DEUTSCHE BUNDESBANK
EUROSYSTEM

Discussion Paper

Deutsche Bundesbank
No 15/2012

Cyclical adjustment in fiscal rules: some evidence on real-time bias for EU-15 countries

Gerhard Kempkes

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0
Telex within Germany 41227, telex from abroad 414431

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet http://www.bundesbank.de

ISBN 978–3–86558–821–0 (Printversion)
ISBN 978–3–86558–822–7 (Internetversion)

Abstract:

Most EU member states will adopt fiscal rules that refer to cyclically-adjusted borrowing limits. Under the standard cyclical adjustment procedure, trend increases in public debt based on cyclical components are prevented if the real-time output gaps used to calculate cyclical components balance over time. We analyse real-time output gaps for EU-15 countries over the 1996-2011 period as estimated by the EU, the IMF and the OECD. Compared to each institution's final estimate, we find that real-time output gaps in our sample period are negatively biased. This bias is observed (i) irrespective of the source of the data, (ii) in all real-time vintages, (iii) basically across the entire cross-section of countries. The magnitude of the bias is considerable: on average, real-time cyclical components as a percentage of GDP are biased downwards by about 0.5 percentage point per year. Our results suggest that fiscal rules should incorporate ex-post checks of the unbiasedness of the cyclical components used within the rule. Potential biases would then decrease or increase future borrowing limits.

Keywords: Public finance, fiscal rules, cyclical adjustment

JEL-Classification: H61, H68, E32

Non technical summary

In the context of the Fiscal Compact most EU member states will adopt national fiscal rules that refer to cyclically-adjusted borrowing limits. Under the standard cyclical adjustment procedure, trend increases in public debt based on cyclical components are prevented if real-time output gaps used to calculate cyclical components balance over time. It is well-known from the literature that real-time output gap estimates are associated with massive uncertainty and subject to strong revisions. We focus on the application of output gap estimates in national fiscal rules and ask whether cyclical components over the last 15 years would have been biased if they had relied on international organisations' real-time estimates of the output gap.

We compile real-time data for the output gaps estimated by the EU, the IMF and the OECD for the EU-15 countries over the 2002-2011, the 2000-2011 and the 1996-2011 periods, respectively and test these estimates for bias. The data suggest that – compared to each institution's final estimate – there is a significant bias towards producing more negative output gaps in real time. For our sample period, we observe this bias (i) irrespective of the source of the estimates (ii) in all examined definitions of real time, ie comparatively early estimations (for example spring of year t-1) or comparatively late estimations (for example autumn of year t), (iii) basically across the entire cross-section of countries. The magnitude of the bias obtained from our samples is clearly relevant: the average bias of real-time cyclical components across all countries and years lies between -0.6% and -0.5% of GDP. Thus, if a national fiscal rule had relied on these output gap estimations, the admissible deficits in real-time would have been allowed to exceed their "final" (autumn 2011) values by 0.5 percentage point per year.

Our results reflect real-time data from 1996 to 2011 and are not necessarily representative for the future. Evidence from the OECD real-time output gaps, which are available for a sufficiently long time period, suggests however that these findings are not merely a peculiarity of the current economic crisis and the preceding boom period. The fact that we detect a similar bias in output gaps produced with the Hodrick-Prescott filter and comprehensive evidence from the literature corroborate that there is no cyclical adjustment model that safeguards against biased real-time estimates. Thus, we advocate dealing with the biased cyclical components within the *fiscal rule*. It could for example be a precautionary default feature of a rule to take account of the cyclical components used in the budgeting processes over time. After roughly one cycle, eg after eight to ten years, the unbiasedness of the cyclical components would be checked and if these turn out to be nonzero on average, the future borrowing limit would be increased or reduced by the amount of the cumulated bias. Although such amendments of the fiscal rule are far from easy to implement in practice – think for example of the dating of business cycles – they appear more realisable than attempts to produce unbiased cyclical components in real time.

Nichttechnische Zusammenfassung

Ein Großteil der EU-Staaten wird im Rahmen des Fiskalpakts nationale Schuldenregeln implementieren, die sich wohl auf konjunkturbereinigte Finanzierungssalden beziehen werden. Im Standard-Konjunkturbereinigungsverfahren, das auch von der EU für den Stabilitäts- und Wachstumspakt verwendet wird, wird ein trendmäßiges Anwachsen des Schuldenstandes aus konjunkturbedingten Defiziten verhindert, sofern sich negative und positive Produktionslücken, die zur Bestimmung der Konjunkturkomponenten verwendet werden, über die Zeit ausgleichen. Dies muss insbesondere auch für die in Echtzeit geschätzten Produktionslücken gelten. Dabei ist aus der Literatur bekannt, dass diese Echtzeit-Schätzungen mit erheblicher Unsicherheit verbunden sind und später noch deutlich revidiert werden. Vor diesem Hintergrund untersuchen wir, inwiefern Konjunkturkomponenten über die letzten 15 Jahre verzerrt gewesen wären, für deren Bestimmung auf Echtzeit-Schätzungen der Produktionslücken internationaler Organisationen zurückgegriffen worden wäre.

Dazu stellen wir einen Datensatz der durch EU, IWF und OECD für die EU-15 Länder in Echtzeit geschätzten Produktionslücken zusammen (2002-2011, 2000-2011 und 1996-2011) und testen diese auf Unverzerrtheit. Gemessen an den letzten Schätzergebnissen vom Herbst 2011 weisen die in Echtzeit geschätzten Produktionslücken eine signifikant negative Verzerrung auf. Für die betrachtete Periode zeigt sich diese Verzerrung (i) in den Schätzergebnissen aller drei Organisationen, (ii) sowohl in frühen Schätzungen (z.B. vom Frühjahr des Jahres t-1) als auch zu späteren Schätzzeitpunkten (z.B. vom Herbst des Jahres t) und (iii) bei nahezu allen betrachteten Ländern. Die daraus resultierende Verzerrung der Echtzeit-Konjunkturkomponenten beträgt pro Jahr und im ungewichteten Durchschnitt über die Länder zwischen -0,6 % und -0,5 % des BIP und ist damit nicht unerheblich: Hätte eine nationale Fiskalregel auf eine der betrachteten Echtzeitschätzungen zurückgegriffen, so wäre der Verschuldungsspielraum pro Jahr um etwa 0,5 Prozentpunkte zu hoch ausgefallen (gemessen an den Ergebnissen vom Herbst 2011).

Die Schätzergebnisse reflektieren selbstverständlich allein den betrachteten Zeitraum von 1996 bis 2011 und müssen für die Zukunft nicht repräsentativ sein. Jedoch sind die

Ergebnisse auch nicht nur eine reine Besonderheit der schweren Wirtschaftskrise 2008/09 und der vorhergehenden Aufschwungphase, wie eine Teilung des Datensatzes zeigt. Die Tatsache, dass auch die mit dem Hodrick-Prescott Filter geschätzten Produktionslücken verzerrt sind und entsprechende Ergebnisse aus der Literatur legen zudem den Schluss nahe, dass verzerrte Echtzeit-Schätzergebnisse nicht das Problem einzelner Verfahren sind. Mögliche Verzerrungen der Echtzeit-Konjunkturkomponenten sollten daher vielmehr im Rahmen der *Fiskalregel* korrigiert werden. Beispielsweise wäre denkbar, dass über die den jeweiligen Haushalten zugrunde gelegten (also in Echtzeit geschätzten) Konjunkturkomponenten standardmäßig Buch geführt wird. Sollte sich nach Ablauf eines Zyklus (z.B. nach acht bis zehn Jahren) zeigen, dass die Konjunkturkomponenten im Durchschnitt nicht ausgeglichen sind, würde die kumulierte Verzerrung den zukünftigen Verschuldensspielraum über den nächsten Zyklus entsprechend vergrößern oder verringern. Die praktische Umsetzung solcher Ergänzungen des fiskalischen Regelwerks ist im Detail sicherlich nicht unkompliziert; im Vergleich zu Ansätzen, die Echtzeit-Konjunkturkomponenten unverzerrt zu schätzen, dürften diese aber leichter realisierbar sein.

Contents

1 Introduction 1

2 Data and descriptive statistics 4

3 Testing for real-time bias 9

3.1 The baseline model 9

3.2 Robustness checks 14

3.2.1 Hodrick-Prescott filter gaps 14

3.2.2 Mincer-Zarnowitz regressions 18

3.2.3 Sample period and subsamples 20

4 Fiscal policy implications 20

5 Conclusions 23

References 25

Appendix 27

Lists of Tables and Figures

Table 1:	Data coverage and sources	4
Table 2:	Summary statistics for output gap estimations	6
Table 3:	Test for bias in real-time output gaps by the OECD	11
Table 4:	Test for bias in real-time output gaps by the EU (HP filter versus production function)	16
Table 5	Mincer-Zarnowitz test for bias in real-time output gaps	19
Table 6	Average real-time bias of cyclical components	22

Figure 1	Definition of real-time output gaps	5
Figure 2	Output gap estimations by the OECD	8

Cyclical adjustment in fiscal rules: some evidence on real-time bias for EU-15 countries[1]

1 Introduction

In the context of the Fiscal Compact and complementing the European Stability and Growth Pact (SGP, for discussions compare eg Balassone and Franco, 2000 and Gennari, Giordano and Momigliano, 2005), most EU member states will adopt national fiscal rules that refer to cyclically-adjusted borrowing limits. Some countries – for example Germany and Spain – have already implemented such rules. For the German central government, the corresponding adjustment method has also been specified and is to be performed in line with the procedure used in the SGP. Under this method, the cyclical component of the budget balance is the product of the output gap and the budget sensitivity, with the latter gauging the impact of cyclical fluctuations in GDP on the budget balance. The output gap is defined as the deviation of GDP from potential output and is a compact measure of the economy's cyclical position. Given the standard assumption of a time-invariant budget sensitivity, it is crucial that the output gaps balance over time, if the accumulation of debt based on cyclical components is to be prevented by the fiscal rule (as stated in the German constitution, for example).[2] Though most production function models meet this balance requirement at least approximately for a given vintage, for the application within a fiscal rule the output gaps estimated at the time when the budget is drafted, *in real time,* must roughly balance over one cycle.

[1] The paper represents the author's personal opinion and does not necessarily reflect the views of the Deutsche Bundesbank or its staff. Address: Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main, Germany. Phone: 00 49 69 9566 3353, e-mail: gerhard.kempkes@bundesbank.de. I thank Jörg Breitung, Anne Ellis, Heinz Herrmann, Thomas Knetsch, Jana Kremer, Christian Schumacher and Karsten Wendorff for helpful remarks. I also thank Cécile Denis for providing details and data on the EU estimations of potential output. All remaining deficiencies are my responsibility alone.

[2] Mayer and Stähler (2011) show in a theoretical paper that the presence of estimation errors and potentially a downward bias significantly hamper the positive welfare effect of such rules.

It is well known that real-time estimates of the output gap are associated with massive uncertainty and heavy revisions as new information becomes available. Marcellino and Musso (2011) test a wide variety of output gap measures for the euro area aggregate and for the U.S. They find that the sign as well as the magnitude of output gaps estimated in real time are highly uncertain. These findings match earlier results for the U.S. (Orphanides and van Norden, 2002), the UK (Nelson and Nikolov, 2003), and Canada (Cayen and van Norden, 2005). Marcellino and Musso (2011) as well as Orphanides and van Norden (2002) offer thorough discussions of the sources of the uncertainty surrounding real-time output gap estimates and compare the performance of various detrending methods.

The goal and scope of our study are different. We focus on the application of output gap estimates in national fiscal rules and ask whether cyclical components would have been biased if they had relied on available estimates of the output gap. Specifically, we test for bias in available real-time estimates of the output gap as estimated by the EU, the IMF and the OECD. Our focus on the application of output gaps in national fiscal rules means that we consider individual EU countries and not the euro area aggregate, which permits us to base our investigation on an entire panel of countries, using yearly data, just as in the national budgeting processes.[3]

We compile real-time data for the output gaps estimated by the EU, the IMF and the OECD for the EU-15 countries over the 2002-2011, the 2000-2011 and the 1996-2011 periods, respectively. These panel data suggest that – compared to each institution's final estimate – there is a significant bias towards producing more negative output gaps in real time. We find that this bias is observed (i) irrespective of the source of the estimates, be it the EU, the IMF or the OECD, (ii) in all definitions of real time, ie comparatively early estimations (for example spring of year t-1) or comparatively late estimations (for example autumn of year t), (iii) basically across the entire cross-section of countries. This finding is robust against a series of robustness checks and the magnitude of the bias is considerable: on average, real-time cyclical components expressed as a percentage of GDP are biased downwards by about 0.5 percentage point per year.

[3] For an overview of the related literature on fiscal forecasting see for example Leal et al (2008).

Our results reflect real-time data from the mid 1990s until the early 2010s and are not necessarily representative for the future. Evidence from the OECD real-time output gaps, which are available for a sufficiently long time period, suggests however that these findings are not merely a peculiarity of the current economic crisis and the preceding boom period.

Given this broad set of empirical evidence and given results from the literature highlighting massive uncertainty surrounding real-time estimates of the output gap, we conclude that there is little hope that cyclical adjustment may be safeguarded against real-time bias. For fiscal policy we suggest dealing with potential biases within the fiscal rule. Fiscal rules could by default take account of the real-time cyclical components over time. If – after roughly one cycle, for example, after eight to ten years – these turn out to be nonzero on average, the bias could be stored on a control account that would reduce or increase future borrowing limits.

The paper is organised as follows: in Section 2 we introduce the data. In Section 3 we present the econometric test for bias and we report our baseline results along with several robustness checks. We sketch the fiscal policy implications of biased real-time output gaps in Section 4 and conclude with Section 5.

2 Data and descriptive statistics

The output gap is defined as the deviation of real GDP from potential output and thus serves as a measure of the economy's cyclical position. Potential output cannot be observed and must be estimated, which is usually accomplished on the basis of an aggregated production function (see D'Auria et al, 2010, for the method used by the European Commission). Potential output is determined by linking the potential values of labour and capital, taking into account a productivity trend which is measured as the trend component of total factor productivity. In the remainder of the paper, the output gap is denoted as a percentage of real potential output. Using the open online databases of the EU, the IMF and the OECD, we set up three real-time data sets for the following vintages, countries and years (see Table 1).

