.1	-0.1	0.2	0.2
	RMSE	1.9	1.5	0.9	0.6	1.8	1.6	1.0	0.6
	MAE	1.5	1.3	0.7	0.4	1.4	1.3	0.8	0.4
	MASE	0.9	0.8	0.4	0.2	0.8	0.8	0.5	0.3
	U2	0.8	0.6	0.4	0.3	0.8	0.7	0.5	0.3
France	ME	-0.6	-0.3	-0.1	0.1	-0.7	-0.5	-0.1	0.0
	RMSE	1.6	1.0	0.6	0.5	1.5	1.2	0.7	0.4
	MAE	1.2	0.8	0.5	0.4	1.1	1.0	0.6	0.3
	MASE	0.9	0.6	0.4	0.3	0.8	0.8	0.4	0.2
	U2	0.8	0.5	0.4	0.3	0.8	0.6	0.4	0.2
Germany	ME	-0.8	-0.4	0.1	0.0	-0.7	-0.5	0.2	0.0
	RMSE	2.4	1.6	0.9	0.5	2.4	2.1	1.1	0.7
	MAE	1.8	1.3	0.7	0.5	1.9	1.7	0.8	0.6
	MASE	0.9	0.7	0.3	0.2	1.0	0.9	0.4	0.3
	U2	0.9	0.6	0.3	0.2	0.9	0.8	0.4	0.2
Italy	ME	-0.9	-0.5	0.0	0.1	-1.0	-0.8	-0.2	0.1
	RMSE	2.1	1.5	0.8	0.5	2.0	1.8	0.9	0.6
	MAE	1.5	1.1	0.7	0.4	1.5	1.3	0.8	0.5
	MASE	0.8	0.6	0.4	0.2	0.8	0.7	0.4	0.3
	U2	0.8	0.6	0.3	0.2	0.8	0.7	0.3	0.2
Japan	ME	-0.9	-0.5	-0.1	-0.1	-1.0	-0.8	0.0	0.0
	RMSE	2.8	2.1	1.0	1.0	2.8	2.3	1.1	1.0
	MAE	2.0	1.4	0.8	1.0	1.9	1.4	0.9	0.8
	MASE	0.9	0.6	0.4	0.4	0.9	0.7	0.4	0.4
	U2	0.9	0.6	0.3	0.3	0.9	0.7	0.4	0.3
UK	ME	0.3	0.5	0.7	0.7	0.3	0.4	0.7	0.7
	RMSE	2.1	1.7	1.5	1.2	2.2	1.9	1.5	1.3
	MAE	1.5	1.4	1.2	0.9	1.5	1.4	1.1	1.0
	MASE	0.9	0.8	0.7	0.5	0.9	0.8	0.6	0.6
	U2	0.8	0.6	0.6	0.5	0.8	0.7	0.6	0.5
US	ME	-0.1	0.1	-0.1	-0.2	0.1	0.1	0.0	-0.1
	RMSE	2.0	1.6	1.0	0.8	2.0	1.7	0.9	0.8
	MAE	1.7	1.3	0.8	0.7	1.7	1.4	0.8	0.7
	MASE	1.2	0.9	0.6	0.5	1.2	1.0	0.5	0.5
	U2	0.9	0.7	0.5	0.4	0.9	0.8	0.4	0.4

