EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1157 / FEBRUARY 2010

REAL TIME ESTIMATES OF THE EURO AREA OUTPUT GAP

RELIABILITY AND FORECASTING PERFORMANCE

by Massimiliano Marcellino
and Alberto Musso

EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1157 / FEBRUARY 2010

REAL TIME ESTIMATES OF THE EURO AREA OUTPUT GAP

RELIABILITY AND FORECASTING PERFORMANCE[1]

by Massimiliano Marcellino[2]

and Alberto Musso[3]

In 2010 all ECB publications feature a motif taken from the €500 banknote.

1 We thank, without implicating, Todd Clark for many useful suggestions and Mary Santoianni for assisting us in collecting the data. We are also grateful to Simon van Norden, Ken West, and participants to an ESCB seminar, the 5th Eurostat Colloquium and a CES-IFO workshop for useful comments. All remaining errors are the sole responsibility of the authors. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the European Central Bank.

2 European University Institute, Badia Fiesolana - Via dei Roccettini 9, I-50014 San Domenico di Fiesole (FI), Italy; Bocconi University and CEPR; e-mail: massimiliano.marcellino@eui.eu

3 European Central Bank, Kaiserstrasse 29, D-60311 Frankfurt am Main, Germany; e-mail: alberto.musso@ecb.europa.eu

Address
Kaiserstrasse 29
60311 Frankfurt am Main, Germany

Postal address
Postfach 16 03 19
60066 Frankfurt am Main, Germany

Telephone
+49 69 1344 0

Website
http://www.ecb.europa.eu

Fax
+49 69 1344 6000

Information on all of the working papers published in the ECB's Working Paper Series can be found on the ECB's website, http://www.ecb.europa.eu/pub/scientific/wps/date/html/index.en.html

ISSN 1725-2806 (online)

CONTENTS

ABSTRACT

This paper provides evidence on the reliability of euro area real-time output gap estimates. A genuine real-time data set for the euro area is used, including vintages of several sets of euro area output gap estimates available from 1999 to 2006. It turns out that real-time estimates of the output gap are characterised by a high degree of uncertainty, much higher than that resulting from model and estimation uncertainty only. In particular, the evidence indicates that both the magnitude and the sign of the real-time estimates of the euro area output gap are very uncertain. The uncertainty is mostly due to parameter instability, while data revisions seem to play a minor role. To benchmark our results, we repeat the analysis for the US over the same sample. It turns out that US real time estimates are much more correlated with final estimates than for the euro area, data revisions play a larger role, but overall the unreliability in real time of the US output gap measures detected in earlier studies is confirmed in the more recent period. Moreover, despite some difference across output gap estimates and forecast horizons, the results point clearly to a lack of any usefulness of real-time output gap estimates for inflation forecasting both in the short term (one-quarter and one-year ahead) and the medium term (two-year and three-year ahead). By contrast, some evidence is provided indicating that several output gap estimates are useful to forecast real GDP growth, particularly in the short term, and some appear also useful in the medium run. No single output gap measure appears superior to all others in all respects.

Key words: Output gap, real-time data, euro area, inflation forecasts, real GDP forecasts, data revisions.

JEL classification codes: E31, E37, E52, E58.

Non-technical summary

Despite their appealing characteristic as a relatively clear summary measure of overall slack in the economy, output gap estimates are problematic and represent a potentially misleading input in monetary policy analysis. The two main problems in interpreting and assessing the implications of the output gap relate to the uncertainty surrounding the corresponding estimates and the uncertain links between these measures and inflation. Although evidence based on real-time data exists for a number of countries, no such evidence exists for the euro area. This paper fills this gap by providing updated evidence on the uncertainty characterising euro area output gap real-time estimates and by comparing it with the case of the US.

A number of studies have addressed the usefulness of euro area output gap estimates in terms of revisions and inflation forecasting performance (including Camba-Méndez and Rodríguez Palenzuela (2003), Proietti, Musso and Westermann (2007) and Rünstler (2002)). However, these papers are based on one specific vintage (the latest available at the time of the study). In contrast, the present study uses a genuine set of real time output gap estimates for the euro area, which allows drawing more robust conclusions regarding the reliability of euro area output gap estimates, as well as comparing results with corresponding ones obtained for the US by Orphanides and van Norden (2002) and others. In addition, we compare a large set of output gap measures, including simple filter based estimates relying on real GDP, measures based on capacity utilization, estimates based on multivariate unobserved component models, and a variety of estimates from international organizations such as the IMF, OECD and European Commission. Finally, we construct gap measures by averaging those described so far.

The first part of the paper provides an assessment of the degree of uncertainty characterising euro area real-time output gap estimates. Consistent with the findings of previous empirical studies for other economic areas, the analysis of the various sources of uncertainty, based on an assessment of alternative estimates, measures of confidence bands around point estimates and past revisions, suggests quite clearly that real-time estimates are characterised by a high degree of uncertainty. In particular, the evidence indicates that both the magnitude and the sign of the real-time estimates of the euro area output gap are very uncertain. For the euro area, changes in the vintages of the time series underlying the gap (e.g., real GDP) explain a minor part of the real time changes in the gap, while recursive computation matters considerably. This finding suggests either the need of a very long estimation sample for reliable gap estimation or, more likely, the presence of parameter changes. Unfortunately, averaging different gap measures does not yield any substantial gains, due to the rather high correlation across alternative gap measures. Real time estimates of the US output gap suffer from similar problems, also in the most recent period, even though they are more correlated with final values with respect to the euro area. In addition, the data revision component of the revision error is larger than for the euro area.

The second part of the paper addresses the question of how useful real time estimates of the euro area output gap are for inflation and growth forecasting purposes. For this purpose, recently developed econometric techniques are used (i.e., the tests of equal predictive ability by Clark and McCraken, 2009), which take into account the real time nature of the data. As regards the predictive content for inflation of alternative measures of the euro area output gap, an out-of-sample forecasting exercise using different sets of real-time estimates points clearly to a lack of any usefulness of output gap real-time estimates for inflation forecasting both in the short term (one-quarter and one-year ahead) and the medium term (two-year and

three-year ahead). These findings are broadly consistent with the empirical literature which shows that while ex post estimates of the output gap tend to have a relatively good in-sample fit in Phillips curve models, the out-of-sample predictive ability of real-time estimates is very limited. As regards the predictive content for real GDP growth of alternative measures of the euro area output gap, a similar out-of-sample forecasting exercise provides mixed results. In particular, some real-time output gap estimates appear to improve significantly the forecasts for real GDP growth, especially when based on capacity utilization in the short run or averaging in the medium term.

Overall, the findings in this paper cast serious doubts on the usefulness of the output gap for structural analysis or economic policy making in the euro area, while there could be some use for forecasting future real economic activity growth.

I. Introduction

Output gap measures are a key component of a conceptual framework which is very useful for the purposes of conjunctural and monetary policy analysis (see for example ECB, 2000, and Mishkin, 2007). Despite their appealing characteristic as a relatively clear summary measure of overall slack in the economy, output gap estimates are problematic and represent a potentially misleading input in monetary policy analysis. The two main problems in interpreting and assessing the implications of broad summary measures of slack such as the output gap relate to the uncertainty surrounding the corresponding estimates and the uncertain links between these measures and inflation. Although evidence based on real-time data exists for a number of countries, no such evidence exists for the euro area. Against this background, the aim of this paper is to provide updated evidence on the uncertainty characterising euro area output gap real-time estimates, and compare it with the case of another large common currency area, namely, the US.

Recent empirical studies for the US, UK and Canada have shown that the problem of output gap measurement uncertainty is particularly severe for real-time estimates (that is, estimates of the output gap for the period when the actual estimation is carried out), which would typically be those of higher interest for conjunctural and policy analysis (Orphanides and van Norden, 2002, Nelson and Nikolov, 2003, and Cayen and van Norden, 2005). It has even been suggested that the mis-measurement of the output gap in real time may have contributed to wrong economic policy decisions in some countries in the past (see for example Orphanides, 2003, for the US and Nelson and Nikolov, 2004, for the UK).

For the euro area the evidence is more limited. A number of studies have addressed the usefulness of euro area output gap estimates in terms of revisions and inflation forecasting performance. Overall, results appear to vary somewhat across study. On the one hand, Mitchell (2007) finds that both point estimates and measures of uncertainty (density estimates) of euro area output gap estimates are clearly unreliable, and Planas and Rossi (2004) find that estimates of the output gap based on bivariate models (the bivariate Phillips curve-based model of Kuttner, 1994) do not exhibit a higher accuracy –i.e. narrower confidence bands- relative to estimates based on univariate methods (the unobserved components model of Watson, 1986). On the other hand, Camba-Méndez and Rodríguez Palenzuela (2003), Proietti, Musso and Westermann (2007) and Rünstler (2002) find that estimates of the euro area output gap based on multivariate methods (mainly multivariate unobserved components models) do not appear to be as unreliable as those for the US. At the same time, it is difficult to assess the results of these papers in terms of usefulness of the euro area output gap for policy purposes as all of them are based on one specific vintage (the latest available at the time of the study). Since only recently have real-time databases become available for the euro area, previous studies could at most be based on pseudo-real time data. In contrast, the present study uses a genuine set of real time output gap estimates for the euro area, which allows drawing more robust conclusions regarding the reliability of euro area output gap estimates, as well as comparing results with corresponding ones obtained for the US by Orphanides and van Norden (2002) and others.

With respect to the previous literature, as mentioned, we present a fully real time evaluation. In addition, we compare a large set of output gap measures, including simple filter based estimates relying on real GDP, measures based on capacity utilization, estimates based on multivariate unobserved component models, and a variety of estimates from international organizations such as the IMF, OECD and European Commission. In addition, we construct gap measures by averaging those described so far. Averaging is a particular way of pooling, and from the forecasting literature it is well known that pooling, and in particular averaging, a set of forecasts can yield substantial gains in terms of mean square forecast error reduction, see e.g. Stock and Watson (1999). Moreover, averaging can reduce problems of parameter instability and it is also a way to take into account method uncertainty, since there is no uniquely accepted or best method to compute a gap, along the lines of Bayesian model averaging.

Next, we assess the performance of the output gap as a leading indicator for inflation and GDP growth. The results for inflation are in general discouraging, in line with the findings for the US by, e.g., Clark and McCracken (2006). Instead, a few output gap measures do significantly improve short and medium term forecasts of GDP growth. Capacity utilization based gap measures perform particularly well. We believe that this positive finding is related to the fact that the gap works as an error correction term in the model for GDP growth. Hence, while gap measures are quite unreliable as coincident indicators of economic activity, some of them could represent useful leading indicators for the euro area.

The paper is structured as follows. Section 2 describes the real time data and gap measures used. Section 3 reports summary measures of the uncertainty characterising euro area real-time output gap estimates. Section 4 compares the results on uncertainty with those for the US over a comparable sample period. Section 5 provides some evidence on the usefulness of euro area real-time output gap estimates for inflation forecasting, assesses the statistical significance of the gains using recently developed tests statistics that are particularly suited for real-time evaluations, and compares the results with those for the US. Section 6 presents a similar exercise for forecasting real GDP growth. Section 7 summarises the conclusions of our analysis. Additional material and more detailed results are presented in Appendices.

II. Data

It is possible to glean some insight on the degree of uncertainty in genuine real-time estimates and projections of the euro area output gap on the basis of estimates published regularly since 1999 by some major international organisations as well as estimates based on euro area real-time data which has become available only recently. In contrast to previous studies, the evidence reported in this paper is based on euro area output gap estimates for which real time vintages for at least a few years are available.

We consider five different types of output gaps, which are compared to real GDP growth in real time. First, measures based on capacity utilization: the deviations from the average value and from a linear trend. Since capacity utilization figures are not revised, changes in the real time vintages are only due to recursive estimation of the mean of the variable, and of the slope of the linear trend. The data are from the European Commission survey on the manufacturing sector. These measures are used as a driving force of the cyclical component of the variables included in some of the more complex output gap models described below. They are included in the analysis as it might be interesting to assess whether it makes a significant difference to use more complex output gap estimates (whether or not based also on these capacity utilization measures) relative to using only these simple measures of slack.

Second, estimates computed on the basis of the multivariate unobserved components (UC) model of Proietti, Musso and Westermann (2007), which combines a production function and a Phillips curve equation. We consider three alternative versions: the common cycle (“CC”) one, where all cyclical components are driven by the cycle in capacity utilisation; the pseudo-integrated cycles ("PIC") one, where all cyclical components are driven also by idiosyncratic cycles; and the bivariate version ("BIV"), where the Kalman filter is applied directly to output rather than to the components of the production function. Appendix I reports some details on the alternative specifications of the UC model used, see Proietti, Musso and Westermann (2007) for additional details. An advantage of these types of measures of output gap is that it is possible to construct and provide confidence intervals around the point estimates.

Third, measures provided by international organizations. These include annual estimates published twice a year by the European Commission (in the context of their annual Spring and Autumn forecasts), the IMF (in the context of the annual Spring and Autumn World Economic Outlook) and the OECD (in the context of the annual June and December OECD Economic Outlook). Note that the EC has two sets of estimates, one based on deviations from a trend derived by applying the HP filter to each euro area country series and then aggregating the result ("EC-T"), and another representing deviations from a trend estimated within a production function approach ("EC-P"), which was started in 2002. The IMF and the OECD gap measures are also based on a production function approach.

Fourth, measures obtained by applying standard filters to the real GDP levels. In particular, we consider the HP filter ("HP"), the Baxter and King (1999) band-pass filter ("BP"), and deviations from a linear trend ("LIN"). In order to reduce the impact of the so-called end-of-sample bias we extend each vintage of real GDP data via a simple AR(4) model (applied to the year-on-year growth rate), apply the filters to the extended levels and finally, as suggested by Baxter and King (1999), we disregard the last three years of filtered data. For the HP filter we use a smoothing coefficient (lambda) of 1600, as was suggested by Hodrick and Prescott (1997) for quarterly data and as is typically done in the literature, while for the band-pass filter we use the cut-off frequencies suggested by Baxter and King (1989), i.e. we keep only the components of the data between the cut-off frequencies between 1.5 and 8 years. Notwithstanding the well-known problems with these filters, they are still fairly common in empirical applications, see e.g. Watson (2007) for a critical review. In our context, they are convenient to isolate the effects of two sources of changes in output gap vintages: recursive estimation and changes in the vintages of real GDP. In particular, we can compute pseudo-real time gaps using recursively the final vintage of real GDP data, in addition to truly real time gaps that are recursively based on the real time vintages of real GDP data. The difference between these two types of gaps is purely due to changes in real time vintages of real GDP.

Fifth, we construct gap measures by averaging some of those in groups 1-4. Averaging is a particular way of pooling, and from the forecasting literature it is well known that pooling, and in particular averaging, a set of forecasts can yield substantial gains in terms of mean square forecast error reduction, see e.g. Stock and Watson (1999). Moreover, averaging is also a way to take into account method uncertainty, since there is no uniquely accepted or best method to compute a gap, along the lines of Bayesian model averaging. We consider five averages: of all gaps in groups 1-4 ("Average All"), of those belonging to the production function approach ("Average PFA", including CC, PIC, EC-P, IMF and OECD), of those from international organizations ("Average Org", including EC-T, EC-P, IMF and OECD), of those from the UC models ("Average UC", including CC, PIC and BIV) and of those from the standard filters ("Average Filters", including HP, BP and LIN).

It is also worth mentioning that, in order to construct a set of quarterly vintages of quarterly estimates, the following steps were undertaken when needed:

- For those vintages for which data before 1991 was not available, estimates were extended backwards using (the changes in) the previously available historical vintage from the same source, or the closest subsequently available historical vintage if previous vintages also lacked historical data.
- Annual data were interpolated to derive quarterly series. We compared alternative approaches, which produced similar results likely because few data points are interpolated and the source data is fairly smooth. In the end, we fitted a local quadratic polynomial for each observation of the annual series, and then used this polynomial to fill in all observations of the quarterly series associated with the period. The quadratic polynomial is formed by taking sets of three adjacent points from the source series (two for end-points) and fitting a quadratic so that the average of the quarterly points matches the annual data actually observed. [4]
- To construct the quarterly database, the latest available biannual vintage was used to represent the quarterly vintage. Thus, for example, the IMF Spring estimates of 2003 (which became available in April 2003) were used to represent the 2003Q2 and 2003Q3 vintages, while the Autumn estimates of 2003 (which became available in October 2003) were used to represent the 2003Q4 and 2004Q1 vintages.

Table 1 summarises the characteristics of the output gap estimates used in the paper. Overall, 19 to 34 vintages are available, depending on the set of estimates. Appendix II shows all vintages of all estimates used.

Table 1 – Vintages of euro area output gap estimates

Data and estimates*	Definition of trend	Sample period**	Frequency***	Vintages		Source
Real GDP		1985Q1-2006Q4	quarterly data	2001Q1-2007Q2	(26)	EABCN
Capacity util. rate	Average	1985Q1-2006Q4	quarterly data	2001Q1-2007Q2	(26)	European Commission
Capacity util. rate	Linear trend	1985Q1-2006Q4	quarterly data	2001Q1-2007Q2	(26)	European Commission
UC - CC	Prod Fn Approach	1985Q1-2006Q4	quarterly data	2002Q3-2007Q2	(20)	ECB
UC - PIC	Prod Fn Approach	1985Q1-2006Q4	quarterly data	2002Q3-2007Q2	(20)	ECB
UC - BIV	Bivariate model	1985Q1-2006Q4	quarterly data	2002Q4-2007Q2	(19)	ECB
EC - Trend	HP trend	1985Q1-2006Q4	annual data	1999Q1-2007Q2	(34)	European Commission
EC - Potential	Prod Fn Approach	1985Q1-2006Q4	annual data	2002Q4-2007Q2	(19)	European Commission
IMF	Prod Fn Approach	1985Q1-2006Q4	annual data	1999Q1-2007Q2	(34)	IMF
OECD	Prod Fn Approach	1985Q1-2006Q4	annual data	1999Q1-2007Q2	(34)	OECD
Band-pass filter	Stochastic trend	1985Q1-2006Q4	quarterly data	2001Q1-2007Q2	(26)	ECB
Hodrick-Prescott filter	Stochastic trend	1985Q1-2006Q4	quarterly data	2001Q1-2007Q2	(26)	ECB
Linear trend filter	Linear trend	1985Q1-2006Q4	quarterly data	2001Q1-2007Q2	(26)	ECB

Source: EABCN, EC, IMF, OECD and own estimates.

Notes: Real GDP data are from the EABCN (see Giannone et al., 2010, for details).

* EC, IMF and OECD publish biannual estimates. To construct the quarterly vintages for each quarter the latest available vintage is used.

** Each vintage available at time T includes data from 1985Q1 to T-2. For those vintages for which no data prior to 1991 was available estimates have been extended backwards using the (changes of the) previously available historical estimate (or if not available the first subsequent estimate).

*** Annual data were interpolation via quadratic match average option of Eviews to derive quarterly estimates.

[4] To evaluate the expected size of the interpolation error, we have aggregated the last vintage of the quarterly CC gaps to annual data, and applied the interpolation method described in the text to obtain interpolated quarterly values of CC. The correlation between the actual and interpolated values of CC is higher than 0.98. Linear or cubic interpolation resulted in correlation values around 0.90.

III. Uncertainty characterising euro area real-time output gap estimates

In this Section we provide a thorough evaluation of the uncertainty characterising euro area output gap estimates, which stems from various sources. In the first subsection we focus on model uncertainty. In the second subsection on parameter estimation uncertainty. In the third subsection on parameter instability. And in the final subsection on the role of data revisions.

III. 1. Model uncertainty

A basic problem in the estimation of the output gap is that several alternative methods have been proposed to estimate it, each with its own advantages and disadvantages, but there is no broad consensus on which approach should be adopted. Moreover, different methods tend to produce significantly different estimates (this source of output gap uncertainty is thus sometimes called "model uncertainty").

Table 2 summarises the main features of the slack measures considered with reference to the final estimate (we take as final estimate the last vintage available in our data set; needless to say, these estimates are likely to be further revised but we follow the convention of using the last available vintage as the closest approximation to what can be thought of as final estimates). All measures exhibit some similarity, notably a high degree of persistence, as indicated by values of the first order autocorrelation index between 0.89 and 0.98. However, clear differences can also be detected. For example, the mean of these estimates, which apart from GDP growth could be expected to be zero, is clearly significantly different from zero in some cases. In part this is due to the fact that we report the mean for the period 1985-2006, while in some cases data is available for a longer period and in other cases the latest estimates are available for a shorter period and had to be extended backwards with previous vintages as explained in the previous section. However, this could also be taken as an indication that some measures may provide less appropriate estimates of the output gap, as for example in the case of deviations of real GDP from a linear trend (given the likely stochastic nature of the underlying trend). Accordingly, also the variability of these estimates tends to differ somewhat, with standard deviation measures in some cases being twice as large as in other cases. This is of course in part related to the different mean of the series. Moreover, the range of fluctuations appears to differ significantly across estimates.

Table 2 – Euro area output gap summary statistics

	mean	st dev	min	max	AR
GDP growth	2.28	1.25	-1.79	4.66	0.89
Cap. util. rate (dev. av.)	0.61	1.86	-4.60	4.20	0.93
Cap. util. rate (dev. lin. trend)	0.34	1.99	-4.60	4.62	0.94
UC-CC	0.04	0.91	-2.33	1.96	0.95
UC-PIC	-0.18	1.15	-2.69	2.45	0.92
UC-BIV	-0.32	1.59	-3.51	2.67	0.95
EC (dev. from trend)	-0.01	1.48	-2.00	2.78	0.98
EC (dev. from potential)	-0.25	1.43	-2.63	2.17	0.98
IMF	-0.20	1.31	-1.93	2.38	0.97
OECD	-0.53	1.59	-2.70	2.99	0.98
Band-Pass Filter	-0.03	0.80	-1.65	1.54	0.95
HP Filter	-0.04	0.89	-1.91	2.04	0.89
Linear Filter	0.26	1.82	-2.69	4.14	0.97
Average All	-0.03	1.27	-2.21	2.70	0.96
Average PFA	-0.22	1.20	-2.09	2.29	0.97
Average UC	-0.15	1.11	-2.11	2.32	0.94
Average Org	-0.25	1.44	-2.29	2.49	0.98
Average Filters	0.06	1.12	-1.69	2.54	0.95

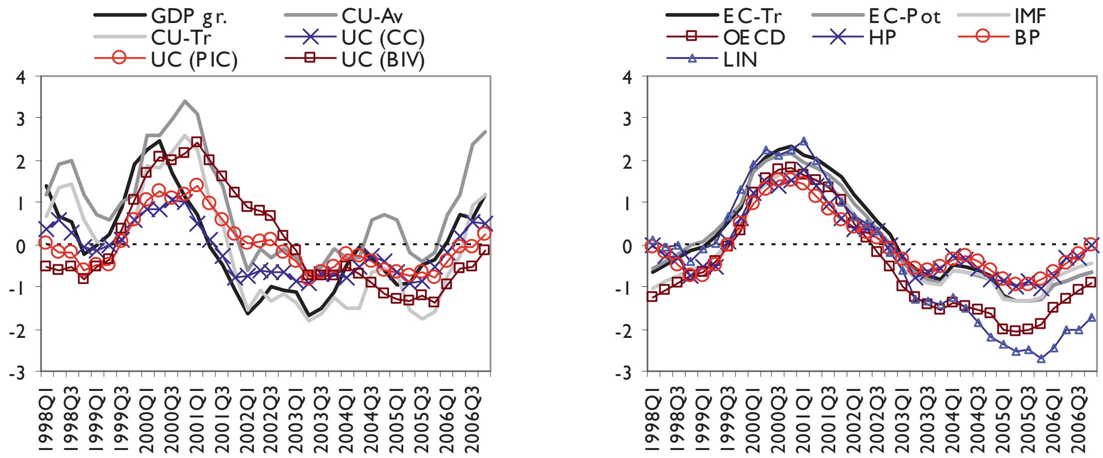
Notes: Sample period is 1985:4 to 2006:4 in all cases. "AR" refers to the first order autocorrelation coefficient.

Differences in estimates can also be significant with regards to specific point estimates (see Chart 1, based on latest vintage data). It is far from rare that some estimates point to a positive output gap in a specific quarter or year, while other estimates point to a negative one. This seems to be the case for both final estimates (Chart 1) and for real time estimates (Chart 2). For example, among the output gap estimates (i.e. those based on the UC and filters and from the EC, IMF and OECD), the average difference between the maximum and the minimum of final estimates from 1998 to 2006 is 1.5 percentage points, with a peak (found in 2006Q3) of 2.5 percentage points.[5] Over the same period, the corresponding average range for real time estimates was 1.6 percentage points, with a peak (found in 2004Q3 and 2004Q4) of 2.6 percentage points.

Moreover, in 42% of the cases for the final estimates from 1998 to 2006 (and 44% from 1985 to 2006) the minimum and the maximum have different signs. For the real time estimates from 1998 to 2006 a different sign is found in 24% of the cases. It should be recognised that the variation across estimates also derives from the different sets of projections for the data used to estimate the gap across institutions, and therefore model uncertainty is not the only source of variation.

Chart 1: Final estimates of euro area output gap and other slack indicators

(percentage deviations from trend/potential output/average)

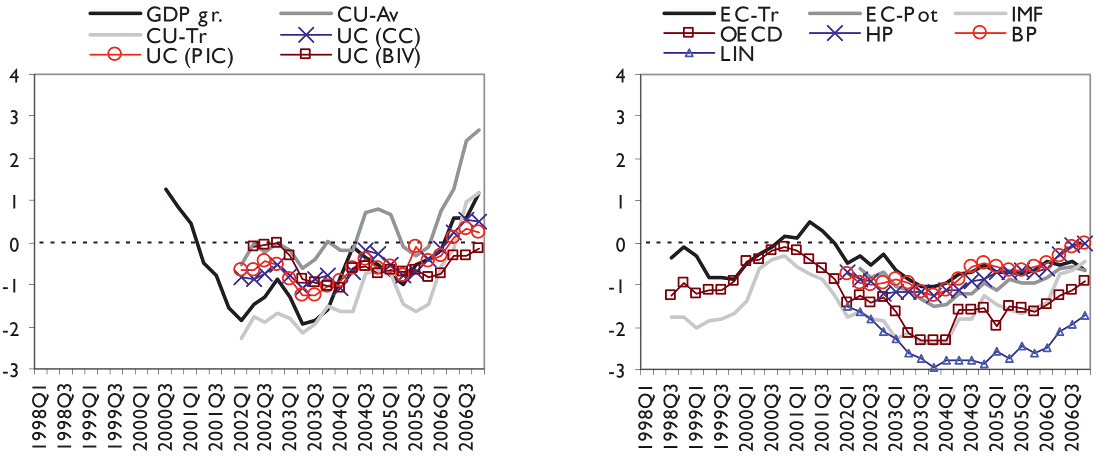


Sources: EABCN, European Commission, IMF, OECD and own estimates.
Note: UC: Estimates from the multivariate unobserved components model of Proietti, Musso and Westermann (2007). The versions of the UC model shown are the common cycles variant (CC), the pseudo-integrated cycles variant (PIC) and the bivariate (BIV) variant respectively.

[5] The range from 1985 to2006 was 1.9 ppt, with a peak of 4.0 ppt in 1992Q1.

Chart 2: Real time estimates of euro area output gap and other slack indicators

(percentage deviations from trend/potential output/average)

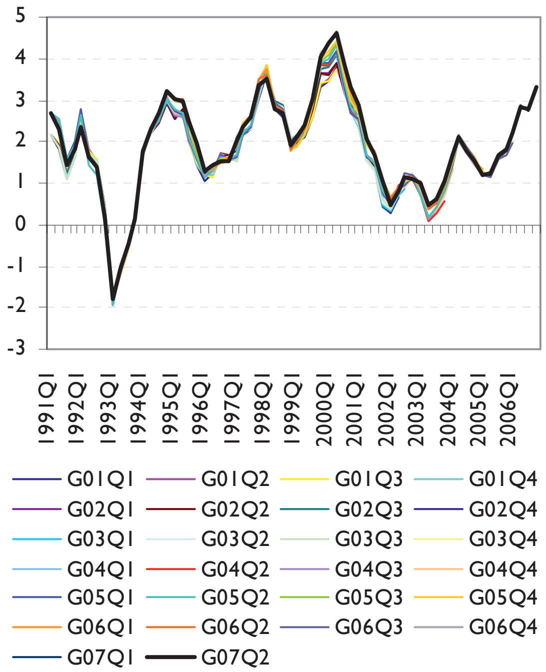


Sources: EABCN, European Commission, IMF, OECD and own estimates.
Note: UC: Estimates from the multivariate unobserved components model of Proietti, Musso and Westermann (2007). The versions of the UC model shown are the common cycles variant (CC), the pseudo-integrated cycles variant (PIC) and the bivariate (BIV) variant respectively.

It can be noticed that uncertainty in output gap estimates tends to be more significant than uncertainty characterising real GDP growth. Although revisions in real GDP growth are occasionally non-negligible (Chart 3) revisions in output gap estimates tend to be clearly more marked (see for example Chart 4 and Appendix II for more examples).

Chart 3: Vintages of euro area real GDP growth

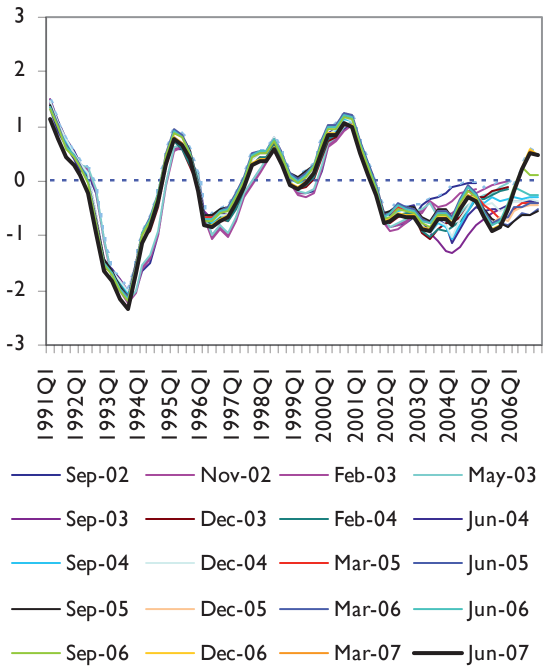
(percentages, year-on-year growth)



Sources: EABCN.

Chart 4: Vintages of euro area output gap estimates by the IMF

(percentage deviations from potential output)



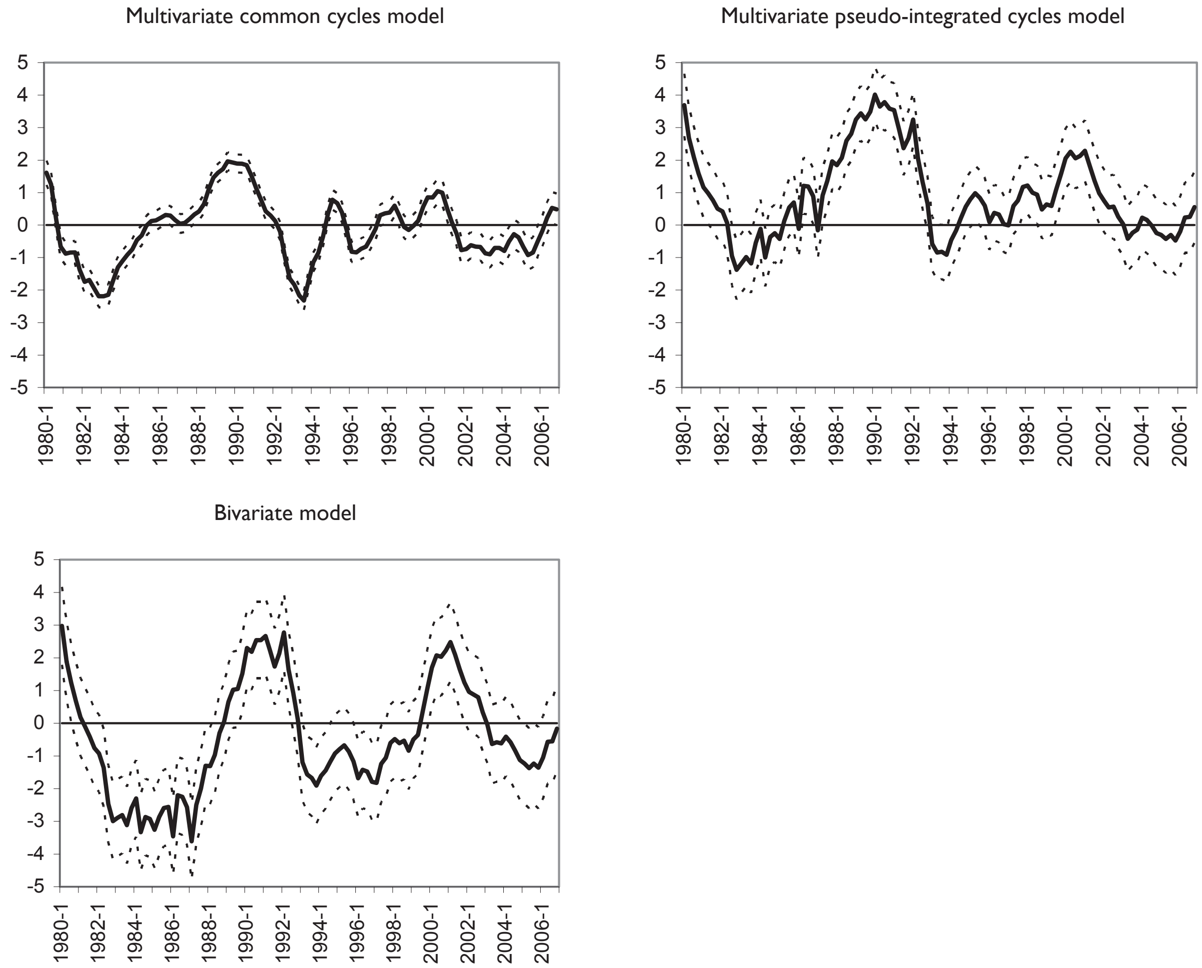
Sources: IMF.