Table 1: Data coverage and sources

Source	Vintages	Years	Countries
EU	2002 autumn - 2011 autumn	2002 - 2011	EU-15
OECD	1996 autumn - 2011 autumn	1996 - 2011	EU-15 without LU
IMF	1999 autumn - 2011 autumn	1999 - 2011	EU-15 without LU, EL and DK

Note: All data is publicly available under http://circa.europa.eu/Public/irc/ecfin/outgaps/library, http://www.oecd.org and http://www.imf.org. EU-15 countries comprise Austria (AT), Belgium (BE), Denmark (DK), Finland (FI), France (FR), Germany (DE), Greece (EL), Ireland (IRL), Italy (IT), Luxembourg (LU), the Netherlands (NL), Portugal (PT), Spain (ES), Sweden (SE) and the UK (UK). The data coverage reports the largest possible sample; for the t-1 real-time vintages, sample size is reduced by one year.

Figure 1 depicts our definition of "real time" as exemplified by the output gaps that the OECD estimated for Germany. The estimation vintages, the point in time when the estimations were produced, are organised in rows. The target years, for which the output gaps were estimated, are listed in the columns. The columns show the organisation's record in estimating the output gaps for a given year (from top to bottom, starting with the black fields with white numerics). For example, the German economy's output gap for 2009 was estimated to be about 0.5% of potential output in spring 2008. In autumn 2008, this value was revised to -1.1%, and in spring 2009 it was further revised to -5.4% and to -3.5% in autumn 2009. The "final" (as of autumn 2011) estimate for the output gap for 2009 was -4.5%. These real-time estimates are then organised into four separate real-time series, which are listed in rows at the bottom of Figure 1. Letting t denote the target year, ie the year for which the output gap estimate is produced, the real-time series estimated in the spring of year t-1 obviously represent the earliest estimates whereas the estimates published in the autumn of year t represent the latest real-time estimations. Note, however, that even the late estimations are often revised substantially in the following vintages.

Figure 1: Definition of real-time output gaps (as a percentage of potential output, the example is for the German economy taken from the OECD Economic Outlooks June 2007 to December 2011)

	2007	2008	2009	2010	2011
2007 spring	0.4	1.0			
2007 autumn	0.0	0.1	0.1		
2008 spring	0.5	0.9	0.5		
2008 autumn	1.2	1.1	-1.1	-1.5	
2009 spring	2.6	1.9	-5.4	-5.7	
2009 autumn	2.6	2.4	-3.5	-2.9	-1.8
2010 spring	1.5	0.9	-5.2	-4.4	-3.6
2010 autumn	1.6	0.6	-5.2	-3.0	-1.9
2011 spring	1.9	1.1	-4.7	-2.5	-0.6
2011 autumn	**2.2**	**1.6**	**-4.5**	**-2.3**	**-0.8**

real time (spring t-1)		1.0	0.5	-5.7	-3.6
real time (autumn t-1)		0.1	-1.1	-2.9	-1.9
real time (spring t)	0.4	0.9	-5.4	-4.4	-0.6
real time (autumn t)	0.0	1.1	-3.5	-3.0	-0.8
final (autumn 2011)	**2.2**	**1.6**	**-4.5**	**-2.3**	**-0.8**

Depending on the national budgeting processes, estimations of the output gap may be of interest at different points in time. For instance, in Germany, the central government budget is prepared and discussed within the government in the spring of year t-1. In the autumn of year t-1, the budget is finally enacted by parliament. In this case we would therefore be interested primarily in the real-time series based on the estimations from autumn (and spring) of the year t-1.

Table 2 shows the overall summary statistics across the three panels of countries. In order to permit comparisons between real-time and final estimations, all series (including the final estimation) are computed based on the sample, for which real-time data is available (for example, 1996-2011 for the OECD, see Table 1). This pooled data can only give a very broad impression; it suggests that the average output gap over one given data vintage (here autumn 2011) is roughly zero for all three institutions. Appendix A1, which reports country-specific averages, shows, that there are important differences across countries and that for a given vintage, too, the output gaps tend to be negative on average for most countries. In fact, the EU production function concept, for example, is not necessarily balanced (see Denis et al 2006). In real time (see Table 2 and Appendix A1), this negative bias seems to be stronger: Table 2 and also Appendix A1 suggest that the cross-country average real-time output gap is about -1.5% of potential output, which is roughly one percentage point more negative than the final vintage.

Table 2 (a): Summary statistics for output gap estimations by the EU (2002-2011)

Variable	Obs	Mean	Std. Dev.	Min	Max
final (autumn 2011)	150	-0.2	2.5	-8.3	4.8
real time autumn t	150	-1.3	1.8	-8.3	2.0
real time spring t	135	-1.6	1.8	-7.7	2.2
real time autumn t-1	135	-1.4	1.7	-7.8	2.0
real time spring t-1	120	-1.5	1.8	-8.5	2.3

(b) IMF (1999-2011)

Variable	Obs	Mean	Std. Dev.	Min	Max
final (autumn 2011)	144	0.1	2.4	-6.5	6.9
real time autumn t	144	-1.3	1.7	-6.8	5.4
real time spring t	144	-1.6	1.9	-6.5	3.3
real time autumn t-1	136	-1.3	1.7	-5.9	5.6
real time spring t-1	136	-1.5	2.0	-7.7	3.1

(c) OECD (1996-2011)

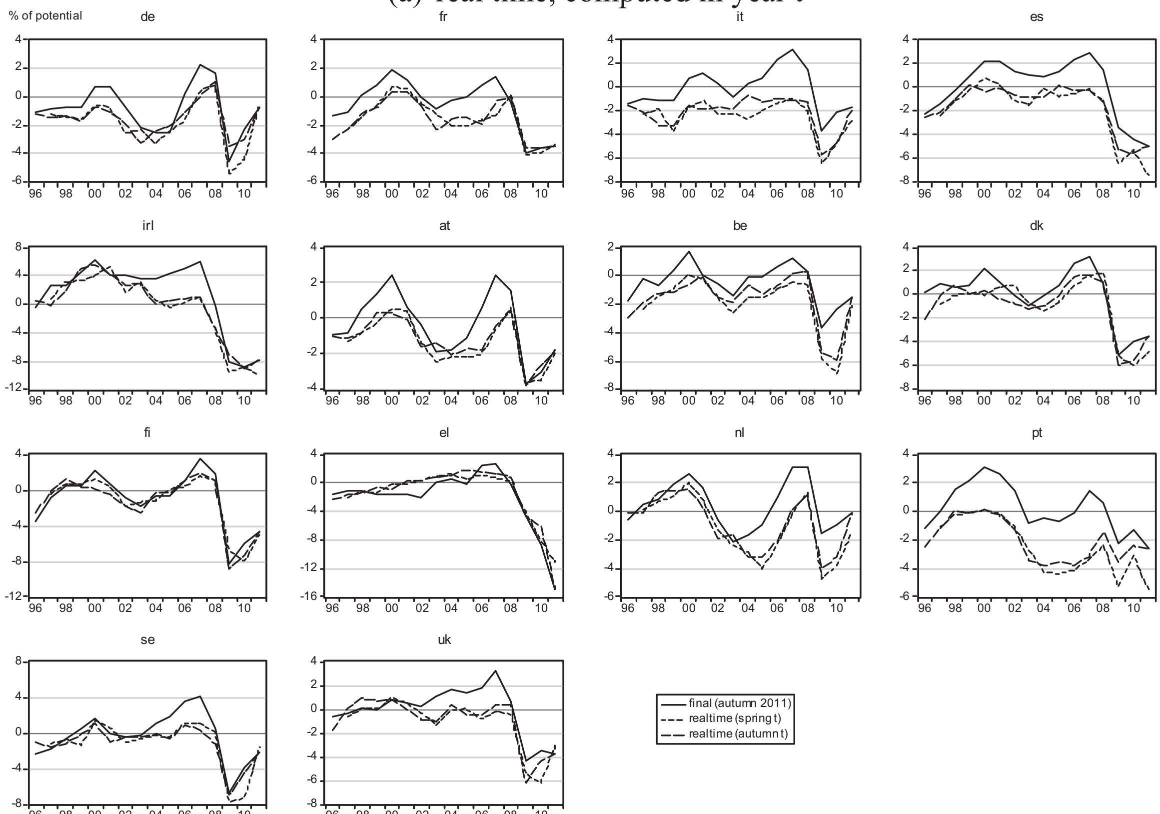
Variable	Obs	Mean	Std. Dev.	Min	Max
final (autumn 2011)	224	-0.3	2.7	-15.0	6.1
real time autumn t	224	-1.3	2.4	-15.0	5.5
real time spring t	210	-1.5	2.6	-11.1	5.2
real time autumn t-1	210	-1.3	2.2	-9.1	5.5
real time spring t-1	196	-1.1	2.4	-10.2	4.8

Figure 2 shows time series of output gap estimations for individual countries based on the estimates by the OECD, while those by the EU and the IMF are reported in Appendices A2 and A3. We show the OECD estimates here because they offer the largest sample. Part (a) of the figure reports the real-time estimations computed in year t, while part (b) reports the earlier estimates of the output gap, ie those computed in the spring and the autumn of year t-1. In line with the country-specific evidence shown in Appendix A1 and also in line with evidence from the pooled data, most country-level time series suggest that real-time estimates are more negative than the final vintage.

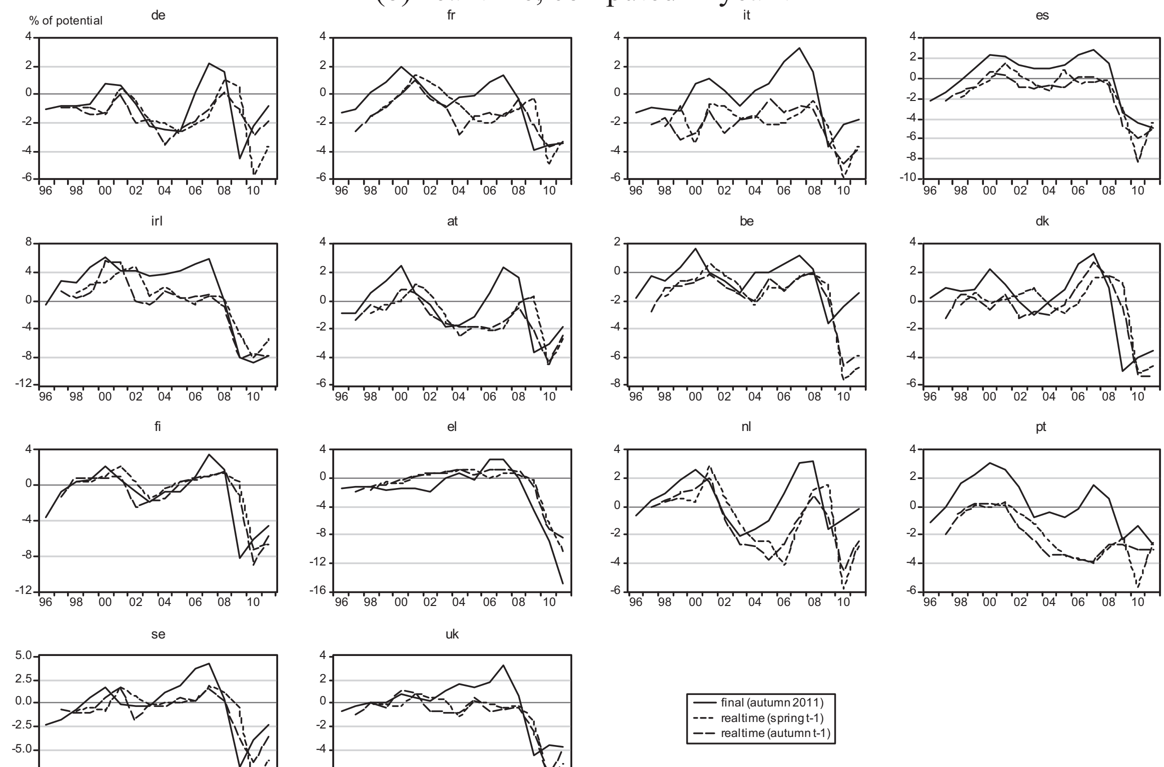
Some limitations, however, apply. The ad-hoc identification of a real-time bias is easier for the later real-time vintages, ie for those computed in year t. Looking at these real-time series, there appears to be an almost constant downward shift across the years. Moreover, comparing the estimates made by the three institutions (Figure 2 and Appendices A2 and A3), most individual-country time series support the impression of a negative shift in the real-time estimates across all three institutions alike (eg Italy and Portugal). The evidence is, however, less uniform in the cases of Sweden and the UK. And for Finland and Greece, there is no immediate reason to suspect the presence of a bias in the real-time series.

Figure 2: Output gap estimations by the OECD (solid line: “final”, ie last available estimation as of autumn 2011, dashed lines: real-time estimations)

(a) real time, computed in year t



(b) real time, computed in year t-1



3 Testing for real-time bias

3.1 The baseline model

Our starting point is the standard regression from the literature on the evaluation of economic forecasts (see Mincer and Zarnowitz, 1969):

(1) $$gap_{it}^{final} = \alpha_i + \beta \cdot gap_{it}^{real-time} + \varepsilon_{it}$$

where gap_{it}^{final} denotes the "final", ie most recent vintage of output gap estimations, which is the 2011 autumn vintage for country i and year t while $gap_{it}^{real\text{-}time}$ denotes the output gap estimated in real time (see the four definitions of real time in the previous section). The α_i capture country-specific biases in the output gaps. If $gap_{it}^{real\text{-}time}$ were an unbiased predictor of gap_{it}^{final}, we would expect $\beta = 1$ and $\alpha_i = 0$, which can be tested jointly using a Wald test. Holden and Peel (1990) show, however, that this test is a sufficient, but not a necessary condition for unbiasedness. The following model is less informative with respect to the properties of the bias, but more robust regarding bias detection:

(2) $$rev_{it}^{gap} = \alpha_i + \varepsilon_{it}$$

where we define $rev^{gap}_{it} = gap^{final}_{it} - gap^{real\text{-}time}_{it}$, ie the revision of the output gap between the final estimation and the real-time estimation. The real-time output gaps are unbiased predictors of the final estimation if we cannot reject the joint null hypothesis of $\alpha_i = 0$. A positive country intercept for country i would imply that for this country on average $gap_{it}^{real\text{-}time} < gap_{it}^{final}$ holds, ie that real-time output gaps are negatively biased compared to the final estimations. Moreover, we highlight that this simple model is highly relevant for policy analysis since it simply yields country-specific averages and thus essentially reflects the bias that would probably have been relevant if the output gaps had been used in a fiscal rule. We use Model (2) as our baseline model and perform robustness checks in Section 3.2.