Table B.1: *Individual statistics: GDP expenditure components, OECD's forecasts*

		Private Consumption (C)				Gov. Consumption (G)				Investment (I)				Exports (X)				Imports (M)			
		4–s	*3–s*	*2–s*	*1–s*	*4–s*	*3–s*	*2–s*	*1–s*	*4–s*	*3–s*	*2–s*	*1–s*	*4–s*	*3–s*	*2–s*	*1–s*	*4–s*	*3–s*	*2–s*	*1–s*
Canada	ME	0.3	0.5	0.4	0.2	0.2	0.2	0.0	0.2	-0.1	0.1	0.4	-0.3	-2.7	-1.9	-1.3	-0.1	-1.4	-1.5	-0.6	0.0
	RMSE	1.2	1.1	0.9	0.5	1.6	1.4	1.3	1.0	6.1	4.9	3.7	3.1	6.5	5.1	3.5	1.6	6.4	5.5	3.7	2.1
	MAE	1.0	0.9	0.8	0.4	1.3	1.2	1.0	0.7	4.4	3.7	3.0	2.5	5.2	3.8	2.8	1.2	4.2	3.6	2.8	1.4
	MASE	0.8	0.7	0.7	0.3	1.2	1.1	0.9	0.6	0.7	0.6	0.5	0.4	1.1	0.8	0.6	0.3	0.7	0.6	0.5	0.2
	U2	0.9	0.7	0.7	0.4	0.9	0.8	0.9	0.7	0.8	0.7	0.4	0.4	1.0	0.7	0.5	0.2	0.8	0.7	0.4	0.2
France	ME	-0.4	0.1	0.0	0.0	0.0	-0.1	-0.1	-0.3	-0.8	-0.1	0.0	0.5	-1.6	-1.0	0.3	0.6	-0.9	-0.3	0.6	0.9
	RMSE	1.1	0.8	0.6	0.5	0.7	0.8	0.8	1.1	3.8	2.7	2.5	1.6	6.6	5.0	2.7	2.1	6.2	5.0	3.3	2.2
	MAE	0.9	0.6	0.4	0.4	0.5	0.6	0.6	0.8	2.9	2.5	2.1	1.2	5.3	3.9	2.2	1.7	5.0	4.0	2.6	1.9
	MASE	1.2	0.8	0.5	0.6	0.6	0.6	0.6	0.8	0.8	0.7	0.6	0.3	0.9	0.7	0.4	0.3	0.9	0.7	0.5	0.3
	U2	0.8	0.6	0.5	0.4	0.5	0.6	0.6	0.8	0.6	0.5	0.5	0.3	1.0	0.8	0.4	0.3	0.9	0.7	0.4	0.3
Germany	ME	-0.7	-0.3	0.1	0.2	0.4	0.7	0.6	0.2	-2.6	-1.3	-0.7	-0.1	0.4	0.8	1.3	1.5	-0.2	0.8	1.0	1.3
	RMSE	1.4	1.2	0.9	0.7	1.1	1.1	1.1	0.9	4.9	3.9	3.2	1.5	6.6	5.3	3.2	2.2	5.9	5.2	3.9	2.4
	MAE	1.3	1.0	0.7	0.5	1.0	0.9	0.8	0.7	3.4	2.9	2.5	1.3	4.7	4.1	2.8	1.8	4.5	4.3	3.4	2.2
	MASE	1.6	1.3	0.8	0.6	0.9	0.8	0.6	0.6	0.7	0.6	0.5	0.3	0.7	0.6	0.4	0.3	0.8	0.7	0.6	0.4
	U2	1.3	1.1	0.8	0.6	0.9	0.9	0.7	0.5	0.8	0.6	0.5	0.2	0.8	0.6	0.3	0.2	0.8	0.7	0.5	0.3
Italy	ME	-0.6	-0.1	0.0	0.1	0.4	0.4	0.4	0.1	-2.4	-1.4	-0.2	0.3	-2.4	-1.6	0.0	0.5	-1.4	-0.5	0.0	1.1