III. 2. Parameter uncertainty

Another source of uncertainty results from the fact that each method requires the estimation of one or more parameters, which are unobserved and may change over time, for example as a result of structural change. Given the limitations of available estimation techniques and the relatively short sample periods available for many variables, parameters tend to be estimated with a significant degree of uncertainty (this source of uncertainty is associated with what is often called "parameter uncertainty"). One way to assess uncertainty which, to some extent, can be associated to parameter uncertainty is by computing and examining confidence bands around point estimates. These are typically not published (and therefore are not available for the estimates from the international organisations). An idea of the magnitude of parameter uncertainty for euro area output gap estimates can be derived from confidence bands of UC estimates, computed as plus and minus twice the standard errors, as shown in Chart 5. For example, for the multivariate UC models, although the width of the confidence bands tends to vary over time and across estimates (with an average between 1980 and 2006 of 0.7pp for the common cycles model and 1.8pp for the pseudo-integrated cycles model), it tends to be particularly high around turning points, for real-time estimates, which are precisely those of highest interest from a policy perspective.

Chart 5: Estimate of euro area output gap and corresponding confidence bands according to a multivariate unobserved components model

(percentage deviations from potential output)



Sources: Own calculations.

Note: Estimates from the multivariate unobserved components model of Proietti, Musso and Westermann (2007). Confidence bands are computed as +/- two standard error.

III. 3. Parameter instability

A third source of uncertainty about output gap measures is represented by parameter instability. To assess its relevance for the reliability of gap measures, we have computed recursively but using the final vintage of data the filter based gaps (namely, HP, Band-pass and linear), what is typically called a pseudo real time evaluation. We have also computed recursively the capacity utilization based gaps. Notice that since this variable is not subject to revisions, the pseudo real time and the fully real time gaps coincide. For the UC model based gaps, a pseudo real time evaluation would be based on the filtered rather than smoothed estimates using the final vintage of data, but unfortunately the filtered values are not available. Similarly, pseudo real time values for the gaps produced by international organizations are not available.

Notice that the correlation between the pseudo real time estimate and the final is very high in the case of the capacity-based measures, while it is insignificant in the three filter-based estimates, with the exception of the linear filter (Table 3). The latter result is likely due to the sensitivity of the filter measures to the end of sample observations and the difficulty of correctly forecasting them.

Table 3 – Pseudo real time estimates of the euro area output gap

		mean	st dev	min	max	AR	corr	sign
Cap. util. rate (dev. av.)	Pseudo RT	0.33	0.91	-0.59	2.70	0.85	1.00	95.0%
	Rev FP	-0.09	0.03	-0.11	0.00	0.98	0.92	
Cap. util. rate (dev. lin. trend)	Peudo RT	-1.18	0.96	-2.30	1.20	0.87	0.96	100.0%
	Rev FP	0.18	0.28	-0.14	0.73	0.99	-0.60	
Band-Pass Filter	Pseudo RT	-0.72	0.35	-1.28	0.00	0.89	-0.02	85.0%
	Rev FP	0.34	0.52	-0.28	1.22	0.94	-0.67	
HP Filter	Pseudo RT	-0.83	0.39	-1.41	-0.01	0.90	0.11	85.0%
	Rev FP	0.42	0.56	-0.32	1.36	0.91	-0.61	
Linear Filter	Pseudo RT	-2.39	0.47	-3.03	-1.44	0.86	0.61	85.0%
	Rev FP	0.97	0.84	-0.07	2.19	0.97	0.20	

Notes: Sample period is 2002:1 to 2006:4 in all cases (20 observations).

"AR" refers to the first order autocorrelation coefficient.

"Rev FP" stands for revision final estimate minus (pseudo) real time estimate.

"sign" refers to the percentage of times the pseudo real time estimate has the same sign as the final estimate

"corr" reports the correlation between pseudo real time estimates and final estimate in the "Pseudo RT" row and the correlation between pseudo real time estimates and the revision final estimate minus (pseudo) real time in the "Rev FP" row.

III. 4. Data uncertainty

Real-time estimates of the output gap tend to be revised significantly over time not only for potential parameter instability but also for a variety of reasons related to data uncertainty. In particular it is worth mentioning the lack of data for the most recent period (for which typically some preliminary estimate based on very limited information is used), revisions of published data (which typically is more substantial for the most recent data), end-of-sample instability (i.e., estimates for the end of the sample period tend to vary significantly with the addition of one or few observations, independently of data revisions) and, for estimates conditional on projections of macroeconomic data for the period ahead, revisions in the projections. Since the effects of data uncertainty can differ across the alternative gap measures, as well as those of parameter instability, we now compare the final estimates evaluated in Section 3.1 with fully real time estimates, computed recursively as in subsection 3.3 but using in each quarter the available vintage of data.

To start with, we consider differences between final and real time (yearly) real GDP growth, which provides an indication of the extent of data revisions in the real GDP series that underlies most gap measures. Although quarterly growth rates are most often the reference measure for conjunctural analysis, we focus here on annual growth rates as the latter have a more pronounced cyclical pattern and are therefore typically the reference measure for business cycle analysis and, accordingly, the rest of the paper. We stress again that in the paper "final" refers to the latest available vintage. From Table 4, there are some differences in the mean, standard deviation and range of final and real time values for growth, which suggest positive revisions of initial values. However, the differences are not marked, and the correlation between the final and real time series is about 0.98 (see also Chart 3). As a consequence, the persistence of the series is similar, 0.88 versus 0.85, and they always have the same sign. These results suggest that revisions to the gap measures are not due to major revisions in the underlying real GDP series, at least over the period under analysis, in line with the graphical evidence provided earlier.

Following Mankiw and Shapiro (1986) and, more recently, Aruoba (2008), one might want to consider whether the revision process in the GDP growth rate is better characterized by the "noise" or "news" models. In particular, in the "noise" model preliminary data are thought of as final data subject to a measurement error, while in the "news" model preliminary data are considered as forecasts of final data. Mankiw and Shapiro (1986) suggest that the two hypotheses can be discriminated by regressing either the final values on a constant and the preliminary values (regression a) or, vice versa, the preliminary values on a constant and the final values (regression b). Under the news hypothesis, the coefficient of preliminary values in the regression a should be equal to one, the constant should be equal to zero, and the coefficient of the final values in the regression b should be smaller than one. Similar restrictions, with the proper changes, should hold under the noise hypothesis.

Unfortunately, our evaluation sample is too short for a formal evaluation of this issue. In particular, when the noise or news hypotheses are tested with a robust F-statistic, they are both rejected. Aruoba (2008) explains that this finding can be due to a non-zero revision error. However, when we remove the mean from the revision error, both the news and the noise restrictions are not rejected. Hence, we take a more informal approach and simply report the correlations between the real time and final values with the revision error (the difference of final and real time values). Under the noise hypothesis, the final value should be uncorrelated with the revision, while under the news hypothesis the real time value should be uncorrelated with the revision. From Table 4, the correlation between final and revision (0.04) is smaller in absolute value than that between real time and revision (-0.16), which provides some evidence in favour of the noise hypothesis. Camacho and Perez-Quiros (2008) find similar values and reach a similar conclusion.

For the UC model based gaps, the results reported in Table 4 suggest that CC has the highest correlation between final and real time values, 0.96, followed by BIV, 0.73, and PIC, 0.51. The ranking in terms of percentage of same signs between final and real time is the same, with 100% for CC and the lowest percentage for PIC, 75%, which means than one out of four quarters the sign of the real time gap is later reversed. The largest revisions are instead for BIV, -0.64 and 0.91. In terms of the revision process, it should be considered that the gaps are in general obtained through complicated procedures so that the revision error can be due to a variety of reasons, as mentioned, in addition to revisions in the underlying GDP data. Hence, the applicability of the news or noise models for the gap is questionable. However, it can still be of interest to consider the correlations between final and real time estimates and the revision error, in particular because this can affect the properties of gap based forecasts (as will be discussed in Sections 5 and 6). It turns out that there are large differences across methods in these correlations, with the lowest value in absolute terms for the correlation between the CC final and revision error (-0.14), and the largest value for that between the BIV final and revision error (0.86).

As regards the output gaps produced by EC, IMF and OECD, results are mixed. On the one hand the highest correlation between final and real time estimates is found for the OECD estimates (0.84) and the lowest for the estimates by the EC based on the production function approach (0.29). However, the highest percentage of same signs between final and real time is found for the latter estimate (84%). The correlation between final estimates and revisions tends to be high in all four cases, ranging from 0.78 (IMF) to 0.96 (EC, deviations from trend). By contrast, the correlation between real time estimates and revisions tends to vary significantly, from the lowest in absolute value (-0.04) for the IMF estimates to the highest (0.56) from the OECD. It can also be observed that the range of real time revisions tends to be larger compared to those found for the UC model based gaps.

Estimates based on filters indicate a relatively high percentage of same sign between real time and final (85% in all three cases considered), but the correlation between these two estimates is either very close to zero (BP and HP) or relatively low compared to the other estimates considered above (0.62 for the linear trend deviations). In all cases for the filter based estimates correlation of final and real time estimates with the revision is relatively large, suggesting that it is difficult to classify these estimates. The range of real time revisions for these estimates is relatively high compared to the UC model based estimates but not relative to the estimates by the EC, IMF and OECD.

Results for the estimates based on pooling some or all of the above mentioned estimates, reported in Appendix III, suggest that there does not appear to be any significant improvement compared to the best set of estimates of each group, either in terms of correlation of real time estimates with the final estimates, percentage of same sign or range of revision of real time estimates.

Table 4 – Revisions to real time euro area output gap estimates

		mean	st dev	min	max	AR	corr	sign
GDP	Final	1.75	0.93	0.48	3.82	0.88	0.98	
	RT	1.57	0.94	0.20	3.42	0.85	-0.16	100.0%
	Rev RT	0.18	0.19	-0.18	0.48	0.51	0.04	
UC-CC	Final	-0.49	0.44	-0.92	0.53	0.88	0.96	
	RT	-0.50	0.48	-1.07	0.54	0.85	-0.43	100.0%
	Rev RT	0.01	0.14	-0.15	0.28	0.60	-0.14	
UC-PIC	Final	-0.35	0.33	-0.83	0.24	0.75	0.51	
	RT	-0.53	0.44	-1.27	0.33	0.80	-0.70	75.0%
	Rev RT	0.18	0.39	-0.53	0.71	0.81	0.26	
UC-BIV	Final	-0.60	0.64	-1.36	0.81	0.90	0.73	
	RT	-0.56	0.34	-1.08	0.00	0.75	0.28	89.5%
	Rev RT	-0.04	0.45	-0.64	0.91	0.86	0.86	
EC (dev. from trend)	Final	0.20	1.20	-1.35	2.32	0.97	0.70	
	RT	-0.47	0.38	-1.04	0.52	0.81	0.47	67.6%
	Rev RT	0.67	0.97	-0.67	2.36	0.96	0.96	
EC (dev. from potential)	Final	-0.68	0.54	-1.34	0.65	0.93	0.29	
	RT	-0.97	0.29	-1.50	-0.56	0.79	-0.25	84.2%
	Rev RT	0.29	0.53	-0.44	1.27	0.93	0.85	
IMF	Final	-0.07	0.98	-1.35	1.73	0.96	0.59	
	RT	-1.47	0.61	-2.42	-0.32	0.88	-0.04	61.8%
	Rev RT	1.40	0.79	0.00	2.53	0.94	0.78	
OECD	Final	-0.43	1.27	-2.06	1.84	0.97	0.84	
	RT	-1.24	0.61	-2.34	-0.10	0.90	0.56	67.6%
	Rev RT	0.81	0.82	-0.56	1.95	0.95	0.92	
Band-Pass Filter	Final	-0.38	0.39	-0.93	0.42	0.89	-0.07	
	RT	-0.71	0.34	-1.25	0.00	0.91	-0.69	85.0%
	Rev RT	0.33	0.54	-0.29	1.26	0.94	0.77	
HP Filter	Final	-0.41	0.45	-1.02	0.37	0.84	0.05	
	RT	-0.80	0.37	-1.24	-0.01	0.92	-0.61	85.0%
	Rev RT	0.39	0.57	-0.32	1.27	0.92	0.76	
Linear Filter	Final	-1.42	1.05	-2.69	0.66	0.97	0.62	
	RT	-2.37	0.46	-2.94	-1.50	0.87	0.23	85.0%
	Rev RT	0.95	0.84	-0.08	2.16	0.98	0.91	

Notes: Sample period is 2002:1 to 2006:4 in all cases (20 observations).

"AR" refers to the first order autocorrelation coefficient.

"Rev RT" stands for revision final estimate minus real time estimate.

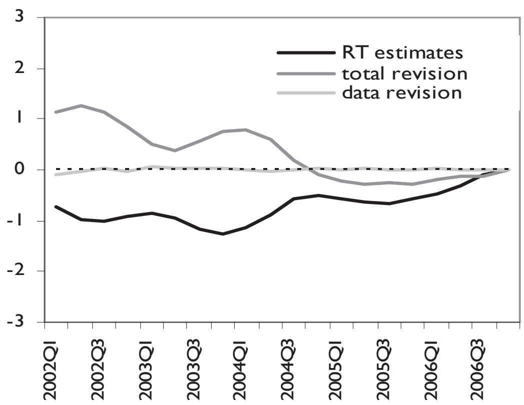
"sign" refers to the percentage of times the real time estimate has the same sign as the final estimate

"corr" reports the correlation between real time estimate and final estimate in the "Final" row, the correlation between real time estimate and the revision (final minus real time) in the "RT" row, and the correlation between final estimate and the revision (final minus real time) in the "Rev RT" row.

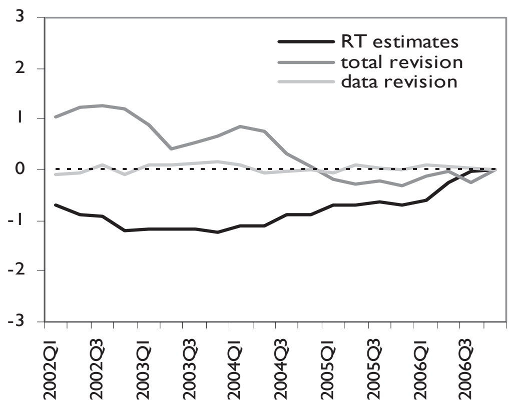
In order to gain some insight into the sources of the revisions in real time estimates it can be useful to undertake a decomposition suggested by Orphanides and van Norden (2002), based on a comparison of genuine real time estimates with pseudo real time estimates, allowing to assess the relative role of parameter instability and data uncertainty. Using our dataset this can be implemented for the three sets of estimates based on filters. The impact of data revision can be assessed by observing the difference between genuine real time estimates and pseudo real time estimates. As suggested by Chart 6, in all cases the contribution to the total revision of data revision is clearly minor over the sample period considered. These results stand in contrast with those for the US reported by Orphanides and van Norden (2002), according to which data revisions are not the major source of revision but appear to play a more significant role compared to what appears to be the case for the euro area. This result could be partly explained by the different sample size and evaluation period, but we will see in the next Section that we find it even over a comparable sample. Thus, for the euro area it appears that the main source of the total revision in real time estimates is represented by the addition of new data points to the data sample over time, rather than revisions to historical data. The latter would appear slightly more important when measured on a quarter on quarter basis.

Chart 6: Real time estimates of euro area output gap, total revision and data revision *(percentage points)*

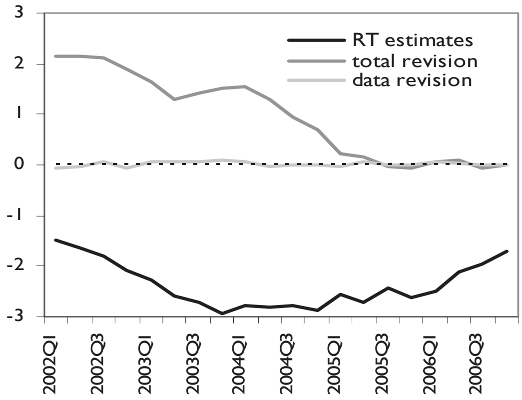
Band-pass filter estimates



HP filter estimates



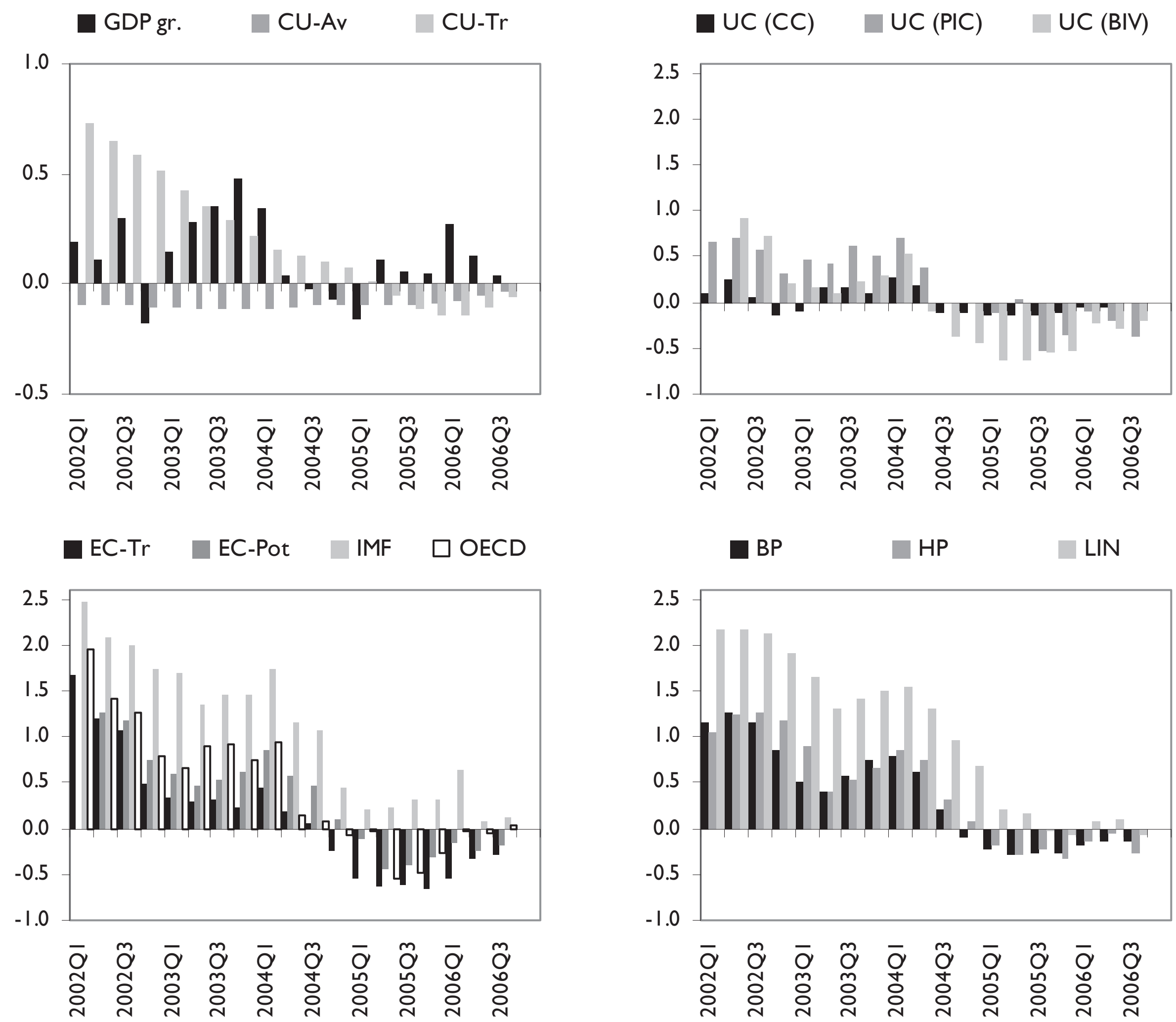
Linear trend filter estimates



Sources: Our calculations.

The revisions of the euro area output gap real time estimates tend to be significant (see Chart 7). Revisions are often of the same (or even higher) magnitude as the estimated gap itself. This appears to be the case particularly for some estimates, such as those by the IMF and those based on the linear trend filter. By contrast, revision of real time estimates based on the UC model appear be more limited, especially those of the CC version. These revisions seem to be larger for the less recent years, but it should be kept in mind that the latest estimates are subject to further changes, reflecting the above-mentioned problem of end-of-sample instability.

Chart 7: Revisions to real-time estimates of euro area output gap
(differences between latest and real-time estimate)



Sources: European Commission, IMF, OECD and own calculations.
Note: EC (P): deviations from potential (available only starting from 2002). EC (T): deviations from Hodrick-Prescott trend. Although capacity utilisation rate data is not revised, revisions to real time estimates of deviations of capacity from average or trend relate to the fact that the average and linear trend change as more data are used to compute them.

In summary, five main results emerge from our analysis of real time measures of the output gap in the euro area. First, there are substantial changes in different vintages of gap data referred to the same quarter; sometimes even the sign of the gap changes, and the size of the revision can be larger than the original value of the gap itself. Second, changes in the vintages of the time series underlying the gap (e.g., real GDP) explain a minor part of the changes in the gap. Third, changes in the vintages of the gap are mostly due to the recursive computation, which suggests either the need of a very long estimation sample or, more likely, the presence of parameter changes. Fourth, averaging different gap measures does not yield any substantial gains. This finding is likely due to the rather high correlation across alternative gap measures. Finally, the UC based gap measures appear to be less subject to revisions over time. However, when confidence bands are computed for the UC based gaps, they are fairly large, in particular around turning points, when precise measurement would be need. This problem is just hidden in the other gap measures, for which confidence bands are either not available or not reported.

IV. Uncertainty: A Comparison with the US experience

In this Section we study the uncertainty characterising US output gap estimates, also in comparison to the euro area. After a short description of the US data, we consider, in turn, the role of model uncertainty, parameter instability, and data uncertainty

IV. 1 Data

For the sake of clarity, in the case of the US we focus on the three filter based estimates of output gaps, namely, the HP filter ("HP"), the Baxter and King (1999) band-pass filter ("BP"), and deviations from a linear trend ("LIN"). The filters are computed using the same specification choices as for the euro area.

Real GDP data from the Federal Reserve Bank of Philadelphia's Real Time Data Set for Macroeconomists (RTDSM) is used. In order to construct a complete set of quarterly vintages of quarterly estimates, for those vintages for which data before 1959 was not available (all those of 1992 and 1996 as well as those of 1997Q1 and 1999Q4 and 2000Q1), estimates were extended backwards using (the changes in) the previously available historical vintage from the same source.

Table 5 summarises the characteristics of the US output gap estimates used in the paper. Overall, 166 vintages are available, depending on the set of estimates. Appendix IV shows all vintages of all estimates used. The availability of so many vintages makes the analysis interesting since we can also assess the effects of the so-called Great Moderation with a longer post 1985 sample, and evaluate whether there have been any substantial changes after the exhaustive analysis of Orphanides and van Norden (2002) whose data stop in 1997.

Table 5 – Vintages of US output gap estimates

Data and estimates	Definition of trend	Sample period	Frequency	Vintages	
Real GDP		1947Q1-2006Q4	quarterly data	1965Q4-2007Q1	(166)
Band-pass filter	Stochastic trend	1985Q1-2006Q4	quarterly data	1965Q4-2007Q1	(166)
Hodrick-Prescott filter	Stochastic trend	1985Q1-2006Q4	quarterly data	1965Q4-2007Q1	(166)
Linear trend filter	Linear trend	1985Q1-2006Q4	quarterly data	1965Q4-2007Q1	(166)

Source: RTDSM and own calculations.

Notes: Real GDP data from the Federal Reserve Bank of Philadelphia's Real Time Data Set for Macroeconomists (RTDSM).

IV.2 Uncertainty measures

IV. 2.1. Model uncertainty

In the absence of a well defined series of actual values, it is difficult to make an a priori choice on the best estimation method for the output gap. In addition, alternative methods tend to produce significantly different estimates of the gap, even within the same class of procedures. We have seen that this is a relevant problem for the euro area, and we now evaluate whether "model uncertainty" matter for filter-based gap estimates for the US.

Table 6 summarises the main features of the slack measures considered, with reference to the final estimate, and Chart 8 reports their temporal evolution. Several comments can be made. First, the post 1985 results are fairly different from the full sample results. In particular, and as expected, the lower volatility of GDP growth associated with the so-called Great Moderation is also reflected in lower volatility of all the gap measures under consideration. Second, the post 1985 results are fairly similar to those for the euro area, in terms of both volatility and persistence of the gap measures. Third, the post 2002 results indicate a further reduction in volatility and persistence of the gap measure. We will come back to this issue in the real time evaluation later on. Finally, as for the euro area, the revisions in output gap estimates tend to be larger than those in real GDP growth (see Chart 9 and Appendix IV for more examples).

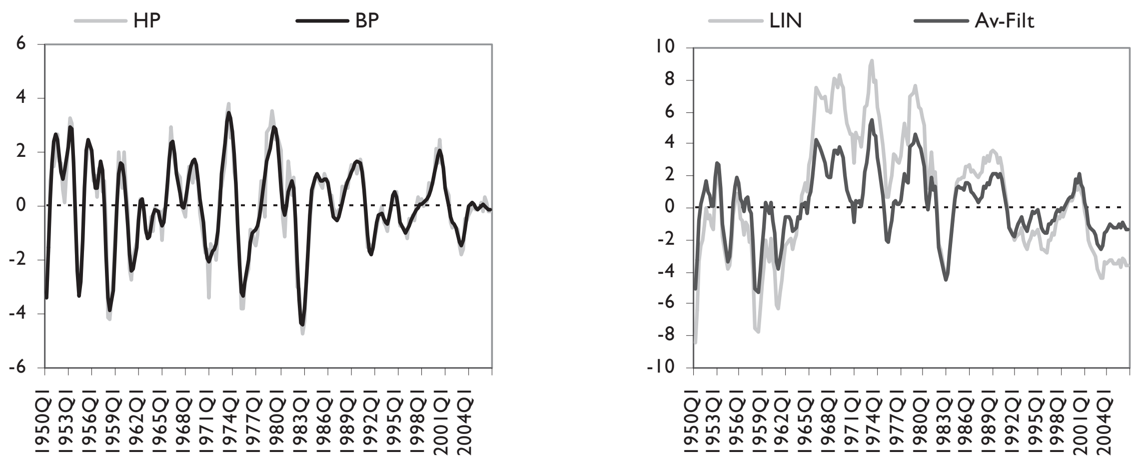
Table 6 – US output gap summary statistics

	mean	st dev	min	max	AR
whole sample period (1947-2006)					
GDP growth	3.23	2.22	-2.71	8.51	0.86
Band-Pass Filter	0.08	1.44	-4.37	3.45	0.93
HP Filter	0.05	1.55	-4.75	3.80	0.87
Linear Filter	1.62	3.70	-4.52	9.22	0.97
Average Filters	0.58	1.97	-4.55	5.49	0.94
1985-2006					
GDP growth	3.13	1.31	-1.00	4.85	0.87
Band-Pass Filter	0.06	0.88	-1.82	2.04	0.94
HP Filter	0.07	0.92	-1.80	2.44	0.88
Linear Filter	-0.63	2.38	-4.36	3.64	0.98
Average Filters	-0.17	1.29	-2.53	2.16	0.96
2002-2006					
GDP growth	2.93	0.98	1.03	4.49	0.83
Band-Pass Filter	-0.41	0.53	-1.44	0.12	0.93
HP Filter	-0.45	0.60	-1.80	0.33	0.81
Linear Filter	-3.44	0.43	-4.36	-2.54	0.57
Average Filters	-1.43	0.45	-2.53	-0.93	0.80

Notes: "AR" refers to the first order autocorrelation coefficient.

Chart 8: Final estimates of US output gap

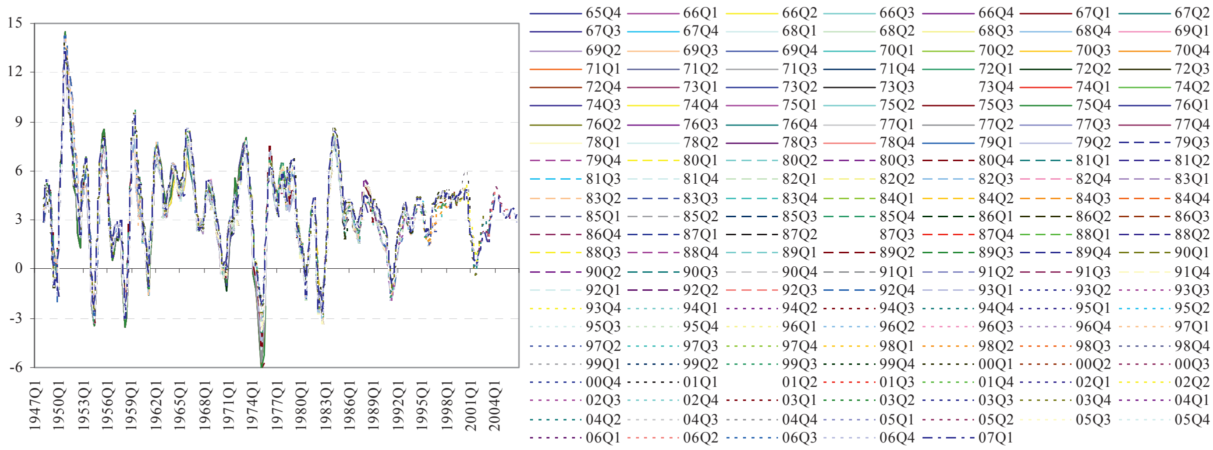
(percentage deviations from trend/potential output/average)



Sources: RTDSM and own calculations.

Chart 9: Vintages of US real GDP growth

(percentages, year-on-year growth)



Sources: RTDSM.

IV. 2.2. Parameter instability

For the euro area, recursive computation of the gap measures (with the final data vintage) revealed a substantial instability in the results, with very low correlation between the pseudo real time and the final gap estimates. To assess whether this is the case also for the US, we have computed the filter based gaps recursively over the period 2002-2006, and Table 7 reports some summary statistics for the alternative measures.

It turns out that the correlation between the pseudo real time estimates and the final is much higher than for the euro area. Moreover, with respect to the final vintage results, there is slightly more volatility and persistence for the HP and BP based measures, less in the case of the linear filter based gap. These findings suggest that also in the most recent period recursive calculation of the US gap is a source of changes in its magnitude and sometimes even in its sign, but that the problem is smaller compared to the euro area.

Table 7 – Pseudo real time estimates of the US output gap

		mean	st dev	min	max	AR	corr	sign
Band-Pass Filter	Peudo RT	-0.54	0.67	-1.68	0.16	0.96	0.91	75.0%
	Rev FP	0.13	0.29	-0.11	0.98	0.85	-0.65	
HP Filter	Peudo RT	-0.63	0.77	-1.84	0.17	0.94	0.89	90.0%
	Rev FP	0.17	0.36	-0.22	1.08	0.75	-0.65	
Linear Filter	Peudo RT	-4.09	0.56	-5.30	-3.36	0.80	0.53	100.0%
	Rev FP	0.66	0.49	0.00	1.55	0.96	-0.68	

Notes: Sample period is 2002:1 to 2006:4 in all cases (20 observations).

"AR" refers to the first order autocorrelation coefficient.

"Rev FP" stands for revision final estimate minus (pseudo) real time estimate.

"sign" refers to the percentage of times the pseudo real time estimate has the same sign as the final estimate

"corr" reports the correlation between pseudo real time estimates and final estimate in the "Pseudo RT" row and the correlation between pseudo real time estimates and the revision final estimate minus (pseudo) real time in the "Rev FP" row.

IV. 2.3. Data uncertainty

We now compare the final estimates with fully real time estimates, computed recursively using in each quarter the available vintage of data.