The ε_{it} in Model (2) are assumed to be independent of the α_i. We do not rule out serial correlation, heteroskedasticity and contemporaneous cross-sectional correlation. In particular, allowing for contemporaneous cross-sectional correlation is important because business cycles in the European Union are – if not synchronised – at least correlated to some degree (see eg Giannone, Lenza and Reichlin, 2008 or Mink, Jacobs and de Haan, 2012). Serial correlation in the ε_{it} could be considered an indication of sluggish adjustment to new information. The Wooldridge (2002) and Drukker (2003) test does indeed indicate the presence of serial correlation, although this tends to be fairly weak. Nevertheless, we prefer estimating the model using the Prais-Winsten estimator, which yields more conservative standard errors. Not surprisingly, for all key results reported below, significance levels tend to be higher when we estimate with OLS.

As an overall measure of the proportion of variance explained by the model, the standard R^2 is not a useful statistic here. The reason is that the standard (centred) R^2 reports the fraction of explained variation in the output gaps *from* their mean value, ie it measures the share of explained deviations from the overall (pooled) bias. What is more relevant here is the uncentred R^2, which gives an indication of the proportion of explained variation in the *observed* output gaps, ie we compute the uncentred $R^2 = \sum(gap^*_{it})^2 / \sum (gap_{it})^2$, where gap^*_{it} denote the predicted output gaps (ie the country effects), whereas gap_{it} denote their true (final) values. The uncentred R^2 thus reports the explained fraction of the raw sum of squared output gaps and accounts for the explanatory power of the overall bias.

Table 3 shows the results for Model (2) estimated with OECD data because the latter offer the largest sample. We report the results for IMF and EU data in Appendices A4 and A5, and we discuss them briefly in the following because the conclusions are fairly similar to those drawn from the OECD data. As explained in Section 2, we have four different definitions of real time at our disposal. The output gap for target year t is first estimated in the spring of year t-1, then in the autumn of year t-1, the spring of year t and finally the autumn of year t. We test for bias in these four real-time vintages within separate regressions, see Columns (1) to (4) of Table 3.

Table 3: Test for bias in real-time output gap estimations by the OECD (for real-time vintages spring of year t-1 to autumn of year t, 1996-2011)

	rev^{gap}_{it} (final – spring t-1)	rev^{gap}_{it} (final – autumn t-1)	rev^{gap}_{it} (final – spring t)	rev^{gap}_{it} (final – autumn t)
	(1)	(2)	(3)	(4)
DE	0.69	0.69	0.92	0.64
	(0.61)	(0.46)	(0.27)***	(0.31)**
FR	0.65	0.98	1.03	1.04
	(0.48)	(0.37)***	(0.30)***	(0.28)***
IT	2.06	2.10	2.39	1.85
	(0.48)***	(0.37)***	(0.30)***	(0.39)***
ES	1.61	1.65	1.98	1.51
	(0.36)***	(0.19)***	(0.26)***	(0.30)***
IRL	1.32	2.01	2.06	1.34
	(0.73)*	(0.61)***	(0.59)***	(0.67)**
AT	0.86	1.00	1.01	0.78
	(0.55)	(0.40)**	(0.28)***	(0.34)**
NL	1.37	1.51	1.56	1.14
	(0.67)**	(0.44)***	(0.34)***	(0.39)***
BE	1.28	1.43	1.51	1.11
	(0.58)**	(0.50)***	(0.34)***	(0.39)***
FI	-0.36	0.17	0.29	0.29
	(0.73)	(0.75)	(0.34)	(0.31)
SE	0.79	0.69	0.97	0.78
	(0.76)	(0.54)	(0.46)**	(0.39)**
PT	2.38	2.40	2.78	2.17
	(0.45)***	(0.36)***	(0.30)***	(0.36)***
UK	0.96	0.98	1.04	0.88
	(0.45)**	(0.45)**	(0.39)***	(0.36)**
DK	0.28	0.66	0.93	0.96
	(0.61)	(0.52)	(0.33)***	(0.28)***
EL	-1.19	-1.25	-0.66	-0.51
	(0.55)**	(0.58)**	(0.44)	(0.48)
N	196	210	210	224
R^2	0.14	0.18	0.26	0.13
uncentred R^2	0.31	0.48	0.69	0.50
Wald test, H0: all intercepts=0	1279.3***	8729.5***	1360.64***	100.39***
Coefficient of a constant in a pooled model	0.91	1.08	1.30	1.06
	(0.45)***	(0.28)***	(0.18)***	(0.17)***
rho	0.06	0.19	0.27	0.38
Wooldridge test, H0: no first-order serial correlation	35.9***	11.8***	6.3**	23.30***

Note: Standard errors (in parentheses) are robust in the presence of heteroskedasticity, contemporaneous cross-sectional correlation and first-order serial correlation. * p<0.1; ** p<0.05; *** p<0.01.

Our findings are as follows. *First*, the Wald test on unbiasedness of the real-time output gaps, which is presented at the bottom of Table 3 is rejected in all regressions. *Second*, the country dummies yield positive coefficients (i) for almost every country (the only exceptions are Finland in the spring t-1 vintage and Greece in all vintages) and (ii) across all real-time vintages. This implies that the real-time output gaps typically fall short of the final output gaps, ie they have a negative bias. *Third*, most country dummies are also individually significantly different from zero. Significance levels tend to be higher for the later real-time vintages. For the autumn of year t-1 vintage, which is used to enact the German central government budget, the coefficients estimated for 9 out of 14 countries yield significantly positive coefficients. *Fourth*, the magnitude of the bias is highly relevant for fiscal policy. Consider the extreme cases of Italy and Portugal. The point estimates for the constant country-level bias of more than two percentage points (obtained from the 'spring t-1' or 'autumn t-1' regressions) imply that an average real-time output gap for these two countries falls short of its final value by at least two percentage points. Had these output gaps been used for cyclical adjustment of the Italian and Portuguese overall government sector, those cyclical components would have undershot their final values by about one percentage point per year.[4] The unweighted average bias in the output gaps across all countries (based on the autumn t-1 vintage) is about 1.1 percentage points. *Fifth* the simple model explains a considerable proportion of the variation in the output gaps, for example in the autumn of year t-1 vintage the model explains about 50% of the variation while in the spring of year t vintage this fraction exceeds 2/3. More than half of the explained variation is due to the overall (pooled) bias. *Sixth*, the Wooldridge (2002) and Drukker (2003) test indicates that the ε_{it} are serially correlated. Serial dependence is, however, rather weak, which suggests that sluggish adaption of new information tends to be a minor issue.

The results obtained from the IMF data strongly support the six conclusions drawn from the OECD data. In particular, the hypothesis of unbiased real-time output gaps is again rejected for every regression. All country intercepts yield positive coefficients and most of them are also significantly different from zero for individual countries. The magnitude of the bias tends to be larger compared to the results obtained from OECD

[4] Assuming the budget sensitivities published by the OECD of 0.53 for Italy and 0.46 for Portugal, see Table 4 of this paper or Girouard and André, 2005, p 22.

data. For example, in the ‘autumn t-1’ vintage the country intercept for Germany is estimated at about 1.4, which is roughly twice as high as the coefficient obtained from the OECD data. The overall and unweighted average magnitude of the downward bias of real-time output gaps lies between 1.3 (autumn of year t-1) and 1.7 (spring of year t) percentage points. The share of variation explained by the model is in line with the estimations based on OECD data. One clear difference between the results obtained from IMF and OECD data is the country intercept estimated for Sweden. While insignificant in the OECD regressions, it yields the largest individual country intercept in the estimations based on IMF data.

With some limitations, the results for the EU data are also in line with the main findings from the OECD estimations. The joint hypothesis of unbiased real-time output gaps is rejected in all regressions. Again, the individual effects are positive across the board with the exception of Greece (all vintages). The significance of the individual country dummies tends to be lower than in the OECD estimations. However, it is not surprising that we tend to obtain imprecise estimates of the country intercepts from the EU data because the corresponding standard errors are estimated based on only eight observations (spring of year t-1). Recall that the efficiency of the estimation of the country effects improves only with an increasing time dimension. The overall average of the downward bias is between 1.0 (autumn of year t) and 1.2 (spring of year t) percentage points, which roughly corresponds to the estimates produced with OECD data. [5]

[5] Given that we focus on the real-time bias, ie the final gap minus the real-time gap, it is irrelevant for our study whether the output gap concepts are symmetric. We would obtain biased estimates if the output gap concept changed during the sample period. Such a shift occurred in the case of the EU production function as from 2004. This shift, which was from balanced to unbalanced output gap concept, does not, however, affect any of our results, which we checked by restricting the sample period to the years 2004-2011. This is remarkable given that the country-specific biases can then only be estimated with even fewer observations.

3.2 Robustness checks

3.2.1 Hodrick-Prescott filter gaps

Given that the bias in the real-time output gaps is observed across all three institutions and across all real-time vintages, it is of some interest whether a similar bias is detected when considering output gaps estimated with a simple statistical filter, such as the Hodrick-Prescott filter. The European Commission has continued to produce output gaps using the HP filter in parallel to using the production function. This provides us with output gap estimates from both methods which are based on a common set of macroeconomic assumptions. Thus, using the EU Commission database, we re-estimate our baseline model with the HP filtered output gaps. The results of this exercise can then easily be compared to the results obtained from the Commission's production function gaps (see Table 4). We find that unbiasedness of the output gaps is rejected for both the HP filter and the production function across all four vintages. Rejection for both is at the 1% significance level. Individual country dummies estimated for the HP filter gaps also yield positive coefficients across the board, although their individual significance tends to be lower than for the country effects obtained from the production function gaps. The overall magnitude of the real-time bias estimated for the HP filter gaps is as strong as for the production function, and the proportion of the variation in the output gaps that is explained by the regression models also tends to be similar to that obtained from the production function estimates.

Similar regressions run on HP filter gaps, which we produced with IMF data covering a limited period (results available from the author upon request), broadly resemble the results obtained from the EU data for the vintages obtained from year t.[6] For the vintages obtained from year t-1, unbiasedness of the output gaps produced with the HP

---

[6] As is well known, the HP trend of a time series is calculated for each year as a weighted average of the original series with the distribution of weights being symmetric only in the middle of the sample period. Towards the end of the sample period, the distribution is more and more skewed and the filtered values are dominated by the actual values. Thus, if we want to compute a robust output gap for target year t in the spring of year t-1, we have to filter a series for which we have a forecast until about year t+4 (see also the discussion in Bouthevillain et al 2001). Looking at our data sets, only the IMF started to publish real GDP series up to year t+5 starting with the spring 2008 projection. We can thus re-estimate Model (2) using the final and the real-time output gaps which we produced using the HP trend of real GDP as published by the IMF for the panel of countries and for the years 2008 to 2011. We set the smoothing parameter λ to 100, which has become a convention in this field following the

filter cannot be rejected, whereas unbiasedness for the production function gaps is rejected, but the overall bias is negative. Individual country dummies show mixed signs for both methods. This evidence is, however, produced only with data for the years 2008 to 2011, in other words 36 observations for the t-1 vintages. One should not read too much into interpretations of these regressions.

Thus, the real-time bias is also prevalent in output gaps produced with simple statistical filters, although the significance and magnitude of the bias tend to be somewhat weaker than those of the bias obtained from production function approaches. This finding suggests that the source of the bias must at least partially relate to the macroeconomic projections.[7]

By controlling for revisions in macroeconomic projections[8] we could – in principle – hope to obtain a bias net of the effect induced by macroeconomic projections. The production functions considered involve a comprehensive set of macroeconomic variables (see for example D'Auria et al, 2010). Obviously, we cannot control for revisions to this entire set of variables. In Appendices A6 and A7, we provide some evidence where we use revisions of the GDP growth rates as proxies for revisions to this whole set of macro variables. While the results confirm all of our findings, they are very difficult to interpret due to severe omitted variables bias.

European Commission. As a benchmark, we also run these regressions based on output gaps produced from the IMF production function for the restricted period.

7 We also cross-check whether a possible (small sample) bias is present in HP trend deviations of real GDP series where macroeconomic projections are excluded (only final data). For these output gap estimates we can indeed clearly accept unbiasedness – irrespective of whether we are dealing with the end-point bias of the HP filter (EU, IMF) or not (OECD). This lends further support to the conclusion that the bias is due to the macroeconomic projections.

8 Macroeconomic projections are typically highly relevant for the assessment of the cyclical position at the current end, and revisions of the macro outlook may account for some fraction of a potential bias. For example, if macroeconomic perspectives deteriorated over a protracted period, potential output would have been repeatedly revised downwards, and output gaps would have been revised upwards.