	RMSE	1.5	1.1	1.2	0.9	1.6	1.6	1.6	1.5	4.3	3.1	2.5	2.1	7.9	6.4	4.4	2.4	6.9	5.8	4.8	2.5
	MAE	1.2	0.9	0.9	0.7	1.2	1.3	1.3	1.1	3.2	2.4	2.0	1.4	6.2	5.3	3.7	2.1	5.6	4.7	3.7	1.9
	MASE	1.1	0.8	0.7	0.6	1.2	1.3	1.2	1.1	0.8	0.6	0.4	0.3	0.8	0.7	0.5	0.3	0.8	0.7	0.5	0.2
	U2	0.8	0.6	0.6	0.5	0.9	0.9	1.1	1.0	0.7	0.5	0.4	0.3	0.8	0.7	0.5	0.3	0.7	0.6	0.5	0.2
Japan	ME	-0.6	-0.4	0.0	-0.2	0.7	0.9	0.6	0.9	-2.2	-1.6	-1.1	-0.7	-1.7	-0.5	0.3	0.0	-1.6	-0.8	-0.4	0.0
	RMSE	1.4	1.4	0.8	0.9	1.8	1.7	1.5	1.5	5.0	4.0	2.3	2.2	10.3	8.1	4.8	1.7	7.4	6.9	3.3	1.8
	MAE	1.0	1.1	0.5	0.8	1.4	1.3	0.9	1.0	3.7	2.7	1.9	1.5	6.8	5.7	3.7	1.4	5.7	5.4	2.7	1.6
	MASE	0.9	1.0	0.5	0.7	1.4	1.3	0.9	1.0	0.9	0.7	0.5	0.4	0.7	0.6	0.4	0.1	0.8	0.8	0.4	0.2
	U2	0.9	0.8	0.5	0.6	1.1	1.0	1.2	1.3	0.9	0.7	0.5	0.5	0.8	0.7	0.3	0.1	0.7	0.7	0.4	0.2
UK	ME	0.6	0.8	0.7	0.4	0.2	0.1	-0.1	0.0	-0.1	0.8	1.3	1.9	-1.0	0.1	1.8	2.6	-0.4	0.7	1.5	2.3
	RMSE	2.2	2.0	1.7	1.3	1.4	1.5	1.7	1.3	7.6	6.4	5.3	4.3	6.2	4.7	3.7	4.1	6.9	5.5	4.1	4.1
	MAE	1.7	1.6	1.2	0.9	1.1	1.2	1.3	0.9	5.0	4.2	3.7	3.1	5.0	3.7	2.4	2.8	5.2	4.2	2.6	2.6
	MASE	1.2	1.1	0.8	0.6	0.8	1.0	1.0	0.7	0.8	0.7	0.6	0.5	0.9	0.7	0.4	0.5	0.9	0.7	0.5	0.5
	U2	0.9	0.8	0.8	0.6	1.0	1.1	1.2	0.8	0.7	0.6	0.6	0.5	0.9	0.7	0.5	0.5	0.9	0.7	0.5	0.5
US	ME	0.4	0.5	-0.1	-0.1	0.3	0.0	0.1	0.0	-1.0	-0.4	-0.3	-1.0	-2.0	-1.4	0.2	0.1	0.0	0.2	0.5	-0.2
	RMSE	1.6	1.3	0.8	0.6	1.3	0.9	1.1	0.6	5.5	3.9	2.2	2.1	6.6	5.5	3.1	1.3	6.6	5.4	2.8	1.2
	MAE	1.3	1.0	0.6	0.5	1.0	0.8	0.9	0.4	4.4	2.9	1.8	1.8	4.7	3.9	2.2	1.1	5.5	4.1	2.3	0.9
	MASE	1.3	1.0	0.7	0.5	1.1	0.9	1.0	0.5	1.1	0.7	0.5	0.5	0.8	0.6	0.4	0.2	1.0	0.7	0.4	0.2
	U2	0.9	0.7	0.6	0.4	0.8	0.6	1.1	0.5	0.8	0.6	0.4	0.4	0.9	0.7	0.4	0.2	0.8	0.7	0.3	0.1

