DISCUSSION PAPER SERIES

No. 7716

THE RELIABILITY OF REAL TIME ESTIMATES OF THE EURO AREA OUTPUT GAP

Massimiliano Marcellino and Alberto Musso

INTERNATIONAL MACROECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP7716.asp**

ISSN 0265-8003

THE RELIABILITY OF REAL TIME ESTIMATES OF THE EURO AREA OUTPUT GAP

Massimiliano Marcellino, European University Institute,
Bocconi University and CEPR
Alberto Musso, European Central Bank

Discussion Paper No. 7716
February 2010

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 7716

February 2010

ABSTRACT

The Reliability of Real Time Estimates of the Euro Area Output Gap

This paper provides evidence on the reliability of euro area real-time output gap estimates, including those provided by the IMF, OECD and EC and a set of model based measures. A genuine real-time data set 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.

JEL Classification: E31, E37, E52 and E58
Keywords: data revisions, euro area, output gap and real-time data

Massimiliano Marcellino
Department of Economics
European University Institute
Via della Piazzuola 43
50133 Florence
ITALY

Email:
massimiliano.marcellino@eui.eu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=139608

Alberto Musso
Euro Area Macroeconomic
Developments Division
European Central Bank
Kaiserstrasse 29
60311 Frankfurt am Main
GERMANY

Email: alberto.musso@ecb.int

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=148610

Submitted 19 February 2010

We thank, without implicating, Todd Clark for many useful suggestions. 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.

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.

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 summarizes 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 estimates 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. [2]

- 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 utilisation rate	Average	1985Q1-2006Q4	quarterly data	2001Q1-2007Q2	(26)	European Commission
Capacity utilisation rate	Linear trend	1985Q1-2006Q4	quarterly data	2001Q1-2007Q2	(26)	European Commission
UC - CC	Prod Fn Approach	1985Q1-2006Q4	quarterly data	2002Q3-2007Q2	(20)	own estimates
UC - PIC	Prod Fn Approach	1985Q1-2006Q4	quarterly data	2002Q3-2007Q2	(20)	own estimates
UC - BIV	Bivariate model	1985Q1-2006Q4	quarterly data	2002Q4-2007Q2	(19)	own estimates
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)	own estimates
Hodrick-Prescott filter	Stochastic trend	1985Q1-2006Q4	quarterly data	2001Q1-2007Q2	(26)	own estimates
Linear trend filter	Linear trend	1985Q1-2006Q4	quarterly data	2001Q1-2007Q2	(26)	own estimates

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

Notes: Real GDP data are from the EABCN (see Giannone et al., 2008, 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.

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.

[2] 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. 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

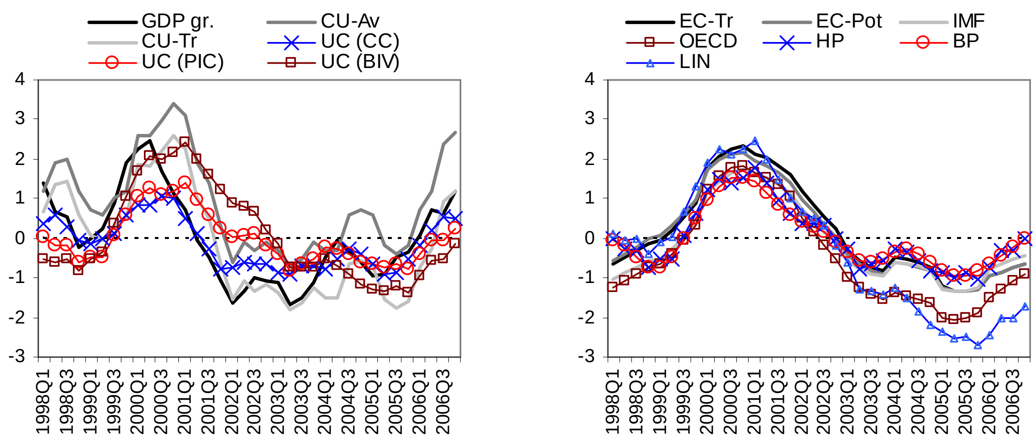
Notes: Sample period is 1985:4 to 2006:4 in all cases. "AR" refers to the first order autocorrelation coefficient.

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.[3] 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)

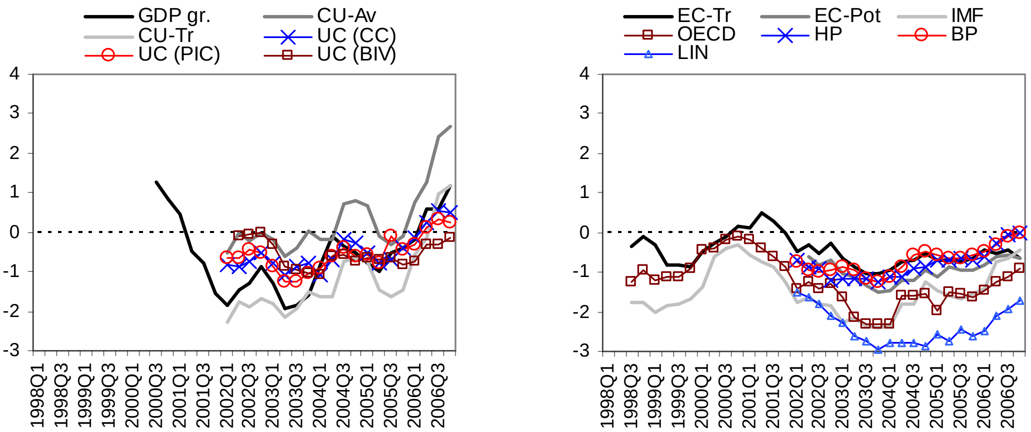


Sources: EABCN, European Commission, IMF, OECD and own estimates.
Note: UC: Estimates from the multivariate unobserved components model of Proietti, Musso and Westermann (2007). The versions of the UC model shown are the common cycles variant (CC), the pseudo-integrated cycles variant (PIC) and the bivariate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