Table 4: Test for bias in real-time output gap estimations produced by the EU: HP filter vs production function (for real-time vintages spring t-1 to autumn t, 2002-2011)

	rev^{gap}_{it} (final – spring t-1)		rev^{gap}_{it} (final – autumn t-1)	
	HP filter	Production function	HP filter	Production function
	(1)	(2)	(3)	(4)
DE	0.03	0.40	0.50	0.14
	(1.21)	(0.98)	(0.80)	(0.80)
FR	0.75	1.55	0.78	1.53
	(0.65)	(0.71)**	(0.54)	(0.59)***
IT	0.77	1.40	1.03	1.13
	(0.84)	(0.77)*	(0.70)	(0.74)
ES	1.71	1.24	1.79	1.05
	(1.11)	(0.95)	(0.78)**	(0.82)
IRL	1.73	2.48	2.30	1.98
	(1.81)	(1.17)**	(1.23)*	(0.86)**
AT	0.46	0.45	0.46	0.29
	(0.83)	(0.81)	(0.63)	(0.65)
NL	1.10	0.90	1.04	0.80
	(1.22)	(0.80)	(0.88)	(0.67)
BE	0.41	1.38	0.63	1.08
	(0.84)	(0.60)**	(0.61)	(0.53)**
FI	1.06	1.09	0.89	0.88
	(1.54)	(1.25)	(1.17)	(1.13)
SE	0.88	1.01	0.74	0.79
	(1.30)	(1.01)	(0.91)	(0.89)
PT	1.68	1.21	1.90	1.17
	(0.68)**	(0.63)*	(0.62)***	(0.48)**
UK	1.05	1.50	1.43	1.83
	(0.95)	(0.88)*	(0.69)**	(0.75)**
DK	0.48	0.68	0.49	0.47
	(0.98)	(0.89)	(0.82)	(0.87)
EL	2.75	-1.33	2.64	-0.75
	(1.17)**	(0.68)*	(0.91)***	(0.53)
LU	1.41	3.37	1.19	2.90
	(1.13)	(1.13)***	(1.03)	(0.92)***
N	120	120	135	135
R^2	0.04	0.13	0.05	0.11
uncentred R^2	0.28	0.40	0.38	0.35
F-test, H0: all intercepts=0 (*)	2.35***	3.06***	3.32***	3.40***
Coefficient of a constant in a pooled model	1.14	1.17	1.24	1.05
	(0.24)***	(0.22)***	(0.20)***	(0.18)***
rho	0.27	0.17	0.19	0.22
Wooldridge test, H0: no first-order serial correlation	80.1***	44.8***	15.9***	82.2***

Note: Standard errors (in parentheses) are robust in the presence of heteroskedasticity, contemporaneous cross-sectional correlation and first-order serial correlation. * p<0.1; ** p<0.05; *** p<0.01. (*) The F-test is based on a regression with Newey-West (1987) corrected standard errors (not reported) because the number of observations per country is smaller than the number of countries and it is thus impossible to base the test on panel-corrected standard errors.

Table 4 continued

	rev^{gap}_{it} (final – spring t)		rev^{gap}_{it} (final – autumn t)	
	HP filter	Production function	HP filter	Production function
	(5)	(6)	(7)	(8)
DE	0.28	0.39	0.55	0.17
	(0.92)	(0.46)	(0.56)	(0.48)
FR	0.86	1.78	0.71	1.52
	(0.52)	(0.57)***	(0.43)	(0.48)***
IT	1.23	1.57	1.24	1.47
	(0.55)**	(0.55)***	(0.53)**	(0.55)***
ES	1.62	1.41	1.15	1.22
	(0.74)**	(0.80)*	(0.59)*	(0.69)*
IRL	2.21	2.08	1.75	1.46
	(1.13)**	(0.65)***	(1.15)	(0.60)**
AT	0.45	0.33	0.44	0.27
	(0.53)	(0.47)	(0.52)	(0.39)
NL	0.76	0.68	0.75	0.51
	(0.92)	(0.50)	(0.48)	(0.41)
BE	0.50	1.13	0.50	0.86
	(0.66)	(0.44)**	(0.41)	(0.40)**
FI	0.98	1.06	0.89	0.98
	(1.02)	(1.05)	(0.92)	(0.81)
SE	0.74	0.98	0.56	0.67
	(0.94)	(0.67)	(0.69)	(0.55)
PT	1.57	1.10	1.47	1.13
	(0.68)**	(0.46)**	(0.58)**	(0.44)**
UK	1.53	2.04	1.32	1.80
	(0.54)***	(0.64)***	(0.66)**	(0.67)***
DK	0.60	0.76	0.70	0.74
	(0.69)	(0.62)	(0.55)	(0.54)
EL	2.56	-0.62	1.71	-0.43
	(0.89)***	(0.59)	(1.08)	(0.68)
LU	1.27	2.55	0.98	2.34
	(0.90)	(0.76)***	(0.97)	(0.81)***
N	135	135	150	150
R^2	0.05	0.14	0.02	0.13
uncentred R^2	0.48	0.54	0.37	0.46
F-test, H0: all intercepts=0 (*)	3.55***	6.82***	4.46***	5.88***
Coefficient of a constant in a pooled model	1.20	1.22	1.09	1.04
	(0.18)***	(0.14)***	(0.14)***	(0.12)***
rho	0.20	0.39	0.41	0.43
Wooldridge test, H0: no first-order serial correlation	10.1***	79.6***	59.0***	64.5***

Note: Standard errors (in parentheses) are robust in the presence of heteroskedasticity, contemporaneous cross-sectional correlation and first-order serial correlation. * p<0.1; ** p<0.05; *** p<0.01. (*) The F-test is based on a regression with Newey-West (1987) corrected standard errors (not reported) because the number of observations per country is smaller than the number of countries and it is thus impossible to base the test on panel-corrected standard errors.

3.2.2 Mincer-Zarnowitz regressions

We used the model proposed by Holden and Peel (1990) as our baseline model because their variant of the test for bias is more robust for bias detection and because it yields estimates that are directly interpretable for policy analysis. We may, however, learn more about the properties of the bias in real-time output gaps by estimating the conventional test proposed by Mincer and Zarnowitz (1969), which is presented in Equation (1). We also run a fully country-specific variant of Model (1) by allowing for country-specific slopes, β_i.

The results for Model (1) are presented in Table 5 while those of the fully country-specific model are shown in Appendix A8. All key results are confirmed. In particular, we reject the joint hypothesis of unbiasedness for every real-time vintage and for both models. The composition of the bias changes over the vintages. In the early 'spring of year t-1' vintage, the β is significantly different from one at the 5% significance level. Specifically, the β is below one, indicating that the output gap estimates from the 'spring of year t-1' vintage are "hysteric" in the sense that they tend to overstate the magnitude of the final output gap estimate (positive or negative). The point estimates of the α_i are positive with the exception of Greece and Finland, but only four of them are significantly different from zero. In the subsequent vintages, the β is no longer significantly different from one and its point estimate converges to one. However, as the β converges to one over the vintages (ie the real-time output gap estimates get more precise), we observe that more α_i are significantly different from zero (ie the constant country-specific biases materialise). This finding probably reflects the difficulties forecasting output gaps at an early point in time, as these forecasts are bad in a symmetric sense, ie they overstate the magnitude of the final output gap, both positively and negatively. The constant country-specific bias is more difficult to explain but Section 3.2.1 suggests that macroeconomic projections are an important factor.

Similar patterns are observed for the IMF and EU data (see Appendix A9 and A10 for the panel models; the fully country-specific models are available from the author upon request). In particular, the tests for unbiasedness are rejected in every regression and the magnitude of the overall bias tends to be larger than in the results obtained from OECD data. The significance of individual country-specific intercepts is somewhat lower.

Table 5: Mincer-Zarnowitz test for bias in real-time output gap estimations by the OECD (for real-time vintages spring of year t-1 to autumn of year t, 1996-2011)

	gap^{final}_{it} (autumn 2011)			
	(1)	(2)	(3)	(4)
DE	0.25	0.58	0.80	0.58
	(0.66)	(0.50)	(0.30)***	(0.33)*
FR	0.32	0.88	0.92	0.97
	(0.52)	(0.42)**	(0.32)***	(0.30)***
IT	1.43	1.95	2.22	1.77
	(0.63)**	(0.47)***	(0.35)***	(0.41)***
ES	1.17	1.54	1.85	1.44
	(0.48)**	(0.29)***	(0.29)***	(0.32)***
IRL	1.28	1.96	2.01	1.33
	(0.92)	(0.64)***	(0.60)***	(0.67)**
AT	0.52	0.91	0.92	0.74
	(0.59)	(0.44)**	(0.31)***	(0.35)**
NL	1.11	1.44	1.48	1.11
	(0.62)*	(0.44)***	(0.33)***	(0.39)***
BE	0.79	1.31	1.39	1.04
	(0.58)	(0.53)**	(0.35)***	(0.39)***
FI	-0.52	0.10	0.20	0.23
	(0.78)	(0.75)	(0.37)	(0.33)
SE	0.55	0.63	0.89	0.74
	(0.76)	(0.55)	(0.46)*	(0.41)*
PT	1.75	2.25	2.61	2.09
	(0.58)***	(0.45)***	(0.34)***	(0.38)***
UK	0.65	0.90	0.96	0.84
	(0.50)	(0.46)*	(0.40)**	(0.36)**
DK	0.17	0.61	0.87	0.91
	(0.63)	(0.51)	(0.34)**	(0.29)***
EL	-1.55	-1.32	-0.78	-0.59
	(0.78)**	(0.63)**	(0.50)	(0.49)
$gap^{realtime}$ (spring t-1)	0.70			
	(0.18)***			
$gap^{realtime}$ (autumn t-1)		0.93		
		(0.13)***		
$gap^{realtime}$ (spring t)			0.93	
			(0.07)***	
$gap^{realtime}$ (autumn t)				0.96
				(0.06)***
N	196	210	210	224
R2	0.42	0.61	0.80	0.79
Wald test, H0: all intercepts=0 & $gap^{realtime}$=1	446.4***	5248.2***	2682.3***	99.4***
Coefficient of a constant in a pooled model	0.64	1.02	1.20	1.05
	(0.17)***	(0.14)***	(0.11)***	(0.10)***
rho	0.16	0.20	0.29	0.38
Wooldridge test, H0: no first-order serial correlation	66.9***	29.2***	21.1***	53.4***

Note: Standard errors (in parentheses) are robust in the presence of heteroskedasticity, contemporaneous cross-sectional correlation and first-order serial correlation. * p<0.1; ** p<0.05; *** p<0.01.

3.2.3 Sample period and subsamples

As in every empirical study our findings are of course subject to the sample period, here the mid-1990s to the early 2010s. But note the evidence presented elsewhere, for example in Gerberding, Seitz and Worms (2005), where the real-time output gaps for Germany from the mid-1970s to the late 1990s, produced by the Bundesbank, also show a clear downward bias.

Using our data-base, we can check whether the results are a peculiarity of the current severe economic downturn and the preceding boom years by restricting the sample period of the OECD data (which is available for a sufficiently long period) to the years 1996 to 2002. All key results are robust against such restriction of the sample period (results available from the author upon request).

We check, moreover, the robustness of the key results against a restriction of the sample period to the years 1996 to 2008. We intend to exclude those "final" values for the output gap that have not yet settled and may potentially be revised to some extent in the future. The results obtained from this sample tend to show even larger bias, and the significance of individual country intercepts also tends to be higher (results available from the author upon request).

4 Fiscal policy implications

Given the biases of the real-time output gaps estimated with the baseline model, we compute the real-time biases of the cyclical components using the budget sensitivities reported in Girouard and André (2005), which are also used in the procedure under the SGP (see Table 6). We choose the real-time vintage produced in the autumn of year t-1 because it is the relevant choice for the German central government deficit rule. It is, however, clear from the previous section that the key messages do not depend on the choice of the real-time vintage.

The unweighted average real-time downward bias of the cyclical components lies between 0.5% and 0.6% of potential output per country and per year, which, taken at face value, implies that if a national fiscal rule had relied on these output gap estimations, deficits in real time would have been allowed to exceed their "final" values by at least 0.5 percentage point per year. Over the course of several years, this magnitude easily translates into significant additional public debt. Note that only one significant bias points towards a *lower* deficit in real time (Greece estimated by the OECD). Note also that there are two countries, Finland and Denmark, for which this pattern does not hold because the biases are never significant, whereas, in the cases of Belgium, France and Portugal, the real-time cyclical components are always significantly biased downwards, irrespective of the estimating institution.[9]

We would like to highlight that these results are of a somewhat illustrative nature because, to our knowledge, in the sample period none of the countries in question based their fiscal rule or budgeting process on one of the examined output gap estimates. However, the preventive arm of the SGP is founded on the European Union's cyclical adjustment procedure.

The fact that trend deviations produced with HP-filtered real GDP series appear to exhibit a similar bias and the evidence presented in Marcellino and Musso (2011) or Orphanides and van Norden (2002) suggest that no method for cyclical adjustment safeguards against biased real-time estimates. Thus, for fiscal policy applications, it seems advisable to adopt a precautionary approach. For example, rules could by default incorporate some account of the cyclical components used in the budgeting processes over time. If – after roughly one cycle, which itself is of course subject to debate[10] – these prove to be biased, the bias could be stored on a control account that would reduce or increase future borrowing limits. Albeit such amendments of the fiscal rule are far from easy to implement in practice, they appear more realisable than attempts to produce unbiased cyclical components in real time.

[9] This result is in line with the literature which highlights the crucial importance of using real-time information, see eg Croushore (2011) or – for a fiscal policy application – Golinelli and Momigliano (2009).

[10] For a practical solution, one could define a fixed period of about eight to ten years.

Table 6: Average real-time bias of cyclical components per year, as a percentage of potential output, obtained from the 'autumn of year t-1' real-time vintage.

	Budget sensitivity	EU	OECD	IMF
DE	*0.51*	ns	ns	0.7 (**)
FR	*0.53*	0.8 (***)	0.5 (***)	0.6 (***)
IT	*0.53*	ns	1.1 (***)	0.9 (***)
ES	*0.44*	ns	0.7 (***)	0.7 (***)
IRL	*0.38*	0.8 (**)	0.8 (***)	ns
AT	*0.47*	ns	0.5 (**)	0.5 (**)
NL	*0.53*	ns	0.8 (***)	ns
BE	*0.52*	0.6 (**)	0.7 (***)	0.4 (*)
FI	*0.48*	ns	ns	ns
SE	*0.55*	ns	ns	1.9 (***)
PT	*0.46*	0.5 (**)	1.1 (***)	0.8 (***)
UK	*0.45*	0.8 (**)	0.4 (**)	ns
DK	*0.59*	ns	ns	/
EL	*0.47*	ns	-0.6 (**)	/
LU	*0.47*	1.4 (***)	/	/
EU average	*0.48*	0.5 (*)	0.5 (***)	0.6 (***)

Note: Following the procedure under the SGP, the budget sensitivities are taken from Girouard and André (2005), Table 9. Biases and statistical significance of the output gaps needed to compute the biases of cyclical components are those estimated for the 'autumn of year t-1 vintage' and have been taken from Table 3 and Appendices A4 and A5. Bias and statistical significance of the EU average coefficients are the point estimates and significance levels of the constant term in pooled regressions. (*) $p<0.1$; (**) $p<0.05$; (***) $p<0.01$ denote the significance levels of the corresponding bias estimates for the output gaps. ns not significant; / not in the sample.