Table B.2: *Individual statistics: GDP expenditure components, IMF's forecasts*

		Private Consumption (C)				Gov. Consumption (G)				Investment (I)				Exports (X)				Imports (M)			
		4–s	*3–s*	*2–s*	*1–s*	*4–s*	*3–s*	*2–s*	*1–s*	*4–s*	*3–s*	*2–s*	*1–s*	*4–s*	*3–s*	*2–s*	*1–s*	*4–s*	*3–s*	*2–s*	*1–s*
Canada	ME	0.5	0.4	0.5	0.1	0.6	0.6	0.5	0.2	0.0	0.4	-0.2	-0.1	-1.2	-0.9	-0.3	0.0	0.2	0.6	0.2	-0.4
	RMSE	1.4	1.2	1.1	0.6	1.7	1.7	1.3	1.0	5.8	4.6	3.9	3.3	6.7	6.1	4.4	2.1	6.7	6.7	4.7	2.9
	MAE	1.1	0.9	1.0	0.5	1.4	1.5	1.1	0.8	4.2	3.7	3.1	2.7	6.0	5.0	3.5	1.6	5.1	5.2	3.4	1.8
	MASE	0.9	0.8	0.8	0.4	1.2	1.3	1.0	0.7	0.6	0.6	0.5	0.4	1.2	1.0	0.7	0.3	0.8	0.9	0.6	0.3
	U2	0.9	0.8	0.8	0.5	1.0	1.0	0.9	0.7	0.8	0.6	0.4	0.4	1.0	0.9	0.7	0.3	0.9	0.9	0.5	0.3
France	ME	-0.4	-0.2	0.0	-0.1	-0.2	-0.2	-0.2	-0.2	-1.0	-0.7	-0.1	0.4	-1.5	-0.8	0.7	1.5	-0.7	-0.1	0.6	1.5
	RMSE	1.0	0.9	0.7	0.6	0.9	0.7	0.7	1.0	4.1	3.6	2.7	1.8	6.3	5.5	3.7	3.1	5.9	5.5	4.3	3.0
	MAE	0.8	0.7	0.5	0.5	0.7	0.5	0.5	0.8	3.1	2.9	2.3	1.5	5.2	4.4	3.1	2.3	4.9	4.4	3.6	2.4
	MASE	1.0	0.9	0.7	0.6	0.7	0.6	0.5	0.8	0.9	0.8	0.6	0.4	0.9	0.7	0.5	0.4	0.8	0.8	0.6	0.4
	U2	0.7	0.7	0.6	0.5	0.6	0.4	0.6	0.8	0.7	0.6	0.6	0.4	1.0	0.8	0.5	0.4	0.8	0.8	0.6	0.4
Germany	ME	-0.8	-0.5	0.1	0.2	0.3	0.4	0.8	0.1	-2.2	-1.7	-0.7	-0.2	1.6	2.1	2.1	1.4	0.9	1.3	1.7	1.5
	RMSE	1.6	1.4	1.0	0.7	1.4	1.3	1.2	1.1	5.8	4.8	3.0	1.8	6.7	6.1	4.4	2.5	5.8	5.5	5.1	2.6
	MAE	1.4	1.3	0.7	0.5	1.2	1.2	0.9	0.8	4.5	3.8	2.4	1.4	4.8	4.6	3.6	2.1	4.4	4.4	4.2	2.3
	MASE	1.7	1.6	0.8	0.6	1.1	1.1	0.7	0.6	1.0	0.8	0.5	0.3	0.7	0.7	0.5	0.3	0.8	0.8	0.7	0.4
	U2	1.5	1.3	0.8	0.6	1.1	1.0	0.7	0.7	0.9	0.8	0.5	0.3	0.8	0.7	0.5	0.3	0.8	0.7	0.6	0.3
Italy	ME	-0.8	-0.5	-0.2	0.0	0.2	0.2	0.4	0.2	-1.6	-1.5	-0.4	-0.1	-1.7	-1.1	0.1	0.7	-0.8	-0.1	0.2	0.9
	RMSE	1.4	1.3	1.1	0.9	1.8	1.8	1.2	1.3	4.1	3.8	2.8	2.4	7.6	7.1	4.3	3.1	6.3	6.1	4.8	4.2
	MAE	1.2	1.1	0.9	0.7	1.3	1.3	1.1	1.0	3.0	2.6	1.8	1.7	6.0	5.7	3.8	3.0	4.8	5.0	3.7	3.3
	MASE	1.1	1.1	0.7	0.5	1.4	1.3	1.0	0.9	0.7	0.6	0.4	0.4	0.8	0.8	0.5	0.4	0.7	0.7	0.5	0.4
	U2	0.8	0.7	0.6	0.5	1.0	1.0	0.9	0.9	0.6	0.6	0.5	0.4	0.8	0.7	0.4	0.3	0.7	0.7	0.5	0.4
Japan	ME	-0.9	-0.6	-0.1	-0.3	0.7	0.7	0.7	0.9	-2.9	-2.9	-1.6	-0.8	0.9	1.0	1.9	0.9	-0.9	-0.3	-0.3	0.1
	RMSE	1.8	1.3	0.8	0.8	1.5	1.5	1.4	1.4	5.0	4.4	3.0	2.1	10.8	8.9	5.5	2.4	7.4	6.8	3.4	2.4
	MAE	1.4	1.1	0.6	0.7	1.1	1.1	1.0	1.0	3.5	3.1	2.5	1.5	7.8	6.2	4.6	2.0	5.7	5.7	3.0	1.9
	MASE	1.2	1.0	0.6	0.6	1.1	1.1	1.1	1.0	0.9	0.8	0.7	0.4	0.8	0.6	0.5	0.2	0.8	0.8	0.5	0.3
	U2	1.1	0.8	0.5	0.5	0.9	0.9	1.1	1.2	0.9	0.8	0.6	0.4	0.9	0.7	0.4	0.2	0.7	0.7	0.4	0.3
UK	ME	0.7	0.6	0.8	0.6	0.2	0.0	0.0	-0.1	0.5	1.0	0.8	1.8	0.6	1.1	2.0	2.2	1.6	1.6	2.0	2.0
	RMSE	2.4	2.0	1.7	1.5	1.6	2.0	1.3	1.3	7.7	6.2	5.1	4.3	5.8	5.3	4.2	3.7	6.7	5.8	4.2	4.1
	MAE	1.9	1.6	1.4	1.1	1.1	1.4	1.1	1.0	4.9	4.4	3.5	3.2	4.8	4.3	2.9	2.8	5.0	4.5	2.8	2.8
	MASE	1.3	1.1	0.9	0.7	0.9	1.1	0.9	0.8	0.8	0.7	0.6	0.5	0.8	0.7	0.5	0.5	0.9	0.8	0.5	0.5
	U2	1.0	0.8	0.9	0.7	1.2	1.4	0.9	0.9	0.7	0.6	0.5	0.5	0.8	0.8	0.6	0.5	0.9	0.8	0.6	0.5
US	ME	0.4	0.6	0.2	0.0	0.5	-0.1	0.1	0.0	-1.0	-0.7	-0.4	-1.2	-1.0	-0.4	0.0	0.8	0.6	1.1	0.9	0.9
	RMSE	1.8	1.5	0.9	0.5	1.5	1.5	1.2	0.8	5.5	4.8	2.2	2.3	6.5	6.6	3.4	2.9	6.9	6.4	3.5	2.3
	MAE	1.3	1.1	0.7	0.4	1.2	1.1	1.0	0.6	4.6	3.9	1.6	2.0	4.9	4.7	2.2	1.9	5.8	5.3	2.5	1.5
	MASE	1.3	1.1	0.7	0.4	1.3	1.2	1.0	0.7	1.2	1.0	0.4	0.5	0.8	0.8	0.4	0.3	1.0	0.9	0.4	0.3
	U2	0.9	0.8	0.6	0.4	1.0	1.0	1.1	0.7	0.8	0.7	0.4	0.4	0.8	0.9	0.4	0.4	0.9	0.8	0.4	0.3