[3] 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)

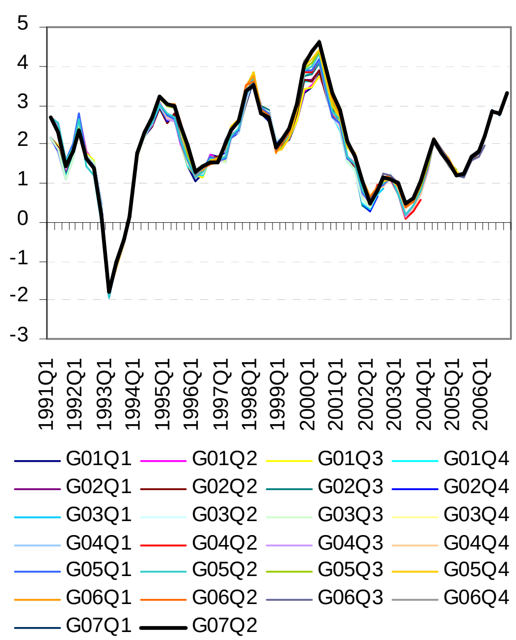


Sources: EABCN, European Commission, IMF, OECD and own estimates.
Note: UC: Estimates from the multivariate unobserved components model of Proietti, Musso and Westermann (2007). The versions of the UC model shown are the common cycles variant (CC), the pseudo-integrated cycles variant (PIC) and the bivariate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

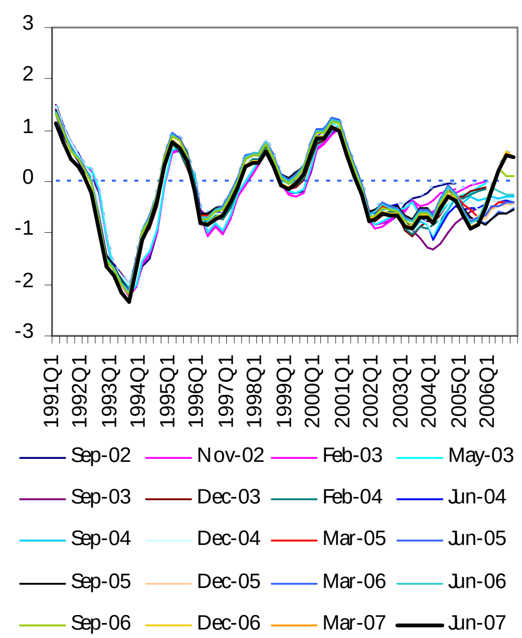
(percentages, year-on-year growth)



Sources: EABCN.

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

(percentage deviations from potential output)



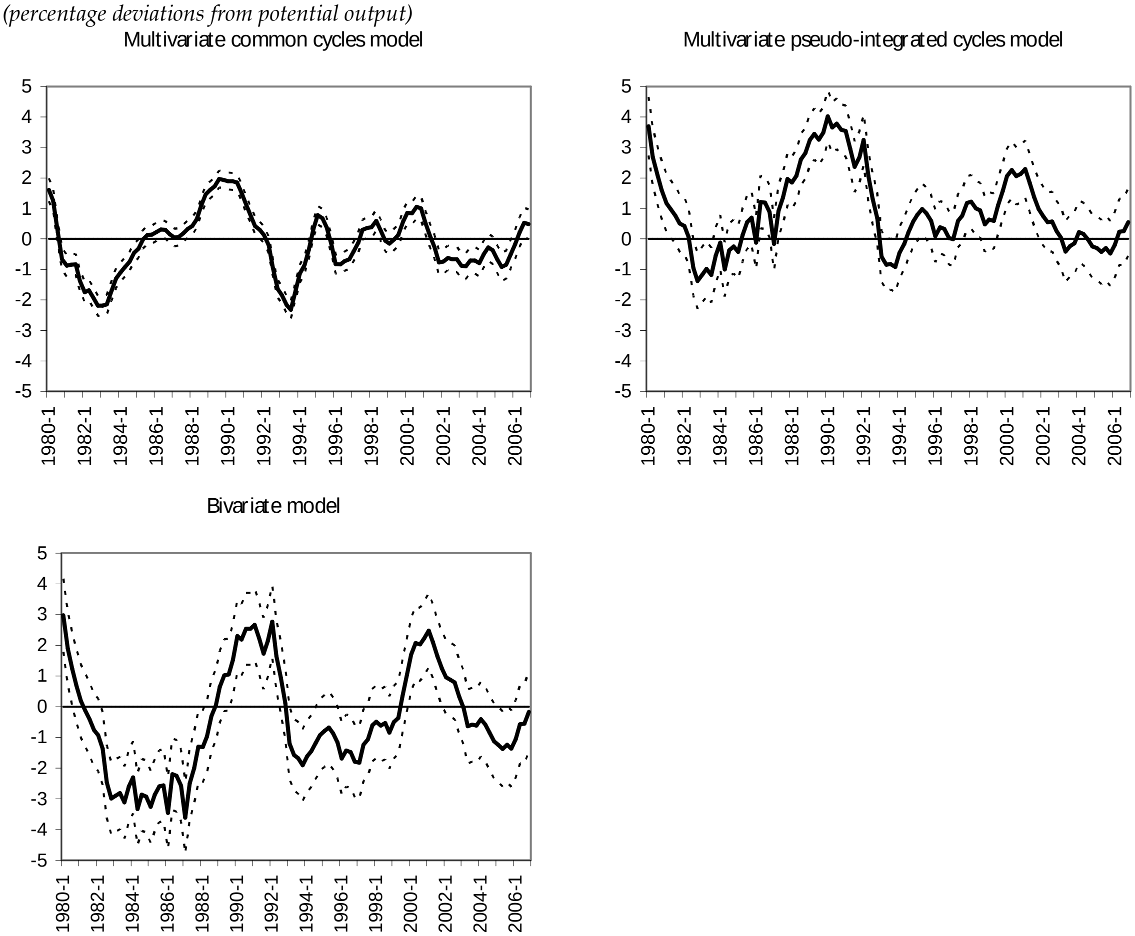
Sources: IMF.