5 Conclusions

We compile real-time data for the output gaps estimated by the EU, the IMF and the OECD for the EU-15 countries over the 2002-2011, the 2000-2011 and the 1996-2011 periods, respectively and test these estimates for bias. The data suggest that – compared to each institution's final estimate – there is a significant bias towards producing more negative output gaps in real time. For our sample period, we observe this bias (i) irrespective of the source of the estimates, be it the EU, IMF or OECD, (ii) in all examined definitions of real time, ie comparatively early estimations (for example spring of year t-1) or comparatively late estimations (for example autumn of year t), (iii) basically across the entire cross-section of countries. The magnitude of the bias obtained from our samples is relevant: the average downward bias of real-time cyclical components across all countries and years lies between 0.5% and 0.6% of potential output. Thus, if a national fiscal rule had relied on these output gap estimations, the admissible deficits in real time would have been allowed to exceed their "final" values (as estimated in autumn 2011) by about 0.5 percentage point per year.

Our results reflect real-time data from the mid 1990s until the early 2010s and are not necessarily representative for the future. Evidence from the OECD real-time output gaps, which are available for a sufficiently long time period, suggests however that these findings are not a mere peculiarity of the current economic crisis and the preceding boom period. The fact that we detect a similar bias in output gaps produced with the Hodrick-Prescott filter and comprehensive evidence from the literature further corroborate that there is no clean cyclical adjustment model that safeguards against biased real-time estimates. For fiscal policy applications this may in the end make a case for using simple and transparent methods for the trend/cycle decomposition, as for example the Hodrick-Prescott filter.

For fiscal rules, we advocate dealing with the biased cyclical components within the fiscal rule. It could for example be a precautionary default feature of a rule that it takes account of the cyclical components used in the budgeting processes over time. After roughly one cycle, eg after eight to ten years, the unbiasedness of the cyclical components would be checked, and if these turn out to be nonzero on average, the future borrowing limit would be reduced or increased by the amount of the cumulated bias. Although such amendments of the fiscal rule are far from easy to implement in practice, they appear more realisable than attempts to produce unbiased cyclical components in real time.

References

Balassone, F. and D. Franco 2000. Public investment, the Stability Pact and the 'Golden Rule'. *Fiscal Studies* 21, 207-229.

Bouthevillain, C., P. Cour-Thimann, G. van den Dool, P. Hernández de Cos, G. Lagenus, M. Mohr, S. Momigliano and M. Tujula 2001. Cyclically adjusted budget balances: an alternative approach. ECB Working Paper No. 77.

Cayen, J.-P. and S. van Norden 2005. The reliability of Canadian output-gap estimates. *North American Journal of Economics and Finance* 16, 373-393.

Croushore, D. 2011. Frontiers of real-time data analysis. *Journal of Economic Literature* 49, 72-100

D'Auria, F., C. Denis, K. Havik, K. Mc Morrow, C. Planas, R. Raciborski, W. Röger and A. Rossi 2010. The production function methodology for calculating potential growth rates and output gaps. European Economy, Economic Papers No. 420.

Denis, C., D. Grenouilleau, K. Mc Morrow and W. Röger 2006. Calculating potential growth rates and output gaps – A revised production function approach. European Economy, Economic Papers No. 247.

Drukker, D.M. 2003. Testing for serial correlation in linear panel-data models. *Stata Journal* 3, 168-177.

Gennari, E., R. Giordano and S. Momigliano 2005. Dealing with unexpected shocks to the budget. *Finanzarchiv* 61, 201-219.

Gerberding, C., F. Seitz and A. Worms 2005. How the Bundesbank really conducted monetary policy. *North American Journal of Economics and Finance* 16, 277-292.

Giannone, D., M. Lenza and L. Reichlin 2008. Business cycles in the euro area. NBER Working Paper No. 14529.

Girouard, N. and C. André 2005. Measuring cyclically-adjusted budget balances for OECD countries. OECD Economics Department Working Paper No. 434.

Golinelli, R. and S. Momigliano 2009. The cyclical reaction of fiscal policies in the euro area: the role of modelling choices and data vintages. *Fiscal Studies* 30, 39-72.

Hodrick, R.J. and E.C. Prescott 1997. Postwar U.S. business cycles: an empirical investigation. *Journal of Money, Credit and Banking* 29, 1-16.

Holden, D. and K.A. Peel 1990. On testing for unbiasedness and efficiency of forecasts. *The Manchester School* 58, 120-127.

Leal, T., J.J. Pérez, M. Tujula and J.-P. Vidal 2008. Fiscal forecasting: lessons from the literature and challenges. *Fiscal Studies* 29, 347-386.

Marcellino, M. and A. Musso 2011. The reliability of real-time estimates of the euro area output gap. *Economic Modelling* 28, 1842-1856.

Mayer, E. and N. Stähler 2011. The debt brake: business cycle and welfare consequences of Germany's new fiscal policy rule. *Empirica,* forthcoming.

Mincer, J.A. and V. Zarnowitz 1969. The evaluation of economic forecasts. In: Mincer, J.A. (ed.). *Economic forecasts and expectations: analysis of forecasting behaviour and performance.* National Bureau of Economic Research, www.nber.org /books/minc69-1, 3-46.

Mink, M., J.P.A.M. Jacobs and J. de Haan 2012. Measuring coherence of output gaps with an application to the euro area. *Oxford Economics Papers* 64, 217-236.

Nelson, E. and K. Nikolov 2003. UK inflation in the 1970s and 1980s: the role of output gap mismeasurement. *Journal of Economics and Business* 55, 353-370.

Newey, W.K. and K.D. West 1987. A simple, positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703-708.

Orphanides, A. and S. van Norden 2002. The unreliability of output-gap estimates in real time. *Review of Economics and Statistics* 84, 569-583.

Prais, G.J. and C.B. Winsten 1954. Trend estimators and serial correlation. Cowles Commission Discussion Paper No. 383

Wooldridge, J.M. 2002. *Econometric analysis of cross section and panel data*. Cambridge, MA, MIT Press.

Appendix

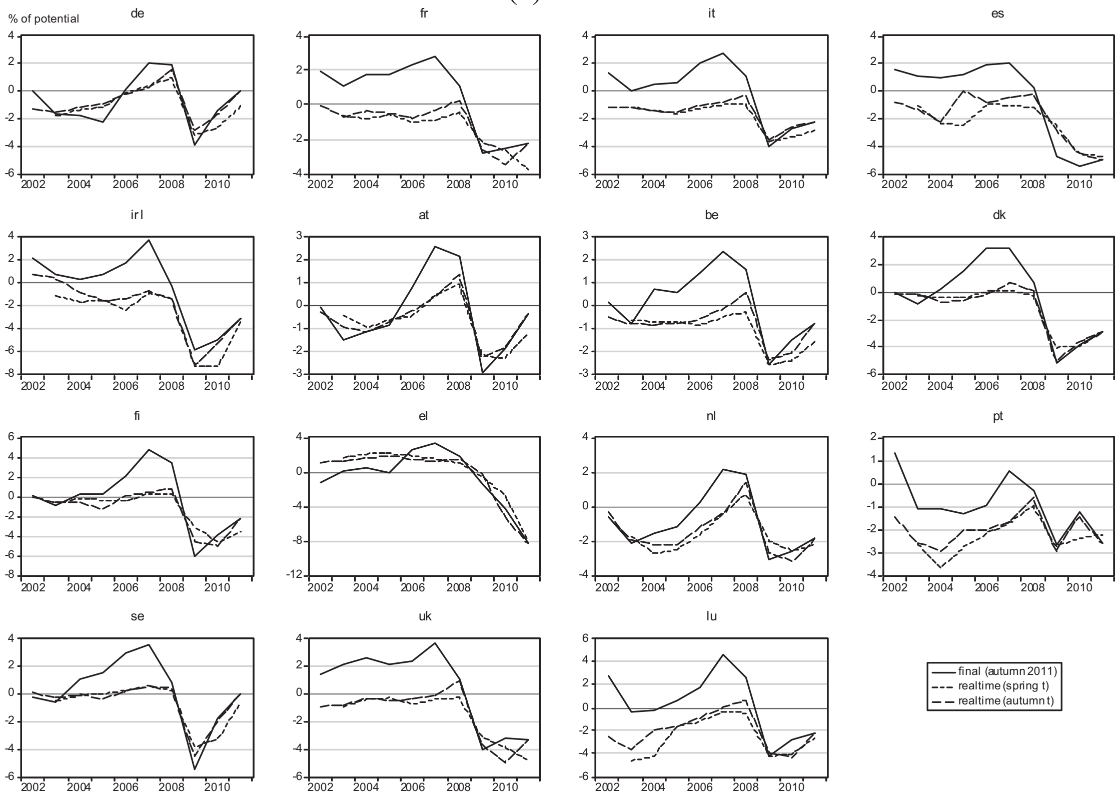
Appendix A1: Summary statistics for average output gap estimates by vintage (EU 1981-2011, OECD 1963-2011 and IMF 1970-2011) and in real time (EU 2002-2011, IMF 1999-2011 and OECD 1996-2011)

	EU		IMF		OECD	
	by vintage	real time (autumn t-1)	by vintage	real time (autumn t-1)	by vintage	real time (autumn t-1)
DE	-0.4	-0.8	-0.8	-1.5	-0.2	-1.5
FR	-0.4	-1.1	-0.2	-1.7	-0.5	-1.5
IT	-0.3	-1.6	-0.3	-2.1	-0.5	-2.2
ES	-1.0	-1.8	-0.5	-1.1	-1.1	-1.5
IRL	-0.4	-2.0	0.2	-0.7	-0.2	-0.6
AT	-0.3	-0.6	-0.4	-1.0	-0.2	-1.3
BE	-0.3	-0.8	0.6	-1.0	-0.4	-1.8
DK	-0.5	-1.3	.	.	-0.3	-0.6
FI	-0.2	-1.3	1.9	-1.3	-1.5	-1.1
EL	-0.7	-0.3	.	.	-0.8	-0.9
NL	-0.4	-1.5	-0.2	-0.7	0.0	-1.0
PT	-0.6	-2.0	-0.6	-2.4	-0.2	-2.1
SE	-0.6	-0.6	0.0	-1.0	-0.5	-0.9
UK	-0.3	-1.4	-0.4	-1.1	-0.4	-1.0
LU	-0.2	-2.0	.	.	.	.
mean	-0.4	-1.3	-0.1	-1.3	-0.5	-1.3

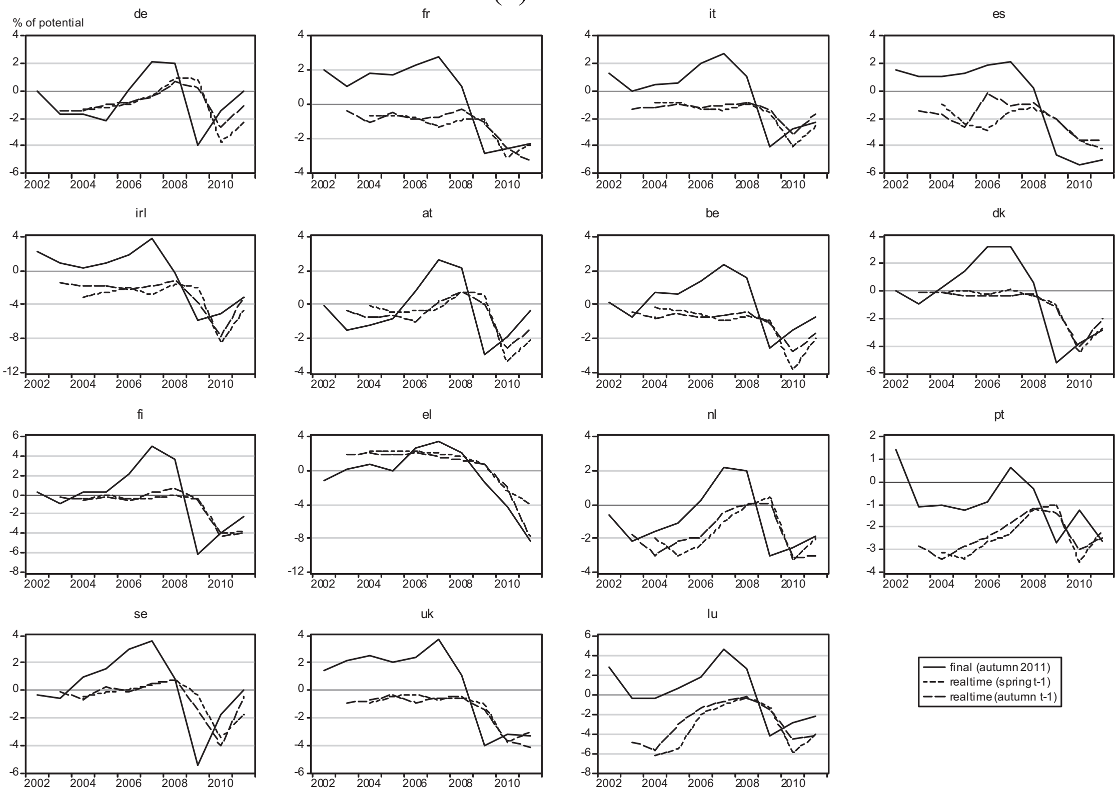
Note: We present output gaps for individual countries which have been averaged across the entire time-series dimension available for a given vintage and across the vintages. For the EU estimations, we have access to 21 vintages (compare with Table 1, we include here the 2008 winter vintage and the 2010 spring Kalman filter estimates), for the OECD 31 and for the IMF 27. For the EU, we typically have estimates for the years 1981 until the forecast horizon, for the OECD from the late 1960s (up to the autumn 2004 vintage and thereafter the late 1980s) until the forecast horizon, for the IMF from 1970 (up to the spring 2003 vintages and thereafter 1980) until the forecast horizon. Note that we do not attempt to date business cycles. Instead, we compute the average output gap starting in the earliest year that is provided until the forecast horizon. Note that for calculating the country-specific average above, we have access to many more target years compared to Table 2 since we are not restricted by the availability of real-time estimates. The evidence presented above confirms that the output gaps estimated for a given vintage yield values that are closer to zero, although there are significant differences across countries. On average, the finding holds that output gaps estimated in real time appear to be about one percentage point more negative. Recall however that sample periods are different. In Table 2 we show real-time and final statistics relying on the same sample, which lead, however, to broadly the same conclusion regarding the cross-country mean.