To start with, we consider differences between final and real time (yearly) real GDP growth, which provides an indication of the extent of data revisions in the GDP series that underlies most gap measures. We stress again that in the paper "final" refers to the latest available vintage.

From Table 8 (and Chart 9), on average real GDP growth is slightly overestimated in real time, and the gap is less negative when based on the BP or HP filters. It is also slightly more volatile and persistent than when computed with the full sample of final data. In addition, the correlations between the real time BP and HP estimates and the revision error are much larger in absolute value than those between the final estimates and the errors, which provides evidence in favour of the news hypothesis. Instead, for the euro area, both corresponding correlations were large, and the results not conclusive.

Table 8 – Revisions to real time US output gap estimates

		mean	st dev	min	max	AR	corr	sign
GDP	Final	2.93	0.98	1.03	4.49	0.83	0.95	
	RT	3.29	0.86	1.58	4.88	0.78	0.21	100.0%
	Rev RT	-0.36	0.33	-0.88	0.22	0.77	0.52	
Band-Pass Filter	Final	-0.41	0.53	-1.44	0.12	0.93	0.87	
	RT	-0.28	0.63	-1.42	0.51	0.94	-0.53	80.0%
	Rev RT	-0.13	0.31	-0.54	0.69	0.77	-0.04	
HP Filter	Final	-0.45	0.60	-1.80	0.33	0.81	0.83	
	RT	-0.35	0.71	-1.71	0.43	0.90	-0.53	75.0%
	Rev RT	-0.10	0.39	-0.76	0.88	0.64	0.03	
Linear Filter	Final	-3.44	0.43	-4.36	-2.54	0.57	0.41	
	RT	-3.58	0.70	-4.93	-2.72	0.75	-0.80	100.0%
	Rev RT	0.15	0.65	-0.73	1.72	0.77	0.22	
Average Filters	Final	-1.43	0.45	-2.53	-0.93	0.80	0.78	
	RT	-1.41	0.64	-2.56	-0.68	0.90	-0.72	100.0%
	Rev RT	-0.03	0.40	-0.49	1.07	0.76	-0.13	

Notes: Sample period is 2002:1 to 2006:4 in all cases (20 observations).

"AR" refers to the first order autocorrelation coefficient.

"Rev RT" stands for revision final estimate minus real time estimate.

"sign" refers to the percentage of times the real time estimate has the same sign as the final estimate

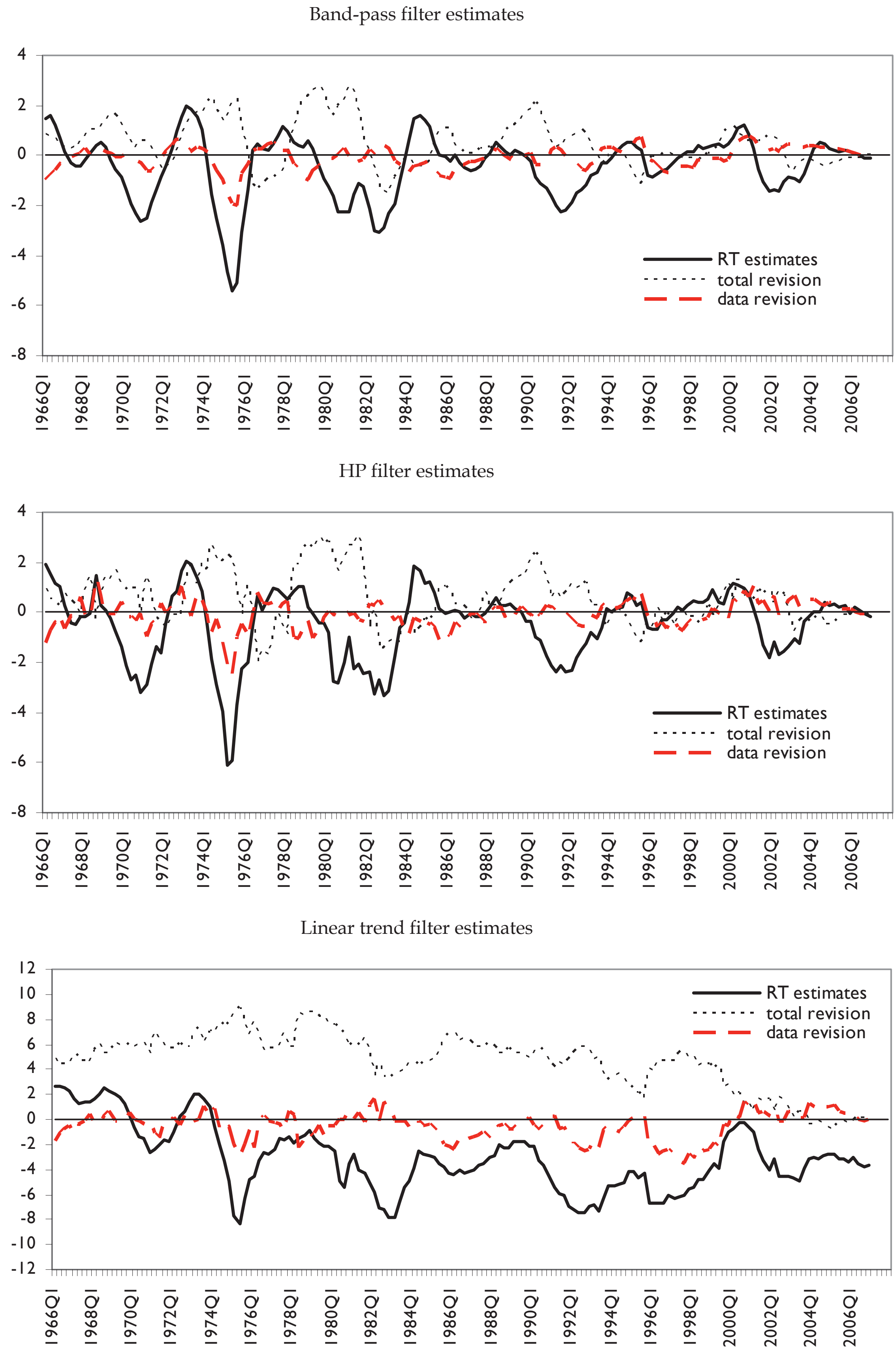
"corr" reports the correlation between real time estimate and final estimate in the "Final" row, the correlation between real time estimate and the revision (final minus real time) in the "RT" row, and the correlation between final estimate and the revision (final minus real time) in the "Rev RT" row.

To disentangle the relative role of recursive computation and real time data, we plot the total revision error and the error purely due to data revisions. Chart 10 presents results for the whole sample, which are useful for comparison with Orphanides and van Norden (2002) who use data up to 1997. From Chart 10, it seems that the total revision error is slightly smaller after 2000, associated with a smaller data revision component. However, such a pattern could change if the post 2000 data will be subject to additional revisions in future releases.

Finally, a direct comparison of the revision process for the US and the euro area is provided in Charts 11 and 12. It turns out that the overall average revision is smaller for the US, and that the real time gap estimates follow more closely the final estimates (a fact which underlies the higher correlation in Table 8). However, the data revision component is larger in the US than in the euro area.

In summary, this Section shows that real time estimates of the US output gap remain unreliable also in the most recent period, even though they are more correlated with final values than for the euro area. In addition, the data revision component of the revision error is larger than for the euro area.

Chart 10: Real time estimates of US output gap, total revision and data revision
(percentage points)

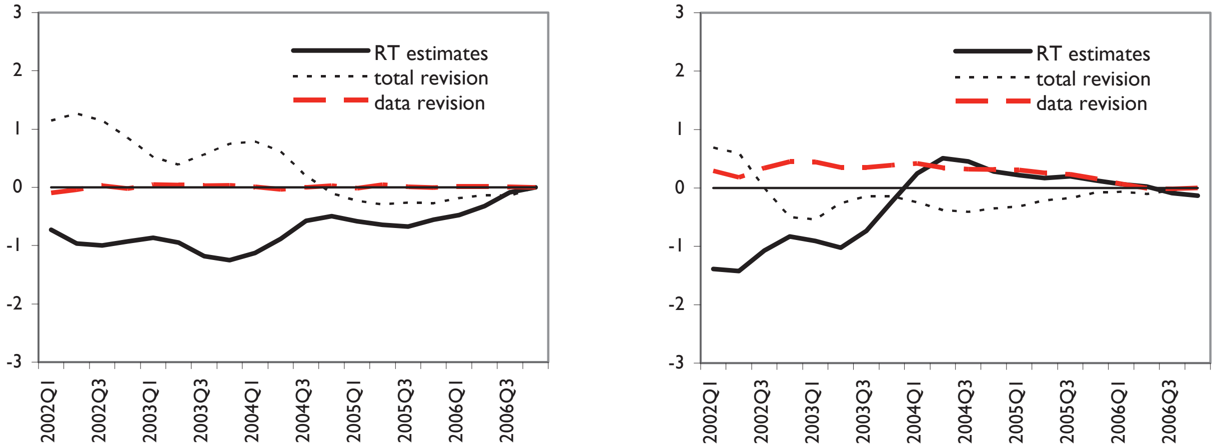


Sources: RTDSM and own calculations.

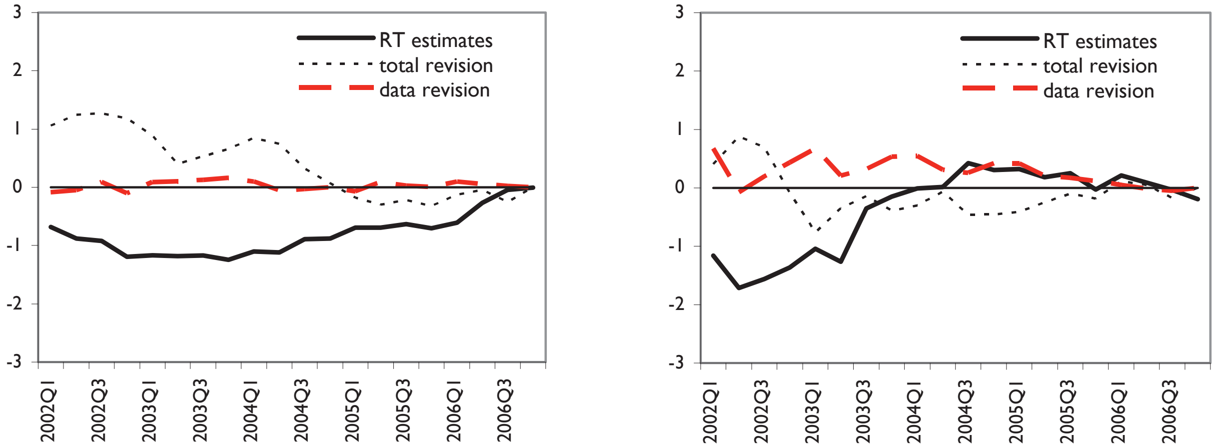
Chart 11: Real time estimates of the output gap, total revision and data revision (2002 onwards)
(percentage points)

Euro Area | US

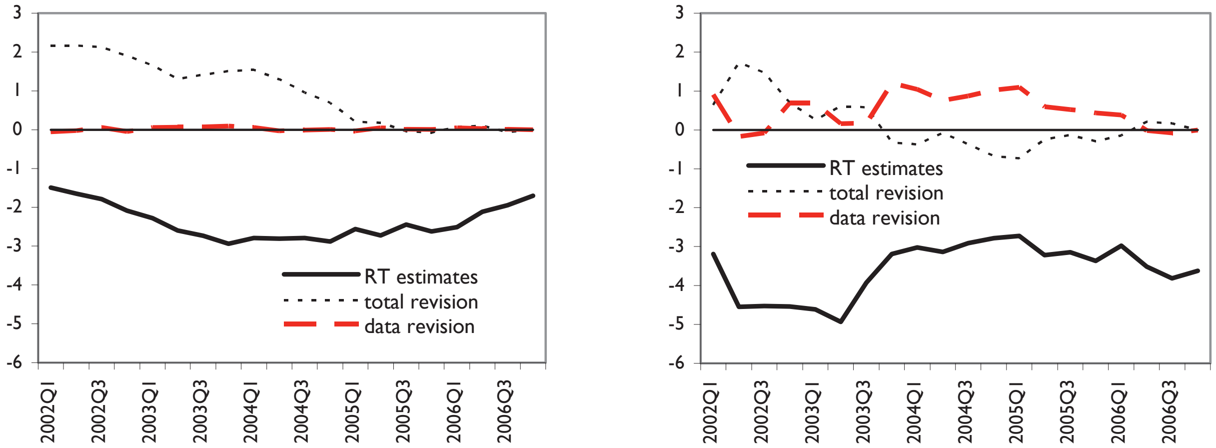
Band-pass filter estimates



HP filter estimates



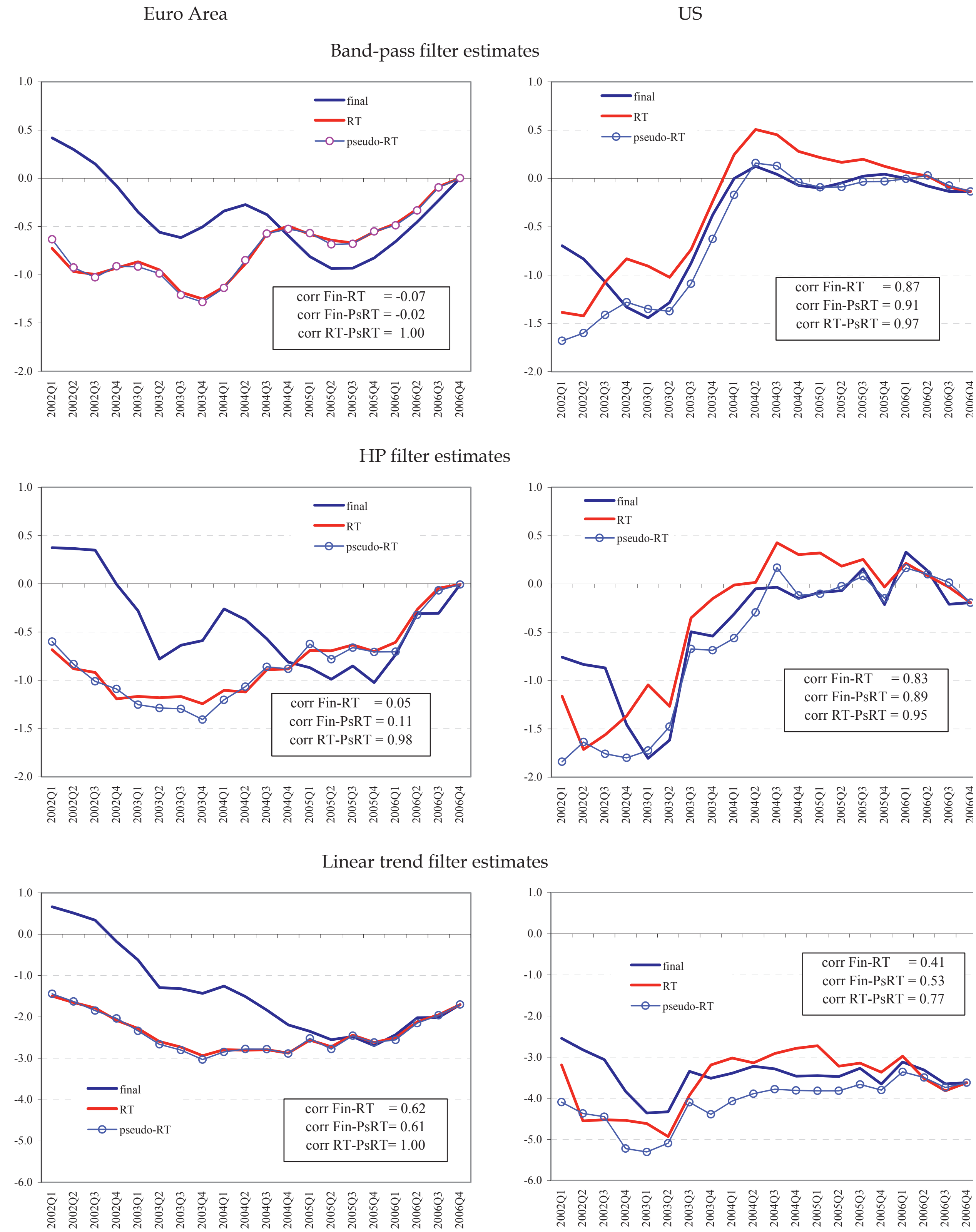
Linear trend filter estimates



Sources: Own calculations.

Chart 12: Final, real time and pseudo real time estimates (2002 onwards)

(percentage points)



Sources: Own calculations.

V. The reliability of inflation forecasts based on output gap estimates in real time

The large empirical literature on the predictive content of the output gap for inflation for the US tends to suggest that models relating inflation to the output gap – typically called Phillips curves – while exhibiting good in sample fits, tend to result in poor out of sample performance, especially in real time (see for example Clark and McCracken, 2006). Against this background, and given the limited informative content of in sample fit analyses, we perform an out of sample assessment in real time of the various output gap measures based on a general benchmark Phillips curve model. As suggested by the empirical literature, starting from Orphanides and van Norden (2002, 2005), it is key to perform such an exercise in real time. Instead of simply comparing the mean squared error (MSE) of the different models, it is important to analyse the statistical significance of the MSE difference, especially in the presence of a relatively limited number of vintages. In this respect it is necessary to bear in mind that standard tests of MSE comparison may be misleading, as they do not take into account the real time nature of the data. In order to evaluate this aspect, we compare results based on standard tests with recent tests proposed by Clark and McCracken (2009) that take into account the real time nature of the data. Next, we perform robustness analyses along some dimensions, such as changing the specification of the forecasting model, the sample period, or the reference series. Finally, we derive comparable results for the US, and assess whether and to what extent the findings are similar to those for the euro area.

We use quarterly data from 1985 onwards, a decision informed by the results of Musso et al (2009) which show that a traditional Phillips curve for the euro area from 1970 onwards is characterised by instability (and some signs of nonlinearity) concentrated in the mid-1980s. Thus, starting from 1985 allows carrying out the analysis with a simple linear Phillips curve. In order to keep the analysis as simple as possible, the inflation measure used will be the GDP deflator. The reason for this choice is that, as indicated by Musso et al (2009), a euro area Phillips curve based on the GDP deflator allows to ignore supply shocks such as oil price or exchange rate changes, which on the contrary appear to play an important role in Phillips curves based on the HICP. An alternative choice would be to use some measure of core inflation, or HICP excluding volatile components. However, apart from the arbitrary choice of such measure among the several available ones (see for example Cristadoro et al., 2005)), the problem is that no real time dataset is available for these measures. Finally, as regards the question of the order of integration of inflation, given the fact that there does not seem to be a widespread consensus on this debated issue (see for example the discussion on this issue with reference to the euro area in the context of the Inflation Persistence Network, as summarised for example by Altissimo et al., 2006), we do not take a stand and follow the approach typically used in the empirical literature of referring to changes in inflation (see, e.g., Stock and Watson, 2003, and Clark and McCracken, 2006).

V. 1. Inflation forecasting assessment: out-of-sample real time evaluation

In this section we carry out an out-of-sample forecasting exercise to assess the predictive content for inflation of alternative measures of the euro area output gap. Note that, in contrast to previous published analyses, we use vintages of real-time output gap estimates produced from 1999 onwards. In order to assess the predictive accuracy of alternative frameworks using these data, the real-time nature of the data needs to be taken into account as well as the possibly nested nature of the alternative models. In order to take into account these features of the exercise, we use the tests proposed recently by Clark and McCracken (2005, 2009), which apply a number of adjustments to standard forecast accuracy tests. The authors show that these adjustments can occasionally imply significantly different indications compared to conventional tests which ignore the real-time nature of the data. Thus, these corrections may be key to reliably assess the predictive content of activity measures for real time inflation forecasting.

Following Stock and Watson (1999, 2003) and Clark and McCracken (2009), we compute forecasts of the change in inflation π_t at horizon τ from reduced-form Phillips curves:

$$\pi_{t+\tau}^{(\tau)} - \pi_t = \alpha + \sum_{k=0}^{3} \beta_k \Delta\pi_{t-k} + \gamma x_t + \delta \Delta x_t + u_{t+\tau} \quad (2)$$

where $\pi_t^{(\tau)} \equiv (400/\tau)\ln(p_t / p_{t-\tau})$, $\pi_t^{(1)} \equiv \pi_t$ and x_t is the output gap (expressed in terms of percentage deviations from trend or potential output).

The benchmark model against which to compare the forecasts of the Phillips curve-based model is an autoregressive (AR) model for inflation:

$$\pi_{t+\tau}^{(\tau)} - \pi_t = \alpha + \sum_{k=0}^{3} \beta_k \Delta\pi_{t-k} + \varepsilon_{t+\tau} \quad (3)$$

which is essentially the same model as (2) but without the output gap measure. We consider four forecast horizons: one quarter ($\tau = 1$), one year ($\tau = 4$), two years ($\tau = 8$), and three years ($\tau = 12$). As discussed, we use quarterly data from 1985 onwards and the GDP deflator to derive the reference inflation measure. For each slack indicator the forecast period covers the sample period for which vintages are available (instead of selecting the same forecast period for all cases, which would result in a loss of several observations, which we prefer to avoid given the already limited sample size). Thus, for example, using the IMF output gap estimates, whose vintages are available from 1999Q1 onwards, for each forecast origin *t* from 1999Q1 onwards we (recursively) estimate the forecast models (2) and (3) with the data that was available in that quarter (reaching up to the previous quarter, *t-1*) and construct forecasts for periods *t* and beyond (for to the four above-mentioned forecast horizons). The starting point of the model estimation sample is always 1985Q1. We then evaluate the forecasts against the latest available vintage of inflation available.

Tables 9 and 10 summarise the results of the exercise comparing the forecast of inflation at different horizons in the short run (one quarter and one year) and the medium run (two and three years) from the AR(4) model and the Phillips curve model with the output gap, based on the mean squared error (MSE).

Tables 11 and 12 report the tests of equal forecast accuracy, based on both the conventional tests and the above-mentioned adjusted tests. In particular, conventional tests to do not consider that model (3) is nested in (2) and that real time data are used (see e.g. Aruoba, 2008), while the corrections proposed by Clark and McCracken (2005, 2009) can handle both features. Specifically, if there were no revisions, one could use the conventional t-test against the critical values simulated in Clark and McCracken (2005). This is called MES-t(conv) in the tables.[6] If instead there were predictable revisions, an adjusted t-statistics (labelled MSE-t (Ω) by Clark and McCracken, 2009), should be compared with normal critical values. However, as shown in the previous section, in our context there are revisions but there is no clear evidence to discriminate the "noise" or "news" hypotheses. Hence, rather than the MSE-t (Ω) statistic we report the MSE-F statistic of Clark and McCracken (2005) that, according to the simulation experiments reported in their later paper, performs well in a variety of situations, including the "news" case. It is reassuring that in most cases the MSE-t (Ω)) and MSE-F statistics provide similar evidence (results available upon request).[7]

As regards the short run, in all cases the MSE of the forecasts based on the AR(4) models are lower than those based on the Phillips curve model, independently on which set of output gaps estimates were used and for both horizons (Table 9). Moreover, in most cases this difference is statistically significant, suggesting that adding the output gap worsens the predictions of inflation (Table 11).

For the medium run results are similar (Table 10). Only in very few cases the MSEs of the forecasts based on the AR(4) models are higher than those based on the Phillips curve model (for the ECT and HP filter cases at both horizons and the band-pass filter case for the three year horizon). However, in all of these few cases the adjusted t-statistics suggest that this difference is not statistically significant (Table 12). Thus, output gaps estimates do not appear to contribute to any significant improvement in forecasting inflation in the medium run. Note that the adjusted statistic provides different indications compared to the conventional statistic in a number of occasions, which suggests that it is important to take into account these adjustments to reduce the probability of deriving misleading results.

There do not appear to be major differences across output gap estimates. For example, it does not appear to be the case that estimates of the EC, IMF and OECD (based on methods which impose some smoothness prior on potential output growth) perform significantly better or worse compared to UC estimates (based on methods which do not impose any smoothness prior). Few minor differences can be detected, as already discussed, but it should be recognised that for the various sets of estimates a different number of vintages is available, implying that results may not be fully comparable.

---

[6] Note that we use two-sided critical values since, while the nesting model should have lower MSE than the nested model, the opposite could also happen, e.g. in the case of parameter instability or marginally significant regressors, see e.g. Clements and Hendry (1999).

[7] We are very grateful to Todd Clark for providing us with a copy of his programmes for his paper Clark and McCracken (2007). All computations for the out-of-sample forecasting exercise have been carried out with WinRats Pro 7.00 (see Estima, 2007).

Table 9 – Inflation forecast accuracy in the short term

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE (AR)	MSE (OG)	diff
		forecast horizon h = 1 quarter			forecast horizon h = 4 quarters		
CAP-AV	2001:01-2006:04	1.570	1.621	-0.051	0.898	0.881	**0.017**
CAP-TR	2001:01-2006:04	1.570	1.626	-0.056	0.898	0.922	-0.024
EC - T	1999:01-2006:04	1.342	1.427	-0.085	0.798	0.903	-0.105
EC - P	2002:04-2006:04	1.447	1.550	-0.103	0.696	0.872	-0.176
IMF	1999:01-2006:04	1.342	1.458	-0.116	0.798	0.965	-0.167
OECD	1999:01-2006:04	1.342	1.472	-0.130	0.798	0.970	-0.172
UC - CC	2002:03-2006:04	1.588	1.784	-0.196	0.751	0.997	-0.246
UC - PIC	2002:03-2006:04	1.588	1.834	-0.246	0.751	0.992	-0.241
UC - BIV	2002:04-2006:04	1.447	1.570	-0.123	0.696	0.803	-0.107
BP	2001:01-2006:04	1.570	1.659	-0.089	0.898	1.009	-0.111
HP	2001:01-2006:04	1.570	1.641	-0.071	0.898	0.961	-0.063
LIN	2001:01-2006:04	1.570	1.623	-0.053	0.898	0.961	-0.063
AV-All	2001:01-2006:04	1.570	1.665	-0.095	0.898	0.951	-0.053
AV-PFA	2001:01-2006:04	1.570	1.726	-0.156	0.898	1.086	-0.188
AV-Org	2001:01-2006:04	1.570	1.683	-0.113	0.898	1.048	-0.150
AV-UC	2001:01-2006:04	1.570	1.688	-0.117	0.898	1.049	-0.151
AV-Fil	2001:01-2006:04	1.570	1.644	-0.073	0.898	0.980	-0.082

Note: Sample period is 1985Q1-2006Q4.

"MSE(AR)" and "MSE(OG)" stand for Mean Squared Error (MSE) of the Autoregressive (AR) and Phillips curve with Output Gap (OG) models, respectively. "diff" refers to the difference between these two MSE.

Table 10 – Inflation forecast accuracy in the medium term

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE (AR)	MSE (OG)	diff
		forecast horizon h = 8 quarters			forecast horizon h = 12 quarters		
CAP-AV	2002:04-2006:04	1.101	1.028	**0.073**	1.028	1.019	**0.010**
CAP-TR	2002:04-2006:04	1.101	1.044	**0.057**	1.028	1.112	-0.084
EC - T	2000:04-2006:04	1.604	1.546	**0.058**	2.205	1.971	**0.234**
EC - P	2004:03-2006:04	0.933	1.070	-0.137	0.853	1.099	-0.247
IMF	2000:04-2006:04	1.604	1.670	-0.067	2.205	2.239	-0.034
OECD	2000:04-2006:04	1.604	1.750	-0.146	2.205	2.353	-0.148
UC - CC	2004:02-2006:04	1.013	1.299	-0.286	0.947	1.463	-0.516
UC - PIC	2004:02-2006:04	1.013	1.154	-0.141	0.947	1.105	-0.158
UC - BIV	2004:03-2006:04	0.933	1.033	-0.101	0.853	1.102	-0.249
BP	2002:04-2006:04	1.101	1.166	-0.065	1.028	1.025	**0.003**
HP	2002:04-2006:04	1.101	1.092	**0.009**	1.028	1.005	**0.023**
LIN	2002:04-2006:04	1.101	1.128	-0.027	1.028	1.041	-0.013
AV-All	2002:04-2006:04	1.101	1.032	**0.069**	1.028	0.982	**0.047**
AV-PFA	2002:04-2006:04	1.101	1.219	-0.118	1.028	1.124	-0.096
AV-Org	2002:04-2006:04	1.101	1.190	-0.089	1.028	1.059	-0.031
AV-UC	2002:04-2006:04	1.101	1.197	-0.096	1.028	1.206	-0.177
AV-Fil	2002:04-2006:04	1.101	1.127	-0.026	1.028	1.025	**0.003**

Note: Sample period is 1985Q1-2006Q4.

"MSE(AR)" and "MSE(OG)" stand for Mean Squared Error (MSE) of the Autoregressive (AR) and Phillips curve with Output Gap (OG) models, respectively. "diff" refers to the difference between these two MSE.

Table 11 – Tests of equal inflation forecast accuracy in the short term

Output gap model	Sample	MSE-t (conv)	MSE-F	MSE-t (conv)	MSE-F
		h = 1 quarter		h = 4 quarters	
CAP-AV	2001:01-2006:04	-0.784 *	-0.758 n.s.	**0.276 n.s.**	**0.415 n.s.**
CAP-TR	2001:01-2006:04	-1.072 *	-0.828 n.s.	-0.528 n.s.	-0.555 n.s.
EC - T	1999:01-2006:04	-2.555 *	-1.910 *	-2.144 *	-3.378 *
EC - P	2002:04-2006:04	-1.334 *	-1.132 *	-2.214 *	-2.823 *
IMF	1999:01-2006:04	-2.465 *	-2.545 *	-2.141 *	-5.019 *
OECD	1999:01-2006:04	-2.977 *	-2.834 *	-2.100 *	-5.138 *
UC - CC	2002:03-2006:04	-2.592 *	-1.980 *	-3.822 *	-3.700 *
UC - PIC	2002:03-2006:04	-3.002 *	-2.416 *	-3.567 *	-3.647 *
UC - BIV	2002:04-2006:04	-2.053 *	-1.335 *	-2.180 *	-1.863 n.s.
BP	2001:01-2006:04	-3.268 *	-1.281 *	-2.441 *	-2.312 n.s.
HP	2001:01-2006:04	-2.924 *	-1.040 n.s.	-2.768 *	-1.380 n.s.
LIN	2001:01-2006:04	-3.332 *	-0.782 n.s.	-2.107 *	-1.383 n.s.
AV-All	2001:01-2006:04	-1.425 *	-1.371 *	-0.663 *	-1.180 n.s.
AV-PFA	2001:01-2006:04	-2.337 *	-2.163 *	-1.880 *	-3.633 *
AV-Org	2001:01-2006:04	-2.093 *	-1.607 *	-1.889 *	-3.009 *
AV-UC	2001:01-2006:04	-1.752 *	-1.670 *	-2.755 *	-3.020 *
AV-Fil	2001:01-2006:04	-3.339 *	-1.084 n.s.	-2.461 *	-1.753 n.s.

Note: Sample period is 1985Q1-2006Q4.

"MSE-t (conv)"reports (non-adjusted) conventional t-statistics. "MSE-F" reports the statistic proposed by Clark and McCracken (2005). In both cases critical values were computed using a Monte Carlo simulation. * (n.s.) = test statistics indicates (no) rejection of the null of equal accuracy at a significance level of 10% or better.

Table 12 – Tests of equal inflation forecast accuracy in the medium term

Output gap model	Sample	MSE-t (conv)	MSE-F	MSE-t (conv)	MSE-F
		h = 8 quarters		h = 12 quarters	
CAP-AV	2002:04-2006:04	**2.102 ***	**1.205 n.s.**	**0.162 n.s.**	**0.122 n.s.**
CAP-TR	2002:04-2006:04	**1.707 ***	**0.930 n.s.**	-1.422 *	-0.977 n.s.
EC - T	2000:04-2006:04	**0.580 ***	**0.930 n.s.**	**1.745 ***	**2.487 n.s.**
EC - P	2004:03-2006:04	-1.751 *	-1.278 n.s.	-3.470 *	-1.346 n.s.
IMF	2000:04-2006:04	-0.767 *	-0.996 n.s.	-0.367 n.s.	-0.317 n.s.
OECD	2000:04-2006:04	-1.316 *	-2.090 n.s.	-1.620 *	-1.324 n.s.
UC - CC	2004:02-2006:04	-5.699 *	-2.422 n.s.	-9.151 *	-2.468 n.s.
UC - PIC	2004:02-2006:04	-4.739 *	-1.346 n.s.	-8.771 *	-1.003 n.s.
UC - BIV	2004:03-2006:04	-2.455 *	-0.973 n.s.	-4.774 *	-1.357 n.s.
BP	2002:04-2006:04	-3.860 *	-0.949 n.s.	**0.433 n.s.**	**0.034 n.s.**
HP	2002:04-2006:04	**0.741 ***	**0.147 n.s.**	**2.819 ***	**0.298 n.s.**
LIN	2002:04-2006:04	-0.677 *	-0.406 n.s.	-0.371 n.s.	-0.160 n.s.
AV-All	2002:04-2006:04	**1.314 ***	**1.145 n.s.**	**0.786 ***	**0.616 n.s.**
AV-PFA	2002:04-2006:04	-1.205 *	-1.645 n.s.	-0.963 *	-1.110 n.s.
AV-Org	2002:04-2006:04	-1.097 *	-1.274 n.s.	-0.355 n.s.	-0.375 n.s.
AV-UC	2002:04-2006:04	-1.802 *	-1.366 n.s.	-2.792 *	-1.913 n.s.
AV-Fil	2002:04-2006:04	-0.904 *	-0.387 n.s.	**0.170 n.s.**	**0.040 n.s.**

Note: Sample period is 1985Q1-2006Q4.