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, Arouba (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. Arouba (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 discussed in Marcellino and Musso (2009)). 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.

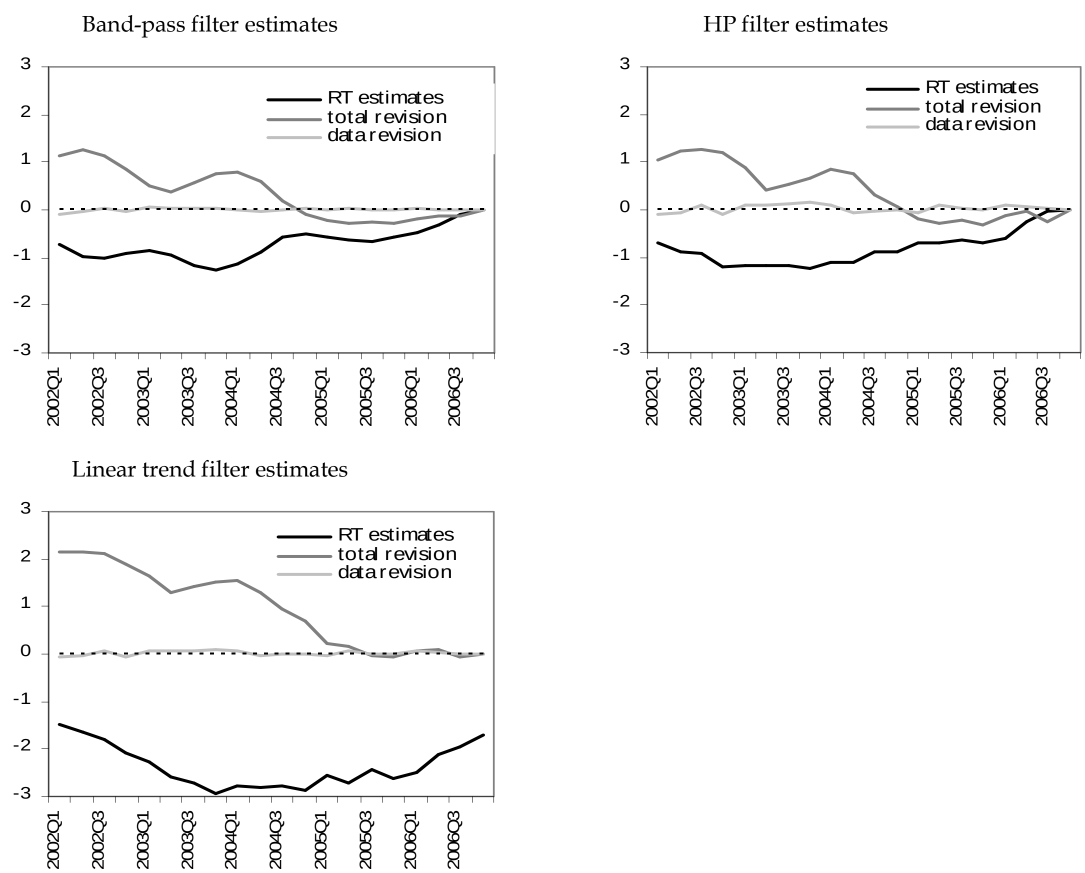
"Rev RT" stands for revision final estimate minus real time estimate.