Appendix A2: Output gap estimations by the EU (solid line: last available estimation as of autumn 2011, dashed lines: real-time estimations)

(a) real time t

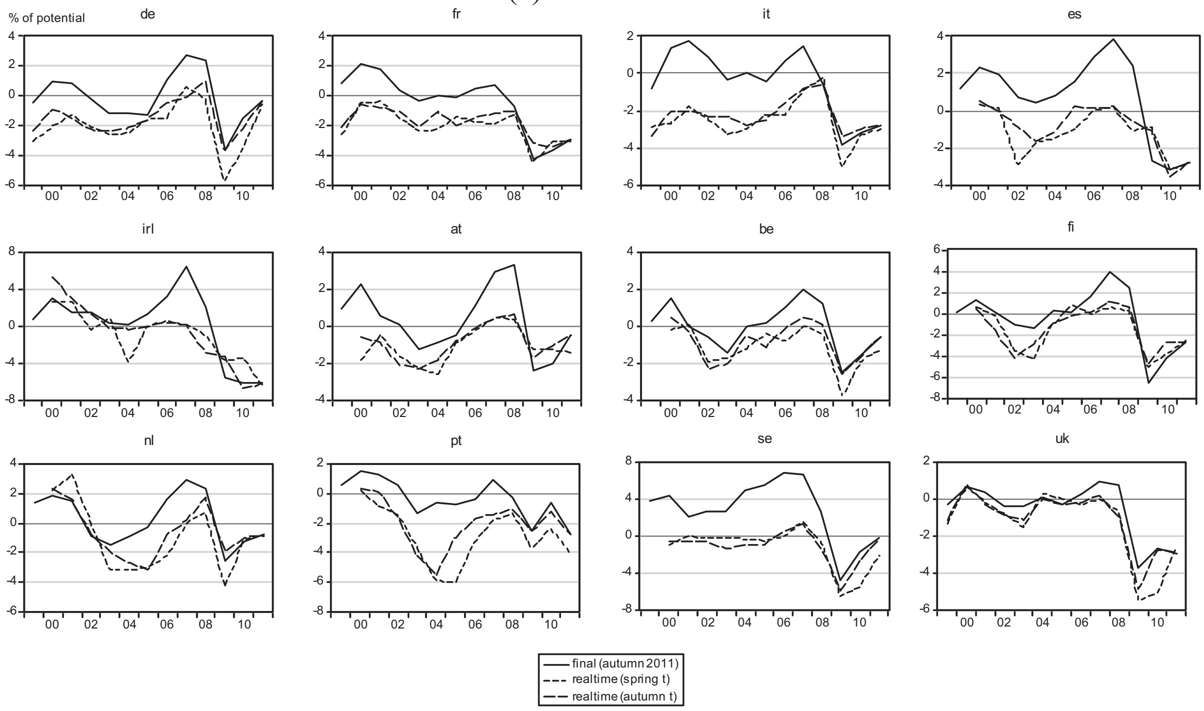


(b) real time t-1

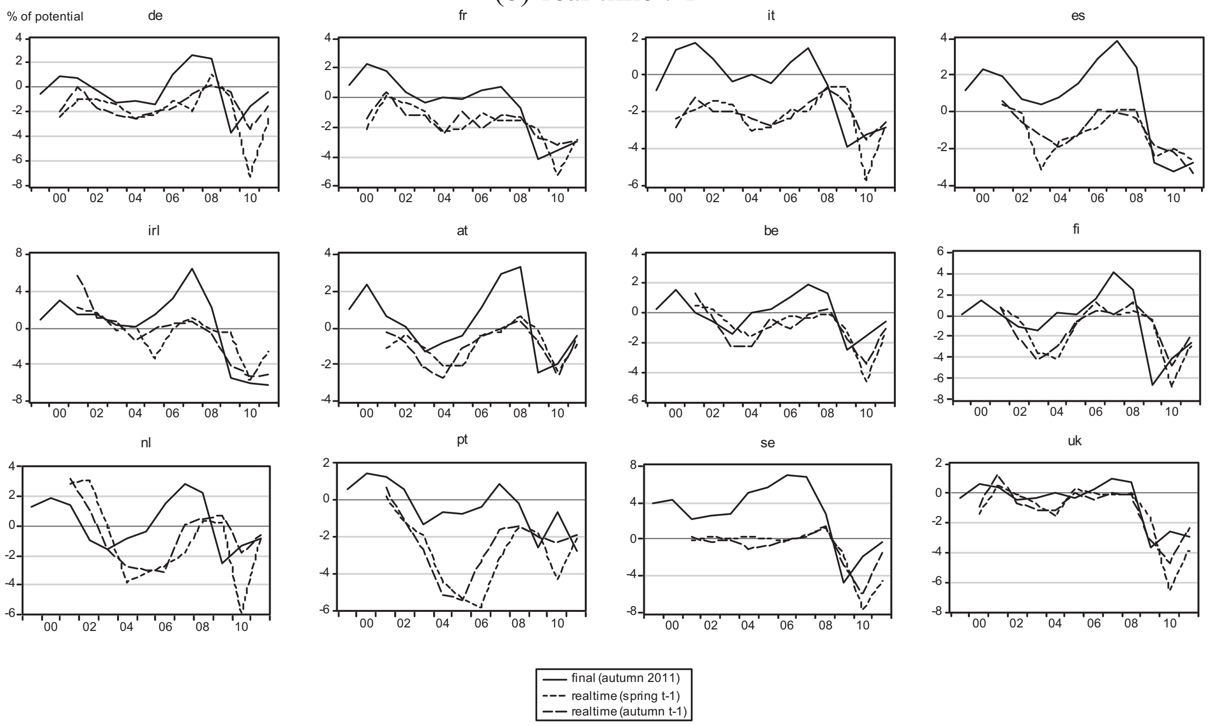


Appendix A3: Output gap estimations by the IMF (solid line: last available estimation as of autumn 2011, dashed lines: real-time estimations)

(a) real time t



(b) real time t-1



Appendix A4: Panel test for bias in real-time output gap estimations by the IMF (for real-time vintages spring of year t-1 to autumn of year t, 2000-2011)

	rev^{gap}_{it} (final – spring t-1)	rev^{gap}_{it} (final – autumn t-1)	rev^{gap}_{it} (final – spring t)	rev^{gap}_{it} (final – autumn t)
	(1)	(2)	(3)	(4)
DE	1.79	1.36	1.82	1.25
	(0.69)***	(0.63)**	(0.35)***	(0.43)***
FR	1.32	1.18	1.64	1.37
	(0.40)***	(0.41)***	(0.35)***	(0.42)***
IT	1.81	1.70	2.03	1.76
	(0.61)***	(0.54)***	(0.51)***	(0.52)***
ES	1.71	1.57	1.81	1.47
	(0.52)***	(0.50)***	(0.66)***	(0.74)**
IRL	0.57	0.47	1.07	0.61
	(0.98)	(0.89)	(1.05)	(1.28)
AT	0.99	1.04	1.44	1.18
	(0.46)**	(0.47)**	(0.62)**	(0.63)*
NL	1.03	0.57	1.00	0.60
	(0.96)	(0.79)	(0.57)*	(0.50)
BE	0.78	0.78	1.09	0.74
	(0.44)*	(0.44)*	(0.32)***	(0.34)**
FI	0.81	0.58	0.94	0.87
	(0.92)	(0.97)	(0.64)	(0.74)
SE	3.62	3.39	3.92	3.56
	(0.88)***	(0.96)***	(0.68)***	(0.85)***
PT	2.20	1.69	2.34	1.48
	(0.66)***	(0.60)***	(0.51)***	(0.62)**
UK	0.67	0.50	0.69	0.52
	(0.47)	(0.37)	(0.35)**	(0.29)*
N	136	136	148	148
R^2	0.11	0.11	0.16	0.12
uncentred R^2	0.43	0.43	0.64	0.53
Wald test, H0: all intercepts=0	831.6***	434.7***	230.2***	45.9***
Coefficient of a constant in a pooled model	1.45	1.28	1.69	1.38
	(0.49)***	(0.42)***	(0.24)***	(0.27)***
rho	0.09	0.23	0.38	0.53
Wooldridge test, H0: no first-order serial correlation	5.4**	22.8***	32.1***	57.9***

Note: Standard errors (in parentheses) are robust in the presence of heteroskedasticity, contemporaneous cross-sectional correlation and first-order serial correlation. * p<0.1; ** p<0.05; *** p<0.01.

Appendix A5: Panel test for bias in real-time output gap estimations by the EU (for real-time vintages spring of year t-1 to autumn of year t, 2002-2011)

	rev^{gap}_{it} (final – spring t-1)	rev^{gap}_{it} (final – autumn t-1)	rev^{gap}_{it} (final – spring t)	rev^{gap}_{it} (final – autumn t)
	(1)	(2)	(3)	(4)
DE	0.40	0.14	0.39	0.17
	(0.98)	(0.80)	(0.46)	(0.48)
FR	1.55	1.53	1.78	1.52
	(0.71)**	(0.59)***	(0.57)***	(0.48)***
IT	1.40	1.13	1.57	1.47
	(0.77)*	(0.74)	(0.55)***	(0.55)***
ES	1.24	1.05	1.41	1.22
	(0.95)	(0.82)	(0.80)*	(0.69)*
IRL	2.48	1.98	2.08	1.46
	(1.17)**	(0.86)**	(0.65)***	(0.60)**
AT	0.45	0.29	0.33	0.27
	(0.81)	(0.65)	(0.47)	(0.39)
NL	0.90	0.80	0.68	0.51
	(0.80)	(0.67)	(0.50)	(0.41)
BE	1.38	1.08	1.13	0.86
	(0.60)**	(0.53)**	(0.44)**	(0.40)**
FI	1.09	0.88	1.06	0.98
	(1.25)	(1.13)	(1.05)	(0.81)
SE	1.01	0.79	0.98	0.67
	(1.01)	(0.89)	(0.67)	(0.55)
PT	1.21	1.17	1.10	1.13
	(0.63)*	(0.48)**	(0.46)**	(0.44)**
UK	1.50	1.83	2.04	1.80
	(0.88)*	(0.75)**	(0.64)***	(0.67)***
DK	0.68	0.47	0.76	0.74
	(0.89)	(0.87)	(0.62)	(0.54)
EL	-1.33	-0.75	-0.62	-0.43
	(0.68)*	(0.53)	(0.59)	(0.68)
LU	3.37	2.90	2.55	2.34
	(1.13)***	(0.92)***	(0.76)***	(0.81)***
N	120	135	135	150
R^2	0.13	0.11	0.14	0.13
uncentred R^2	0.40	0.35	0.54	0.46
F-test, H0: all intercepts=0 (*)	3.06***	3.40***	6.82***	5.88***
Coefficient of a constant in a pooled model	1.17	1.05	1.22	1.04
	(0.68)*	(0.55)*	(0.36)***	(0.31)***
rho	0.17	0.22	0.39	0.43
Wooldridge test, H0: no first-order serial correlation	44.8***	82.2***	79.6***	64.5***

Note: Standard errors (in parentheses) are robust in the presence of heteroskedasticity, contemporaneous cross-sectional correlation and first-order serial correlation. * p<0.1; ** p<0.05; *** p<0.01. (*) The F-test is based on a regression with Newey-West (1987) corrected standard errors (not reported) because the number of observations per country is smaller than the number of countries and it is thus impossible to base the test on panel-corrected standard errors.

Appendix A6: Controlling for revisions in GDP growth

Taking the revisions of the growth rate of real GDP as proxies for the revisions of the macroeconomic variables and augmenting Model (2) yields:

(3) $$rev_{it}^{gap} = \alpha_i + \gamma \cdot rev_{it}^{\Delta GDP} + \varepsilon_{it}$$

We define: $rev^{\Delta GDP}_{it} = \Delta GDP^{final}_{it} - \Delta GDP^{real\text{-}time}_{it}$ where ΔGDP^{final}_{it} denotes the growth rate of real GDP of country i and year t, reported in the final vintage, while $\Delta GDP^{real\text{-}time}_{it}$ denotes the corresponding value reported at the time that the real-time output gap was estimated. Irrespective of the result for γ, a rejection of the joint null of $\alpha_i = 0$ would still indicate that the real-time output gaps are biased. We could also account for country-specific differences in the relationship between the revision in output gaps and GDP growth, ie estimate a fully country-specific model (γ_i). [11] We report the results for Model (3) in the following table and the results for the model with country-specific γ in Appendix A7. All six key results hold. The significance of the country-specific bias is higher, in particular in the spring t-1 vintage, and the overall bias tends to be larger.