more volatile and thus harder to predict. Thus, one should look at MASE to compare the performance of forecast models across GDP components. According to this statistic, forecast models perform comparatively worse when predicting private consumption and government consumption, although specific conclusions vary from country to country. In particular, forecast models perform worse when predicting Germany's private consumption, and Italy's and Japan's government consumption. For other components, the performance is lower in predictions for US's investment and imports, and for Canada's exports. In the remaining cases, the performance of forecast models very similar.

The U2 statistic reveals that forecast models perform generally better than the *naïve forecast*. The major exceptions are Germany's forecasts for private consumption (at 4– and 3–period spans), and Japan's, UK's and US's forecasts for government consumption.

C Decomposing GDP Forecast Errors into Component Contributions

For some countries and time spans, forecast errors in component contributions are comparatively large and tend to add up, leading to large biases in forecasted GDP (*e.g.* France, Germany, Italy and Japan for 1–year ahead forecasts, and the UK for same–year forecasts). For other countries and/or time spans, forecast errors in component contributions are also comparatively large but tend to cancel out, resulting in a relatively small GDP forecast bias (*e.g.* Canada, the UK and the US for 1–year ahead forecasts, and Canada, Germany and Japan for same–year forecasts). Finally, in other cases, forecast errors in component contributions are comparatively small and result in small GDP prediction biases (*e.g.* France, Italy and the US for same–year forecasts). Naturally, this latter outcome is more common for same–year predictions.

Canada and Italy have the largest external sector contributions to the GDP forecast error at 1–year spans. For these countries, exports and imports are generally overestimated, but the (positive) contribution of the former to the GDP forecast error dominates the (negative) contribution of the latter. In France, exports and imports are also overestimated, but the contribution of the former to the GDP forecast error is to a great extent offset by the latter. External sector forecasts issued by IMF for Germany also display a large contribution to the GDP forecast error, but in this case there is an underestimation of both variables. For other countries, there is no identifiable pattern. In Germany, Italy and Japan, private consumption and investment are substantially overestimated at 1–year horizons, contributing to an overestimated GDP. In the case of Italy, this adds up with the effects of an overestimated external sector. On the opposite direction, private consumption for Canada, the UK and the US is underestimated, but this effect interacts with that of overestimated components, thus originating a comparatively small bias in GDP forecasts.

For same–year predictions, component contributions to the GDP forecast error are substantially smaller except for the UK, explaining more accurate GDP forecasts. Nevertheless, the contributions of exports and imports are still large for some countries, most noticeably Germany and the UK. The US presents the lowest component contributions to the GDP forecast error.

Table C.1: *Contributions of expenditure components (in percentage points) to the average GDP forecast error.*