"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)*

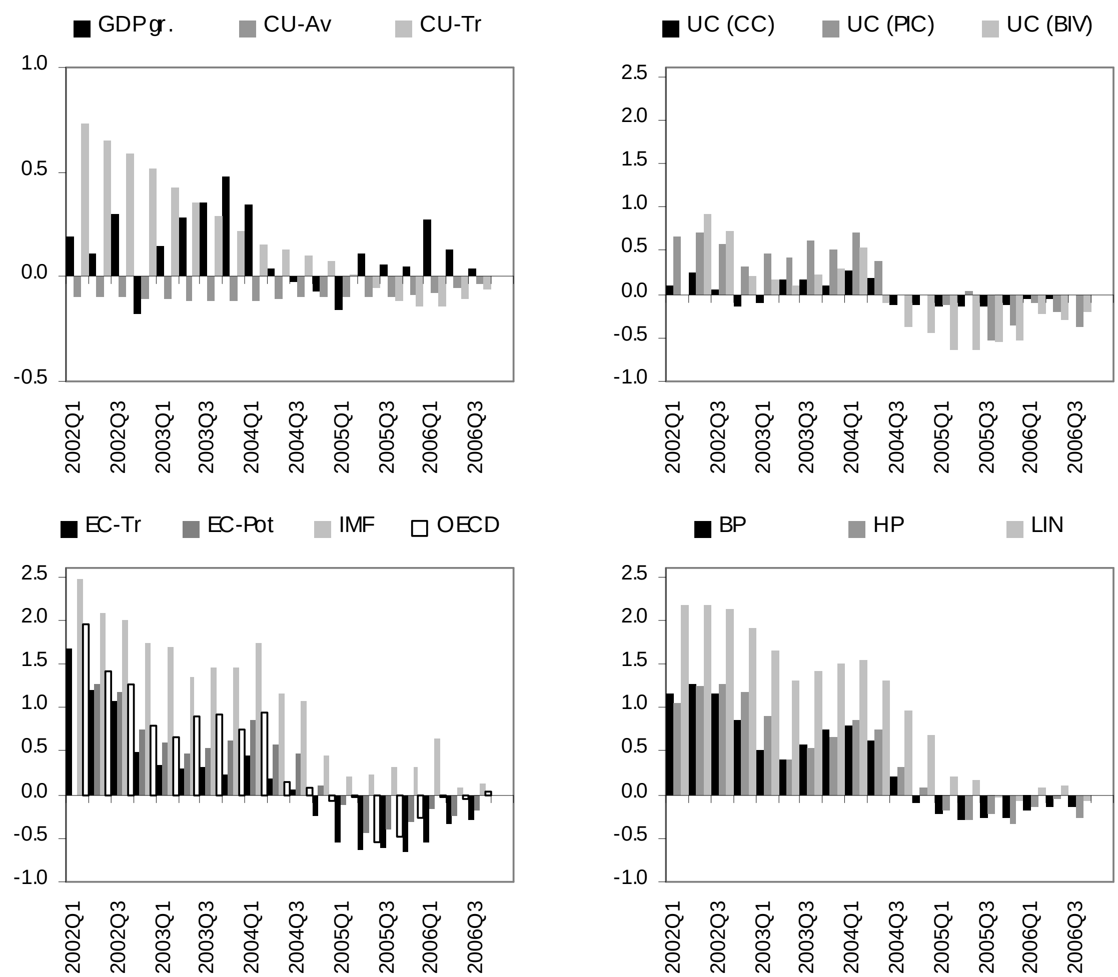


Sources: Our calculations.

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		Source
Real GDP		1947Q1-2006Q4	quarterly data	1965Q4-2007Q1	(166)	RTDSM
Band-pass filter	Stochastic trend	1985Q1-2006Q4	quarterly data	1965Q4-2007Q1	(166)	own estimates
Hodrick-Prescott filter	Stochastic trend	1985Q1-2006Q4	quarterly data	1965Q4-2007Q1	(166)	own estimates
Linear trend filter	Linear trend	1985Q1-2006Q4	quarterly data	1965Q4-2007Q1	(166)	own estimates

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.

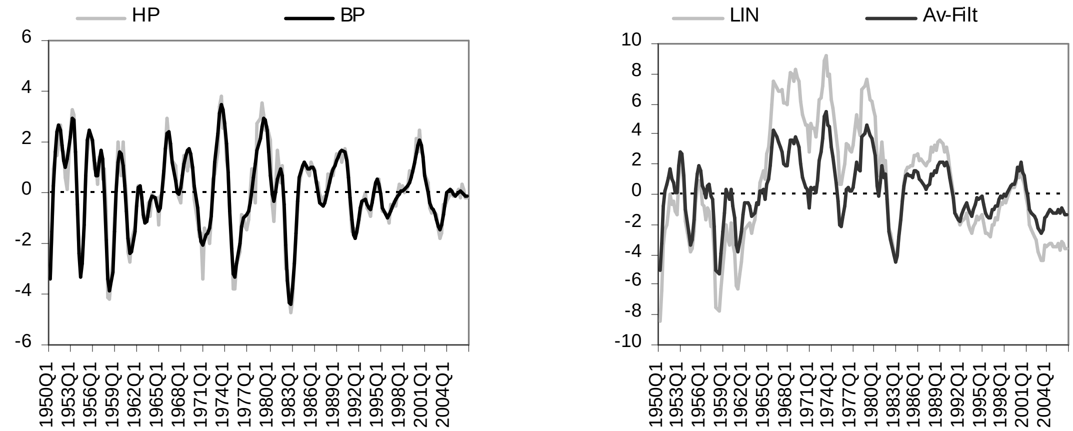
Table 6 summarises the main features of the slack measures considered, with reference to the final estimate, and Chart 8 reports their temporal evolution. Several comments can be made. First, the post 1985 results are fairly different from the full sample results. In particular, and as expected, the 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