However, given the large set of variables in the production function not included in Model (3), this may merely reflect omitted variables bias. Recall that the output gap is defined as (GDP/POT-1), where GDP denotes real GDP while POT denotes (real) potential output. If the revision in GDP was perfectly correlated with the revisions in potential output, we could not expect to obtain unbiased estimates of γ (or γ_i), but we could still hope to obtain unbiased estimates of the country-specific intercepts α_i, which are our main interest. Omitted revisions of potential output that are not perfectly correlated with revisions of real GDP would, however, be absorbed by the country-specific intercepts. Assume that the revisions in GDP as well as the revisions in potential output are both downward biased (ie forecasters expected higher growth in real time than in the final outcome), which reflects the empirical evidence, at least for the

[11] If we think of the properties of a statistical filter, such as the HP filter, including lagged and leaded revisions of real GDP growth as control variables seems straightforward, which we have also tested. While lagged revisions of real GDP growth revisions have virtually no significant effect on the revision of the output gap, leaded revisions play a role in the regressions based on data from the later real-time vintages (from year t). The key results remain virtually unchanged.

revisions of GDP growth. Assume also that the revisions of GDP and potential output are not perfectly correlated, which we would also expect to hold in reality. Controlling only for the revisions in GDP growth (while omitting the revisions in the other macro variables used to estimate potential output) would then result in negatively biased estimates. Thus, the potential omitted variables bias makes interpreting of the results reported below and in Appendix A7 virtually impossible.

Appendix A6 continued: Test for bias in real-time output gap estimations by the OECD (for real-time vintages spring of year t-1 to autumn of year t, 1996-2011)

	rev^{gap}_{it} (final – spring t-1)	rev^{gap}_{it} (final – autumn t-1)	rev^{gap}_{it} (final – spring t)	rev^{gap}_{it} (final – autumn t)
	(1)	(2)	(3)	(4)
DE	1.28	0.89	0.95	0.65
	(0.36)***	(0.36)**	(0.28)***	(0.29)**
FR	1.05	1.13	1.06	1.02
	(0.38)***	(0.31)***	(0.27)***	(0.25)***
IT	2.79	2.40	2.36	1.81
	(0.36)***	(0.34)***	(0.32)***	(0.40)***
ES	1.46	1.39	1.79	1.41
	(0.47)***	(0.32)***	(0.31)***	(0.31)***
IRL	1.93	1.76	1.87	1.26
	(0.92)**	(0.85)**	(0.64)***	(0.69)*
AT	0.86	0.81	0.86	0.69
	(0.41)**	(0.36)**	(0.29)***	(0.35)**
NL	1.62	1.42	1.42	1.02
	(0.47)***	(0.40)***	(0.33)***	(0.40)**
BE	1.76	1.53	1.44	1.06
	(0.45)***	(0.42)***	(0.36)***	(0.38)***
FI	0.06	0.15	0.23	0.18
	(0.45)	(0.41)	(0.31)	(0.32)
SE	0.80	0.53	0.85	0.70
	(0.56)	(0.43)	(0.48)*	(0.40)*
PT	2.84	2.55	2.66	2.09
	(0.45)***	(0.40)***	(0.31)***	(0.37)***
UK	0.94	0.83	0.87	0.78
	(0.52)*	(0.52)	(0.44)*	(0.36)**
DK	0.99	1.14	1.08	0.99
	(0.48)**	(0.38)***	(0.29)***	(0.27)***
EL	-0.44	-0.84	-0.59	-0.48
	(0.51)	(0.58)	(0.41)	(0.47)
$rev^{\Delta GDP}$ (final - spring t-1)	0.76			
	(0.07)***			
$rev^{\Delta GDP}$ (final - autumn t-1)		0.64		
		(0.08)**		
$rev^{\Delta GDP}$ (final - spring t)			0.34	
			(0.10)**	
$rev^{\Delta GDP}$ (final - autumn t)				0.22
				(0.13)*
N	196	210	210	224
R2	0.73	0.59	0.34	0.15
Wald test, H0: all intercepts=0	19225***	891***	1282***	91***
Coefficient of a constant in a pooled model	1.28	1.16	1.26	1.03
	(0.22)***	(0.18)***	(0.17)***	(0.17)***
rho	0.42	0.42	0.36	0.39
Wooldridge test, H0: no first-order serial correlation	44.9***	51.7***	9.4***	25.7***

Note: Standard errors (in parentheses) are robust in the presence of heteroskedasticity, contemporaneous cross-sectional correlation and first-order serial correlation. * p<0.1; ** p<0.05; *** p<0.01.

Appendix A7: Test for bias in real-time output gap estimations by the OECD, fully country-specific model (for real-time vintages of year t-1, 1996-2011)

	rev^{gap}_{it} (final – spring t-1)		rev^{gap}_{it} (final – autumn t-1)	
	Intercept	$rev^{\Delta GDP}$ (final-spring t-1)	Intercept	$rev^{\Delta GDP}$ (final-autumnt-1)
DE	1.33	0.85	0.92	0.74
	(0.32)***	(0.08)***	(0.36)**	(0.12)***
FR	1.17	0.97	1.23	1.07
	(0.34)***	(0.15)***	(0.28)***	(0.18)***
IT	2.84	0.80	2.41	0.67
	(0.35)***	(0.11)***	(0.36)***	(0.15)***
ES	1.50	0.65	1.54	0.12
	(0.43)***	(0.18)***	(0.25)***	(0.15)
IRL	1.64	0.43	1.85	0.28
	(0.75)**	(0.14)***	(0.75)**	(0.16)*
AT	0.87	0.81	0.81	0.65
	(0.39)**	(0.12)***	(0.37)**	(0.14)***
NL	1.71	1.05	1.41	0.68
	(0.37)***	(0.11)***	(0.41)***	(0.16)***
BE	1.81	0.94	1.53	0.88
	(0.40)***	(0.13)***	(0.41)***	(0.16)***
FI	0.05	0.74	0.14	0.79
	(0.41)	(0.08)***	(0.38)	(0.07)***
SE	0.79	0.85	0.51	0.68
	(0.51)	(0.11)***	(0.44)	(0.12)***
PT	2.83	0.71	2.49	0.49
	(0.44)***	(0.15)***	(0.41)***	(0.18)***
UK	0.94	0.59	0.83	0.63
	(0.48)**	(0.15)***	(0.53)	(0.22)***
DK	1.19	1.01	1.35	0.96
	(0.42)***	(0.14)***	(0.33)***	(0.12)***
EL	-0.53	0.66	-0.84	0.66
	(0.51)	(0.15)***	(0.61)	(0.20)***
N	196		210	
R^2	0.76		0.65	
Wald test, H0: all intercepts=0	747***		787***	
Coefficient of a constant in a pooled model	1.29 (0.20)***		1.20 (0.17)***	
Rho	0.37		0.44	
Wooldridge test, H0: no first-order serial correlation	78.7***		108.5***	

Note: Standard errors (in parentheses) are robust in the presence of heteroskedasticity, contemporaneous cross-sectional correlation and first-order serial correlation. * p<0.1; ** p<0.05; *** p<0.01.

Appendix A7 continued

	rev^{gap}_{it} (final – spring t)		rev^{gap}_{it} (final – autumn t)	
	Intercept	$rev^{\Delta GDP}$ (final-spring t)	Intercept	$rev^{\Delta GDP}$ (final-autumn t)
DE	0.97	0.39	0.67	0.67
	(0.25)***	(0.22)*	(0.26)***	(0.29)**
FR	1.12	0.88	0.94	1.18
	(0.21)***	(0.22)***	(0.18)***	(0.27)***
IT	2.39	0.21	1.85	0.07
	(0.29)***	(0.24)	(0.39)***	(0.59)
ES	2.11	-0.25	1.56	-0.09
	(0.33)***	(0.39)	(0.36)***	(0.55)
IRL	1.97	0.17	1.35	0.03
	(0.59)***	(0.18)	(0.65)**	(0.29)
AT	0.83	0.43	0.72	0.16
	(0.29)***	(0.26)*	(0.36)**	(0.41)
NL	1.42	0.36	1.09	0.11
	(0.32)***	(0.21)*	(0.42)***	(0.36)
BE	1.43	0.38	1.02	0.45
	(0.32)***	(0.21)*	(0.35)***	(0.28)
FI	0.22	0.40	0.25	0.09
	(0.28)	(0.13)***	(0.31)	(0.20)
SE	0.85	0.44	0.78	0.04
	(0.44)*	(0.27)	(0.40)*	(0.38)
PT	2.71	0.23	2.08	0.30
	(0.29)***	(0.21)	(0.40)***	(0.56)
UK	1.00	0.07	0.75	0.29
	(0.41)**	(0.24)	(0.36)**	(0.27)
DK	1.25	0.78	1.02	0.46
	(0.22)***	(0.16)***	(0.25)***	(0.27)*
EL	-0.57	0.39	-0.49	0.21
	(0.39)	(0.18)**	(0.44)	(0.24)
N	210		224	
R^2	0.39		0.19	
Wald test, H0: all intercepts=0	1009***		66***	
Coefficient of a constant in a pooled model	1.30 (0.17)***		1.03 (0.16)***	
Rho	0.29		0.35	
Wooldridge test, H0: no first-order serial correlation	9.1**		21.5***	

Note: Standard errors (in parentheses) are robust in the presence of heteroskedasticity, contemporaneous cross-sectional correlation and first-order serial correlation. * p<0.1; ** p<0.05; *** p<0.01.

Appendix A8: Mincer-Zarnowitz test for bias in real-time output gap estimations by the OECD, fully country-specific model (for real-time vintages spring of year t-1 to autumn of year t, 1996-2011)

	gap^{final}_{it}		gap^{final}_{it}	
	Intercept	$gap^{realtime}$ (spring t-1)	Intercept	$gap^{realtime}$ (autumn t-1)
DE	-0.54	0.19	0.58	0.93
	(0.68)	(0.27)	(0.74)	(0.40)**
FR	0.07	0.48	0.91	0.95
	(0.56)	(0.27)*	(0.52)*	(0.26)***
IT	0.91	0.46	2.15	1.02
	(0.81)	(0.31)	(0.68)***	(0.26)***
ES	1.35	0.83	1.97	1.19
	(0.45)***	(0.15)***	(0.18)***	(0.07)***
IRL	1.30	1.26	2.10	1.12
	(0.73)*	(0.20)***	(0.57)***	(0.13)***
AT	0.11	0.35	1.00	1.00
	(0.65)	(0.32)	(0.53)*	(0.28)***
NL	0.66	0.22	1.09	0.60
	(0.52)	(0.19)	(0.40)***	(0.18)***
BE	-0.04	0.23	0.23	0.36
	(0.45)	(0.14)	(0.43)	(0.16)**
FI	-0.56	0.63	-0.16	0.71
	(0.79)	(0.26)**	(0.68)	(0.21)***
SE	0.30	0.42	0.65	0.95
	(0.74)	(0.23)*	(0.54)	(0.22)***
PT	1.29	0.48	2.05	0.83
	(0.65)**	(0.23)**	(0.59)***	(0.23)***
UK	0.62	0.68	0.77	0.80
	(0.52)	(0.21)***	(0.46)*	(0.21)***
DK	0.09	0.53	0.48	0.76
	(0.63)	(0.28)*	(0.48)	(0.20)***
EL	-0.86	1.30	-0.82	1.40
	(0.55)	(0.15)***	(0.47)*	(0.15)***
N	196		210	
R^2	0.53		0.68	
Wald test, H0: all intercepts=0 & $gap^{realtime}$=1	616***		29980***	
Coefficient of a constant in a pooled model	0.39 (0.45)		0.85 (0.31)***	
Rho	0.16		0.13	
Wooldridge test, H0: no first-order serial correlation	315.2***		161.6***	

Note: Standard errors (in parentheses) are robust in the presence of heteroskedasticity, contemporaneous cross-sectional correlation and first-order serial correlation. * p<0.1; ** p<0.05; *** p<0.01.

Appendix A8 continued

	gap^{final}_{it}		gap^{final}_{it}	
	Intercept	$gap^{realtime}$ (spring t)	Intercept	$gap^{realtime}$ (autumn t)
DE	0.76	0.91	1.00	1.23
	(0.33)**	(0.12)***	(0.37)***	(0.16)***
FR	0.97	0.96	1.07	1.02
	(0.40)**	(0.17)***	(0.37)***	(0.16)***
IT	2.38	0.99	1.65	0.90
	(0.51)***	(0.17)***	(0.56)***	(0.20)***
ES	1.82	0.92	1.85	1.19
	(0.30)***	(0.09)***	(0.34)***	(0.12)***
IRL	2.03	0.96	1.38	1.05
	(0.57)***	(0.11)***	(0.63)**	(0.13)***
AT	1.30	1.20	1.10	1.25
	(0.35)***	(0.17)***	(0.39)***	(0.20)***
NL	1.23	0.70	0.92	0.71
	(0.30)***	(0.11)***	(0.35)***	(0.14)***
BE	0.70	0.58	0.41	0.59
	(0.30)**	(0.10)***	(0.34)	(0.12)***
FI	0.31	1.02	0.26	0.98
	(0.35)	(0.10)***	(0.32)	(0.08)***
SE	0.79	0.84	1.04	1.20
	(0.45)*	(0.14)***	(0.39)***	(0.14)***
PT	2.18	0.76	1.95	0.89
	(0.40)***	(0.12)***	(0.55)***	(0.21)***
UK	0.84	0.82	0.68	0.81
	(0.41)**	(0.16)***	(0.35)*	(0.13)***
DK	0.80	0.88	0.88	0.93
	(0.33)**	(0.12)***	(0.28)***	(0.09)***
EL	-0.26	1.23	-0.64	0.94
	(0.39)	(0.09)***	(0.47)	(0.09)***
N	210		224	
R^2	0.83		0.81	
Wald test, H0: all intercepts=0 & $gap^{realtime}$=1	857.2***		247.1***	
Coefficient of a constant in a pooled model	1.08 (0.20)***		0.96 (0.18)***	
Rho	0.24		0.34	
Wooldridge test, H0: no first-order serial correlation	25.9***		88.9***	

Note: Standard errors (in parentheses) are robust in the presence of heteroskedasticity, contemporaneous cross-sectional correlation and first-order serial correlation. * p<0.1; ** p<0.05; *** p<0.01.