"MSE-t (conv)"reports (non-adjusted) conventional t-statistics. "MSE-F" reports the statistic proposed by Clark and McCracken (2005). In both cases critical values were computed using a Monte Carlo simulation. * (n.s.) = test statistics indicates (no) rejection of the null of equal accuracy at a significance level of 10% or better.

Looking at different sub-samples, there are broadly no indications that the forecast performance of Phillips curve models with the output gap may have improved much in the more recent period, say from 2003 onwards, relative to the AR(4) model (see Appendix V). The only noticeable differences with respect to the results in the previous subsection are that there are larger gains in terms of MSE for two-tear ahead EC-T gap based forecasts when evaluation is conducted over 2000:4-2002:4, and for this sample there are gains also for two-year ahead EC-P, OECD, and UC-BIV gap based forecasts. However, none of these gains are statistically significant according to the modified t-statistics by Clark and McCracken (2009). Moreover, the two-year ahead EC-T gap based forecast no longer beats the AR over the more recent subsample 2003:1-2006:4.

Finally, to assess the impact of real-time data we compare results based on the three simple filters considered (BP, HP, LIN) applied to the real time vintages to those based on the filters applied to the pseudo-real time estimates. For the other sets of estimates this comparison is not possible, either because the real time data for some series needed to estimate the gap is not available or because often judgment is also used to occasionally adjust estimates. As shown in Appendix VI, results for the three simple filters tend to be very similar, which confirms the in-sample result that data revisions in the underlying series seem to play a minor role for the euro area, at least over the sample under analysis.

V. 2. Inflation forecasting assessment: robustness analysis

In order to assess whether results depend on the rather simple Phillips curve specification adopted, we perform a similar out-of-sample exercise with a more general Phillips curve. In particular, taking as reference the Phillips curve included in the Area Wide Model (Fagan et al, 2005), we include among the regressors also import prices (y_t) and unit labour costs (z_t):

$$\pi_{t+\tau}^{(\tau)} - \pi_t = \alpha + \sum_{k=0}^{3} \beta_k \Delta\pi_{t-k} + \gamma_X x_t + \delta_X \Delta x_t +$$

$$+ \gamma_Y y_t + \delta_Y \Delta y_t + \gamma_Z z_t + \delta_Z \Delta z_t + u_{t+\tau}$$

The benchmark is also adjusted to take into account these additional factors, and corresponds to the same equation without the output gap:

$$\pi_{t+\tau}^{(\tau)} - \pi_t = \alpha + \sum_{k=0}^{3} \beta_k \Delta\pi_{t-k} + \qquad \text{(2B)}$$

$$+ \gamma_Y y_t + \delta_Y \Delta y_t + \gamma_Z z_t + \delta_Z \Delta z_t + \varepsilon_{t+\tau}$$

Overall, forecasting results based on these generalised functions tend to be very similar (all results reported in Appendix VII). In particular, while the MSE of the equations with the output gap is occasionally lower, especially in the medium term, it is never the case that the output gap coefficient is significant for the cases when the equation with the gap appears to perform better.

A second robustness check aims at assessing the role of the reference "final" vintage considered. Following Clark and McCracken (2009) as well as Romer and Romer (2000), we consider an alternative definition of reference series compared to the latest available vintage. More precisely, taking into consideration the revisions which affect in particular the first release of the data, we consider as reference series the second release of the GDP deflator.

Results based on the second release are again very similar to those based on the latest available vintage (see Appendix VIII for all results).

In summary, despite some differences across output gap estimates and forecast horizon, the results in this section point clearly to a lack of any usefulness of real-time output gap estimates for inflation forecasting in the euro area, both in the short term and the medium term.

V. 3. Inflation forecasting assessment: a comparison with the US

For comparison with the euro area, we have replicated the analysis for the US using two data sets: first, data over the sample 1950:1-2006:4 with vintages from 1970 (as in Clark and McCracken (2009)); second, data over the sample 1985:1-2006:4 with vintages from 2001 (as for most euro area cases). For each case, we have considered the performance of simple filter based output gap measures in a real time out of sample context. The results are summarized in Tables 13 and 14.

A clear difference between the full sample and post 1985 results for the US emerges. Over the longer sample, all gap measures have predictive content, the gains are fairly large and increase with the forecast horizons, and in most cases they are statistically significant, see Tables 13 and 14. However, in the after 1985 sample, the only gap measure that preserves some predictive gains is BP, but the gains are very small and never statistically significant, when evaluated with the proper Clark and McCracken (2009) statistic. Appendix IX reports more details on the sub-sample analysis.

In summary, focusing on the more relevant post 1985 real time forecasting results, the findings for the US are qualitatively similar to those for the euro area, and overall support the lack of significant predictive content of output gap measures for inflation, both in the short and in the medium term.

Table 13 – Inflation forecast accuracy in the short term in the US

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 1 (one-quarter ahead)				
			1965-2006 (vintages from 1970Q1 onwards)				
BP	1.928	1.759	**0.169**	*0.953*	*	*15.747*	*
HP	1.928	1.763	**0.165**	*0.686*	*	*15.304*	*
LIN	1.928	1.754	**0.174**	*0.726*	*	*16.269*	*
AV-Fil	1.928	1.765	**0.163**	*0.725*	*	*15.155*	*
			1985-2006 (vintages from 2001Q1 onwards)				
BP - 2001	0.845	0.808	**0.038**	*0.365*	*n.s.*	*1.074*	*n.s.*
HP - 2001	0.845	0.863	-0.018	*-0.102*	*n.s.*	*-0.473*	*n.s.*
LIN - 2001	0.845	0.890	-0.044	*-0.326*	*n.s.*	*-1.145*	*n.s.*
AV-Fil - 2001	0.845	0.867	-0.021	*-0.138*	*n.s.*	*-0.569*	*n.s.*
			forecast horizon h = 4 (one-year ahead)				
			1965-2006 (vintages from 1970Q1 onwards)				
BP	1.993	1.632	**0.361**	*1.257*	*	*35.575*	*
HP	1.993	1.690	**0.302**	*1.136*	*	*28.803*	*
LIN	1.993	1.802	**0.191**	*0.555*	*	*17.056*	*
AV-Fil	1.993	1.722	**0.271**	*0.787*	*	*25.334*	*
			1985-2006 (vintages from 2001Q1 onwards)				
BP - 2001	0.522	0.484	**0.038**	*0.363*	*n.s.*	*1.568*	*n.s.*
HP - 2001	0.522	0.647	-0.125	*-0.858*	*	*-3.864*	*
LIN - 2001	0.522	0.724	-0.203	*-1.216*	*	*-5.594*	*
AV-Fil - 2001	0.522	0.683	-0.162	*-0.938*	*	*-4.728*	*

Note: Sample period is 1950Q1-2006Q4.

"MSE(AR)" and "MSE(OG)" stand for Mean Squared Error (MSE) of the Autoregressive (AR) and Phillips curve with Output Gap (OG) models, respectively. "diff" refers to the difference between these two MSE. "MSE-t (conv)"reports (non-adjusted) conventional t-statistics. "MSE-F" reports the statistic proposed by Clark and McCracken (2005). In both cases critical values were computed using a Monte Carlo simulation. * (n.s.) = test statistics indicates (no) rejection of the null of equal accuracy at a significance level of 10% or better.

Table 14 – Inflation forecast accuracy in the medium term in the US

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
forecast horizon h = 8 (two-years ahead)							
1965-2006 (vintages from 1970Q1 onwards)							
BP	4.776	3.416	**1.360**	*1.621*	*	*62.520*	*
HP	4.776	3.355	**1.421**	*1.781*	*	*66.481*	*
LIN	4.776	3.891	**0.885**	*0.860*	*	*35.716*	*
AV-Fil	4.776	3.415	**1.361**	*1.400*	*	*62.555*	*
1985-2006 (vintages from 2001Q1 onwards)							
BP - 2001	0.891	0.685	**0.206**	*6.044*	*	*4.816*	*n.s.*
HP - 2001	0.891	0.949	-0.058	*-0.923*	*	*-0.983*	*n.s.*
LIN - 2001	0.891	1.456	-0.565	*-3.279*	*	*-6.207*	*
AV-Fil - 2001	0.891	1.146	-0.255	*-2.429*	*	*-3.556*	*n.s.*
forecast horizon h = 12 (three-years ahead)							
1965-2006 (vintages from 1970Q1 onwards)							
BP	7.384	6.977	**0.406**	*0.797*	*	*8.911*	*
HP	7.384	6.541	**0.842**	*1.559*	*	*19.702*	*
LIN	7.384	6.303	**1.081**	*0.894*	*	*26.236*	*
AV-Fil	7.384	6.217	**1.167**	*1.415*	*	*28.720*	*
1985-2006 (vintages from 2001Q1 onwards)							
BP - 2001	1.409	1.374	**0.035**	*0.174*	*n.s.*	*0.304*	*n.s.*
HP - 2001	1.409	1.611	-0.202	*-2.117*	*	*-1.504*	*n.s.*
LIN - 2001	1.409	2.332	-0.923	*-8.464*	*	*-4.749*	*n.s.*
AV-Fil - 2001	1.409	1.938	-0.529	*-6.124*	*	*-3.275*	*n.s.*

Note: Sample period is 1950Q1-2006Q4.

"MSE(AR)" and "MSE(OG)" stand for Mean Squared Error (MSE) of the Autoregressive (AR) and Phillips curve with Output Gap (OG) models, respectively. "diff" refers to the difference between these two MSE. "MSE-t (conv)"reports (non-adjusted) conventional t-statistics. "MSE-F" reports the statistic proposed by Clark and McCracken (2005). In both cases critical values were computed using a Monte Carlo simulation. * (n.s.) = test statistics indicates (no) rejection of the null of equal accuracy at a significance level of 10% or better.

VI. Do real time output gap estimates help forecasting real GDP growth?

Despite the fact that most often the usefulness of the output gap for forecasting is assessed with respect to inflation, it has been suggested, for example by Giannone and Reichlin (2006), that also the ability to predict real GDP growth can be a useful criterion to assess the output gap. We propose a possible justification in terms of cointegration analysis and error correction models. More precisely, potential output can be seen as a stochastic trend of output, implying a role of the output gap in the context of error correction models (speed of adjustment). Since changes in potential output take place slowly, real GDP growth should adjust to reduce deviations from potential, and the gap could therefore provide useful information for forecasting real GDP growth.

We undertake a similar exercise as that for inflation, using a very similar setup. The exercise is carried out using vintages form 2001Q1 onwards, since this is the first available real time vintage for euro area real GDP growth.

VI. 1. Real GDP growth forecasting assessment: out-of-sample real time evaluation

We now carry out an out-of-sample forecasting exercise to evaluate the predictive content for real GDP growth of alternative measures of the euro area output gap, similar to that presented for inflation in section V. Thus, using the above-mentioned vintages of real-time output gap estimates, we use the tests proposed recently by Clark and McCracken (2009) and similar equations as for inflation.

Following Stock and Watson (2003), we compute forecasts of real GDP growth Y_t at horizon τ from the following equations:

$$Y_{t+\tau}^{(\tau)} = \alpha + \sum_{k=0}^{3} \beta_k Y_{t-k} + \gamma x_t + \delta \Delta x_t + u_{t+\tau} \qquad (5)$$

where $Y_t^{(\tau)} \equiv (400/\tau)\ln(y_t / y_{t-\tau})$, $y_t^{(1)} \equiv y_t$ and x_t is the output gap (expressed in terms of percentage deviations from trend or potential output). Again, we consider four forecast horizons: one quarter ($\tau = 1$), one year ($\tau = 4$), two years ($\tau = 8$), and three years ($\tau = 12$).

The benchmark model against which to compare the forecasts of this model is an autoregressive (AR) model for real GDP growth:

$$Y_{t+\tau}^{(\tau)} = \alpha + \sum_{k=0}^{3} \beta_k Y_{t-k} + \varepsilon_{t+\tau} \qquad (6)$$

Tables 15 and 16 contain the results of the exercise comparing the forecast of real GDP growth at different horizons in the short run (one quarter and one year) and the medium run (two and three years) from the AR(4) model and the model with the output gap.

As regards the short run, in several cases the MSE of the forecasts based on the AR(4) models are lower than those based on the model with the output gap. But there are some output gap estimates that do improve the forecasts, in particular those based on capacity utilization (Table 15).

Table 15 – Real GDP growth forecast accuracy in the short term

Output gap model	sample	MSE (AR)	MSE (OG)	diff	MSE (AR)	MSE (OG)	diff
		forecast horizon h = 1 quarter			forecast horizon h = 4 quarters		
CAP-AV	2001:01-2006:04	1.574	1.346	**0.228**	0.838	0.635	**0.203**
CAP-TR	2001:01-2006:04	1.574	1.324	**0.250**	0.838	0.789	**0.049**
EC - T	2001:01-2006:04	1.574	1.879	-0.305	0.838	1.410	-0.572
EC - P	2002:04-2006:04	1.556	1.388	**0.167**	0.560	0.879	-0.319
IMF	2001:01-2006:04	1.574	1.853	-0.278	0.838	1.468	-0.629
OECD	2001:01-2006:04	1.574	1.643	-0.069	0.838	1.001	-0.163
UC - CC	2002:03-2006:04	1.469	2.244	-0.775	0.638	0.665	-0.027
UC - PIC	2002:03-2006:04	1.469	10.244	-8.775	0.638	4.859	-4.221
UC - BIV	2002:04-2006:04	1.556	5.339	-3.783	0.560	0.852	-0.292
BP	2001:01-2006:04	1.574	3.291	-1.717	0.838	0.842	-0.003
HP	2001:01-2006:04	1.574	7.195	-5.621	0.838	7.750	-6.912
LIN	2001:01-2006:04	1.574	1.517	**0.058**	0.838	0.753	**0.085**
AV-All	2001:01-2006:04	1.574	3.756	-2.182	0.838	0.651	**0.187**
AV-PFA	2001:01-2006:04	1.574	2.265	-0.691	0.838	1.095	-0.257
AV-Org	2001:01-2006:04	1.574	1.745	-0.171	0.838	1.232	-0.394
AV-UC	2001:01-2006:04	1.574	1.527	**0.047**	0.838	0.537	**0.301**
AV-Fil	2001:01-2006:04	1.574	12.289	-10.715	0.838	1.708	-0.870

Note: Sample period is 1985Q1-2006Q4.
"MSE(AR)" and "MSE(OG)" stand for Mean Squared Error (MSE) of the Autoregressive (AR) and models with the Output Gap (OG), respectively. "diff" refers to the difference between these two MSE.

For the medium run, the results are qualitatively similar but now a few other gap measures seem to yield lower MSEs, in particular those based on the BP or linear filters and the average of all gap measures (Table 16).

The test results reported in Tables 17 and 18 suggest that the forecast gains arising from the selected gap measures either in the short or in the medium run are statistically significant.

Table 16 –Real GDP growth forecast accuracy in the medium term

Output gap model	sample	MSE (AR)	MSE (OG)	diff	MSE (AR)	MSE (OG)	diff
		forecast horizon h = 8 quarters			forecast horizon h = 12 quarters		
CAP-AV	2002:04-2006:04	0.707	0.234	**0.472**	0.626	0.128	**0.498**
CAP-TR	2002:04-2006:04	0.707	0.645	**0.062**	0.626	0.576	**0.050**
EC - T	2002:04-2006:04	0.707	0.907	-0.200	0.626	0.832	-0.206
EC - P	2004:03-2006:04	0.240	0.715	-0.475	0.219	0.559	-0.340
IMF	2002:04-2006:04	0.707	1.477	-0.770	0.626	1.145	-0.519
OECD	2002:04-2006:04	0.707	0.650	**0.056**	0.626	0.384	**0.242**
UC - CC	2004:02-2006:04	0.262	0.224	**0.038**	0.285	0.272	**0.013**
UC - PIC	2004:02-2006:04	0.262	2.282	-2.019	0.285	0.720	-0.435
UC - BIV	2004:03-2006:04	0.240	0.769	-0.529	0.219	0.641	-0.421
BP	2002:04-2006:04	0.707	0.598	**0.108**	0.626	0.574	**0.052**
HP	2002:04-2006:04	0.707	4.109	-3.402	0.626	1.363	-0.736
LIN	2002:04-2006:04	0.707	0.599	**0.107**	0.626	0.579	**0.047**
AV-All	2002:04-2006:04	0.707	0.314	**0.392**	0.626	0.277	**0.350**
AV-PFA	2002:04-2006:04	0.707	0.696	**0.011**	0.626	0.509	**0.117**
AV-Org	2002:04-2006:04	0.707	0.905	-0.198	0.626	0.710	-0.083
AV-UC	2002:04-2006:04	0.707	0.888	-0.182	0.626	0.540	**0.086**
AV-Fil	2002:04-2006:04	0.707	2.964	-2.257	0.626	1.399	-0.772

Note: Sample period is 1985Q1-2006Q4.

"MSE(AR)" and "MSE(OG)" stand for Mean Squared Error (MSE) of the Autoregressive (AR) and models with the Output Gap (OG), respectively. "diff" refers to the difference between these two MSE.

Table 17 – Tests of equal real GDP growth forecast accuracy in the short term

Output gap model	sample	MSE-t (conv)	MSE-F	MSE-t (conv)	MSE-F
		h = 1 quarter		h = 4 quarters	
CAP-AV	2001:01-2006:04	0.617 n.s.	**4.061** *	0.878 *	**6.716** *
CAP-TR	2001:01-2006:04	0.594 n.s.	**4.539** *	0.675 *	1.299 n.s.
EC - T	2001:01-2006:04	-0.615 n.s.	-3.893 *	-1.921 *	-8.520 *
EC - P	2002:04-2006:04	0.324 n.s.	**2.048** *	-0.938 *	-5.080 *
IMF	2001:01-2006:04	-0.550 n.s.	-3.607 *	-1.922 *	-9.007 *
OECD	2001:01-2006:04	-0.173 n.s.	-1.005 n.s.	-0.927 *	-3.411 *
UC - CC	2002:03-2006:04	-1.099 *	-6.214 *	-0.272 n.s.	-0.611 n.s.
UC - PIC	2002:03-2006:04	-2.524 *	-15.418 *	-1.861 *	-13.030 *
UC - BIV	2002:04-2006:04	-2.719 *	-12.046 *	-2.071 *	-4.797 *
BP	2001:01-2006:04	-1.980 *	-12.520 *	-0.024 n.s.	-0.085 n.s.
HP	2001:01-2006:04	-3.981 *	-18.749 *	-3.784 *	-18.729 *
LIN	2001:01-2006:04	0.866 *	0.910 n.s.	1.622 *	2.375 n.s.
AV-All	2001:01-2006:04	-2.235 *	-13.941 *	1.270 *	**6.019** *
AV-PFA	2001:01-2006:04	-1.119 *	-7.323 *	-1.239 *	-4.926 *
AV-Org	2001:01-2006:04	-0.373 n.s.	-2.352 *	-1.591 *	-6.710 *
AV-UC	2001:01-2006:04	0.104 n.s.	0.746 n.s.	0.854 *	**11.769** *
AV-Fil	2001:01-2006:04	-4.221 *	-1.294 *	-1.075 *	-10.693 *

Note: Sample period is 1985Q1-2006Q4.

"MSE-t (conv)"reports (non-adjusted) conventional t-statistics. "MSE-F" reports the statistic proposed by Clark and McCracken (2005). In both cases critical values were computed using a Monte Carlo simulation. * (n.s.) = test statistics indicates (no) rejection of the null of equal accuracy at a significance level of 10% or better.

Table 18 – Tests of equal real GDP growth forecast accuracy in the medium term

Output gap model	sample	MSE-t (conv)	MSE-F	MSE-t (conv)	MSE-F
		h = 8 quarters		h = 12 quarters	
CAP-AV	2002:04-2006:04	1.877 *	**34.319** *	3.207 *	**50.579** *
CAP-TR	2002:04-2006:04	1.092 *	1.631 n.s.	0.883 *	1.137 n.s.
EC - T	2002:04-2006:04	-5.018 *	-3.753 n.s.	-5.025 *	-3.213 n.s.
EC - P	2004:03-2006:04	-6.873 *	-6.641 n.s.	-9.647 *	-3.647 n.s.
IMF	2002:04-2006:04	-2.958 *	-8.867 *	-4.217 *	-5.891 n.s.
OECD	2002:04-2006:04	0.391 n.s.	1.465 n.s.	2.540 *	8.201 n.s.
UC - CC	2004:02-2006:04	0.828 *	1.886 n.s.	0.705 n.s.	0.335 n.s.
UC - PIC	2004:02-2006:04	-1.998 *	-9.735 *	-3.338 *	-4.230 n.s.
UC - BIV	2004:03-2006:04	-1.499 *	-6.877 n.s.	-2.454 *	-3.947 n.s.
BP	2002:04-2006:04	0.986 *	3.077 n.s.	0.469 n.s.	1.187 n.s.
HP	2002:04-2006:04	-2.639 *	-14.077 *	-1.421 *	-7.025 n.s.
LIN	2002:04-2006:04	4.123 *	3.047 n.s.	1.052 *	1.059 n.s.
AV-All	2002:04-2006:04	1.684 *	**21.240** *	2.583 *	**16.444** *
AV-PFA	2002:04-2006:04	0.088 n.s.	0.269 n.s.	1.932 *	2.988 n.s.
AV-Org	2002:04-2006:04	-1.566 *	-3.722 n.s.	-1.436 *	-1.523 n.s.
AV-UC	2002:04-2006:04	-0.692 *	-3.479 n.s.	0.392 n.s.	2.070 n.s.
AV-Fil	2002:04-2006:04	-2.311 *	-12.947 *	-1.431 *	-7.177 n.s.

Note: Sample period is 1985Q1-2006Q4.

"MSE-t (conv)"reports (non-adjusted) conventional t-statistics. "MSE-F" reports the statistic proposed by Clark and McCracken (2005). In both cases critical values were computed using a Monte Carlo simulation. * (n.s.) = test statistics indicates (no) rejection of the null of equal accuracy at a significance level of 10% or better.

To assess the impact of real-time data we compare results based on the three simple filters considered (BP, HP, LIN) applied to the real time vintages to those based on the filters applied to the pseudo-real time estimates. For the other sets of estimates this comparison is not possible, either because the real time data for some series needed to estimate the gap is not available or because often judgment is also used to occasionally adjust estimates. As for the case of inflation, results for the three simple filters tend to be very similar (see Appendix X).

Furthermore, and similar to the case of inflation, if we consider as reference series the second release of real GDP, rather than the first one, results are very similar (see Appendix XII for all results).

Finally, looking at different sub-samples, there are broad indications that the forecast performance of the models with the output gap may have improved in the more recent period, say from 2004 onwards, relative to the AR(4) model, especially in the short run but in several cases also for the medium run (results reported in Appendix XII). However, it should be considered that the evaluation samples in this case become very short.

Overall, the results suggest that selected output gap estimates may improve real GDP growth forecasts even in real-time. Measures based on capacity utilization, linear filters, and averages of all gap measures perform particularly well at most horizons.

VI. 3. Real GDP growth forecasting assessment: a comparison with the US

We now assess whether the positive role of selected gap measures for forecasting real GDP growth we have detected for the euro area is present for the US as well. As for inflation, we consider two sample periods for the US: first, data over 1950:1-2006:4 with vintages from 1970 (as in Clark and McCracken, 2009); second, data over 1985:1-2006:4 with vintages from 2001 (as for the euro area). For each case, we have considered the performance of simple filter based output gap measures in a real time out of sample context. The results are summarized in Tables 19 and 20.

The real-time out of sample results reported in Tables 19 and 20 for, respectively, the short run and the medium run confirm that there are minor differences across the two subsamples. Moreover, the gap measures under evaluation are in general useless for forecasting real GDP growth. In the case of the euro area, the results were better for the linear filter based gap, and also for the BP gap in the medium run. A possible explanation for the different results for the US and the euro area is the fairly different growth path of the two areas.

Finally, an evaluation for the most recent period, after 2004, finds basically the same results (see Appendix XIII for details, also on the statistical significance of the MSE gains and losses).

Table 19 – Real GDP growth forecast accuracy in the short term in the US

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 1 (one-quarter ahead)				
			1965-2006 (vintages from 1970Q1 onwards)				
BP	11.575	13.795	-2.221	*-0.935*	*	*-26.402*	*
HP	11.575	11.584	-0.010	*-0.010*	*	*-0.141*	*
LIN	11.575	11.638	-0.064	*-0.108*	*	*-0.901*	*
AV-Fil	11.575	13.224	-1.650	*-0.817*	*	*-20.458*	*
			1985-2006 (vintages from 2001Q1 onwards)				
BP - 2001	3.320	4.099	-0.779	*-0.430*	*n.s.*	*-4.373*	*
HP - 2001	3.320	4.004	-0.684	*-0.835*	*	*-3.929*	*
LIN - 2001	3.320	3.491	-0.171	*-0.574*	*n.s.*	*-1.130*	*
AV-Fil - 2001	3.320	3.752	-0.432	*-0.364*	*n.s.*	*-2.648*	*
			forecast horizon h = 4 (one-year ahead)				
			1965-2006 (vintages from 1970Q1 onwards)				
BP	5.457	5.613	-0.155	*-0.196*	*	*-4.460*	*
HP	5.457	5.292	**0.165**	*0.274*	*	***5.016***	*
LIN	5.457	6.199	-0.742	*-1.050*	*	*-19.260*	*
AV-Fil	5.457	6.692	-1.235	*-1.357*	*	*-29.712*	*
			1985-2006 (vintages from 2001Q1 onwards)				
BP - 2001	1.282	1.630	-0.348	*-0.873*	*	*-4.272*	*n.s.*
HP - 2001	1.282	2.403	-1.121	*-1.364*	*	*-9.330*	*
LIN - 2001	1.282	1.314	-0.032	*-0.224*	*n.s.*	*-0.490*	*n.s.*
AV-Fil - 2001	1.282	1.722	-0.440	*-1.774*	*	*-5.112*	*n.s.*

Note: Sample period is 1950Q1-2006Q4.

"MSE(AR)" and "MSE(OG)" stand for Mean Squared Error (MSE) of the Autoregressive (AR) and Phillips curve with Output Gap (OG) models, respectively. "diff" refers to the difference between these two MSE. "MSE-t (conv)"reports (non-adjusted) conventional t-statistics. "MSE-F" reports the statistic proposed by Clark and McCracken (2005). In both cases critical values were computed using a Monte Carlo simulation. * (n.s.) = test statistics indicates (no) rejection of the null of equal accuracy at a significance level of 10% or better.

Table 20 – Real GDP growth forecast accuracy in the medium term in the US

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 8 (two-years ahead)				
			1965-2006 (vintages from 1970Q1 onwards)				
BP	3.471	3.138	**0.333**	*1.071*	*	***16.684***	*
HP	3.471	3.337	**0.135**	*0.371*	*	***6.331***	*
LIN	3.471	4.810	-1.338	*-2.226*	*	*-43.691*	*
AV-Fil	3.471	4.583	-1.112	*-1.765*	*	*-38.088*	*
			1985-2006 (vintages from 2001Q1 onwards)				
BP - 2001	0.719	0.886	-0.166	*-1.407*	*	*-3.006*	*n.s.*
HP - 2001	0.719	1.630	-0.911	*-1.952*	*	*-8.942*	*n.s.*
LIN - 2001	0.719	0.752	-0.032	*-0.436*	*n.s.*	*-0.689*	*n.s.*
AV-Fil - 2001	0.719	1.249	-0.529	*-1.974*	*	*-6.784*	*n.s.*
			forecast horizon h = 12 (three-years ahead)				
			1965-2006 (vintages from 1970Q1 onwards)				
BP	2.191	2.277	-0.086	*-0.568*	*	*-5.802*	*
HP	2.191	2.511	-0.320	*-1.227*	*	*-19.513*	*
LIN	2.191	3.726	-1.535	*-3.139*	*	*-63.026*	*
AV-Fil	2.191	2.995	-0.804	*-1.842*	*	*-41.056*	*
			1985-2006 (vintages from 2001Q1 onwards)				
BP - 2001	0.269	0.432	-0.164	*-2.288*	*	*-4.541*	*n.s.*
HP - 2001	0.269	1.061	-0.792	*-3.792*	*	*-8.961*	*n.s.*
LIN - 2001	0.269	0.385	-0.117	*-3.542*	*	*-3.629*	*n.s.*
AV-Fil - 2001	0.269	0.734	-0.465	*-3.201*	*	*-7.607*	*n.s.*

Note: Sample period is 1950Q1-2006Q4.

"MSE(AR)" and "MSE(OG)" stand for Mean Squared Error (MSE) of the Autoregressive (AR) and Phillips curve with Output Gap (OG) models, respectively. "diff" refers to the difference between these two MSE. "MSE-t (conv)"reports (non-adjusted) conventional t-statistics. "MSE-F" reports the statistic proposed by Clark and McCracken (2005). In both cases critical values were computed using a Monte Carlo simulation. * (n.s.) = test statistics indicates (no) rejection of the null of equal accuracy at a significance level of 10% or better.

VII. Conclusions

This paper has provided a thorough evaluation of the reliability of output gap measures for the euro area computed in real time. Consistent with the findings of previous empirical studies for other economic areas, the analysis of the various sources of uncertainty, based on an assessment of alternative estimates, measures of confidence bands around point estimates and past revisions, suggests quite clearly that real-time estimates are characterised by a high degree of uncertainty. In particular, the evidence indicates that both the magnitude and the sign of the real-time estimates of the euro area output gap are very uncertain.

For the euro area, changes in the vintages of the time series underlying the gap (e.g., real GDP) explain a minor part of the real time changes in the gap, while recursive computation matters considerably. This finding suggests either the need of a very long estimation sample for reliable gap estimation or, more likely, the presence of parameter changes. Unfortunately, averaging different gap measures does not yield any substantial gains, due to the rather high correlation across alternative gap measures.

Real time estimates of the US output gap suffer from similar problems, also in the most recent period, even though they are more correlated with final values with respect to the euro area. In addition, the data revision component of the revision error is larger than for the euro area.

As regards the predictive content for inflation of alternative measures of the euro area output gap, an out-of-sample forecasting exercise using different sets of real-time estimates points clearly to a lack of any usefulness of output gap real-time estimates for inflation forecasting both in the short term (one-quarter and one-year ahead) and the medium term (two-year and three-year ahead). These findings are broadly consistent with the empirical literature which shows that while ex post estimates of the output gap tend to have a relatively good in-sample fit in Phillips curve models, the out-of-sample predictive ability of real-time estimates is very limited.

As regards the predictive content for real GDP growth of alternative measures of the euro area output gap, a similar out-of-sample forecasting exercise provides mixed results. In particular, some real-time output gap estimates appear to improve significantly the forecasts for real GDP growth, especially when based on capacity utilization in the short run or averaging in the medium term.

It is worthwhile to notice that estimates of the output gap based on multivariate models do not seem to be systematically superior to univariate estimates both as regards reliability and as regards forecasting performance. This may appear to be somewhat surprising as ex ante more information included in an estimation model could be expected to result in improved estimates along some dimension. However, it appears that the uncertainty characterising all of these estimates is such that the contribution of more information to the output gap estimates may be of second order, such that for the assessment criteria considered it does not seem to make any significant difference.

Overall, the findings in this paper cast serious doubts on the usefulness of the output gap for structural analysis or economic policy making in the euro area, while there could be some use for forecasting future real economic activity growth.