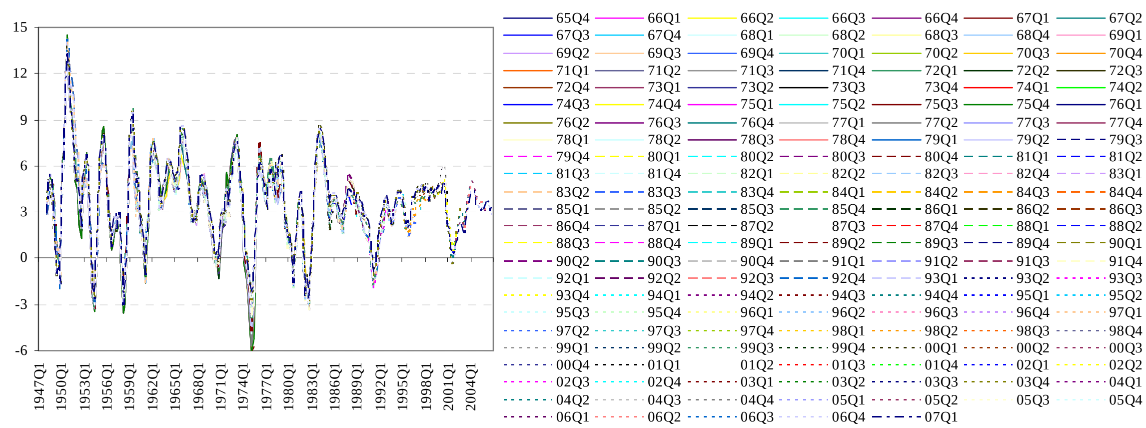
Notes: "AR" refers to the first order autocorrelation coefficient.

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

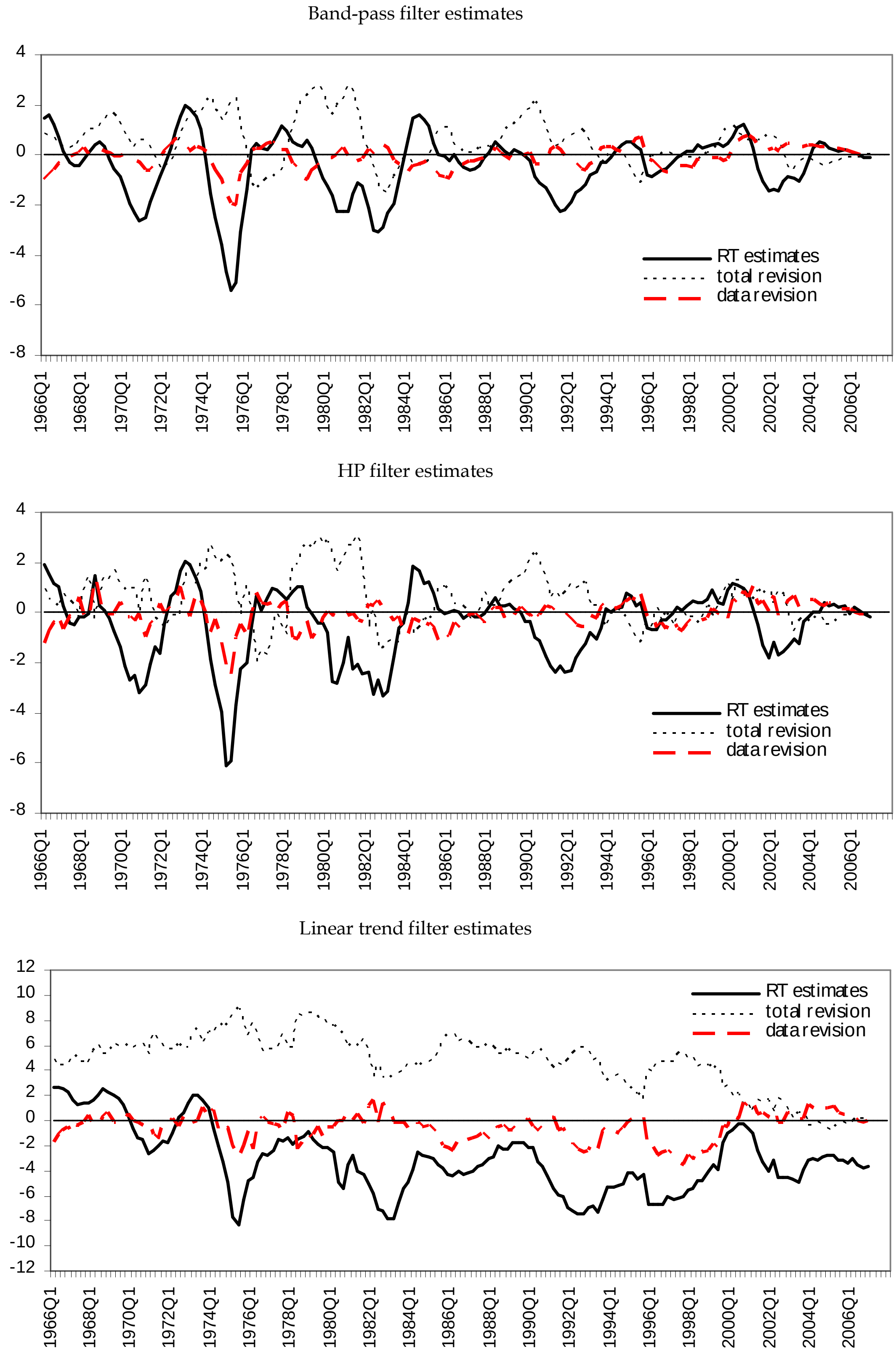
"corr" reports the correlation between real time estimate and final estimate in the "Final" row, the correlation between real time estimate and the revision (final minus real time) in the "RT" row, and the correlation between final estimate and the revision (final minus real time) in the "Rev RT" row.