		OECD				IMF			
		4–step Spr $t-1$	*3–step* Aut $t-1$	*2–step* Spr t	*1–step* Aut t	*4–step* Spr $t-1$	*3–step* Aut $t-1$	*2–step* Spr t	*1–step* Aut t
Canada	C	0.15	0.31	0.21	0.13	0.29	0.24	0.28	0.06
	G	0.02	0.02	0.00	0.04	0.11	0.12	0.10	0.04
	I	0.02	0.07	0.20	0.06	0.04	0.14	0.08	0.13
	X	-1.13	-0.80	-0.53	-0.06	-0.57	-0.43	-0.19	-0.05
	M	-0.64	-0.64	-0.22	0.00	-0.02	0.11	0.03	-0.16
	υ	-0.08	-0.21	-0.04	0.13	-0.03	-0.03	-0.05	-0.10
	GDP	-0.38	0.03	0.06	0.30	-0.15	-0.07	0.18	0.24
France	C	-0.22	0.03	-0.02	-0.03	-0.24	-0.11	0.00	-0.07
	G	0.00	-0.03	-0.03	-0.08	-0.04	-0.05	-0.05	-0.04
	I	-0.10	0.05	0.02	0.11	-0.12	-0.07	-0.01	0.09
	X	-0.48	-0.32	0.03	0.08	-0.46	-0.29	0.09	0.27
	M	-0.32	-0.16	0.10	0.15	-0.30	-0.14	0.09	0.29
	υ	-0.16	-0.18	0.05	0.14	-0.13	-0.09	0.01	0.04
	GDP	-0.63	-0.30	-0.07	0.07	-0.71	-0.46	-0.05	0.01
Germany	C	-0.42	-0.16	0.07	0.14	-0.50	-0.33	0.04	0.10
	G	0.08	0.13	0.11	0.03	0.06	0.07	0.14	0.01
	I	-0.52	-0.27	-0.17	-0.04	-0.45	-0.36	-0.18	-0.07
	X	-0.03	0.13	0.46	0.40	0.39	0.59	0.78	0.43
	M	-0.17	0.12	0.25	0.30	0.17	0.33	0.57	0.41
	υ	-0.07	-0.10	-0.10	-0.26	-0.09	-0.19	-0.03	-0.06
	GDP	-0.80	-0.39	0.12	-0.02	-0.75	-0.54	0.19	0.00
Italy	C	-0.35	-0.08	-0.02	0.08	-0.44	-0.32	-0.12	0.00
	G	0.06	0.07	0.07	0.01	0.04	0.03	0.08	0.03
	I	-0.43	-0.23	0.01	0.12	-0.27	-0.27	-0.03	0.03
	X	-0.67	-0.46	-0.03	0.14	-0.49	-0.35	-0.02	0.16
	M	-0.41	-0.17	0.01	0.28	-0.25	-0.07	0.04	0.23
	υ	0.06	0.01	-0.06	0.01	-0.05	0.09	-0.05	0.06
	GDP	-0.93	-0.53	-0.03	0.09	-0.97	-0.76	-0.18	0.05
Japan	C	-0.35	-0.22	-0.02	-0.13	-0.52	-0.32	-0.04	-0.15
	G	0.11	0.14	0.10	0.14	0.12	0.11	0.11	0.14
	I	-0.57	-0.42	-0.32	-0.18	-0.74	-0.71	-0.42	-0.21
	X	-0.28	-0.12	0.04	-0.01	0.03	0.05	0.23	0.08
	M	-0.19	-0.11	-0.05	-0.01	-0.13	-0.07	-0.03	0.01
	υ	-0.01	0.03	0.05	0.08	-0.01	0.02	0.09	0.10
	GDP	-0.91	-0.48	-0.10	-0.09	-0.98	-0.78	0.01	-0.04

Continued on next page

Table C.1: *Contributions of expenditure components to the average GDP forecast error (continued).*

		OECD				IMF			
		4–step Spr $t-1$	*3–step* Aut $t-1$	*2–step* Spr t	*1–step* Aut t	*4–step* Spr $t-1$	*3–step* Aut $t-1$	*2–step* Spr t	*1–step* Aut t
UK	C	0.40	0.50	0.42	0.25	0.45	0.38	0.47	0.35
	G	0.04	0.03	-0.03	-0.01	0.04	0.01	0.00	-0.02
	I	-0.08	0.06	0.17	0.28	0.02	0.11	0.10	0.25
	X	-0.36	-0.07	0.42	0.63	0.07	0.19	0.47	0.56
	M	-0.28	0.07	0.36	0.52	0.28	0.31	0.49	0.48
	υ	0.02	0.07	0.11	0.04	0.01	0.04	0.13	0.06
	GDP	0.30	0.52	0.74	0.67	0.31	0.41	0.68	0.73
US	C	0.26	0.31	-0.07	-0.06	0.27	0.42	0.15	0.00
	G	0.05	0.01	0.02	0.00	0.08	-0.01	0.01	0.00
	I	-0.18	-0.05	-0.04	-0.16	-0.20	-0.14	-0.06	-0.20
	X	-0.22	-0.17	0.03	0.01	-0.13	-0.08	0.00	0.07
	M	-0.08	-0.03	0.04	-0.02	-0.02	0.06	0.08	0.12
	υ	-0.05	-0.01	0.01	0.03	0.02	0.01	0.02	0.16
	GDP	-0.07	0.11	-0.09	-0.17	0.05	0.14	0.04	-0.08

Notes: (i) **GDP** in the table represents the mean of e_t^{gdp}, which is equal to the Mean Error statistic in Table 3, by definition; (ii) The sum of the contributions in the table differs from **GDP** by an error υ whose source is explained in Section 2.5.