Appendix I – Main features of the UC model of Proietti, Musso and Westermann (2007)

The bivariate model of output and inflation is based on the decomposition of output y_t, into a trend component, y_t^T, and a cyclical component, y_t^C, as proposed by Harvey and Jäger (1993):

$$y_t = y_t^T + y_t^C$$

where the trend component is modelled as a local linear trend (with an IMA(2,1) reduced form):

$$y_t^T = y_{t-1}^T + \beta_{t-1} + \eta_t, \quad \eta_t \sim NID(0, \sigma_\eta^2)$$

$$\beta_t^T = \phi\beta_{t-1}^T + \varsigma_t, \quad \varsigma_t \sim NID(0, \sigma_\varsigma^2)$$

and the cyclical component has the following stochastic specification (with an ARMA(2,1) reduced form):

$$y_t^C = \rho\cos\lambda_c y_{t-1}^C + \rho\sin\lambda_c y_{t-1}^{C*} + \kappa_t, \quad \kappa_t \sim NID(0, \sigma_\kappa^2)$$

$$y_t^{C*} = -\rho\sin\lambda_c y_{t-1}^C + \rho\cos\lambda_c y_{t-1}^{C*} + \kappa_t^*, \quad \kappa_t^* \sim NID(0, \sigma_\kappa^2)$$

A Phillips-type relationship, relating the changes in inflation to the output gap, is included in order to ensure coherence with the definition of potential output as the non-inflationary level of output in the medium-term:[8]

$$\varphi(L)\Delta\pi_t = \theta_\Pi(L) y_{t-1}^C + \theta_Z(L)' z_t + e_t \tag{1}$$

where z_t represents cost factors, i.e. changes in the commodity prices, including energy, and the nominal effective exchange rate of the euro.

The multivariate model is based on the production function approach, where output growth is driven by increases in labour and capital inputs and by technological progress.[9] Denoting by y_t, l_t and k_t respectively the logarithms of output, employment and capital stock of an economic sector, and assuming a Cobb-Douglas technology exhibiting constant returns to scale, the aggregate production function takes the form:

$$y_t = f_t + \beta l_t + (1-\beta)k_t$$

where f_t represents total factor productivity (TFP) and β is the elasticity of output with respect to labour. TFP is computed as a residual from the above equation.

All the variables in the production function are decomposed into their trend and cycle components:

$$f_t = f_t^T + f_t^C$$

$$l_t = l_t^T + l_t^C$$

$$k_t = k_t^T$$

it is common practice to set the trend values of the capital stock equal to the actual values.

The labour input is further decomposed into working-age population, participation rate and the employment rate.

---

8 This is the so-called triangle equation, explaining the change in inflation by three sources, i.e. a measure of the gap, cost factors and additional inflationary dynamics.

9 See Proietti, Musso and Westermann (2007).

Potential output is the value corresponding to the trend values of factor inputs and of TFP:

$$y_t^T = f_t^T + \beta l_t^T + (1-\beta)k_t^T$$

Trend and cyclical components are modelled along the lines of an extended version of the multivariate structural time series model proposed by Harvey and Koopman (1997). Basically, this model belongs to the seemingly unrelated time series models class, i.e. it does not contain interactions between the particular variables. However, the model allows for correlation among cyclical components of the particular series, while the trend components are assumed to be uncorrelated, according to long-run balanced growth assumptions.

Turning to modelling the particular components, all trends of the endogenous variables are specified as a local linear trend. Thus, for example, for TFP:

$$f_t^T = \delta_t + f^T{}_{t-1} + \eta^f{}_t$$

$$\delta_t = \delta_{t-1} + \varsigma^f{}_t$$

where the innovations $\eta^f{}_t$ and $\varsigma^f{}_t$ are white noise. That is, f_t^T is assumed to follow a random walk with drift. The drift δ_t itself follows a random walk.

Cyclical components are expressed as function of autoregressive processes of second order with complex roots. Thus, for example, for TFP:

$$f_t^C = \theta_f(L)\psi_t + \kappa_t^f$$

In addition, as for the bivariate model, the Phillips curve (1) is added.

The multivariate variants of the model considered are the common cycles model and the pseudo-integrated cycles model. In the common cycles model it is assumed that the cycle in capacity utilisation rates drives the cyclical component in all series. In particular, it is assumed that capacity cap_t is given by

$$cap_t = m(t) + \psi_{CAP,t}$$

where $m(t)$ is a deterministic trend with a slope change in 1975:1 and the cyclical component $\psi_{CAP,t}$ follows an AR(2) process. Then, the transitory components of TFP, the participation rate and the employment rate are expressed as a linear combination of current and lagged values of $\psi_{CAP,t}$. For example, for TFP:

$$f_t^C = \theta_f(L)\psi_{CAP,t} \text{ with } \theta_f(L) = \theta_{f,0} + \theta_{f,1}L$$

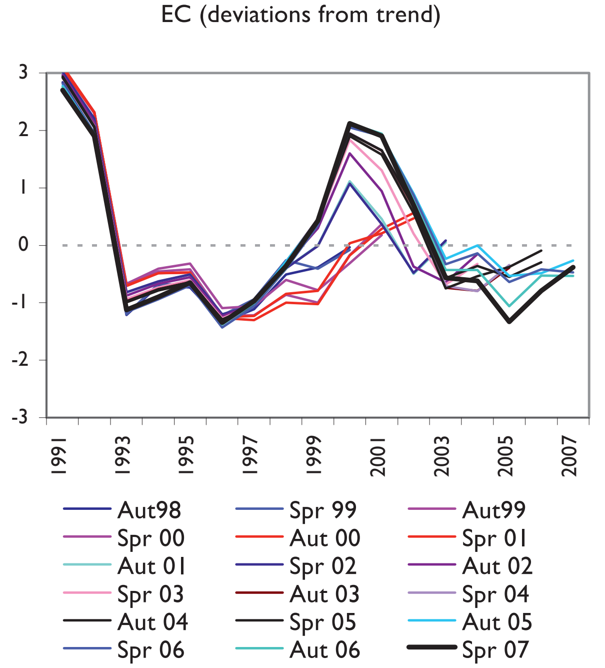
There are some indications that labour market variables could follow a cyclical pattern more persistent than that of other variables including capacity, largely due to specific frictions existing in the labour markets. Therefore, a variant of the model, the pseudo-integrated cycles representation, was developed to attempt to capture this specificity. More precisely, it is assumed that the cyclical component of each series is driven by a combination of autonomous forces (an specific, or idiosyncratic, cycle) and by a common cycle driven by capacity utilisation, with a transmission mechanism of the impulses represented by an autoregressive process.

Estimation is been carried out with the Kalman filter. The standard errors of the parameters have been estimated via a Monte Carlo simulation following the method suggested by Hamilton (1986). For more details see Proietti, Musso and Westermann (2007).

Appendix II – Vintages of output gap estimates

Chart A: Vintages of annual estimates of euro area output gap by the EC (dev. from trend)

(percentage deviations from trend output)

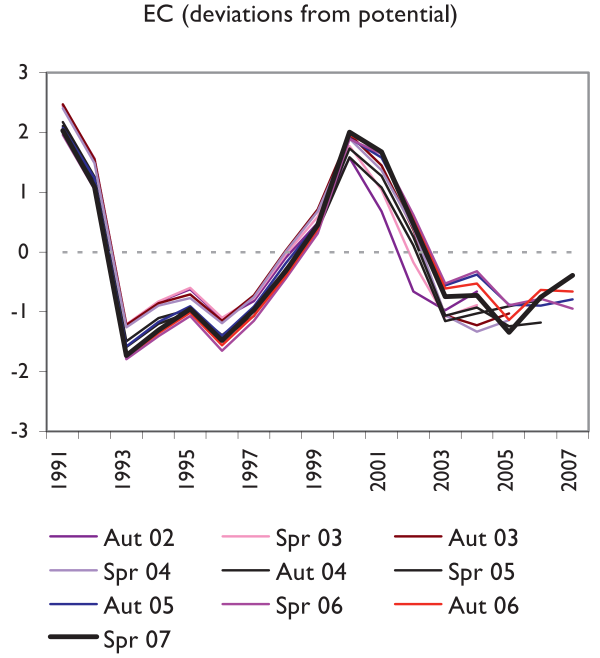


Sources: European Commission.

Note: Estimates are deviations from trend computed via the Hodrick-Prescott filter.

Chart B: Vintages of annual estimates of euro area output gap by the EC (dev. from potential)

(percentage deviations from potential output)

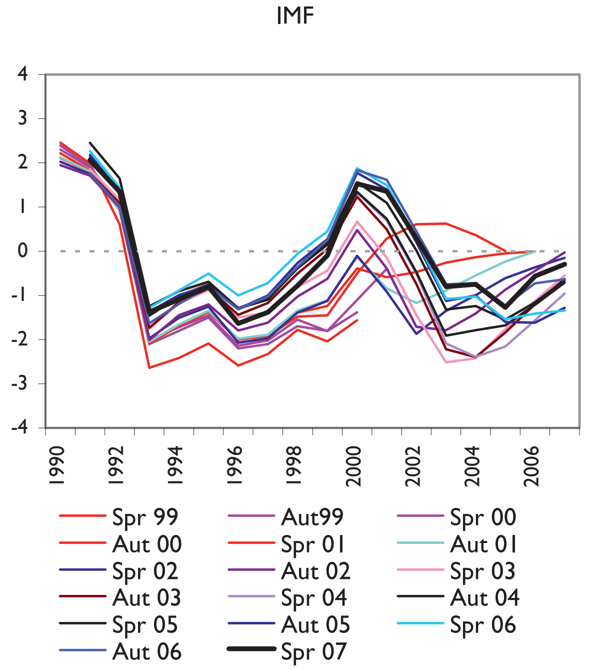


Sources: European Commission.

Note: Estimates by the EC of the output gap as deviations from potential start in Autumn 2002.

Chart C: Vintages of annual estimates of euro area output gap by the IMF

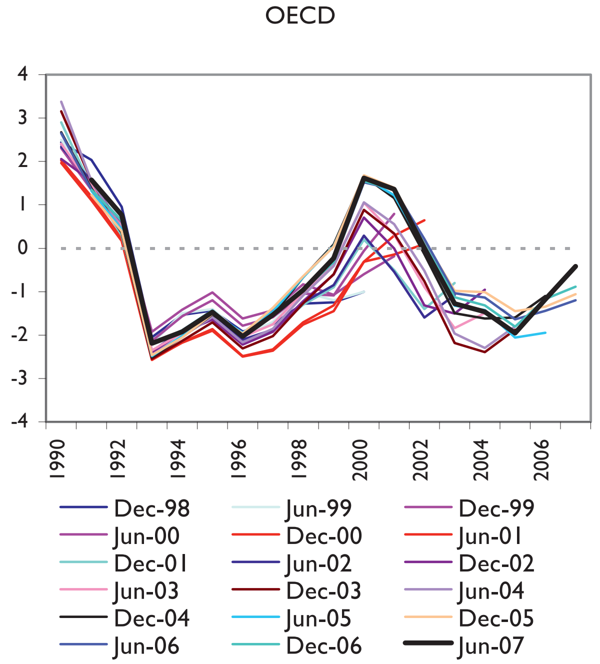
(percentage deviations from potential output)



Sources: IMF.

Chart D: Vintages of annual estimates of euro area output gap by the OECD

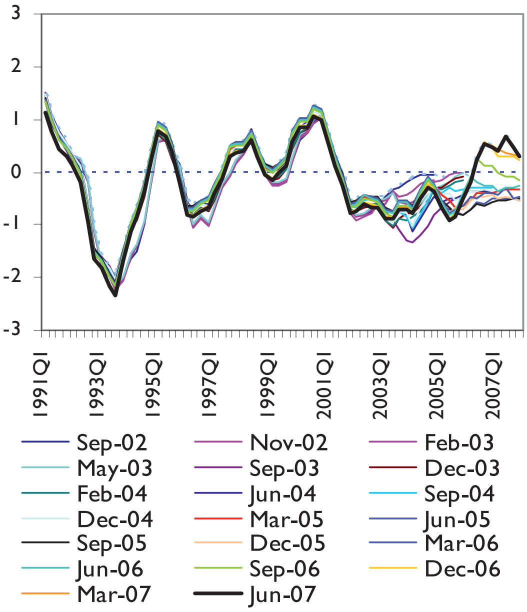
(percentage deviations from potential output)



Sources: OECD.

Chart E: Vintages of quarterly estimates of euro area output gap (UC CC)

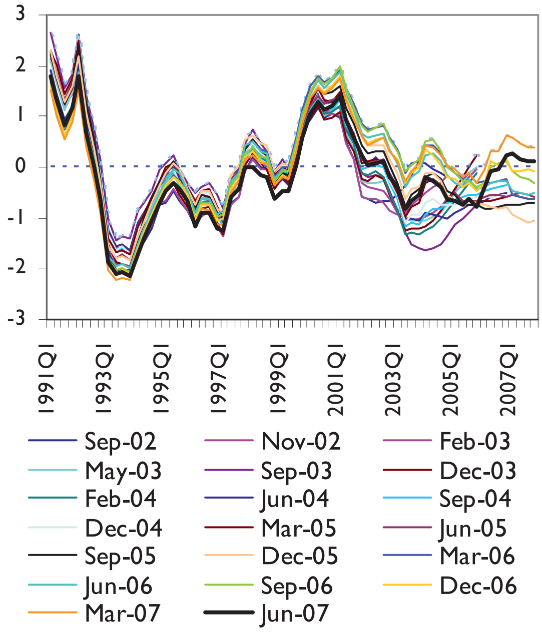
(percentage deviations from potential output)



Sources: Own calculations.

Chart F: Vintages of quarterly estimates of euro area output gap (UC PIC)

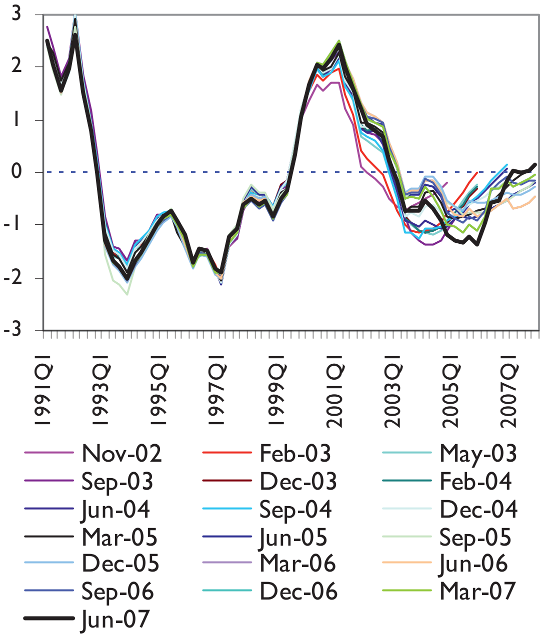
(percentage deviations from potential output)



Sources: Own calculations.

Chart G: Vintages of quarterly estimates of euro area output gap (UC BIV)

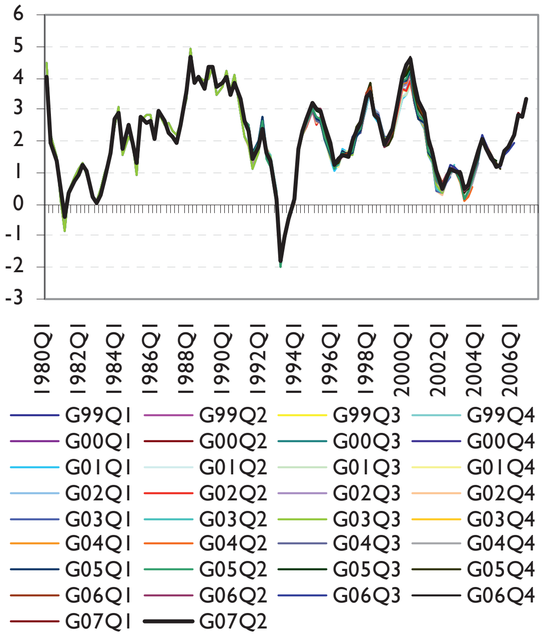
(percentage deviations from potential output)



Sources: Own calculations.

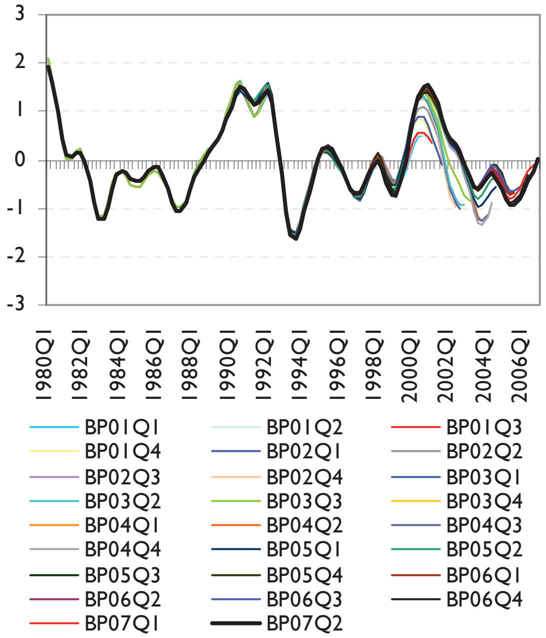
Chart H: Vintages of quarterly estimates of euro area real GDP growth

(annual percentage changes)



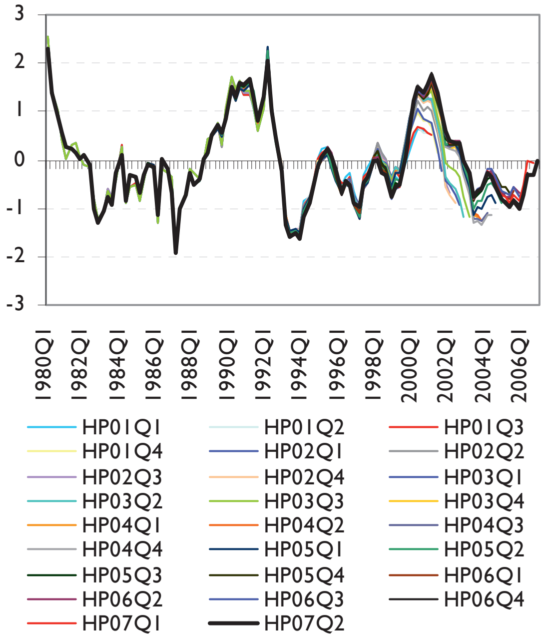
Sources: EABCN.

Chart I: Vintages of quarterly estimates of euro area output gap (band pass filter)
(percentage deviations from potential output)



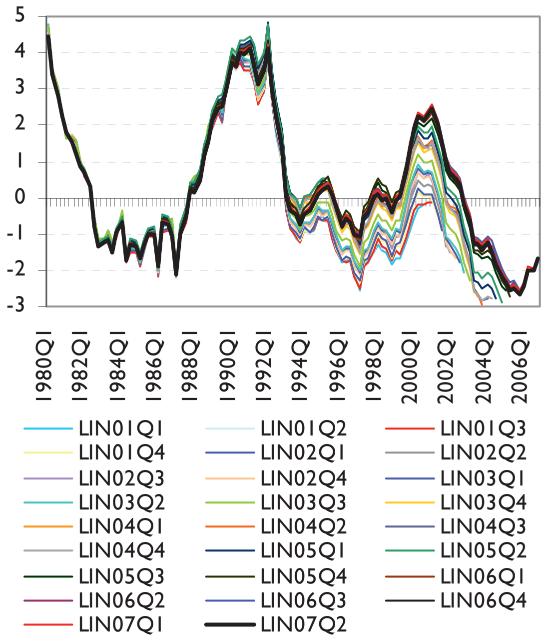
Sources: EABCN and own calculations.

Chart J: Vintages of quarterly estimates of euro area output gap (HP filter)
(percentage deviations from potential output)



Sources: EABCN and own calculations.

Chart K: Vintages of quarterly estimates of euro area output gap (linear trend filter)
(percentage deviations from potential output)



Sources: EABCN and own calculations.

Appendix III – Revisions for average output gap estimates

Table – Revisions to real time euro area output gap estimates

		mean	st dev	min	max	AR	corr	sign
Average All	Final	-0.61	0.47	-1.27	0.23	0.88	0.39	
	RT	-0.94	0.41	-1.46	0.01	0.91	-0.45	85.0%
	Rev RT	0.33	0.49	-0.26	1.32	0.96	0.64	
Average PFA	Final	-0.65	0.45	-1.28	0.29	0.90	0.29	
	RT	-1.06	0.40	-1.63	-0.25	0.88	-0.53	90.0%
	Rev RT	0.40	0.50	-0.25	1.46	0.94	0.65	
Average Org	Final	-0.74	0.65	-1.51	0.83	0.94	0.17	
	RT	-1.22	0.33	-1.81	-0.65	0.80	-0.33	85.0%
	Rev RT	0.49	0.68	-0.35	2.07	0.95	0.87	
Average UC	Final	-0.45	0.38	-1.00	0.19	0.85	0.69	
	RT	-0.50	0.38	-1.05	0.19	0.76	-0.39	75.0%
	Rev RT	0.04	0.30	-0.41	0.62	0.75	0.40	
Average Filters	Final	-0.74	0.61	-1.51	0.49	0.93	0.16	
	RT	-1.30	0.32	-1.81	-0.57	0.90	-0.36	85.0%
	Rev RT	0.56	0.64	-0.22	1.56	0.96	0.87	
Average Pseudo	Final	-0.74	0.61	-1.51	0.49	0.93	0.19	
	RT	-1.31	0.35	-1.91	-0.57	0.87	-0.36	85.0%
	Rev RT	0.58	0.64	-0.22	1.57	0.95	0.85	

Notes: Sample period is 2002:1 to 2006:4 in all cases (20 observations).

"AR" refers to the first order autocorrelation coefficient.

"Rev RT" stands for revision final estimate minus real time estimate.

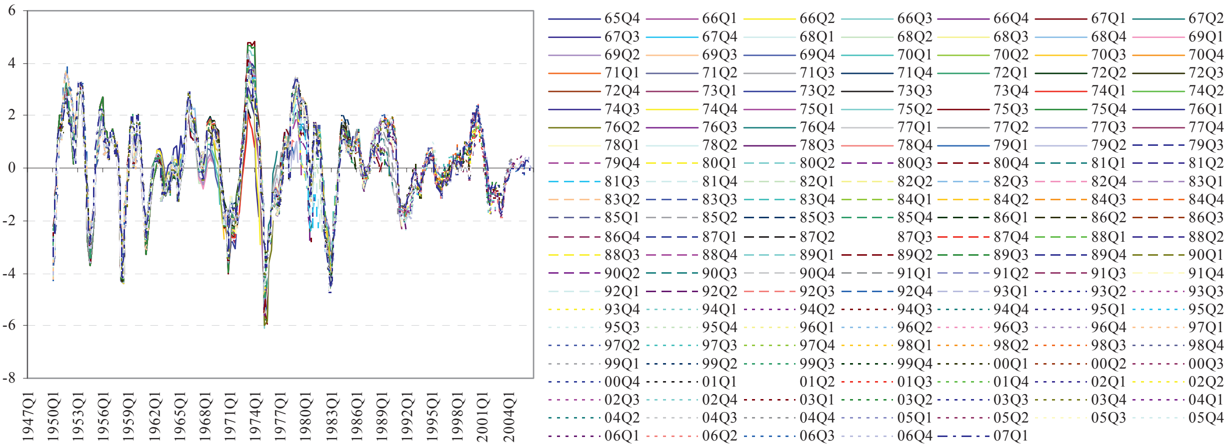
"sign" refers to the percentage of times the real time estimate has the same sign as the final estimate

"corr" reports the correlation between real time estimate and final estimate in the "Final" row, the correlation between real time estimate and the revision (final minus real time) in the "RT" row, and the correlation between final estimate and the revision (final minus real time) in the "Rev RT" row.

Appendix IV – Vintages of US output gap estimates

Chart A: Vintages of annual estimates of US output gap (deviations from HP trend)

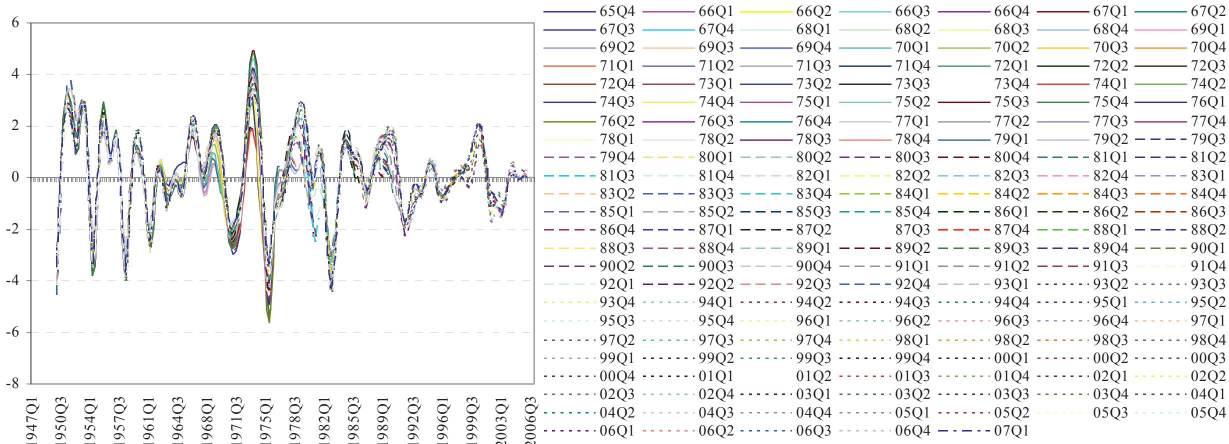
(percentage deviations from trend output)



Sources: RTDSM and own calculations. Note: Estimates are deviations from trend computed via the HP filter.

Chart B: Vintages of annual estimates of US output gap (band-pass cycles)

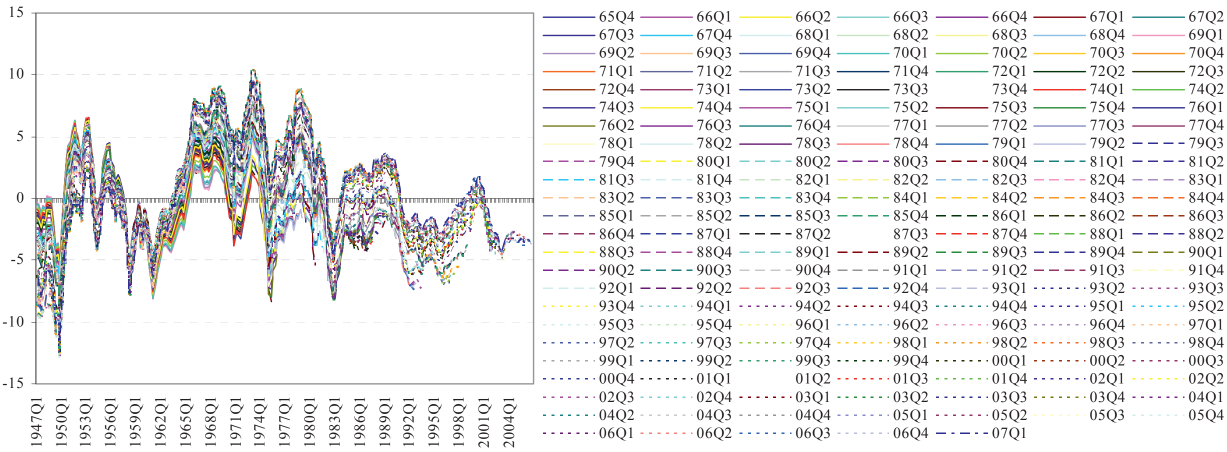
(percentage deviations from trend output)



Sources: RTDSM and own calculations. Note: Estimates are the cycles extracted via the band-pass filter.

Chart C: Vintages of annual estimates of US output gap (deviations from linear trend)

(percentage deviations from trend output)



Sources: RTDSM and own calculations. Note: Estimates are deviations from a linear trend.

Appendix V – Results for sub-sample inflation forecasting analysis

Table A – Tests of equal inflation forecast accuracy: one quarter ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
	forecast horizon h = 1 (one-quarter ahead)							
CAP-AV	2001:01-2006:04	1.570	1.621	-0.051	*-0.784*	*	*-0.758*	n.s.
	2001:01-2004:04	1.802	1.827	-0.024	*-0.256*	n.s.	*-0.213*	n.s.
	2005:01-2006:04	1.106	1.211	-0.105	*-2.614*	*	*-0.693*	*n.s.*
CAP-TR	2001:01-2006:04	1.570	1.626	-0.056	*-1.072*	*	*-0.828*	n.s.
	2001:01-2004:04	1.802	1.848	-0.046	*-0.622*	n.s.	*-0.396*	n.s.
	2005:01-2006:04	1.106	1.183	-0.077	*-1.538*	*	*-0.521*	*n.s.*
EC - T	1999:01-2006:04	1.342	1.427	-0.085	*-2.555*	*	*-1.910*	*
	1999:01-2002:04	1.191	1.290	-0.099	*-2.153*	*	*-1.223*	*
	2003:01-2006:04	1.492	1.564	-0.072	*-1.163*	*	*-0.734*	*n.s.*
EC - P	2002:04-2006:04	1.447	1.550	-0.103	*-1.334*	*	*-1.132*	*
	2002:04-2005:04	1.318	1.454	-0.136	*-1.339*	*	*-1.212*	*
	2006:01-2006:04	1.864	1.862	**0.002**	*0.017*	*n.s.*	*0.004*	*n.s.*
IMF	1999:01-2006:04	1.342	1.458	-0.116	*-2.465*	*	*-2.545*	*
	1999:01-2002:04	1.191	1.279	-0.088	*-1.729*	*	*-1.104*	*
	2003:01-2006:04	1.492	1.636	-0.144	*-1.637*	*	*-1.404*	*
OECD	1999:01-2006:04	1.342	1.472	-0.130	*-2.977*	*	*-2.834*	*
	1999:01-2002:04	1.191	1.289	-0.098	*-1.410*	*	*-1.214*	*
	2003:01-2006:04	1.492	1.655	-0.163	*-2.550*	*	*-1.575*	*
UC - CC	2002:03-2006:04	1.588	1.784	-0.196	*-2.592*	*	*-1.980*	*
	2002:03-2005:04	1.509	1.766	-0.256	*-2.847*	*	*-2.034*	*
	2006:01-2006:04	1.864	1.849	**0.015**	*0.120*	*n.s.*	*0.032*	*n.s.*
UC - PIC	2002:03-2006:04	1.588	1.834	-0.246	*-3.002*	*	*-2.416*	*
	2002:03-2005:04	1.509	1.775	-0.266	*-2.651*	*	*-2.098*	*
	2006:01-2006:04	1.864	2.041	-0.177	*-1.311*	*	*-0.346*	*n.s.*
UC - BIV	2002:04-2006:04	1.447	1.570	-0.123	*-2.053*	*	*-1.335*	*
	2002:04-2005:04	1.318	1.444	-0.126	*-1.806*	*	*-1.136*	*
	2006:01-2006:04	1.864	1.978	-0.114	*-0.937*	*n.s.*	*-0.230*	*n.s.*
BP	2001:01-2006:04	1.570	1.659	-0.089	*-3.268*	*	*-1.281*	*
	2001:01-2004:04	1.802	1.907	-0.104	*-2.980*	*	*-0.877*	n.s.
	2006:01-2006:04	1.864	1.978	-0.114	*-0.937*	*n.s.*	*-0.230*	*n.s.*
HP	2001:01-2006:04	1.570	1.641	-0.071	*-2.924*	*	*-1.040*	n.s.
	2001:01-2004:04	1.802	1.886	-0.083	*-2.632*	*	*-0.707*	n.s.
	2005:01-2006:04	1.106	1.153	-0.047	*-1.449*	*	*-0.325*	*n.s.*
LIN	2001:01-2006:04	1.570	1.623	-0.053	*-3.332*	*	*-0.782*	n.s.
	2001:01-2004:04	1.802	1.858	-0.056	*-2.894*	*	*-0.483*	n.s.
	2005:01-2006:04	1.106	1.152	-0.046	*-1.671*	*	*-0.321*	*n.s.*
AV-All	2001:01-2006:04	1.570	1.665	-0.095	*-1.425*	*	*-1.371*	*
	2001:01-2004:04	1.802	1.891	-0.089	*-0.921*	*	*-0.750*	n.s.
	2005:01-2006:04	1.106	1.214	-0.108	*-2.083*	*	*-0.713*	*n.s.*
AV-PFA	2001:01-2006:04	1.570	1.726	-0.156	*-2.337*	*	*-2.163*	*
	2001:01-2004:04	1.802	1.965	-0.162	*-1.742*	*	*-1.322*	*
	2005:01-2006:04	1.106	1.247	-0.142	*-1.959*	*	*-0.909*	*
AV-Org	2001:01-2006:04	1.570	1.683	-0.113	*-2.093*	*	*-1.607*	*
	2001:01-2004:04	1.802	1.916	-0.114	*-1.526*	*	*-0.950*	n.s.
	2005:01-2006:04	1.106	1.216	-0.110	*-1.766*	*	*-0.726*	*n.s.*
AV-UC	2001:01-2006:04	1.570	1.688	-0.117	*-1.752*	*	*-1.670*	*
	2001:01-2004:04	1.802	1.933	-0.131	*-1.317*	*	*-1.080*	*
	2005:01-2006:04	1.106	1.197	-0.091	*-2.965*	*	*-0.609*	*n.s.*
AV-Fil	2001:01-2006:04	1.570	1.644	-0.073	*-3.339*	*	*-1.071*	n.s.
	2001:01-2004:04	1.802	1.885	-0.082	*-2.965*	*	*-0.698*	n.s.
	2005:01-2006:04	1.106	1.161	-0.056	*-1.616*	*	*-0.383*	*n.s.*

Note: Sample period is 1985Q1-2006Q4. See Tables 4 and 12 in the main text for explanation of terms.