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)



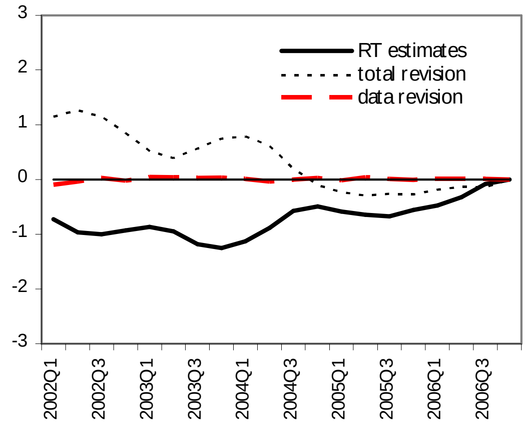
Sources: RTDSM and own calculations.

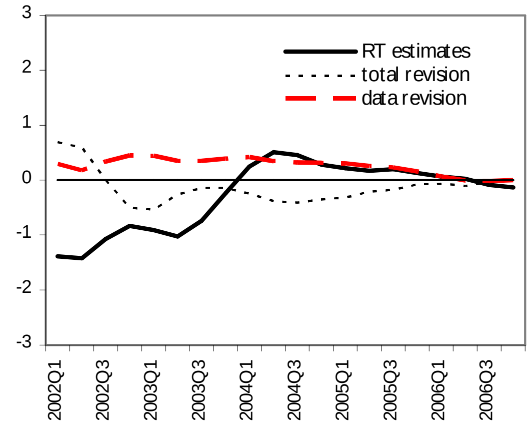
Chart 11: Real time estimates of the output gap, total revision and data revision (2002 onwards)

(percentage points)