The MTWAE and the MTWSE statistics, detailed by country, are presented in Tables C.2 and C.3. At 1–year horizons, MTWAE fluctuates between 6.5 (IMF's forecasts for Canada at the 4–period span) and 2.7 (OECD's forecasts for the US at the 3–period span) percentage points. The overall accuracy of component predictions increases at same–year spans, leading to comparatively more accurate GDP forecasts (MTWAE fluctuates between 4.3 percentage points for Canada at the 2–period span and 1.1 percentage points for the US at the 1–period span). OECD's forecasts are more accurate than IMF's forecasts for most countries.

Table C.2: *The* MTWAE *statistic and its decomposition (in percentage points).*

		OECD				IMF			
		4–step	*3–step*	*2–step*	*1–step*	*4–step*	*3–step*	*2–step*	*1–step*
		Spr t−1	*Aut t−1*	*Spr t*	*Aut t*	*Spr t−1*	*Aut t−1*	*Spr t*	*Aut t*
Canada	MTWAE	5.81	4.77	3.51	2.01	6.48	5.92	4.26	2.49
	C	0.60	0.51	0.45	0.22	0.66	0.55	0.57	0.29
	G	0.27	0.24	0.22	0.15	0.29	0.30	0.23	0.16
	I	1.24	1.10	0.73	0.64	1.23	1.10	0.85	0.72
	X	2.05	1.49	1.05	0.46	2.33	1.95	1.34	0.62
	M	1.66	1.42	1.06	0.55	1.97	2.01	1.27	0.69
France	MTWAE	4.08	3.16	2.24	1.88	4.05	3.55	2.77	2.10
	C	0.53	0.34	0.24	0.25	0.43	0.38	0.30	0.27
	G	0.13	0.15	0.15	0.19	0.17	0.13	0.12	0.19
	I	0.87	0.75	0.72	0.63	0.97	0.88	0.78	0.63
	X	1.32	0.96	0.54	0.39	1.29	1.09	0.74	0.50
	M	1.22	0.96	0.60	0.41	1.18	1.07	0.83	0.51
Germany	MTWAE	5.04	4.42	3.23	2.11	5.28	4.98	3.77	2.27
	C	0.74	0.60	0.41	0.32	0.81	0.74	0.43	0.31
	G	0.19	0.16	0.16	0.13	0.23	0.22	0.17	0.15
	I	0.87	0.80	0.70	0.48	1.07	0.94	0.70	0.47
	X	1.76	1.51	0.98	0.57	1.77	1.68	1.23	0.68
	M	1.48	1.36	0.98	0.62	1.41	1.40	1.24	0.65
Italy	MTWAE	4.68	3.89	3.16	2.14	4.46	4.32	3.06	2.57
	C	0.70	0.51	0.56	0.42	0.69	0.65	0.55	0.39
	G	0.24	0.26	0.26	0.23	0.27	0.27	0.21	0.20
	I	0.85	0.71	0.60	0.50	0.83	0.77	0.56	0.51
	X	1.56	1.32	0.92	0.53	1.51	1.43	0.90	0.73
	M	1.33	1.10	0.83	0.46	1.16	1.21	0.83	0.74
Japan	MTWAE	3.39	2.96	1.75	1.48	3.63	3.13	2.02	1.49
	C	0.58	0.61	0.29	0.42	0.78	0.61	0.36	0.36
	G	0.23	0.21	0.14	0.16	0.18	0.18	0.17	0.16
	I	1.18	0.93	0.61	0.61	1.16	1.02	0.68	0.57
	X	0.86	0.70	0.45	0.15	0.95	0.77	0.53	0.22
	M	0.54	0.51	0.25	0.14	0.55	0.55	0.28	0.17
UK	MTWAE	4.78	3.91	2.72	2.49	4.70	4.21	2.90	2.70
	C	1.10	0.98	0.78	0.58	1.18	0.99	0.86	0.70
	G	0.24	0.27	0.29	0.20	0.25	0.32	0.25	0.23
	I	0.64	0.54	0.41	0.39	0.64	0.58	0.38	0.35
	X	1.33	0.96	0.59	0.68	1.24	1.09	0.70	0.70
	M	1.48	1.15	0.65	0.64	1.40	1.24	0.71	0.72
US	MTWAE	3.44	2.71	1.67	1.12	3.57	3.18	1.70	1.36
	C	0.89	0.71	0.45	0.32	0.90	0.76	0.47	0.27
	G	0.17	0.13	0.15	0.07	0.19	0.18	0.16	0.11
	I	1.10	0.87	0.54	0.50	1.15	1.02	0.53	0.59
	X	0.51	0.42	0.23	0.11	0.51	0.50	0.23	0.19
	M	0.78	0.58	0.31	0.11	0.81	0.72	0.32	0.20