Table B – Tests of equal inflation forecast accuracy: one year ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
	forecast horizon h = 4 (one-year ahead)							
CAP-AV	2001:04-2006:04	0.898	0.881	**0.017**	*0.276*	n.s.	*0.415*	n.s.
	2001:04-2004:04	1.162	1.113	**0.049**	*0.507*	n.s.	*0.572*	n.s.
	2005:01-2006:04	0.468	0.502	-0.034	*-0.808*	*n.s.*	*-0.540*	*n.s.*
CAP-TR	2001:04-2006:04	0.898	0.922	-0.024	*-0.528*	n.s.	*-0.555*	n.s.
	2001:04-2004:04	1.162	1.165	-0.003	*-0.037*	n.s.	*-0.029*	n.s.
	2005:01-2006:04	0.468	0.528	-0.060	*-1.312*	*	*-0.906*	*n.s.*
EC - T	1999:04-2006:04	0.798	0.903	-0.105	*-2.144*	*	*-3.378*	*
	1999:04-2002:04	0.856	0.945	-0.089	*-1.575*	*	*-1.223*	*n.s.*
	2003:01-2006:04	0.751	0.869	-0.118	*-1.754*	*	*-2.179*	*n.s.*
EC - P	2003:03-2006:04	0.696	0.872	-0.176	*-2.214*	*	*-2.823*	*
	2003:03-2005:04	0.829	1.050	-0.221	*-2.368*	*	*-2.105*	*n.s.*
	2006:01-2006:04	0.364	0.426	-0.063	*-0.928*	*n.s.*	*-0.588*	*n.s.*
IMF	1999:04-2006:04	0.798	0.965	-0.167	*-2.141*	*	*-5.019*	*
	1999:04-2002:04	0.856	0.976	-0.120	*-1.153*	*	*-1.592*	*n.s.*
	2003:01-2006:04	0.751	0.957	-0.206	*-2.308*	*	*-3.440*	*
OECD	1999:04-2006:04	0.798	0.970	-0.172	*-2.100*	*	*-5.138*	*
	1999:04-2002:04	0.856	0.921	-0.065	*-0.636*	*n.s.*	*-0.920*	*n.s.*
	2003:01-2006:04	0.751	1.010	-0.259	*-4.487*	*	*-4.097*	*
UC - CC	2003:02-2006:04	0.751	0.997	-0.246	*-3.822*	*	*-3.700*	*
	2003:02-2005:04	0.892	1.186	-0.294	*-4.761*	*	*-2.729*	*
	2006:01-2006:04	0.364	0.477	-0.113	*-1.454*	*	*-0.948*	*n.s.*
UC - PIC	2003:02-2006:04	0.751	0.992	-0.241	*-3.567*	*	*-3.647*	*
	2003:02-2005:04	0.892	1.197	-0.305	*-4.785*	*	*-2.799*	*
	2006:01-2006:04	0.364	0.431	-0.067	*-1.347*	*	*-0.626*	*n.s.*
UC - BIV	2003:03-2006:04	0.696	0.803	-0.107	*-2.180*	*	*-1.863*	*n.s.*
	2003:03-2005:04	0.829	0.989	-0.160	*-4.221*	*	*-1.620*	*n.s.*
	2006:01-2006:04	0.364	0.337	**0.027**	*0.616*	*n.s.*	*0.319*	*n.s.*
BP	2001:04-2006:04	0.898	1.009	-0.111	*-2.441*	*	*-2.312*	n.s.
	2001:04-2004:04	1.162	1.329	-0.167	*-4.821*	*	*-1.629*	n.s.
	2005:01-2006:04	0.468	0.489	-0.021	*-2.587*	*	*-0.343*	*n.s.*
HP	2001:04-2006:04	0.898	0.961	-0.063	*-2.768*	*	*-1.380*	n.s.
	2001:04-2004:04	1.162	1.255	-0.093	*-4.543*	*	*-0.959*	n.s.
	2005:01-2006:04	0.468	0.484	-0.015	*-2.483*	*	*-0.254*	*n.s.*
LIN	2001:04-2006:04	0.898	0.961	-0.063	*-2.107*	*	*-1.383*	n.s.
	2001:04-2004:04	1.162	1.255	-0.092	*-2.337*	*	*-0.957*	n.s.
	2005:01-2006:04	0.468	0.484	-0.016	*-1.713*	*	*-0.265*	*n.s.*
AV-All	2001:04-2006:04	0.898	0.951	-0.053	*-0.663*	*	*-1.180*	n.s.
	2001:04-2004:04	1.162	1.221	-0.058	*-0.445*	n.s.	*-0.621*	n.s.
	2005:01-2006:04	0.468	0.514	-0.046	*-1.417*	*	*-0.709*	*n.s.*
AV-PFA	2001:04-2006:04	0.898	1.086	-0.188	*-1.880*	*	*-3.633*	*
	2001:04-2004:04	1.162	1.355	-0.193	*-1.283*	*	*-1.849*	n.s.
	2005:01-2006:04	0.468	0.648	-0.180	*-2.519*	*	*-2.220*	*
AV-Org	2001:04-2006:04	0.898	1.048	-0.150	*-1.889*	*	*-3.009*	*
	2001:04-2004:04	1.162	1.339	-0.177	*-1.403*	*	*-1.717*	n.s.
	2005:01-2006:04	0.468	0.575	-0.107	*-2.092*	*	*-1.485*	*n.s.*
AV-UC	2001:04-2006:04	0.898	1.049	-0.151	*-2.755*	*	*-3.020*	*
	2001:04-2004:04	1.162	1.333	-0.171	*-1.821*	*	*-1.664*	n.s.
	2005:01-2006:04	0.468	0.587	-0.119	*-5.122*	*	*-1.616*	*n.s.*
AV-Fil	2001:04-2006:04	0.898	0.980	-0.082	*-2.461*	*	*-1.753*	n.s.
	2001:04-2004:04	1.162	1.283	-0.120	*-3.254*	*	*-1.217*	n.s.
	2005:01-2006:04	0.468	0.488	-0.020	*-2.295*	*	*-0.321*	*n.s.*

Note: Sample period is 1985Q1-2006Q4. See Tables 4 and 12 in the main text for explanation of terms.

Table C – Tests of equal inflation forecast accuracy: two years ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 8 (two-years ahead)					
CAP-AV	2002:04-2006:04	1.101	1.028	**0.073**	*2.102*	*	*1.205*	n.s.
	2002:04-2004:04	1.301	1.16	**0.141**	*3.690*	*	*1.092*	n.s.
	2005:01-2006:04	0.876	0.880	-0.003	*-0.079*	*n.s.*	*-0.031*	*n.s.*
CAP-TR	2002:04-2006:04	1.101	1.044	**0.057**	*1.707*	*	*0.93*	n.s.
	2002:04-2004:04	1.301	1.157	**0.145**	*3.577*	*	*1.124*	n.s.
	2005:01-2006:04	0.876	0.917	-0.041	*-0.726*	*n.s.*	*-0.360*	*n.s.*
EC - T	2000:04-2006:04	1.604	1.546	**0.058**	*0.580*	*	*0.930*	*n.s.*
	2000:04-2002:04	2.828	2.520	**0.308**	*2.132*	*	*1.101*	*n.s.*
	2003:01-2006:04	0.915	0.999	-0.084	*-1.734*	*	*-1.338*	*n.s.*
EC - P	2004:03-2006:04	0.933	1.070	-0.137	*-1.751*	*	*-1.278*	*n.s.*
	2004:03-2005:04	1.215	1.427	-0.213	*-1.565*	*	*-0.894*	*n.s.*
	2006:01-2006:04	0.510	0.533	-0.023	*-1.226*	*	*-0.172*	*n.s.*
IMF	2000:04-2006:04	1.604	1.670	-0.067	*-0.767*	*	*-0.996*	*n.s.*
	2000:04-2002:04	2.828	2.867	-0.040	*-0.167*	*n.s.*	*-0.124*	*n.s.*
	2003:01-2006:04	0.915	0.997	-0.082	*-1.518*	*	*-1.312*	*n.s.*
OECD	2000:04-2006:04	1.604	1.750	-0.146	*-1.316*	*	*-2.090*	*n.s.*
	2000:04-2002:04	2.828	2.760	**0.068**	*0.293*	*n.s.*	*0.222*	*n.s.*
	2003:01-2006:04	0.915	1.182	-0.267	*-4.470*	*	*-3.613*	*n.s.*
UC - CC	2004:02-2006:04	1.013	1.299	-0.286	*-5.699*	*	*-2.422*	*n.s.*
	2004:02-2005:04	1.300	1.589	-0.289	*-4.833*	*	*-1.273*	*n.s.*
	2006:01-2006:04	0.510	0.791	-0.281	*-3.131*	*	*-1.420*	*n.s.*
UC - PIC	2004:02-2006:04	1.013	1.154	-0.141	*-4.739*	*	*-1.346*	*n.s.*
	2004:02-2005:04	1.300	1.458	-0.158	*-3.543*	*	*-0.759*	*n.s.*
	2006:01-2006:04	0.510	0.622	-0.112	*-3.847*	*	*-0.719*	*n.s.*
UC - BIV	2004:03-2006:04	0.933	1.033	-0.101	*-2.455*	*	*-0.973*	*n.s.*
	2004:03-2005:04	1.215	1.330	-0.116	*-1.682*	*	*-0.522*	*n.s.*
	2006:01-2006:04	0.510	0.588	-0.078	*-2.494*	*	*-0.528*	*n.s.*
BP	2002:04-2006:04	1.101	1.166	-0.065	*-3.860*	*	*-0.949*	n.s.
	2002:04-2004:04	1.301	1.400	-0.099	*-5.439*	*	*-0.635*	n.s.
	2005:01-2006:04	0.876	0.903	-0.027	*-4.133*	*	*-0.241*	*n.s.*
HP	2002:04-2006:04	1.101	1.092	**0.009**	*0.741*	*	*0.147*	n.s.
	2002:04-2004:04	1.301	1.298	**0.004**	*0.179*	n.s.	*0.024*	n.s.
	2005:01-2006:04	0.876	0.860	**0.016**	*1.148*	*	*0.150*	*n.s.*
LIN	2002:04-2006:04	1.101	1.128	-0.027	*-0.677*	*	*-0.406*	n.s.
	2002:04-2004:04	1.301	1.384	-0.083	*-1.817*	*	*-0.538*	n.s.
	2005:01-2006:04	0.876	0.840	**0.036**	*2.257*	*	*0.340*	*n.s.*
AV-All	2002:04-2006:04	1.101	1.032	**0.069**	*1.314*	*	*1.145*	n.s.
	2002:04-2004:04	1.301	1.137	**0.165**	*2.724*	*	*1.303*	n.s.
	2005:01-2006:04	0.876	0.914	-0.037	*-0.675*	*n.s.*	*-0.328*	*n.s.*
AV-PFA	2002:04-2006:04	1.101	1.219	-0.118	*-1.205*	*	*-1.645*	n.s.
	2002:04-2004:04	1.301	1.205	**0.096**	*1.199*	*	*0.716*	n.s.
	2005:01-2006:04	0.876	1.235	-0.358	*-3.698*	*	*-2.323*	*n.s.*
AV-Org	2002:04-2006:04	1.101	1.190	-0.089	*-1.097*	*	*-1.274*	n.s.
	2002:04-2004:04	1.301	1.203	**0.098**	*1.465*	*	*0.733*	n.s.
	2005:01-2006:04	0.876	1.176	-0.300	*-2.277*	*	*-2.040*	*n.s.*
AV-UC	2002:04-2006:04	1.101	1.197	-0.096	*-1.802*	*	*-1.366*	n.s.
	2002:04-2004:04	1.301	1.230	**0.071**	*1.321*	*	*0.523*	n.s.
	2005:01-2006:04	0.876	1.161	-0.285	*-4.490*	*	*-1.963*	*n.s.*
AV-Fil	2002:04-2006:04	1.101	1.127	-0.026	*-0.904*	*	*-0.387*	n.s.
	2002:04-2004:04	1.301	1.367	-0.065	*-2.004*	*	*-0.431*	n.s.
	2005:01-2006:04	0.876	0.857	**0.019**	*1.317*	*	*0.179*	*n.s.*

Note: Sample period is 1985Q1-2006Q4. See Tables 4 and 12 in the main text for explanation of terms.

Table D – Tests of equal inflation forecast accuracy: three years ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 12 (three-years ahead)					
CAP-AV	2003:04-2006:04	1.028	1.019	**0.010**	*0.162*	n.s.	*0.122*	n.s.
	2003:04-2004:04	1.312	1.134	**0.178**	*2.016*	*	*0.784*	n.s.
	2005:01-2006:04	0.851	0.946	-0.096	*-3.275*	*	*-0.809*	*n.s.*
CAP-TR	2003:04-2006:04	1.028	1.112	-0.084	*-1.422*	*	*-0.977*	n.s.
	2003:04-2004:04	1.312	1.213	**0.099**	*0.961*	*	*0.408*	n.s.
	2005:01-2006:04	0.851	1.048	-0.198	*-5.352*	*	*-1.508*	*n.s.*
EC - T	2001:04-2006:04	2.205	1.971	**0.234**	*1.745*	*	*2.487*	*n.s.*
	2001:04-2002:04	6.221	5.572	**0.649**	*5.188*	*	*0.583*	*n.s.*
	2003:01-2006:04	0.950	0.846	**0.104**	*1.034*	*	*1.959*	*n.s.*
EC - P	2005:03-2006:04	0.853	1.099	-0.247	*-3.470*	*	*-1.346*	*n.s.*
	2005:03-2005:04	0.648	0.470	**0.178**	*3.645*	*	*0.756*	*n.s.*
	2006:01-2006:04	0.955	1.414	-0.459	*-3.803*	*	*-1.298*	*n.s.*
IMF	2001:04-2006:04	2.205	2.239	-0.034	*-0.367*	*n.s.*	*-0.317*	*n.s.*
	2001:04-2002:04	6.221	6.712	-0.491	*-2.509*	*	*-0.366*	*n.s.*
	2003:01-2006:04	0.950	0.841	**0.109**	*1.326*	*	*2.078*	*n.s.*
OECD	2001:04-2006:04	2.205	2.353	-0.148	*-1.620*	*	*-1.324*	*n.s.*
	2001:04-2002:04	6.221	6.622	-0.401	*-2.713*	*	*-0.303*	*n.s.*
	2003:01-2006:04	0.950	1.019	-0.069	*-0.523*	*n.s.*	*-1.090*	*n.s.*
UC - CC	2005:02-2006:04	0.947	1.463	-0.516	*-9.151*	*	*-2.468*	*n.s.*
	2005:02-2005:04	0.937	1.324	-0.387	*-3.484*	*	*-0.878*	*n.s.*
	2006:01-2006:04	0.955	1.567	-0.612	*-12.187*	*	*-1.563*	*
UC - PIC	2005:02-2006:04	0.947	1.105	-0.158	*-8.771*	*	*-1.003*	*n.s.*
	2005:02-2005:04	0.937	1.013	-0.076	*-2.492*	*	*-0.226*	*n.s.*
	2006:01-2006:04	0.955	1.175	-0.220	*-6.679*	*	*-0.748*	*n.s.*
UC - BIV	2005:03-2006:04	0.853	1.102	-0.249	*-4.774*	*	*-1.357*	*n.s.*
	2005:03-2005:04	0.648	0.516	**0.131**	*2.964*	*	*0.509*	*n.s.*
	2006:01-2006:04	0.955	1.395	-0.439	*-5.716*	*	*-1.260*	*n.s.*
BP	2003:04-2006:04	1.028	1.025	**0.003**	*0.433*	n.s.	*0.034*	n.s.
	2003:04-2004:04	1.312	1.298	**0.014**	*5.128*	*	*0.055*	n.s.
	2005:01-2006:04	0.851	0.855	-0.005	*-0.647*	*n.s.*	*-0.043*	*n.s.*
HP	2003:04-2006:04	1.028	1.005	**0.023**	*2.819*	*	*0.298*	n.s.
	2003:04-2004:04	1.312	1.313	-0.001	*-0.119*	n.s.	*-0.004*	n.s.
	2005:01-2006:04	0.851	0.813	**0.038**	*3.839*	*	*0.376*	*n.s.*
LIN	2003:04-2006:04	1.028	1.041	-0.013	*-0.371*	n.s.	*-0.160*	n.s.
	2003:04-2004:04	1.312	1.465	-0.153	*-7.805*	*	*-0.522*	n.s.
	2005:01-2006:04	0.851	0.776	**0.075**	*4.839*	*	*0.771*	*n.s.*
AV-All	2003:04-2006:04	1.028	0.982	**0.047**	*0.786*	*	*0.616*	n.s.
	2003:04-2004:04	1.312	1.070	**0.242**	*2.944*	*	*1.132*	n.s.
	2005:01-2006:04	0.851	0.927	-0.076	*-3.629*	*	*-0.654*	*n.s.*
AV-PFA	2003:04-2006:04	1.028	1.124	-0.096	*-0.963*	*	*-1.110*	n.s.
	2003:04-2004:04	1.312	1.063	**0.249**	*4.205*	*	*1.171*	n.s.
	2005:01-2006:04	0.851	1.162	-0.311	*-7.398*	*	*-2.144*	*n.s.*
AV-Org	2003:04-2006:04	1.028	1.059	-0.031	*-0.355*	n.s.	*-0.375*	n.s.
	2003:04-2004:04	1.312	1.047	**0.264**	*5.499*	*	*1.263*	n.s.
	2005:01-2006:04	0.851	1.066	-0.215	*-4.105*	*	*-1.613*	*n.s.*
AV-UC	2003:04-2006:04	1.028	1.206	-0.177	*-2.792*	*	*-1.913*	n.s.
	2003:04-2004:04	1.312	1.380	-0.068	*-0.711*	n.s.	*-0.246*	n.s.
	2005:01-2006:04	0.851	1.097	-0.246	*-3.657*	*	*-1.793*	*n.s.*
AV-Fil	2003:04-2006:04	1.028	1.025	**0.003**	*0.170*	n.s.	*0.040*	n.s.
	2003:04-2004:04	1.312	1.386	-0.074	*-5.751*	*	*-0.266*	n.s.
	2005:01-2006:04	0.851	0.800	**0.051**	*5.011*	*	*0.512*	*n.s.*

Note: Sample period is 1985Q1-2006Q4. See Tables 4 and 12 in the main text for explanation of terms.

Appendix VI –Inflation forecasting analysis: real time versus pseudo real time vintages

Table – Tests of equal inflation forecast accuracy: real time versus pseudo real time vintages

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)	MSE-F
forecast horizon h = 1 (one-quarter ahead)					
BP	1.570	1.659	-0.089	*-3.268 **	*-1.281 **
HP	1.570	1.641	-0.071	*-2.924 **	*-1.040 n.s.*
LIN	1.570	1.623	-0.053	*-3.332 **	*-0.782 n.s.*
BP pseudo	1.698	1.750	-0.052	*-2.723 **	*-0.716 n.s.*
HP pseudo	1.698	1.719	-0.021	*-2.145 **	*-0.299 n.s.*
LIN pseudo	1.698	1.710	-0.012	*-1.280 **	*-0.166 n.s.*

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)	MSE-F
forecast horizon h = 4 (one-year ahead)					
BP	0.898	1.009	-0.111	*-2.441 **	*-2.312 n.s.*
HP	0.898	0.961	-0.063	*-2.768 **	*-1.380 n.s.*
LIN	0.898	0.961	-0.063	*-2.107 **	*-1.383 n.s.*
BP pseudo	1.021	1.109	-0.088	*-2.964 **	*-1.669 n.s.*
HP pseudo	1.021	1.035	-0.014	*-1.397 **	*-0.289 n.s.*
LIN pseudo	1.021	1.034	-0.013	*-0.619* n.s.	*-0.255 n.s.*

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)	MSE-F
forecast horizon h = 8 (two-years ahead)					
BP	1.101	1.166	-0.065	*-3.860 **	*-0.949 n.s.*
HP	1.101	1.092	**0.009**	*0.741 **	*0.147 n.s.*
LIN	1.101	1.128	-0.027	*-0.677 **	*-0.406 n.s.*
BP pseudo	1.178	1.239	-0.061	*-5.116 **	*-0.836 n.s.*
HP pseudo	1.178	1.137	**0.041**	*5.768 **	*0.619 n.s.*
LIN pseudo	1.178	1.133	0.045	*1.512 **	*0.680 n.s.*

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)	MSE-F
forecast horizon h = 12 (three-years ahead)					
BP	1.028	1.025	**0.003**	*0.433 n.s.*	*0.034 n.s.*
HP	1.028	1.005	**0.023**	*2.819 **	*0.298 n.s.*
LIN	1.028	1.041	-0.013	*-0.371 n.s.*	*-0.160 n.s.*
BP pseudo	1.039	1.066	-0.028	*-2.067 **	*-0.337 n.s.*
HP pseudo	1.039	0.993	**0.046**	*3.924 **	*0.599 n.s.*
LIN pseudo	1.039	0.956	0.083	*2.707 **	*1.124 n.s.*

Note: Sample period is 1985Q1-2006Q4.
See Tables 4 and 12 in the main text for explanation of terms.

Appendix VII –Inflation forecasting analysis: extended Phillips curve

Table A – Inflation forecast accuracy with extended Phillips curve: 1-quarter ahead

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
		forecast horizon h = 1 (one-quarter ahead)					
CAP-AV	1.412	1.502	-0.090	*-1.182*	*	*-1.442*	*
CAP-TR	1.412	1.513	-0.101	*-1.651*	*	*-1.607*	*
EC - T	1.252	1.331	-0.080	*-1.906*	*	*-1.918*	*
EC - P	1.317	1.428	-0.111	*-1.309*	*	*-1.317*	*
IMF	1.252	1.356	-0.104	*-1.915*	*	*-2.460*	*
OECD	1.252	1.386	-0.135	*-2.646*	*	*-3.109*	*
UC - CC	1.411	1.632	-0.220	*-2.909*	*	*-2.430*	*
UC - PIC	1.411	1.661	-0.250	*-3.289*	*	*-2.707*	*
UC - BIV	1.317	1.425	-0.108	*-1.932*	*	*-1.286*	*
BP	1.412	1.502	-0.090	*-3.759*	*	*-1.437*	*
HP	1.412	1.471	-0.060	*-2.531*	*	*-0.971*	*n.s.*
LIN	1.412	1.456	-0.044	*-2.638*	*	*-0.723*	*n.s.*
AV-All	1.412	1.509	-0.097	*-1.421*	*	*-1.544*	*
AV-PFA	1.412	1.559	-0.147	*-2.075*	*	*-2.261*	*
AV-Org	1.412	1.518	-0.107	*-1.715*	*	*-1.684*	*
AV-UC	1.412	1.517	-0.105	*-1.678*	*	*-1.659*	*
AV-Fil	1.412	1.474	-0.062	*-3.002*	*	*-1.672*	*

Table B – Inflation forecast accuracy with extended Phillips curve: 4-quarters ahead

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
		forecast horizon h = 4 (one-year ahead)					
CAP-AV	0.713	0.752	-0.039	*-0.685*	*	*-1.102*	n.s.
CAP-TR	0.713	0.796	-0.083	*-2.056*	*	*-2.198*	n.s.
EC - T	0.654	0.749	-0.095	*-1.820*	*	*-3.663*	*
EC - P	0.525	0.698	-0.173	*-2.304*	*	*-3.463*	*
IMF	0.654	0.797	-0.143	*-1.901*	*	*-5.201*	*
OECD	0.654	0.815	-0.161	*-1.883*	*	*-5.727*	*
UC - CC	0.548	0.810	-0.262	*-3.747*	*	*-4.852*	*
UC - PIC	0.548	0.719	-0.172	*-3.968*	*	*-3.578*	*
UC - BIV	0.525	0.598	-0.073	*-1.843*	*	*-1.714*	*n.s.*
BP	0.713	0.823	-0.110	*-2.550*	*	*-2.816*	*
HP	0.713	0.739	-0.026	*-0.644*	*	*-0.751*	*n.s.*
LIN	0.713	0.748	-0.035	*-0.743*	*	*-0.987*	*n.s.*
AV-All	0.713	0.758	-0.045	*-0.642*	n.s.	*-1.259*	n.s.
AV-PFA	0.713	0.861	-0.148	*-1.575*	*	*-3.618*	*
AV-Org	0.713	0.842	-0.129	*-1.609*	*	*-3.211*	*
AV-UC	0.713	0.835	-0.122	*-2.851*	*	*-3.067*	*
AV-Fil	0.713	0.763	-0.050	*-1.131*	*	*-1.388*	n.s.

Table C – Inflation forecast accuracy with extended Phillips curve: 8-quarters ahead

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 8 (two-years ahead)				
CAP-AV	0.833	0.815	**0.017**	0.515	n.s.	*0.364*	n.s.
CAP-TR	0.833	0.835	-0.002	-0.064	n.s.	*-0.041*	n.s.
EC - T	1.392	1.329	**0.062**	*0.620*	*	*1.174*	*n.s.*
EC - P	0.653	0.813	-0.161	*-2.108*	*	*-1.978*	*n.s.*
IMF	1.392	1.442	-0.051	*-0.565*	*	*-0.877*	*n.s.*
OECD	1.392	1.523	-0.131	*-1.120*	*	*-2.156*	*n.s.*
UC - CC	0.697	1.040	-0.343	*-7.128*	*	*-3.625*	*
UC - PIC	0.697	0.700	-0.003	*-0.063*	*n.s.*	*-0.041*	*n.s.*
UC - BIV	0.653	0.685	-0.033	*-0.550*	*n.s.*	*-0.478*	*n.s.*
BP	0.833	0.925	-0.092	*-7.926*	*	*-1.693*	*n.s.*
HP	0.833	0.764	**0.068**	*1.742*	*	*1.521*	*n.s.*
LIN	0.833	0.792	**0.041**	*0.756*	*	*0.875*	*n.s.*
AV-All	0.833	0.747	**0.086**	1.685	*	*1.950*	n.s.
AV-PFA	0.833	0.901	-0.068	-0.704	*	*-1.290*	n.s.
AV-Org	0.833	0.905	-0.072	-0.815	*	*-1.352*	n.s.
AV-UC	0.833	0.848	-0.015	-0.265	n.s.	*-0.298*	n.s.
AV-Fil	0.833	0.803	**0.030**	0.780	*	*0.641*	n.s.

Table D – Inflation forecast accuracy with extended Phillips curve: 12-quarters ahead

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 12 (three-years ahead)				
CAP-AV	0.712	0.834	-0.121	*-2.418*	*	*-1.893*	*n.s.*
CAP-TR	0.712	0.964	-0.252	*-4.160*	*	*-3.396*	*n.s.*
EC - T	1.938	1.766	**0.173**	*1.416*	*	*2.057*	*n.s.*
EC - P	0.467	0.788	-0.320	*-4.495*	*	*-2.439*	*n.s.*
IMF	1.938	2.021	-0.083	*-0.722*	*	*-0.862*	*n.s.*
OECD	1.938	2.113	-0.174	*-1.707*	*	*-1.733*	*n.s.*
UC - CC	0.508	0.989	-0.481	*-12.266*	*	*-3.404*	*
UC - PIC	0.508	0.504	**0.004**	*0.143*	*n.s.*	*0.057*	*n.s.*
UC - BIV	0.467	0.693	-0.225	*-3.702*	*	*-1.953*	*n.s.*
BP	0.712	0.736	-0.024	*-2.247*	*	*-0.415*	*n.s.*
HP	0.712	0.706	**0.006**	*0.130*	*n.s.*	*0.114*	*n.s.*
LIN	0.712	0.713	-0.001	*-0.011*	*n.s.*	*-0.015*	*n.s.*
AV-All	0.712	0.691	**0.021**	*0.449*	n.s.	*0.398*	*n.s.*
AV-PFA	0.712	0.763	-0.051	*-0.611*	n.s.	*-0.865*	*n.s.*
AV-Org	0.712	0.764	-0.052	*-0.584*	n.s.	*-0.881*	*n.s.*
AV-UC	0.712	0.919	-0.207	*-2.712*	*	*-2.928*	*n.s.*
AV-Fil	0.712	0.711	**0.001**	*0.026*	n.s.	*0.024*	*n.s.*

Appendix VIII –Inflation forecasting analysis: alternative reference series

Table A – Inflation forecast accuracy with alternative reference series: 1-quarter ahead

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 1 (one-quarter ahead)				
CAP-AV	1.709	1.743	-0.034	*-0.518*	*n.s.*	*-0.468*	*n.s.*
CAP-TR	1.709	1.755	-0.045	*-0.790*	*	*-0.620*	*n.s.*
EC - T	1.528	1.606	-0.077	*-1.828*	*	*-1.540*	*
EC - P	1.292	1.489	-0.197	*-2.912*	*	*-2.253*	*
IMF	1.528	1.627	-0.099	*-1.599*	*	*-1.945*	*
OECD	1.528	1.662	-0.133	*-2.425*	*	*-2.567*	*
UC - CC	1.357	1.562	-0.205	*-3.023*	*	*-2.362*	*
UC - PIC	1.357	1.651	-0.294	*-3.812*	*	*-3.204*	*
UC - BIV	1.292	1.453	-0.161	*-2.970*	*	*-1.880*	*
BP	1.709	1.784	-0.075	*-2.687*	*	*-1.009*	*n.s.*
HP	1.709	1.769	-0.060	*-2.499*	*	*-0.817*	*n.s.*
LIN	1.709	1.759	-0.050	*-2.249*	*	*-0.677*	*n.s.*
AV-All	1.709	1.819	-0.110	*-1.546*	*	*-1.454*	*
AV-PFA	1.709	1.892	-0.183	*-2.442*	*	*-2.320*	*
AV-Org	1.709	1.853	-0.144	*-2.246*	*	*-1.862*	*
AV-UC	1.709	1.838	-0.129	*-2.245*	*	*-1.679*	*
AV-Fil	1.709	1.775	-0.065	*-2.535*	*	*-1.155*	*n.s.*

Table B – Inflation forecast accuracy with alternative reference series: 4-quarters ahead

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 4 (one-year ahead)				
CAP-AV	0.940	0.969	-0.030	*-0.845*	*	*-0.646*	*n.s.*
CAP-TR	0.940	0.967	-0.028	*-0.673*	*	*-0.604*	*n.s.*
EC - T	0.786	0.910	-0.123	*-2.081*	*	*-3.935*	*
EC - P	0.340	0.507	-0.167	*-2.118*	*	*-4.610*	*
IMF	0.786	0.928	-0.142	*-1.779*	*	*-4.422*	*
OECD	0.786	0.976	-0.190	*-2.494*	*	*-5.636*	*
UC - CC	0.481	0.697	-0.216	*-2.411*	*	*-4.651*	*
UC - PIC	0.481	0.722	-0.242	*-2.183*	*	*-5.018*	*
UC - BIV	0.340	0.436	-0.096	*-1.616*	*	*-3.085*	*
BP	0.940	1.069	-0.129	*-2.100*	*	*-2.541*	*n.s.*
HP	0.940	1.016	-0.077	*-2.541*	*	*-1.585*	*n.s.*
LIN	0.940	1.013	-0.074	*-2.310*	*	*-1.526*	*n.s.*
AV-All	0.940	1.063	-0.123	*-1.847*	*	*-2.431*	*
AV-PFA	0.940	1.151	-0.212	*-2.186*	*	*-3.862*	*
AV-Org	0.940	1.121	-0.181	*-2.104*	*	*-3.398*	*
AV-UC	0.940	1.144	-0.204	*-5.301*	*	*-3.745*	*
AV-Fil	0.940	1.035	-0.096	*-2.384*	*	*-1.943*	*n.s.*

Table C – Inflation forecast accuracy with alternative reference series: 8-quarters ahead

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 8 (two-years ahead)				
CAP-AV	0.951	0.894	**0.057**	*2.370*	*	*1.090*	*n.s.*
CAP-TR	0.951	0.907	**0.044**	*1.343*	*	*0.822*	*n.s.*
EC - T	1.162	1.080	**0.082**	*1.538*	*	*1.892*	*n.s.*
EC - P	0.496	0.555	-0.059	*-1.410*	*	*-1.070*	*n.s.*
IMF	1.162	1.159	**0.003**	*0.044*	*n.s.*	*0.063*	*n.s.*
OECD	1.162	1.257	-0.095	*-1.375*	*	*-1.895*	*n.s.*
UC - CC	0.555	0.809	-0.254	*-4.927*	*	*-3.455*	*
UC - PIC	0.555	0.657	-0.102	*-4.119*	*	*-1.711*	*n.s.*
UC - BIV	0.496	0.549	-0.053	*-2.106*	*	*-0.966*	*n.s.*
BP	0.951	1.022	-0.070	*-2.802*	*	*-1.172*	*n.s.*
HP	0.951	0.934	**0.018**	*1.579*	*	*0.324*	*n.s.*
LIN	0.951	0.969	-0.018	*-0.649*	*	*-0.315*	*n.s.*
AV-All	0.951	0.872	**0.079**	*2.267*	*	*1.545*	*n.s.*
AV-PFA	0.951	1.006	-0.054	*-0.794*	*	*-0.917*	*n.s.*
AV-Org	0.951	0.965	-0.014	*-0.306*	*n.s.*	*-0.245*	*n.s.*
AV-UC	0.951	1.045	-0.093	*-2.541*	*	*-1.518*	*n.s.*
AV-Fil	0.951	0.970	-0.018	*-0.931*	*	*-0.321*	*n.s.*