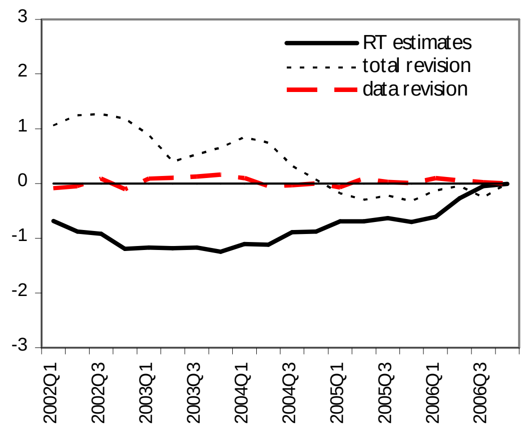
Euro Area | US

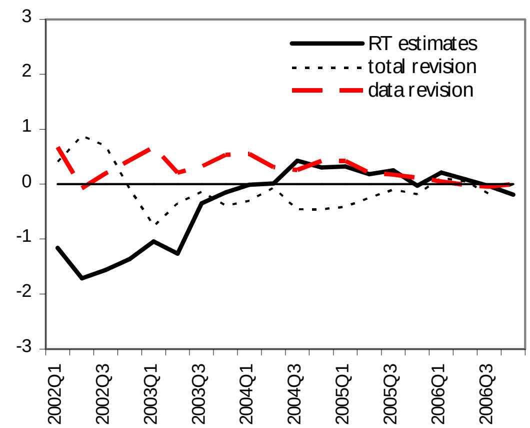
Band-pass filter estimates



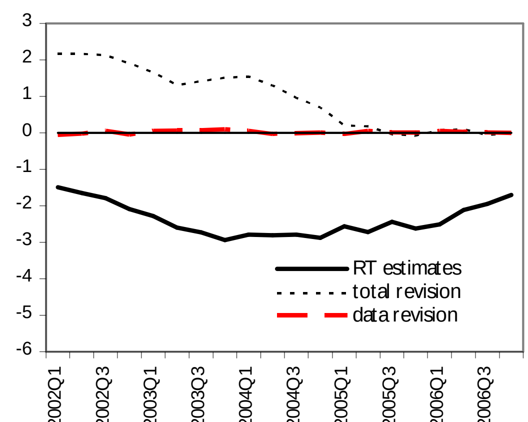


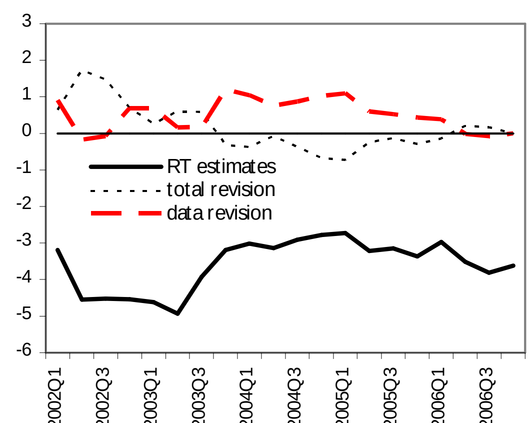
HP filter estimates





Linear trend filter estimates



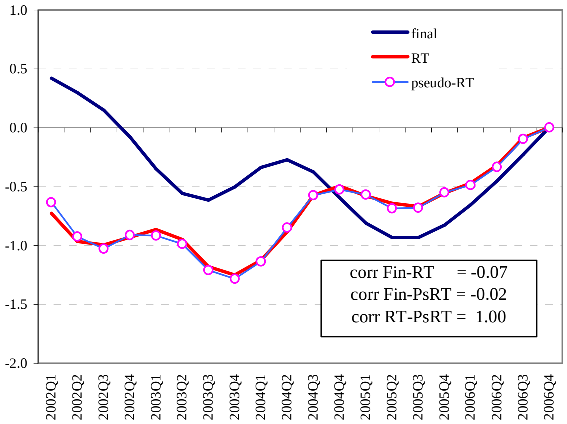


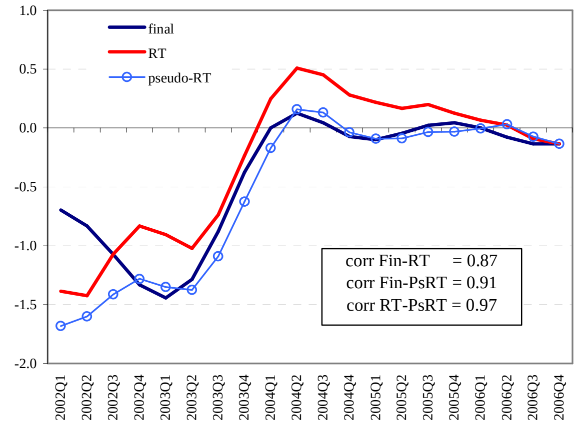
Sources: Own calculations.

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

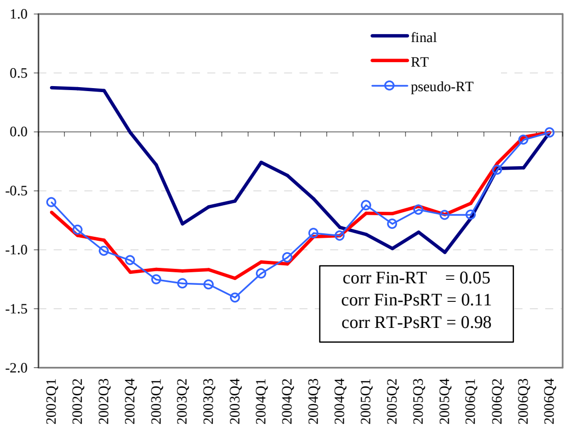
Euro Area | US

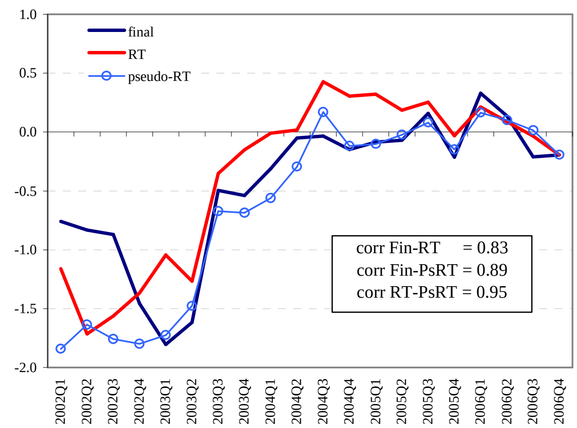
Band-pass filter estimates



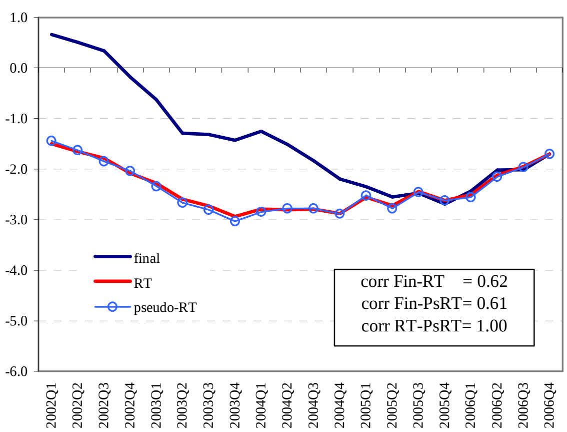


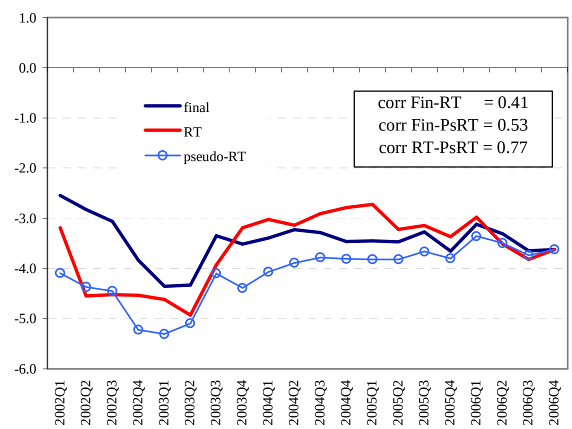
HP filter estimates





Linear trend filter estimates





Sources: Own calculations.