Table C.3: *The* MTWSE *statistic and its decomposition.*

		OECD				IMF			
		4–step	*3–step*	*2–step*	*1–step*	*4–step*	*3–step*	*2–step*	*1–step*
		Spr $t-1$	*Aut* $t-1$	*Spr* t	*Aut* t	*Spr* $t-1$	*Aut* $t-1$	*Spr* t	*Aut* t
Canada	MTWSE	16.76	11.63	5.19	1.79	17.24	15.07	7.83	2.97
	C	0.55	0.38	0.30	0.09	0.65	0.46	0.46	0.14
	G	0.11	0.09	0.07	0.04	0.13	0.12	0.07	0.04
	I	2.68	1.89	0.96	0.68	2.58	1.82	1.16	0.82
	X	6.56	4.12	1.86	0.34	6.81	5.59	2.93	0.70
	M	6.85	5.16	1.99	0.64	7.07	7.08	3.21	1.27
France	MTWSE	7.15	4.30	1.83	1.16	6.76	5.38	2.83	1.53
	C	0.42	0.21	0.11	0.09	0.34	0.28	0.14	0.10
	G	0.03	0.04	0.04	0.07	0.05	0.03	0.03	0.06
	I	1.37	0.91	0.71	0.56	1.52	1.27	0.83	0.56
	X	2.84	1.60	0.44	0.21	2.53	1.89	0.77	0.42
	M	2.48	1.55	0.53	0.24	2.32	1.91	1.06	0.39
Germany	MTWSE	14.93	9.54	3.68	1.37	15.36	11.97	6.16	1.78
	C	0.69	0.52	0.30	0.15	0.95	0.73	0.32	0.16
	G	0.05	0.05	0.04	0.03	0.07	0.07	0.05	0.05
	I	1.57	1.10	0.72	0.31	1.98	1.47	0.62	0.32
	X	8.01	4.71	1.41	0.43	7.96	6.03	2.58	0.68
	M	4.62	3.16	1.20	0.45	4.39	3.68	2.60	0.56
Italy	MTWSE	9.62	6.32	3.51	1.51	8.74	7.73	3.27	2.16
	C	0.75	0.43	0.53	0.28	0.71	0.58	0.45	0.26
	G	0.10	0.10	0.10	0.08	0.13	0.13	0.06	0.06
	I	1.43	0.95	0.56	0.38	1.41	1.28	0.53	0.41
	X	4.24	2.81	1.17	0.39	3.93	3.33	1.07	0.61
	M	3.10	2.03	1.15	0.38	2.56	2.42	1.17	0.82
Japan	MTWSE	5.87	4.05	1.25	0.96	6.18	4.49	1.50	0.83
	C	0.61	0.58	0.20	0.26	0.95	0.55	0.18	0.18
	G	0.08	0.08	0.05	0.06	0.06	0.06	0.05	0.05
	I	2.53	1.75	0.53	0.58	2.46	1.91	0.71	0.48
	X	2.09	1.19	0.39	0.04	2.13	1.52	0.47	0.08
	M	0.55	0.46	0.09	0.02	0.57	0.45	0.10	0.04
UK	MTWSE	9.52	5.87	3.22	2.76	8.87	6.70	3.43	3.01
	C	1.90	1.51	1.10	0.64	2.28	1.66	1.17	0.89
	G	0.09	0.11	0.15	0.08	0.13	0.19	0.08	0.09
	I	0.80	0.53	0.30	0.29	0.77	0.47	0.27	0.23
	X	2.86	1.51	0.77	0.91	2.36	1.87	0.93	0.85
	M	3.87	2.21	0.90	0.83	3.32	2.52	0.97	0.95
US	MTWSE	4.69	3.02	1.10	0.65	5.01	4.09	1.31	0.90
	C	1.25	0.78	0.30	0.15	1.46	0.99	0.34	0.12
	G	0.04	0.02	0.03	0.01	0.06	0.06	0.04	0.02
	I	1.89	1.23	0.50	0.45	1.94	1.65	0.58	0.58
	X	0.56	0.38	0.11	0.02	0.53	0.55	0.13	0.08
	M	0.95	0.61	0.15	0.02	1.02	0.84	0.23	0.11