Table D – Inflation forecast accuracy with alternative reference series: 12-quarters ahead

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 12 (three-years ahead)				
CAP-AV	1.430	1.397	**0.033**	*0.496*	*n.s.*	*0.303*	*n.s.*
CAP-TR	1.430	1.459	-0.030	*-0.375*	*n.s.*	*-0.264*	*n.s.*
EC - T	1.921	1.649	**0.272**	*2.273*	*	*3.461*	*n.s.*
EC - P	1.101	1.348	-0.247	*-2.939*	*	*-1.101*	*n.s.*
IMF	1.921	1.857	**0.065**	*0.775*	*	*0.732*	*n.s.*
OECD	1.921	1.943	-0.022	*-0.237*	*n.s.*	*-0.237*	*n.s.*
UC - CC	1.056	1.539	-0.483	*-8.621*	*	*-2.198*	*n.s.*
UC - PIC	1.056	1.199	-0.143	*-6.228*	*	*-0.835*	*n.s.*
UC - BIV	1.101	1.346	-0.245	*-3.851*	*	*-1.091*	*n.s.*
BP	1.430	1.432	-0.002	*-0.346*	*n.s.*	*-0.017*	*n.s.*
HP	1.430	1.380	**0.049**	*4.591*	*	*0.465*	*n.s.*
LIN	1.430	1.425	**0.005**	*0.163*	*n.s.*	*0.046*	*n.s.*
AV-All	1.430	1.322	**0.108**	*1.486*	*	*1.062*	*n.s.*
AV-PFA	1.430	1.418	**0.012**	*0.103*	*n.s.*	*0.108*	*n.s.*
AV-Org	1.430	1.364	**0.066**	*0.657*	*	*0.630*	*n.s.*
AV-UC	1.430	1.641	-0.211	*-3.103*	*	*-1.675*	*n.s.*
AV-Fil	1.430	1.411	**0.019**	*1.104*	*	*0.173*	*n.s.*

Appendix IX – Sub-sample inflation forecasting analysis for the US

Table A – Tests of equal inflation forecast accuracy: one quarter ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
				forecast horizon h = 1 (one-quarter ahead)				
BP	1965:04-2006:03	1.928	1.759	**0.169**	*0.953*	*	*15.747*	*
	1965:04-1984:04	3.299	2.982	**0.316**	*0.848*	*	*8.172*	*
	1985:01-2006:03	0.715	0.677	**0.038**	*0.906*	*	*4.923*	*
HP	1965:04-2006:03	1.928	1.763	**0.165**	*0.686*	*	*15.304*	*
	1965:04-1984:04	3.299	2.973	**0.325**	*0.641*	*	*8.428*	*
	1985:01-2006:03	0.715	0.693	**0.022**	*0.535*	*	*2.786*	*
LIN	1965:04-2006:03	1.928	1.754	**0.174**	*0.726*	*	*16.269*	*
	1965:04-1984:04	3.299	2.848	**0.451**	*0.903*	*	*12.196*	*
	1985:01-2006:03	0.715	0.786	-0.071	*-1.139*	*	*-7.880*	*
AV-Fil	1965:04-2006:03	1.928	1.765	**0.163**	*0.725*	*	*15.155*	*
	1965:04-1984:04	3.299	2.902	**0.396**	*0.844*	*	*10.517*	*
	1985:01-2006:03	0.715	0.758	-0.043	*-0.682*	*	*-4.979*	*
BP - 2001	2001:01-2006:03	0.845	0.808	**0.038**	*0.365*	*n.s.*	*1.074*	*n.s.*
	2001:01-2004:04	0.928	0.897	**0.031**	*0.208*	*n.s.*	*0.549*	*n.s.*
	2005:01-2006:03	0.658	0.604	**0.054**	*2.073*	*	*0.621*	*n.s.*
HP - 2001	2001:01-2006:03	0.845	0.863	-0.018	*-0.102*	*n.s.*	*-0.473*	*n.s.*
	2001:01-2004:04	0.928	0.978	-0.051	*-0.203*	*n.s.*	*-0.826*	*n.s.*
	2005:01-2006:03	0.658	0.601	**0.057**	*1.829*	*	*0.665*	*n.s.*
LIN - 2001	2001:01-2006:03	0.845	0.890	-0.044	*-0.326*	*n.s.*	*-1.145*	*n.s.*
	2001:01-2004:04	0.928	1.015	-0.087	*-0.456*	*n.s.*	*-1.375*	*n.s.*
	2005:01-2006:03	0.658	0.604	**0.054**	*0.784*	*n.s.*	*0.624*	*n.s.*
AV-Fil - 2001	2001:01-2006:03	0.845	0.867	-0.021	*-0.138*	n.s.	*-0.569*	n.s.
	2001:01-2004:04	0.928	0.990	-0.062	*-0.283*	n.s.	*-1.009*	n.s.
	2006:01-2006:04	1.864	1.978	-0.114	*-0.937*	*n.s.*	*-0.230*	*n.s.*

Note: Sample period is 1965Q1-2006Q4.
See Tables 4 and 12 in the main text for explanation of terms.

Table B – Tests of equal inflation forecast accuracy: one year ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 4 (one-year ahead)					
BP	1966:03-2006:03	1.993	1.632	**0.361**	*1.257*	*	*35.575*	*
	1966:03-1984:04	3.726	2.863	**0.864**	*1.550*	*	*22.327*	*
	1985:01-2006:03	0.518	0.585	-0.067	*-0.525*	*	*-9.990*	*
HP	1966:03-2006:03	1.993	1.690	**0.302**	*1.136*	*	*28.803*	*
	1966:03-1984:04	3.726	2.979	**0.747**	*1.442*	*	*18.558*	*
	1985:01-2006:03	0.518	0.594	-0.076	*-0.674*	*	*-11.109*	*
LIN	1966:03-2006:03	1.993	1.802	**0.191**	*0.555*	*	*17.056*	*
	1966:03-1984:04	3.726	3.017	**0.709**	*1.034*	*	*17.384*	*
	1985:01-2006:03	0.518	0.768	-0.250	*-1.715*	*	*-28.286*	*
AV-Fil	1966:03-2006:03	1.993	1.722	**0.271**	*0.787*	*	*25.334*	*
	1966:03-1984:04	3.726	2.899	**0.828**	*1.212*	*	*21.131*	*
	1985:01-2006:03	0.518	0.721	-0.203	*-1.708*	*	*-24.454*	*
BP - 2001	2001:04-2006:03	0.522	0.484	**0.038**	*0.363*	*n.s.*	*1.568*	*n.s.*
	2001:04-2004:04	0.736	0.677	**0.059**	*0.389*	*n.s.*	*1.133*	*n.s.*
	2005:01-2006:03	0.124	0.125	-0.001	*-0.051*	*n.s.*	*-0.067*	*n.s.*
HP - 2001	2001:04-2006:03	0.522	0.647	-0.125	*-0.858*	*	*-3.864*	*
	2001:04-2004:04	0.736	0.926	-0.191	*-0.891*	*	*-2.676*	*
	2005:01-2006:03	0.124	0.127	-0.003	*-0.393*	*n.s.*	*-0.150*	*n.s.*
LIN - 2001	2001:04-2006:03	0.522	0.724	-0.203	*-1.216*	*	*-5.594*	*
	2001:04-2004:04	0.736	1.032	-0.296	*-1.183*	*	*-3.730*	*
	2005:01-2006:03	0.124	0.153	-0.029	*-3.017*	*	*-1.326*	*n.s.*
AV-Fil - 2001	2001:04-2006:03	0.522	0.683	-0.162	*-0.938*	*	*-4.728*	*
	2001:04-2004:04	0.736	0.980	-0.244	*-0.942*	*	-3.240	n.s.
	2005:01-2006:03	0.124	0.132	-0.008	*-1.443*	*	*-0.415*	*n.s.*

Note: Sample period is 1965Q1-2006Q4.
See Tables 4 and 12 in the main text for explanation of terms.

Table C – Tests of equal inflation forecast accuracy: two years ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 8 (two-years ahead)					
BP	1967:03-2006:03	4.776	3.416	**1.360**	*1.621*	*	*62.520*	*
	1967:03-1984:04	9.473	5.878	**3.595**	*2.456*	*	*42.810*	*
	1985:01-2006:03	0.997	1.434	-0.438	*-0.958*	*	*-26.554*	*
HP	1967:03-2006:03	4.776	3.355	**1.421**	*1.781*	*	*66.481*	*
	1967:03-1984:04	9.473	5.892	**3.581**	*2.580*	*	*42.540*	*
	1985:01-2006:03	0.997	1.314	-0.317	*-0.952*	*	*-21.006*	*
LIN	1967:03-2006:03	4.776	3.891	**0.885**	*0.860*	*	*35.716*	*
	1967:03-1984:04	9.473	6.067	**3.406**	*1.883*	*	*39.302*	*
	1985:01-2006:03	0.997	2.140	-1.143	*-2.979*	*	*-46.483*	*
AV-Fil	1967:03-2006:03	4.776	3.415	**1.361**	*1.400*	*	*62.555*	*
	1967:03-1984:04	9.473	5.581	**3.892**	*2.278*	*	*48.817*	*
	1985:01-2006:03	0.997	1.673	-0.676	*-1.932*	*	*-35.160*	*
BP - 2001	2002:04-2006:03	0.891	0.685	**0.206**	*6.044*	*	*4.816*	*n.s.*
	2002:04-2004:04	1.125	0.870	**0.255**	*3.898*	*	*2.635*	*n.s.*
	2005:01-2006:03	0.591	0.447	**0.144**	*3.490*	*	*2.251*	*n.s.*
HP - 2001	2002:04-2006:03	0.891	0.949	-0.058	*-0.923*	*	*-0.983*	*n.s.*
	2002:04-2004:04	1.125	1.318	-0.193	*-1.721*	*	*-1.320*	*n.s.*
	2005:01-2006:03	0.591	0.476	**0.115**	*2.666*	*	*1.696*	*n.s.*
LIN - 2001	2002:04-2006:03	0.891	1.456	-0.565	*-3.279*	*	*-6.207*	*
	2002:04-2004:04	1.125	1.797	-0.672	*-2.151*	*	*-3.366*	*n.s.*
	2005:01-2006:03	0.591	1.018	-0.427	*-3.021*	*	*-2.937*	*n.s.*
AV-Fil - 2001	2002:04-2006:03	0.891	1.146	-0.255	*-2.429*	*	*-3.556*	n.s.
	2002:04-2004:04	1.125	1.498	-0.373	*-1.955*	*	*-2.242*	n.s.
	2005:01-2006:03	0.591	0.693	-0.102	*-1.187*	*	*-1.034*	*n.s.*

Note: Sample period is 1965Q1-2006Q4.
See Tables 4 and 12 in the main text for explanation of terms.

Table D – Tests of equal inflation forecast accuracy: three years ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 12 (three-years ahead)					
BP	1968:03-2006:03	7.384	6.977	**0.406**	*0.797*	*	*8.911*	*
	1968:03-1984:04	13.018	12.004	**1.014**	*1.360*	*	*5.576*	*
	1985:01-2006:03	3.110	3.164	-0.055	*-0.111*	*n.s.*	*-1.505*	*n.s.*
HP	1968:03-2006:03	7.384	6.541	**0.842**	*1.559*	*	*19.702*	*
	1968:03-1984:04	13.018	11.160	**1.859**	*2.066*	*	*10.992*	*
	1985:01-2006:03	3.110	3.038	**0.071**	*0.246*	*n.s.*	*2.046*	*n.s.*
LIN	1968:03-2006:03	7.384	6.303	**1.081**	*0.894*	*	*26.236*	*
	1968:03-1984:04	13.018	10.383	**2.635**	*1.675*	*	*16.753*	*
	1985:01-2006:03	3.110	3.208	-0.099	*-0.099*	*n.s.*	*-2.672*	*n.s.*
AV-Fil	1968:03-2006:03	7.384	6.217	**1.167**	*1.415*	*	*28.720*	*
	1968:03-1984:04	13.018	10.561	**2.457**	*2.032*	*	*15.352*	*
	1985:01-2006:03	3.110	2.921	**0.189**	*0.314*	*n.s.*	*5.618*	*
BP - 2001	2003:04-2006:03	1.409	1.374	**0.035**	*0.174*	*n.s.*	*0.304*	*n.s.*
	2003:04-2004:04	0.453	1.187	-0.734	*-16.708*	*	*-3.093*	*n.s.*
	2005:01-2006:03	2.092	1.508	**0.584**	*12.854*	*	*2.711*	*n.s.*
HP - 2001	2003:04-2006:03	1.409	1.611	-0.202	*-2.117*	*	*-1.504*	*n.s.*
	2003:04-2004:04	0.453	0.924	-0.471	*-16.232*	*	*-2.549*	*
	2005:01-2006:03	2.092	2.102	-0.010	*-0.106*	*n.s.*	*-0.033*	*n.s.*
LIN - 2001	2003:04-2006:03	1.409	2.332	-0.923	*-8.464*	*	*-4.749*	*n.s.*
	2003:04-2004:04	0.453	1.215	-0.762	*-6.939*	*	*-3.136*	*n.s.*
	2005:01-2006:03	2.092	3.130	-1.038	*-7.523*	*	*-2.322*	*n.s.*
AV-Fil - 2001	2003:04-2006:03	1.409	1.938	-0.529	*-6.124*	*	*-3.275*	n.s.
	2003:04-2004:04	0.453	1.241	-0.788	*-9.849*	*	*-3.175*	n.s.
	2005:01-2006:03	2.092	2.436	-0.344	*-3.633*	*	*-0.988*	*n.s.*

Note: Sample period is 1965Q1-2006Q4.
See Tables 4 and 12 in the main text for explanation of terms.

Appendix X –Real GDP growth forecasting analysis: real time versus pseudo real time vintages

Table – Tests of equal real GDP growth forecast accuracy: real time versus pseudo real time vintages

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
forecast horizon h = 1 (one-quarter ahead)							
BP	1.574	3.291	-1.717	*-1.980*	*	*-12.520*	*
HP	1.574	7.195	-5.621	*-3.981*	*	*-18.749*	*
LIN	1.574	1.517	**0.058**	*0.866*	*	*0.910*	*n.s.*
BP pseudo	1.476	3.254	-1.778	*-2.146*	*	*-13.110*	*
HP pseudo	1.476	4.770	-3.293	*-3.049*	*	*-16.571*	*
LIN pseudo	1.476	1.394	**0.082**	*0.129*	n.s.	*1.415*	*n.s.*
forecast horizon h = 4 (one-year ahead)							
BP	0.838	0.842	-0.003	*-0.024*	*n.s.*	*-0.085*	*n.s.*
HP	0.838	7.750	-6.912	*-3.784*	*	*-18.729*	*
LIN	0.838	0.753	**0.085**	*1.622*	*	*2.375*	*n.s.*
BP pseudo	0.892	0.931	-0.039	*-0.230*	n.s.	*-0.882*	*n.s.*
HP pseudo	0.892	2.964	-2.073	*-1.200*	*	*-14.684*	*
LIN pseudo	0.892	1.043	-0.151	*-0.493*	n.s.	*-3.048*	*n.s.*
forecast horizon h = 8 (two-years ahead)							
BP	0.707	0.598	**0.108**	*0.986*	*	*3.077*	*n.s.*
HP	0.707	4.109	-3.402	*-2.639*	*	*-14.077*	*
LIN	0.707	0.599	**0.107**	*4.123*	*	*3.047*	*n.s.*
BP pseudo	0.933	0.678	**0.255**	*2.536*	*	*6.400*	*n.s.*
HP pseudo	0.933	0.815	**0.118**	*0.293*	n.s.	*2.464*	*n.s.*
LIN pseudo	0.933	2.531	-1.598	*-9.390*	*	*-10.734*	*
forecast horizon h = 12 (three-years ahead)							
BP	0.626	0.574	**0.052**	*0.469*	*n.s.*	*1.187*	*n.s.*
HP	0.626	1.363	-0.736	*-1.421*	*	*-7.025*	*n.s.*
LIN	0.626	0.579	**0.047**	*1.052*	*	*1.059*	*n.s.*
BP pseudo	0.820	0.646	**0.174**	*1.737*	*	*3.500*	*n.s.*
HP pseudo	0.820	0.539	**0.281**	*2.415*	*	*6.794*	*n.s.*
LIN pseudo	0.820	1.644	-0.824	*-7.558*	*	*-6.516*	*n.s.*

Note: Sample period is 1985Q1-2006Q4.

Appendix XI –Real GDP forecasting analysis: alternative reference series

Table A – Real GDP growth forecast accuracy with alternative reference series: 1-quarter ahead

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 1 (one-quarter ahead)				
CAP-AV	1.454	1.457	-0.003	*-0.008*	*n.s.*	*-0.046*	*n.s.*
CAP-TR	1.454	1.495	-0.041	*-0.105*	*n.s.*	*-0.662*	*n.s.*
EC - T	1.454	2.014	-0.560	*-1.334*	*	*-6.673*	*
EC - P	1.226	1.278	-0.052	*-0.121*	*n.s.*	*-0.686*	*n.s.*
IMF	1.454	1.990	-0.536	*-1.253*	*	*-6.461*	*
OECD	1.454	1.745	-0.291	*-0.818*	*	*-4.002*	*
UC - CC	1.162	2.346	-1.184	*-1.816*	*	*-9.082*	*
UC - PIC	1.162	9.627	-8.465	*-2.429*	*	*-15.827*	*
UC - BIV	1.226	4.887	-3.661	*-2.785*	*	*-12.734*	*
BP	1.454	3.626	-2.172	*-2.519*	*	*-14.377*	*
HP	1.454	6.258	-4.804	*-3.538*	*	*-18.423*	*
LIN	1.454	1.384	**0.070**	*1.183*	*	*1.209*	*n.s.*
AV-All	1.454	4.286	-2.832	*-3.219*	*	*-15.858*	*
AV-PFA	1.454	2.492	-1.038	*-1.987*	*	*-9.995*	*
AV-Org	1.454	1.869	-0.415	*-1.068*	*	*-5.331*	*
AV-UC	1.454	1.652	-0.198	*-0.476*	*n.s.*	*-2.880*	*
AV-Fil	1.454	13.262	-11.808	*-4.794*	*	*-1.355*	*

Table B – Real GDP growth forecast accuracy with alternative reference series: 4-quarters ahead

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 4 (one-year ahead)				
CAP-AV	1.085	0.816	**0.270**	*1.023*	*	**6.947**	*
CAP-TR	1.085	1.063	**0.022**	*0.265*	*n.s.*	*0.444*	*n.s.*
EC - T	1.085	1.785	-0.699	*-2.222*	*	*-8.229*	*
EC - P	0.671	1.085	-0.414	*-1.225*	*	*-5.342*	*
IMF	1.085	1.849	-0.763	*-2.277*	*	*-8.671*	*
OECD	1.085	1.319	-0.234	*-1.268*	*	*-3.721*	*
UC - CC	0.795	0.848	-0.053	*-0.493*	*n.s.*	*-0.932*	*n.s.*
UC - PIC	0.795	4.201	-3.406	*-1.974*	*	*-12.161*	*
UC - BIV	0.671	0.704	-0.033	*-0.284*	*n.s.*	*-0.665*	*n.s.*
BP	1.085	1.060	**0.026**	*0.146*	*n.s.*	**0.511**	*n.s.*
HP	1.085	7.155	-6.070	*-2.951*	*	*-17.815*	*
LIN	1.085	0.987	**0.098**	*1.497*	*	**2.084**	*n.s.*
AV-All	1.085	0.896	**0.189**	*1.070*	*	**4.426**	*
AV-PFA	1.085	1.431	-0.345	*-1.648*	*	*-5.070*	*
AV-Org	1.085	1.583	-0.497	*-1.955*	*	*-6.600*	*
AV-UC	1.085	0.579	**0.507**	*1.052*	*	**18.385**	*
AV-Fil	1.085	1.604	-0.519	*-0.567*	*n.s.*	*-6.792*	*

Table C – Real GDP growth forecast accuracy with alternative reference series: 8-quarters ahead

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
		forecast horizon h = 8 (two-years ahead)					
CAP-AV	0.909	0.321	**0.588**	*1.998*	*	**31.148**	*
CAP-TR	0.909	0.84	**0.069**	*1.06*	*	*1.398*	*n.s.*
EC - T	0.909	1.134	-0.225	*-5.476*	*	*-3.370*	*n.s.*
EC - P	0.280	0.803	-0.523	*-7.737*	*	*-6.514*	*n.s.*
IMF	0.909	1.732	-0.823	*-2.960*	*	*-8.075*	*n.s.*
OECD	0.909	0.844	**0.065**	*0.400*	*n.s.*	*1.318*	*n.s.*
UC - CC	0.323	0.281	**0.042**	*0.724*	*n.s.*	*1.662*	*n.s.*
UC - PIC	0.323	2.105	-1.782	*-1.979*	*	*-9.312*	*
UC - BIV	0.280	0.726	-0.446	*-1.478*	*	*-6.143*	*n.s.*
BP	0.909	0.684	**0.225**	*1.602*	*	*5.598*	*n.s.*
HP	0.909	3.781	-2.872	*-2.009*	*	*-12.912*	*
LIN	0.909	0.780	**0.129**	*3.665*	*	*2.821*	*n.s.*
AV-All	0.909	0.425	**0.484**	*1.763*	*	**19.363**	*
AV-PFA	0.909	0.886	**0.024**	*0.162*	*n.s.*	*0.451*	*n.s.*
AV-Org	0.909	1.126	-0.217	*-1.529*	*	*-3.273*	*n.s.*
AV-UC	0.909	0.793	**0.116**	*0.326*	*n.s.*	**2.497**	*n.s.*
AV-Fil	0.909	2.693	-1.784	*-1.622*	*	*-11.260*	*n.s.*

Table D – Real GDP growth forecast accuracy with alternative reference series: 12-quarters ahead

Output gap model	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
		forecast horizon h = 12 (three-years ahead)					
CAP-AV	0.774	0.180	**0.594**	*3.198*	*	**42.940**	*
CAP-TR	0.774	0.715	**0.060**	*0.917*	*	*1.086*	*n.s.*
EC - T	0.774	1.007	-0.233	*-4.785*	*	*-3.007*	*n.s.*
EC - P	0.243	0.610	-0.367	*-9.982*	*	*-3.606*	*n.s.*
IMF	0.774	1.329	-0.555	*-4.268*	*	*-5.427*	*n.s.*
OECD	0.774	0.499	**0.275**	*2.472*	*	*7.164*	*n.s.*
UC - CC	0.318	0.307	**0.011**	*0.562*	*n.s.*	*0.251*	*n.s.*
UC - PIC	0.318	0.713	-0.394	*-3.318*	*	*-3.873*	*n.s.*
UC - BIV	0.243	0.647	-0.404	*-2.331*	*	*-3.743*	*n.s.*
BP	0.774	0.691	**0.084**	*0.762*	*	*1.576*	*n.s.*
HP	0.774	1.262	-0.488	*-0.864*	*	*-5.026*	*n.s.*
LIN	0.774	0.731	**0.043**	*0.834*	*	*0.763*	*n.s.*
AV-All	0.774	0.338	**0.436**	*2.554*	*	**16.756**	*
AV-PFA	0.774	0.647	**0.127**	*1.840*	*	*2.558*	*n.s.*
AV-Org	0.774	0.866	-0.091	*-1.442*	*	*-1.372*	*n.s.*
AV-UC	0.774	0.479	**0.295**	*1.055*	*	*8.018*	*n.s.*
AV-Fil	0.774	1.314	-0.540	*-0.933*	*	*-5.340*	*n.s.*

Appendix XII – Results for sub-sample real GDP forecasting analysis

Table A – Tests of equal real GDP growth forecast accuracy: one quarter ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 1 (one-quarter ahead)					
CAP-AV	2001:01-2006:04	1.574	1.346	**0.228**	0.617	n.s.	4.061	*
	2001:01-2003:04	1.593	1.525	**0.068**	0.138	n.s.	0.539	n.s.
	2004:01-2006:04	1.555	1.168	**0.387**	0.724	n.s.	3.976	*
CAP-TR	2001:01-2006:04	1.574	1.324	**0.250**	0.594	n.s.	4.539	*
	2001:01-2003:04	1.593	1.622	-0.029	-0.054	n.s.	-0.217	n.s.
	2004:01-2006:04	1.555	1.025	**0.530**	0.857	n.s.	6.203	*
EC - T	2001:01-2006:04	1.574	1.879	-0.305	-0.615	n.s.	-3.893	*
	2001:01-2003:04	1.593	2.567	-0.974	-2.646	*	-4.554	*
	2004:01-2006:04	1.555	1.191	**0.365**	0.467	n.s.	**3.674**	*
EC - P	2002:04-2006:04	1.556	1.388	**0.167**	0.324	n.s.	**2.048**	*
	2002:04-2005:04	1.062	1.459	-0.397	-0.927	*	-3.537	*
	2006:01-2006:04	3.161	1.160	**2.001**	4.063	*	**6.901**	*
IMF	2001:01-2006:04	1.574	1.853	-0.278	-0.550	n.s.	-3.607	*
	2001:01-2003:04	1.593	2.501	-0.908	-2.063	*	-4.356	*
	2004:01-2006:04	1.555	1.205	**0.351**	0.449	n.s.	**3.494**	*
OECD	2001:01-2006:04	1.574	1.643	-0.069	-0.173	n.s.	-1.005	n.s.
	2001:01-2003:04	1.593	2.371	-0.778	-2.239	*	-3.938	*
	2004:01-2006:04	1.555	0.915	**0.641**	1.287	*	**8.402**	*
UC - CC	2002:03-2006:04	1.469	2.244	-0.775	-1.099	*	-6.214	*
	2002:03-2005:04	0.986	2.512	-1.526	-2.638	*	-8.504	*
	2006:01-2006:04	3.161	1.307	**1.854**	2.689	*	**5.671**	*
UC - PIC	2002:03-2006:04	1.469	10.244	-8.775	-2.524	*	-15.418	*
	2002:03-2005:04	0.986	9.156	-8.170	-1.872	*	-12.492	*
	2006:01-2006:04	3.161	14.055	-10.894	-7.099	*	-3.100	*
UC - BIV	2002:04-2006:04	1.556	5.339	-3.783	-2.719	*	-12.046	*
	2002:04-2005:04	1.062	3.234	-2.173	-2.591	*	-8.732	*
	2006:01-2006:04	3.161	12.179	-9.018	-8.414	*	-2.962	*
BP	2001:01-2006:04	1.574	3.291	-1.717	-1.980	*	-12.520	*
	2001:01-2003:04	1.593	5.110	-3.517	-4.616	*	-8.259	*
	2006:01-2006:04	1.864	1.978	-0.114	-0.937	n.s.	-0.230	n.s.
HP	2001:01-2006:04	1.574	7.195	-5.621	-3.981	*	-18.749	*
	2001:01-2003:04	1.593	4.798	-3.205	-2.067	*	-8.016	*
	2004:01-2006:04	1.555	9.593	-8.037	-5.463	*	-10.054	*
LIN	2001:01-2006:04	1.574	1.517	**0.058**	0.866	*	0.910	n.s.
	2001:01-2003:04	1.593	1.496	**0.097**	1.382	*	0.777	n.s.
	2004:01-2006:04	1.555	1.537	**0.018**	0.197	n.s.	0.142	n.s.
AV-All	2001:01-2006:04	1.574	3.756	-2.182	-2.235	*	-13.941	*
	2001:01-2003:04	1.593	4.791	-3.198	-3.237	*	-8.010	*
	2004:01-2006:04	1.555	2.720	-1.165	-0.776	n.s.	-5.139	*
AV-PFA	2001:01-2006:04	1.574	2.265	-0.691	-1.119	*	-7.323	*
	2001:01-2003:04	1.593	3.121	-1.528	-2.740	*	-5.875	*
	2004:01-2006:04	1.555	1.410	**0.146**	0.158	n.s.	1.240	n.s.
AV-Org	2001:01-2006:04	1.574	1.745	-0.171	-0.373	n.s.	-2.352	*
	2001:01-2003:04	1.593	2.414	-0.821	-2.334	*	-4.082	*
	2004:01-2006:04	1.555	1.076	**0.479**	0.684	n.s.	**5.345**	*
AV-UC	2001:01-2006:04	1.574	1.527	**0.047**	0.104	n.s.	0.746	n.s.
	2001:01-2003:04	1.593	2.004	-0.411	-0.614	n.s.	-2.461	*
	2004:01-2006:04	1.555	1.049	**0.506**	0.963	*	**5.786**	*
AV-Fil	2001:01-2006:04	1.574	12.289	-10.715	-4.221	*	-20.926	*
	2001:01-2003:04	1.593	17.402	-15.809	-7.348	*	-10.902	*
	2004:01-2006:04	1.555	7.176	-5.621	-2.026	*	-9.399	*

Note: Sample period is 1965Q1-2006Q4. See Tables 4 and 12 in the main text for explanation of terms.

Table B – Tests of equal real GDP growth forecast accuracy: one year ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 4 (one-year ahead)					
CAP-AV	2001:04-2006:04	0.838	0.635	**0.203**	*0.878*	*	**6.716**	*
	2001:04-2003:04	1.313	0.614	**0.699**	*6.639*	*	**10.255**	*
	2004:01-2006:04	0.482	0.651	-0.169	*-4.539*	*	*-3.116*	*
CAP-TR	2001:04-2006:04	0.838	0.789	**0.049**	*0.675*	*	*1.299*	n.s.
	2001:04-2003:04	1.313	1.258	**0.055**	*0.477*	n.s.	*0.395*	n.s.
	2004:01-2006:04	0.482	0.438	**0.044**	*0.613*	*n.s.*	**1.207**	*
EC - T	2001:04-2006:04	0.838	1.410	-0.572	*-1.921*	*	*-8.520*	*
	2001:04-2003:04	1.313	2.419	-1.106	*-8.319*	*	*-4.114*	*
	2004:01-2006:04	0.482	0.654	-0.172	*-0.456*	*n.s.*	*-3.158*	*n.s.*
EC - P	2003:03-2006:04	0.560	0.879	-0.319	*-0.938*	*	*-5.080*	*
	2003:03-2005:04	0.281	1.121	-0.839	*-5.545*	*	*-7.490*	*
	2006:01-2006:04	1.257	0.275	**0.982**	*8.772*	*	**14.298**	*
IMF	2001:04-2006:04	0.838	1.468	-0.629	*-1.922*	*	*-9.007*	*
	2001:04-2003:04	1.313	2.151	-0.838	*-3.446*	*	*-3.505*	*
	2004:01-2006:04	0.482	0.955	-0.473	*-0.943*	*	*-5.946*	*
OECD	2001:04-2006:04	0.838	1.001	-0.163	*-0.927*	*	*-3.411*	*
	2001:04-2003:04	1.313	1.778	-0.465	*-4.027*	*	*-2.354*	*
	2004:01-2006:04	0.482	0.418	**0.064**	*0.291*	*n.s.*	*1.851*	*n.s.*
UC - CC	2003:02-2006:04	0.638	0.665	-0.027	*-0.272*	*n.s.*	*-0.611*	*n.s.*
	2003:02-2005:04	0.413	0.559	-0.146	*-1.649*	*	*-2.868*	*
	2006:01-2006:04	1.257	0.958	**0.299**	*2.155*	*	**1.248**	*
UC - PIC	2003:02-2006:04	0.638	4.859	-4.221	*-1.861*	*	*-13.030*	*
	2003:02-2005:04	0.413	5.187	-4.774	*-1.649*	*	*-10.124*	*
	2006:01-2006:04	1.257	3.957	-2.700	*-19.278*	*	*-2.729*	*
UC - BIV	2003:03-2006:04	0.560	0.852	-0.292	*-2.071*	*	*-4.797*	*
	2003:03-2005:04	0.281	0.536	-0.255	*-1.235*	*	*-4.756*	*
	2006:01-2006:04	1.257	1.641	-0.384	*-5.315*	*	*-0.936*	*n.s.*
BP	2001:04-2006:04	0.838	0.842	-0.003	*-0.024*	n.s.	*-0.085*	n.s.
	2001:04-2003:04	1.313	1.073	0.240	*0.949*	*	*2.017*	n.s.
	2004:01-2006:04	0.482	0.668	-0.186	*-6.452*	*	*-3.344*	*n.s.*
HP	2001:04-2006:04	0.838	7.750	-6.912	*-3.784*	*	*-18.729*	*
	2001:04-2003:04	1.313	4.274	-2.961	*-8.839*	*	*-6.235*	*
	2004:01-2006:04	0.482	10.357	-9.875	*-7.166*	*	*-11.442*	*
LIN	2001:04-2006:04	0.838	0.753	**0.085**	*1.622*	*	*2.375*	n.s.
	2001:04-2003:04	1.313	1.127	**0.186**	*2.567*	*	*1.485*	n.s.
	2004:01-2006:04	0.482	0.472	**0.010**	*0.220*	*n.s.*	*0.242*	*n.s.*
AV-All	2001:04-2006:04	0.838	0.651	**0.187**	*1.270*	*	**6.019**	*
	2001:04-2003:04	1.313	0.969	**0.344**	*2.080*	*	**3.195**	*
	2004:01-2006:04	0.482	0.413	**0.069**	*0.494*	*n.s.*	*1.997*	*n.s.*
AV-PFA	2001:04-2006:04	0.838	1.095	-0.257	*-1.239*	*	*-4.926*	*
	2001:04-2003:04	1.313	1.757	-0.444	*-4.561*	*	*-2.274*	n.s.
	2004:01-2006:04	0.482	0.598	-0.117	*-0.350*	*n.s.*	*-2.336*	*n.s.*
AV-Org	2001:04-2006:04	0.838	1.232	-0.394	*-1.591*	*	*-6.710*	*
	2001:04-2003:04	1.313	2.031	-0.718	*-4.607*	*	*-3.181*	*
	2004:01-2006:04	0.482	0.632	-0.150	*-0.417*	*n.s.*	*-2.855*	*n.s.*
AV-UC	2001:04-2006:04	0.838	0.537	**0.301**	*0.854*	*	**11.769**	*
	2001:04-2003:04	1.313	0.214	**1.099**	*7.391*	*	**46.201**	*
	2004:01-2006:04	0.482	0.779	-0.297	*-2.002*	*	*-4.580*	*n.s.*
AV-Fil	2001:04-2006:04	0.838	1.708	-0.870	*-1.075*	*	*-10.693*	*
	2001:04-2003:04	1.313	0.345	**0.968**	*3.916*	*	**25.237**	*
	2004:01-2006:04	0.482	2.730	-2.248	*-4.483*	*	*-9.881*	*

Note: Sample period is 1965Q1-2006Q4. See Tables 4 and 12 in the main text for explanation of terms.