Appendix A9: Mincer-Zarnowitz test for bias in real-time output gap estimations by the IMF (for real-time vintages spring of year t-1 to autumn of year t, 2000-2011)

	gap^{final}_{it} (autumn 2011)			
	(1)	(2)	(3)	(4)
DE	0.69	0.93	1.55	1.30
	(0.77)	(0.72)	(0.42)***	(0.45)***
FR	0.25	0.68	1.36	1.43
	(0.64)	(0.59)	(0.45)***	(0.47)***
IT	0.50	1.07	1.67	1.84
	(0.77)	(0.73)	(0.59)***	(0.59)***
ES	0.97	1.23	1.65	1.51
	(0.71)	(0.64)*	(0.69)**	(0.74)**
IRL	0.07	0.23	0.92	0.64
	(1.18)	(0.97)	(1.06)	(1.27)
AT	0.45	0.74	1.29	1.21
	(0.62)	(0.58)	(0.65)**	(0.62)*
NL	0.40	0.39	0.89	0.61
	(0.74)	(0.72)	(0.52)*	(0.50)
BE	0.21	0.48	0.94	0.76
	(0.47)	(0.47)	(0.35)***	(0.35)**
FI	-0.09	0.18	0.74	0.91
	(1.04)	(1.03)	(0.70)	(0.74)
SE	2.93	3.08	3.74	3.60
	(1.05)***	(1.04)***	(0.74)***	(0.84)***
PT	0.54	0.99	1.96	1.55
	(0.82)	(0.75)	(0.57)***	(0.68)**
UK	-0.07	0.17	0.51	0.57
	(0.52)	(0.44)	(0.35)	(0.33)*
$gap^{realtime}$ (spring t-1)	0.41			
	(0.23)*			
$gap^{realtime}$ (autumn t-1)		0.70		
		(0.23)***		
$gap^{realtime}$ (spring t)			0.86	
			(0.12)***	
$gap^{realtime}$ (autumn t)				1.03
				(0.13)***
N	136	136	148	148
R2	0.21	0.33	0.56	0.61
Wald test, H0: all intercepts=0 & $gap^{realtime}$=1	429.2***	604.4***	331.3***	50.2***
Coefficient of a constant in a pooled model	0.81	1.03	1.56	1.40
	(0.23)***	(0.22)***	(0.11)***	(0.17)***
rho	0.24	0.27	0.39	0.52
Wooldridge test, H0: no first-order serial correlation	70.1***	35.8***	185.2***	50.5***

Note: Standard errors (in parentheses) are robust in the presence of heteroskedasticity, contemporaneous cross-sectional correlation and first-order serial correlation. * p<0.1; ** p<0.05; *** p<0.01.

Appendix A10: Mincer-Zarnowitz test for bias in real-time output gap estimations by the EU (for real-time vintages spring of year t-1 to autumn of year t, 2002-2011)

	gap^{final}_{it} (autumn 2011)			
	(1)	(2)	(3)	(4)
DE	0.12	0.12	0.72	0.29
	(1.03)	(0.80)	(0.39)*	(0.38)
FR	1.21	1.51	2.23	1.75
	(0.91)	(0.69)**	(0.54)***	(0.43)***
IT	0.95	1.10	2.18	1.79
	(1.06)	(0.83)	(0.51)***	(0.50)***
ES	0.62	1.01	2.13	1.58
	(1.39)	(1.00)	(0.79)***	(0.64)**
IRL	1.59	1.93	2.99	1.86
	(1.82)	(1.17)*	(0.77)***	(0.59)***
AT	0.28	0.28	0.57	0.38
	(0.82)	(0.65)	(0.39)	(0.30)
NL	0.46	0.76	1.21	0.81
	(1.03)	(0.82)	(0.50)**	(0.41)**
BE	1.08	1.06	1.50	1.03
	(0.76)	(0.59)*	(0.39)***	(0.35)***
FI	0.77	0.86	1.50	1.25
	(1.39)	(1.13)	(0.87)*	(0.65)*
SE	0.85	0.79	1.24	0.81
	(1.06)	(0.86)	(0.53)**	(0.45)*
PT	0.58	1.12	1.82	1.50
	(1.15)	(0.85)	(0.57)***	(0.48)***
UK	1.14	1.80	2.55	2.11
	(1.09)	(0.86)**	(0.57)***	(0.60)***
DK	0.38	0.46	1.21	1.02
	(1.05)	(0.89)	(0.52)**	(0.49)**
EL	-1.22	-0.74	-0.51	-0.28
	(0.88)	(0.52)	(0.64)	(0.68)
LU	2.48	2.83	3.36	2.72
	(1.77)	(1.26)**	(0.79)***	(0.75)***
$gap^{realtime}$ (spring t-1)	0.74			
	(0.41)*			
$gap^{realtime}$ (autumn t-1)		0.98		
		(0.30)***		
$gap^{realtime}$ (spring t)			1.30	
			(0.16)***	
$gap^{realtime}$ (autumn t)				1.19
				(0.14)***
N	120	135	135	150
R2	0.21	0.36	0.71	0.72
Wald test, H0: all intercepts=0 & $gap^{realtime}$=1	3.37***	3.24***	9.8***	7.43***
Coefficient of a constant in a pooled model	0.70	0.88	1.43	1.22
	(0.28)**	(0.23)***	(0.18)***	(0.14)***
rho	0.19	0.18	0.29	0.32
Wooldridge test, H0: no first-order serial correlation	119.4***	121.3***	52.8***	50.4***

Note: Standard errors (in parentheses) are robust in the presence of heteroskedasticity, contemporaneous cross-sectional correlation and first-order serial correlation. * p<0.1; ** p<0.05; *** p<0.01. (*) The F-test is based on a regression with Newey-West (1987) corrected standard errors (not reported) because the number of observations per country is smaller than the number of countries and it is thus impossible to base the test on panel-corrected standard errors.

The following Discussion Papers have been published since 2012:

01	2012	A user cost approach to capital measurement in aggregate production functions	Thomas A. Knetsch
02	2012	Assessing macro-financial linkages: a model comparison exercise	Gerke, Jonsson, Kliem Kolasa, Lafourcade, Locarno Makarski, McAdam
03	2012	Executive board composition and bank risk taking	A. N. Berger T. Kick, K. Schaeck
04	2012	Stress testing German banks against a global cost-of-capital shock	Klaus Duellmann Thomas Kick
05	2012	Regulation, credit risk transfer with CDS, and bank lending	Thilo Pausch Peter Welzel
06	2012	Maturity shortening and market failure	Felix Thierfelder
07	2012	Towards an explanation of cross-country asymmetries in monetary transmission	Georgios Georgiadis
08	2012	Does Wagner's law ruin the sustainability of German public finances?	Christoph Priesmeier Gerrit B. Koester
09	2012	Bank regulation and stability: an examination of the Basel market risk framework	Gordon J. Alexander Alexandre M. Baptista Shu Yan
10	2012	Capital regulation, liquidity requirements and taxation in a dynamic model of banking	Gianni De Nicolò Andrea Gamba Marcella Lucchetta
11	2012	Credit portfolio modelling and its effect on capital requirements	Dilek Bülbül Claudia Lambert

12	2012	Trend growth expectations and U.S. house prices before and after the crisis	Mathias Hoffmann, Michael U. Krause, Thomas Laubach
13	2012	The PHF: a comprehensive panel survey on household finances and wealth in Germany	Ulf von Kalckreuth, Martin Eisele, Julia Le Blanc, Tobias Schmidt, Junyi Zhu
14	2012	The effectiveness of monetary policy in steering money market rates during the financial crisis	Puriya Abbassi, Tobias Linzert
15	2012	Cyclical adjustment in fiscal rules: some evidence on real-time bias for EU-15 countries	Gerhard Kempkes

The following Discussion Papers have been published since 2011:

Series 1: Economic Studies

01	2011	Long-run growth expectations and "global imbalances"	M. Hoffmann, M. Krause, T. Laubach
02	2011	Robust monetary policy in a New Keynesian model with imperfect interest rate pass-through	Rafael Gerke, Felix Hammermann
03	2011	The impact of fiscal policy on economic activity over the business cycle – evidence from a threshold VAR analysis	Anja Baum, Gerrit B. Koester
04	2011	Classical time-varying FAVAR models – estimation, forecasting and structural analysis	S. Eickmeier, W. Lemke, M. Marcellino

05	2011	The changing international transmission of financial shocks: evidence from a classical time-varying FAVAR	Sandra Eickmeier Wolfgang Lemke Massimiliano Marcellino
06	2011	FiMod – a DSGE model for fiscal policy simulations	Nikolai Stähler Carlos Thomas
07	2011	Portfolio holdings in the euro area – home bias and the role of international, domestic and sector-specific factors	Axel Jochem Ute Volz
08	2011	Seasonality in house prices	F. Kajuth, T. Schmidt
09	2011	The third pillar in Europe: institutional factors and individual decisions	Julia Le Blanc
10	2011	In search for yield? Survey-based evidence on bank risk taking	C. M. Buch S. Eickmeier, E. Prieto
11	2011	Fatigue in payment diaries – empirical evidence from Germany	Tobias Schmidt
12	2011	Currency blocs in the 21st century	Christoph Fischer
13	2011	How informative are central bank assessments of macroeconomic risks?	Malte Knüppel Guido Schultefrankenfeld
14	2011	Evaluating macroeconomic risk forecasts	Malte Knüppel Guido Schultefrankenfeld
15	2011	Crises, rescues, and policy transmission through international banks	Claudia M. Buch Cathérine Tahmee Koch Michael Koetter
16	2011	Substitution between net and gross settlement systems – A concern for financial stability?	Ben Craig Falko Fecht

17	2011	Recent developments in quantitative models of sovereign default	Nikolai Stähler
18	2011	Exchange rate dynamics, expectations, and monetary policy	Qianying Chen
19	2011	An information economics perspective on main bank relationships and firm R&D	D. Hoewer T. Schmidt, W. Sofka
20	2011	Foreign demand for euro banknotes issued in Germany: estimation using direct approaches	Nikolaus Bartzsch Gerhard Rösl Franz Seitz
21	2011	Foreign demand for euro banknotes issued in Germany: estimation using indirect approaches	Nikolaus Bartzsch Gerhard Rösl Franz Seitz
22	2011	Using cash to monitor liquidity – implications for payments, currency demand and withdrawal behavior	Ulf von Kalckreuth Tobias Schmidt Helmut Stix
23	2011	Home-field advantage or a matter of ambiguity aversion? Local bias among German individual investors	Markus Baltzer Oscar Stolper Andreas Walter
24	2011	Monetary transmission right from the start: on the information content of the eurosystem's main refinancing operations	Puriya Abbassi Dieter Nautz
25	2011	Output sensitivity of inflation in the euro area: indirect evidence from disaggregated consumer prices	Annette Fröhling Kirsten Lommatzsch
26	2011	Detecting multiple breaks in long memory: the case of U.S. inflation	Uwe Hassler Barbara Meller

27	2011	How do credit supply shocks propagate internationally? A GVAR approach	Sandra Eickmeier Tim Ng
28	2011	Reforming the labor market and improving competitiveness: an analysis for Spain using FiMod	Tim Schwarzmüller Nikolai Stähler
29	2011	Cross-border bank lending, risk aversion and the financial crisis	Cornelia Düwel, Rainer Frey Alexander Lipponer
30	2011	The use of tax havens in exemption regimes	Anna Gumpert James R. Hines, Jr. Monika Schnitzer
31	2011	Bank-related loan supply factors during the crisis: an analysis based on the German bank lending survey	Barno Blaes
32	2011	Evaluating the calibration of multi-step-ahead density forecasts using raw moments	Malte Knüppel
33	2011	Optimal savings for retirement: the role of individual accounts and disaster expectations	Julia Le Blanc Almuth Scholl
34	2011	Transitions in the German labor market: structure and crisis	Michael U. Krause Harald Uhlig
35	2011	U-MIDAS: MIDAS regressions with unrestricted lag polynomials	C. Foroni M. Marcellino, C. Schumacher

Series 2: Banking and Financial Studies

01	2011	Contingent capital to strengthen the private safety net for financial institutions: Cocos to the rescue?	George M. von Furstenberg
02	2011	Gauging the impact of a low-interest rate environment on German life insurers	Anke Kablau Michael Wedow
03	2011	Do capital buffers mitigate volatility of bank lending? A simulation study	Frank Heid Ulrich Krüger
04	2011	The price impact of lending relationships	Ingrid Stein
05	2011	Does modeling framework matter? A comparative study of structural and reduced-form models	Yalin Gündüz Marliese Uhrig-Homburg
06	2011	Contagion at the interbank market with stochastic LGD	Christoph Memmel Angelika Sachs, Ingrid Stein
07	2011	The two-sided effect of financial globalization on output volatility	Barbara Meller
08	2011	Systemic risk contributions: a credit portfolio approach	Klaus Düllmann Natalia Puzanova
09	2011	The importance of qualitative risk assessment in banking supervision before and during the crisis	Thomas Kick Andreas Pfingsten
10	2011	Bank bailouts, interventions, and moral hazard	Lammertjan Dam Michael Koetter
11	2011	Improvements in rating models for the German corporate sector	Till Förstemann

12	2011	The effect of the interbank network structure on contagion and common shocks	Co-Pierre Georg
13	2011	Banks' management of the net interest margin: evidence from Germany	Christoph Memmel Andrea Schertler
14	2011	A hierarchical Archimedean copula for portfolio credit risk modelling	Natalia Puzanova
15	2011	Credit contagion between financial systems	Natalia Podlich Michael Wedow
16	2011	A hierarchical model of tail dependent asset returns for assessing portfolio credit risk	Natalia Puzanova
17	2011	Contagion in the interbank market and its determinants	Christoph Memmel Angelika Sachs
18	2011	Does it pay to have friends? Social ties and executive appointments in banking	A. N. Berger, T. Kick M. Koetter, K. Schaeck

Visiting researcher at the Deutsche Bundesbank

The Deutsche Bundesbank in Frankfurt is looking for a visiting researcher. Among others under certain conditions visiting researchers have access to a wide range of data in the Bundesbank. They include micro data on firms and banks not available in the public. Visitors should prepare a research project during their stay at the Bundesbank. Candidates must hold a PhD and be engaged in the field of either macroeconomics and monetary economics, financial markets or international economics. Proposed research projects should be from these fields. The visiting term will be from 3 to 6 months. Salary is commensurate with experience.

Applicants are requested to send a CV, copies of recent papers, letters of reference and a proposal for a research project to:

Deutsche Bundesbank
Personalabteilung
Wilhelm-Epstein-Str. 14

60431 Frankfurt
GERMANY