Table C – Tests of equal real GDP growth forecast accuracy: two years ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 8 (two-years ahead)					
CAP-AV	2002:04-2006:04	0.707	0.234	**0.472**	*1.877*	*	**34.319**	*
	2002:04-2003:04	1.758	0.458	**1.300**	*17.567*	*	**14.197**	*
	2004:01-2006:04	0.269	0.141	**0.128**	*1.486*	*	**10.885**	*
CAP-TR	2002:04-2006:04	0.707	0.645	**0.062**	*1.092*	*	*1.631*	n.s.
	2002:04-2003:04	1.758	1.494	**0.264**	*6.679*	*	*0.882*	n.s.
	2004:01-2006:04	0.269	0.291	-0.022	*-0.933*	*	*-0.918*	*n.s.*
EC - T	2002:04-2006:04	0.707	0.907	-0.200	*-5.018*	*	*-3.753*	*n.s.*
	2002:04-2003:04	1.758	1.929	-0.171	*-1.909*	*	*-0.443*	*n.s.*
	2004:01-2006:04	0.269	0.481	-0.212	*-4.587*	*	*-5.300*	*n.s.*
EC - P	2004:03-2006:04	0.240	0.715	-0.475	*-6.873*	*	*-6.641*	*n.s.*
	2004:03-2005:04	0.314	0.887	-0.573	*-22.091*	*	*-3.874*	*n.s.*
	2006:01-2006:04	0.129	0.457	-0.328	*-3.166*	*	*-2.871*	*n.s.*
IMF	2002:04-2006:04	0.707	1.477	-0.770	*-2.958*	*	*-8.867*	*
	2002:04-2003:04	1.758	1.693	**0.065**	*3.037*	*	*0.191*	n.s.
	2004:01-2006:04	0.269	1.387	-1.118	*-8.750*	*	*-9.676*	*
OECD	2002:04-2006:04	0.707	0.650	**0.056**	*0.391*	*n.s.*	*1.465*	n.s.
	2002:04-2003:04	1.758	1.253	**0.505**	*8.022*	*	*2.015*	n.s.
	2004:01-2006:04	0.269	0.400	-0.131	*-2.244*	*	*-3.934*	*n.s.*
UC - CC	2004:02-2006:04	0.262	0.224	**0.038**	*0.828*	*	*1.886*	*n.s.*
	2004:02-2005:04	0.339	0.240	**0.098**	*2.284*	*	*2.866*	*n.s.*
	2006:01-2006:04	0.129	0.195	-0.067	*-2.046*	*	*-1.364*	*n.s.*
UC - PIC	2004:02-2006:04	0.262	2.282	-2.019	*-1.998*	*	*-9.735*	*
	2004:02-2005:04	0.339	3.492	-3.153	*-2.965*	*	*-6.321*	*
	2006:01-2006:04	0.129	0.164	-0.035	*-1.875*	*	*-0.856*	*n.s.*
UC - BIV	2004:03-2006:04	0.240	0.769	-0.529	*-1.499*	*	*-6.877*	*n.s.*
	2004:03-2005:04	0.314	1.232	-0.918	*-2.087*	*	*-4.469*	*n.s.*
	2006:01-2006:04	0.129	0.074	**0.055**	*1.792*	*	*2.963*	*n.s.*
BP	2002:04-2006:04	0.707	0.598	**0.108**	*0.986*	*	*3.077*	n.s.
	2002:04-2003:04	1.758	1.439	**0.319**	*2.341*	*	*1.109*	n.s.
	2004:01-2006:04	0.269	0.248	**0.020**	*0.204*	*n.s.*	*0.988*	*n.s.*
HP	2002:04-2006:04	0.707	4.109	-3.402	*-2.639*	*	*-14.077*	*
	2002:04-2003:04	1.758	1.417	**0.340**	*1.795*	*	*1.201*	n.s.
	2004:01-2006:04	0.269	5.230	-4.962	*-6.967*	*	*-11.384*	*
LIN	2002:04-2006:04	0.707	0.599	**0.107**	*4.123*	*	*3.047*	n.s.
	2002:04-2003:04	1.758	1.686	**0.072**	*1.193*	*	*0.212*	n.s.
	2004:01-2006:04	0.269	0.146	**0.122**	*4.291*	*	**10.040**	*
AV-All	2002:04-2006:04	0.707	0.314	**0.392**	*1.684*	*	**21.240**	*
	2002:04-2003:04	1.758	0.577	**1.181**	*10.198*	*	**10.229**	*
	2004:01-2006:04	0.269	0.205	**0.064**	*0.942*	*	*3.757*	*n.s.*
AV-PFA	2002:04-2006:04	0.707	0.696	**0.011**	*0.088*	n.s.	*0.269*	n.s.
	2002:04-2003:04	1.758	1.413	**0.345**	*9.668*	*	*1.221*	n.s.
	2004:01-2006:04	0.269	0.397	-0.128	*-1.680*	*	*-3.877*	*n.s.*
AV-Org	2002:04-2006:04	0.707	0.905	-0.198	*-1.566*	*	*-3.722*	n.s.
	2002:04-2003:04	1.758	1.521	**0.237**	*10.304*	*	*0.777*	n.s.
	2004:01-2006:04	0.269	0.648	-0.379	*-6.633*	*	*-7.025*	*
AV-UC	2002:04-2006:04	0.707	0.888	-0.182	*-0.692*	*	*-3.479*	n.s.
	2002:04-2003:04	1.758	0.855	**0.903**	*7.012*	*	*5.283*	n.s.
	2004:01-2006:04	0.269	0.902	-0.634	*-6.141*	*	*-8.429*	*n.s.*
AV-Fil	2002:04-2006:04	0.707	2.964	-2.257	*-2.311*	*	*-12.947*	*
	2002:04-2003:04	1.758	0.707	**1.051**	*4.426*	*	*7.435*	n.s.
	2004:01-2006:04	0.269	3.904	-3.636	*-11.241*	*	*-11.175*	*

Note: Sample period is 1965Q1-2006Q4. See Tables 4 and 12 in the main text for explanation of terms.

Table D – Tests of equal real GDP growth forecast accuracy: three years ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)	MSE-F
		forecast horizon h = 12 (three-years ahead)				
CAP-AV	2003:04-2006:04	0.626	0.128	**0.498**	*3.207* *	**50.579** *
	2003:04-2003:04	1.342	0.165	**1.178**	*0.000* n.s.	**7.145** *
	2004:01-2006:04	0.567	0.125	**0.442**	*3.123* *	**42.398** *
CAP-TR	2003:04-2006:04	0.626	0.576	**0.05**	*0.883* *	*1.137* n.s.
	2003:04-2003:04	1.342	1.169	**0.173**	*0.000* n.s.	*0.148* n.s.
	2004:01-2006:04	0.567	0.527	**0.040**	*0.707 n.s.*	*0.915 n.s.*
EC - T	2003:04-2006:04	0.626	0.832	-0.206	*-5.025* *	*-3.213 n.s.*
	2003:04-2003:04	1.342	2.276	-0.934	*0.000 n.s.*	*-0.410 n.s.*
	2004:01-2006:04	0.567	0.712	-0.145	*-7.937* *	*-2.445 n.s.*
EC - P	2005:03-2006:04	0.219	0.559	-0.340	*-9.647* *	*-3.647 n.s.*
	2005:03-2005:04	0.541	1.037	-0.496	*-11.677* *	*-0.956 n.s.*
	2006:01-2006:04	0.058	0.320	-0.262	*-9.925* *	*-3.274 n.s.*
IMF	2003:04-2006:04	0.626	1.145	-0.519	*-4.217* *	*-5.891 n.s.*
	2003:04-2003:04	1.342	1.446	-0.103	*0.000 n.s.*	*-0.071 n.s.*
	2004:01-2006:04	0.567	1.120	-0.554	*-4.684* *	*-5.930 n.s.*
OECD	2003:04-2006:04	0.626	0.384	**0.242**	*2.540* *	*8.201 n.s.*
	2003:04-2003:04	1.342	0.764	**0.578**	*0.000 n.s.*	*0.756 n.s.*
	2004:01-2006:04	0.567	0.352	**0.214**	*2.391* *	*7.298 n.s.*
UC - CC	2005:02-2006:04	0.285	0.272	**0.013**	*0.705 n.s.*	*0.335 n.s.*
	2005:02-2005:04	0.588	0.534	**0.054**	*1.722* *	*0.301 n.s.*
	2006:01-2006:04	0.058	0.075	-0.017	*-9.518* *	*-0.922 n.s.*
UC - PIC	2005:02-2006:04	0.285	0.720	-0.435	*-3.338* *	*-4.230 n.s.*
	2005:02-2005:04	0.588	1.501	-0.913	*-6.206* *	*-1.825 n.s.*
	2006:01-2006:04	0.058	0.135	-0.077	*-3.680* *	*-2.278 n.s.*
UC - BIV	2005:03-2006:04	0.219	0.641	-0.421	*-2.454* *	*-3.947 n.s.*
	2005:03-2005:04	0.541	1.886	-1.345	*-5.209* *	*-1.426 n.s.*
	2006:01-2006:04	0.058	0.018	**0.040**	*2.699* *	**9.107** *
BP	2003:04-2006:04	0.626	0.574	**0.052**	*0.469* n.s.	*1.187* n.s.
	2003:04-2003:04	1.342	2.018	-0.676	*0.000* n.s.	*-0.335* n.s.
	2004:01-2006:04	0.567	0.454	**0.113**	*0.857 n.s.*	*2.992 n.s.*
HP	2003:04-2006:04	0.626	1.363	-0.736	*-1.421* *	*-7.025* n.s.
	2003:04-2003:04	1.342	0.168	**1.175**	*0.000* n.s.	*7.013* n.s.
	2004:01-2006:04	0.567	1.462	-0.896	*-1.823* *	*-7.350 n.s.*
LIN	2003:04-2006:04	0.626	0.579	**0.047**	*1.052* *	*1.059* n.s.
	2003:04-2003:04	1.342	1.935	-0.592	*0.000* n.s.	*-0.306* n.s.
	2004:01-2006:04	0.567	0.466	**0.100**	*4.027* *	*2.586 n.s.*
AV-All	2003:04-2006:04	0.626	0.277	**0.350**	*2.583* *	**16.444** *
	2003:04-2003:04	1.342	0.002	**1.340**	*0.000* n.s.	**669.840** *
	2004:01-2006:04	0.567	0.299	**0.267**	*2.553* *	*10.711 n.s.*
AV-PFA	2003:04-2006:04	0.626	0.509	**0.117**	*1.932* *	*2.988* n.s.
	2003:04-2003:04	1.342	1.066	**0.276**	*0.000* n.s.	*0.259* n.s.
	2004:01-2006:04	0.567	0.463	**0.104**	*1.733* *	*2.691 n.s.*
AV-Org	2003:04-2006:04	0.626	0.710	-0.083	*-1.436* *	*-1.523* n.s.
	2003:04-2003:04	1.342	1.408	-0.066	*0.000* n.s.	*-0.047* n.s.
	2004:01-2006:04	0.567	0.651	-0.085	*-1.363* *	*-1.558 n.s.*
AV-UC	2003:04-2006:04	0.626	0.540	**0.086**	*0.392* n.s.	*2.070* n.s.
	2003:04-2003:04	1.342	0.922	**0.421**	*0.000* n.s.	*0.457* n.s.
	2004:01-2006:04	0.567	0.509	**0.058**	*0.256 n.s.*	*1.372 n.s.*
AV-Fil	2003:04-2006:04	0.626	1.399	-0.772	*-1.431* *	*-7.177* n.s.
	2003:04-2003:04	1.342	0.010	**1.333**	*0.000* n.s.	**135.894** *
	2004:01-2006:04	0.567	1.514	-0.948	*-1.870* *	*-7.509 n.s.*

Note: Sample period is 1965Q1-2006Q4. See Tables 4 and 12 in the main text for explanation of terms.

Appendix XIII – Sub-sample real GDP forecasting analysis in the US

Table A – Tests of equal real GDP growth forecast accuracy: one quarter ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 1 (one-quarter ahead)					
BP	1965:04-2006:03	11.575	13.795	-2.221	*-0.935*	*	*-26.402*	*
	1965:04-1984:04	20.377	23.560	-3.184	*-0.639*	*	*-10.405*	*
	1985:01-2006:03	3.784	5.153	-1.369	*-1.802*	*	*-23.109*	*
HP	1965:04-2006:03	11.575	11.584	-0.010	*-0.010*	*	*-0.141*	*
	1965:04-1984:04	20.377	19.962	**0.415**	*0.197*	*	*1.600*	*
	1985:01-2006:03	3.784	4.170	-0.386	*-1.232*	*	*-8.050*	*
LIN	1965:04-2006:03	11.575	11.638	-0.064	*-0.108*	*	*-0.901*	*
	1965:04-1984:04	20.377	20.107	**0.270**	*0.224*	*	*1.034*	*
	1985:01-2006:03	3.784	4.144	-0.359	*-1.097*	*	*-7.547*	*
AV-Fil	1965:04-2006:03	11.575	13.224	-1.650	*-0.817*	*	*-20.458*	*
	1965:04-1984:04	20.377	22.160	-1.783	*-0.422*	*	*-6.195*	*
	1985:01-2006:03	3.784	5.316	-1.532	*-2.189*	*	*-25.068*	*
BP - 2001	2001:01-2006:03	3.320	4.099	-0.779	*-0.430*	*n.s.*	*-4.373*	*
	2001:01-2003:04	5.042	6.174	-1.132	*-0.329*	*n.s.*	*-2.200*	*
	2004:01-2006:03	1.440	1.835	-0.395	*-1.116*	*	*-2.367*	*
HP - 2001	2001:01-2006:03	3.320	4.004	-0.684	*-0.835*	*	*-3.929*	*
	2001:01-2003:04	5.042	6.455	-1.413	*-0.975*	*	*-2.626*	*
	2004:01-2006:03	1.440	1.329	**0.111**	*0.846*	*n.s.*	*0.921*	*n.s.*
LIN - 2001	2001:01-2006:03	3.320	3.491	-0.171	*-0.574*	*n.s.*	*-1.130*	*
	2001:01-2003:04	5.042	5.089	-0.047	*-0.084*	*n.s.*	*-0.111*	*n.s.*
	2004:01-2006:03	1.440	1.747	-0.307	*-1.453*	*	*-1.934*	*
AV-Fil - 2001	2001:01-2006:03	3.320	3.752	-0.432	*-0.364*	n.s.	*-2.648*	*
	2001:01-2003:04	5.042	5.497	-0.454	*-0.200*	n.s.	*-0.992*	*
	2006:01-2006:04	1.864	1.978	-0.114	*-0.937*	*n.s.*	*-0.230*	*n.s.*

Note: Sample period is 1965Q1-2006Q4.
See Tables 4 and 12 in the main text for explanation of terms.

Table B – Tests of equal real GDP growth forecast accuracy: one year ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 4 (one-year ahead)					
BP	1966:03-2006:03	5.457	5.613	-0.155	*-0.196*	*	*-4.460*	*
	1966:03-1984:04	9.888	9.702	**0.185**	*0.109*	*	*1.414*	*
	1985:01-2006:03	1.689	2.134	-0.445	*-1.556*	*	*-18.157*	*
HP	1966:03-2006:03	5.457	5.292	**0.165**	*0.274*	*	*5.016*	*
	1966:03-1984:04	9.888	8.753	**1.134**	*0.944*	*	*9.590*	*
	1985:01-2006:03	1.689	2.348	-0.660	*-2.505*	*	*-24.441*	*
LIN	1966:03-2006:03	5.457	6.199	-0.742	*-1.050*	*	*-19.260*	*
	1966:03-1984:04	9.888	10.092	-0.205	*-0.137*	*	*-1.502*	*
	1985:01-2006:03	1.689	2.887	-1.198	*-3.071*	*	*-36.107*	*
AV-Fil	1966:03-2006:03	5.457	6.692	-1.235	*-1.357*	*	*-29.712*	*
	1966:03-1984:04	9.888	10.260	-0.372	*-0.199*	*	*-2.683*	*
	1985:01-2006:03	1.689	3.658	-1.969	*-3.635*	*	*-46.835*	*
BP - 2001	2001:04-2006:03	1.282	1.630	-0.348	*-0.873*	*	*-4.272*	*n.s.*
	2001:04-2003:04	2.112	3.162	-1.050	*-1.053*	*	*-2.988*	*n.s.*
	2004:01-2006:03	0.603	0.377	**0.226**	*0.615*	*n.s.*	*6.602*	*
HP - 2001	2001:04-2006:03	1.282	2.403	-1.121	*-1.364*	*	*-9.330*	*
	2001:04-2003:04	2.112	5.037	-2.925	*-6.476*	*	*-5.226*	*n.s.*
	2004:01-2006:03	0.603	0.247	**0.355**	*1.129*	*	*15.785*	*
LIN - 2001	2001:04-2006:03	1.282	1.314	-0.032	*-0.224*	*n.s.*	*-0.490*	*n.s.*
	2001:04-2003:04	2.112	2.228	-0.116	*-0.658*	*n.s.*	*-0.469*	*n.s.*
	2004:01-2006:03	0.603	0.566	**0.036**	*0.116*	*n.s.*	*0.708*	*n.s.*
AV-Fil - 2001	2001:04-2006:03	1.282	1.722	-0.440	*-1.774*	*	*-5.112*	n.s.
	2001:04-2003:04	2.112	3.144	-1.032	*-4.981*	*	*-2.954*	n.s.
	2004:01-2006:03	0.603	0.559	**0.044**	*0.116*	*n.s.*	*0.867*	*n.s.*

Note: Sample period is 1965Q1-2006Q4.
See Tables 4 and 12 in the main text for explanation of terms.

Table C – Tests of equal real GDP growth forecast accuracy: two years ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
				forecast horizon h = 8 (two-years ahead)				
BP	1967:03-2006:03	3.471	3.138	**0.333**	*1.071*	*	*16.684*	*
	1967:03-1984:04	6.116	5.532	**0.585**	*0.905*	*	*7.397*	*
	1985:01-2006:03	1.343	1.211	**0.131**	*0.642*	*	*9.438*	*n.s.*
HP	1967:03-2006:03	3.471	3.337	**0.135**	*0.371*	*	*6.331*	*
	1967:03-1984:04	6.116	5.260	**0.857**	*1.425*	*	*11.399*	*
	1985:01-2006:03	1.343	1.789	-0.446	*-2.308*	*	*-21.703*	*
LIN	1967:03-2006:03	3.471	4.810	-1.338	*-2.226*	*	*-43.691*	*
	1967:03-1984:04	6.116	7.597	-1.481	*-1.145*	*	*-13.643*	*
	1985:01-2006:03	1.343	2.567	-1.224	*-2.708*	*	*-41.487*	*
AV-Fil	1967:03-2006:03	3.471	4.583	-1.112	*-1.765*	*	*-38.088*	*
	1967:03-1984:04	6.116	6.652	-0.535	*-0.450*	*	*-5.633*	*
	1985:01-2006:03	1.343	2.919	-1.576	*-3.001*	*	*-46.971*	*
BP - 2001	2002:04-2006:03	0.719	0.886	-0.166	*-1.407*	*	*-3.006*	*n.s.*
	2002:04-2003:04	1.710	2.009	-0.299	*-2.787*	*	*-0.744*	*n.s.*
	2004:01-2006:03	0.269	0.375	-0.106	*-0.787*	*n.s.*	*-3.114*	*n.s.*
HP - 2001	2002:04-2006:03	0.719	1.630	-0.911	*-1.952*	*	*-8.942*	*n.s.*
	2002:04-2003:04	1.710	3.994	-2.284	*-19.845*	*	*-2.859*	*n.s.*
	2004:01-2006:03	0.269	0.556	-0.287	*-1.165*	*	*-5.682*	*n.s.*
LIN - 2001	2002:04-2006:03	0.719	0.752	-0.032	*-0.436*	*n.s.*	*-0.689*	*n.s.*
	2002:04-2003:04	1.710	1.608	**0.102**	*1.490*	*	*0.318*	*n.s.*
	2004:01-2006:03	0.269	0.362	-0.094	*-1.201*	*	*-2.841*	*n.s.*
AV-Fil - 2001	2002:04-2006:03	0.719	1.249	-0.529	*-1.974*	*	*-6.784*	n.s.
	2002:04-2003:04	1.710	3.195	-1.485	*-13.305*	*	*-2.324*	n.s.
	2004:01-2006:03	0.269	0.364	-0.095	*-1.402*	*	*-2.873*	*n.s.*

Note: Sample period is 1965Q1-2006Q4.
See Tables 4 and 12 in the main text for explanation of terms.

Table D – Tests of equal real GDP growth forecast accuracy: three years ahead

Output gap model	Sample	MSE (AR)	MSE (OG)	diff	MSE-t (conv)		MSE-F	
			forecast horizon h = 12 (three-years ahead)					
BP	1968:03-2006:03	2.191	2.277	-0.086	*-0.568*	*	*-5.802*	*
	1968:03-1984:04	3.942	4.125	-0.184	*-0.533*	*	*-2.937*	*
	1985:01-2006:03	0.863	0.876	-0.013	*-0.202*	*n.s.*	*-1.251*	*n.s.*
HP	1968:03-2006:03	2.191	2.511	-0.320	*-1.227*	*	*-19.513*	*
	1968:03-1984:04	3.942	3.847	**0.095**	*0.208*	*	*1.633*	*
	1985:01-2006:03	0.863	1.498	-0.635	*-3.628*	*	*-36.894*	*
LIN	1968:03-2006:03	2.191	3.726	-1.535	*-3.139*	*	*-63.026*	*
	1968:03-1984:04	3.942	5.583	-1.641	*-1.622*	*	*-19.404*	*
	1985:01-2006:03	0.863	2.317	-1.454	*-4.226*	*	*-54.595*	*
AV-Fil	1968:03-2006:03	2.191	2.995	-0.804	*-1.842*	*	*-41.056*	*
	1968:03-1984:04	3.942	4.286	-0.344	*-0.423*	*	*-5.304*	*
	1985:01-2006:03	0.863	2.015	-1.152	*-3.078*	*	*-49.737*	*
BP - 2001	2003:04-2006:03	0.269	0.432	-0.164	*-2.288*	*	*-4.541*	*n.s.*
	2003:04-2003:04	0.952	1.038	-0.086	*0.000*	*n.s.*	*-0.083*	*n.s.*
	2004:01-2006:03	0.207	0.377	-0.171	*-2.126*	*	*-4.976*	*n.s.*
HP - 2001	2003:04-2006:03	0.269	1.061	-0.792	*-3.792*	*	*-8.961*	*n.s.*
	2003:04-2003:04	0.952	1.834	-0.882	*0.000*	*n.s.*	*-0.481*	*n.s.*
	2004:01-2006:03	0.207	0.991	-0.784	*-3.486*	*	*-8.706*	*n.s.*
LIN - 2001	2003:04-2006:03	0.269	0.385	-0.117	*-3.542*	*	*-3.629*	*n.s.*
	2003:04-2003:04	0.952	0.778	**0.174**	*0.000*	*n.s.*	*0.223*	*n.s.*
	2004:01-2006:03	0.207	0.350	-0.143	*-5.455*	*	*-4.496*	*n.s.*
AV-Fil - 2001	2003:04-2006:03	0.269	0.734	-0.465	*-3.201*	*	*-7.607*	n.s.
	2003:04-2003:04	0.952	2.055	-1.103	*0.000*	n.s.	*-0.537*	n.s.
	2004:01-2006:03	0.207	0.614	-0.407	*-3.075*	*	*-7.298*	*n.s.*

Note: Sample period is 1965Q1-2006Q4.
See Tables 4 and 12 in the main text for explanation of terms.

References

- Altissimo F., L. Bilke, A. Levin, T. Mathă, B. Mojon (2006): "Sectoral and Aggregate Inflation Dynamics in the Euro Area", *Journal of the European Economic Association*. Vol. 4, Numbers 2 - 3, April 2006, pp. 585-593.
- Andrews, D. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, vol. 59, issue 3, pp. 817-58, May.
- Aruoba S. B. (2008): "Data Revisions are not Well-Behaved", *Journal of Money, Credit and Banking*, vol. 40, issue 2-3, pages 319-340, March 2008.
- Baxter, M. and R. G. King (1999): "Measuring business cycles: Approximate band-pass filters for economic time series", *Review of Economics and Statistics*, vol. 81, pages 575-593.
- Camacho, M. and G. Perez-Quiros (2008): "Introducing the euro-sting: short term indicator of euro area growth", WP 0807, Banco de España.
- Camba-Méndez, G. and D. Rodríguez Palenzuela (2003): "Assessment Criteria for Output Gap Estimates", *Economic Modelling*, May 2003, vol. 20, n. 3, pages 528-561.
- Cayen, J.-Ph. and S. van Norden (2005): "The Reliability of Canadian Output Gap Estimates in Real Time", *North American Journal of Economics and Finance*, vol. 16, n. 3, pp. 373-393.
- Clark, T. E. and M. W. McCracken (2005): "The Power of Tests of Predictive Ability in the Presence of Structural Breaks", *Journal of Econometrics*, vol. 124, pp. 1-31.
- Clark, T. E. and M. W. McCracken (2006): "The Predictive Content of the Output Gap for Inflation: Resolving In-Sample and Out-of-Sample Evidence", *Journal of Money, Credit, and Banking*, vol. 38, August 2006; pp. 1127-1148.
- Clark, T. E. and M. W. McCracken (2009): "Test of Equal Predictive Ability with Real-Time Data", *Journal of Business and Economic Statistics*, vol. 27, n. 4, pp. 441-454, October 2009.
- Clements, M. and D.F. Hendry (1999). *Forecasting Non-Stationary Economic Time Series*. Cambridge, Massachusetts: MIT Press.
- Cristadoro, R., M. Forni, L. Reichlin and G.Veronese (2005): "A core inflation measure for the euro area", *Journal of Money, Credit and Banking*, Vol. 37, No.3, pp. 539-560.
- ECB (2000): "Potential output and output gaps: concept, uses and estimates", ECB Monthly Bulletin, October 2000.
- Estima (2007): Rats Version 7, User's Guide.
- Fagan G., J. Henry and R. Mestre (2005): "An area-wide model for the euro area", *Economic Modelling*, vol. 22(1), pages 39-59, January.
- Giannone, D., J. Henry, M. Lalik and M. Modugno (2010): "An area-wide real-time database for the euro area", *ECB Working Paper* 1145, January 2010.
- Giannone, D. and L. Reichlin (2006): "Trends and cycles in the Euro area: how much heterogeneity and should we worry about it? ", *ECB Working Paper* 595, March 2006.
- Hamilton, J. (1986): "A standard error for the estimated state vector of a state space model", *Journal of Econometrics*, vol. 33, issue 3, December 1986, pages 387-397.
- Harvey, A. and A. Jäger (1993): "Detrending, stylized facts and the business cycle", *Journal of Applied Econometrics*, vol. 8, pages 231-247.
- Harvey, A. and S. Koopman (1997): "Multivariate structural time series models", in C. Heij, H. Schumacher, B. Hanzon and C. Praagman (eds), *Systematic dynamics in economic and financial models*, Chichester: Wiley and Sons.

- Hendry and Doornik (2001). PcGive Volume 1.
- Hodrick, R. J. and E. Prescott (1997): "Postwar U.S. Business Cycles: An Empirical Investigation", *Journal of Money, Credit and Banking*, vol. 29 (1), pages 1-16.
- Kilian, L. (1999): "Exchange Rates and Monetary Fundamentals: What Do We Learn From Long-Horizon Regressions?", *Journal of Applied Econometrics*, Vol. 14(5), pages. 491-510.
- Kuttner, K. (1994): "Estimating Potential Output as a Latent Variable", *Journal of Business and Economic Statistics*, vol. 12(3), pages 361-68, July.
- Mankiw, N. G. and M.D. Shapiro (1986): "News or Noise: An Analysis of GNP Revisions", *Survey of Current Business*, pages 20-25.
- Mishkin F. (2007): "Estimating Potential Output", speech given at the Conference on Price Measurement for Monetary Policy, Federal Reserve Bank of Dallas, 24 May 2007.
- Mitchell J. (2007): "Density Estimates for Real-time Eurozone Output Gap Estimates", in G. L. Mazzi and G. Savio (eds), *Growth and Cycle in the Eurozone*, Palgrave, Basingstoke.
- Musso A., L. Stracca and D. van Dijk (2009): "Instability and nonlinearity in the euro area Phillips curve", *International Journal of Central Banking*, June 2009, vol. 5 (2), pages 181-212.
- Nelson, E. and K. Nikolov (2003): "UK inflation in the 1970s and 1980s: the role of output gap mismeasurement", *Journal of Economics and Business*, 2003, vol. 55, pages 353-370.
- Nelson, E. and K. Nikolov (2004): "Monetary policy and stagflation in the UK", *Journal of Money, Credit and Banking*, 2004, vol. 36(3), pages 293-318.
- Orphanides, A. (2003): "The quest for prosperity without inflation", *Journal of Monetary Economics*, vol. 50 (3), pages 633-663.
- Orphanides, A. and S. van Norden (2002): "The unreliability of output-gap estimates in real time", *The Review of Economics and Statistics*, November 2002, v. 84 (4), pages 569-583.
- Orphanides, A. and S. van Norden (2005): "The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time," *Journal of Money, Credit and Banking*, vol. 37(3), pages 583-601, June.
- Planas, C. and A. Rossi (2004): "Can inflation data improve the real-time reliability of output gap estimates?", *Journal of Applied Econometrics*, vol. 19, no. 1, 2004, pages 121-133.
- Proietti, T., A. Musso and T. Westermann (2007): "Estimating potential output and the output gap for the euro area: a model-based production function approach", *Empirical Economics,* vol. 33, pages 85-113.
- Romer C. D. and D. H. Romer (2000): "Federal Reserve Information and the Behavior of Interest Rates", *American Economic Review*, vol. 90, n. 3, pp. 429-457, June.
- Rünstler, G. (2002): "The information content of output gaps in real time: an application to the euro area", *ECB working paper* 182, September 2002.
- Stock, J. and M. Watson (1999): "Forecasting Inflation", *Journal of Monetary Economics*, 1999, vol. 44, no. 2.
- Stock J. and M. Watson (2003): "Forecasting Output and Inflation: The Role of Asset Prices", *Journal of Economic Literature*, 2003
- Watson M. (1986): "Univariate detrending methods with stochastic trends", *Journal of Monetary Economics*, vol. 18(1), pages 49-75, July.
- Watson, M. (2007): "How Accurate are Real-Time Estimates of Output Trends and Gaps?", Federal Reserve Bank of Richmond *Economic Quarterly*, Spring 2007.