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.

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. 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. The results in Marcellino and Musso (2009) suggest that the gap is also not useful for forecasting inflation over the short or medium term. However, they find that some gap measures can improve forecasts of future real economic activity growth, and in this respect gap measures based on capacity utilization perform particularly well.

References

- 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.
- Estima (2007): Rats Version 7, User's Guide.
- 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.
- 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.
- Marcellino, M. and Musso, A. (2009), "The Forecasting Performance of Real Time Estimates of the Euro Area Output Gap", mimeo.
- Mankiw, N. G. and M.D. Shapiro (1986): "News or Noise: An Analysis of GNP Revisions", *Survey of Current Business*, pages 20-25.
- 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.
- 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.
- 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.
- 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.
- Watson, M. (2007): "How Accurate are Real-Time Estimates of Output Trends and Gaps?", Federal Reserve Bank of Richmond *Economic Quarterly*, Spring 2007.

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):

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

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

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

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

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

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: [4]

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

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.[5] 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.

---

[4] 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.

[5] 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 <u>common cycles model</u> 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:

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

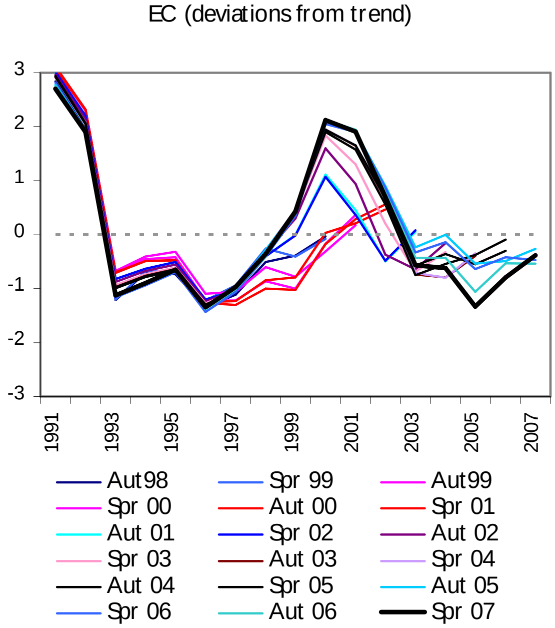
There are some indications that labour market variables could follow a cyclical pattern more persistent than that of other variables including capacity, largely due to specific frictions existing in the labour markets. Therefore, a variant of the model, <u>the pseudo-integrated cycles representation</u>, 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).

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)

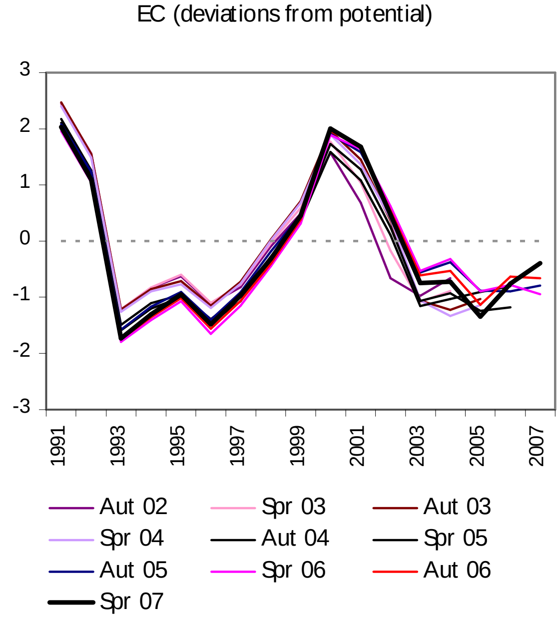


Sources: European Commission.

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)

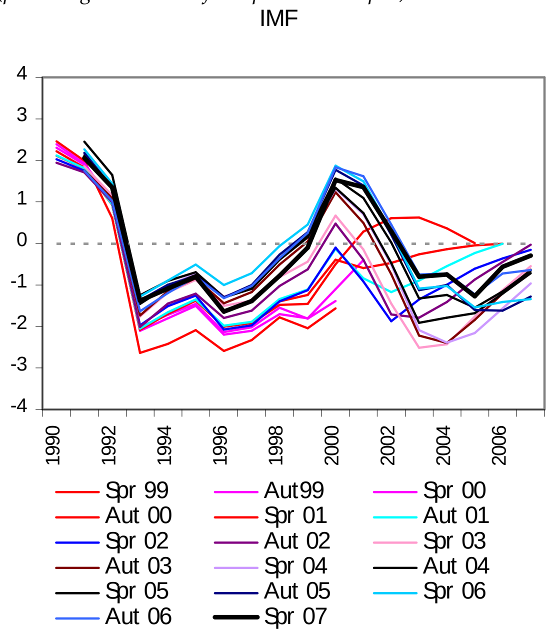


Sources: European Commission.

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

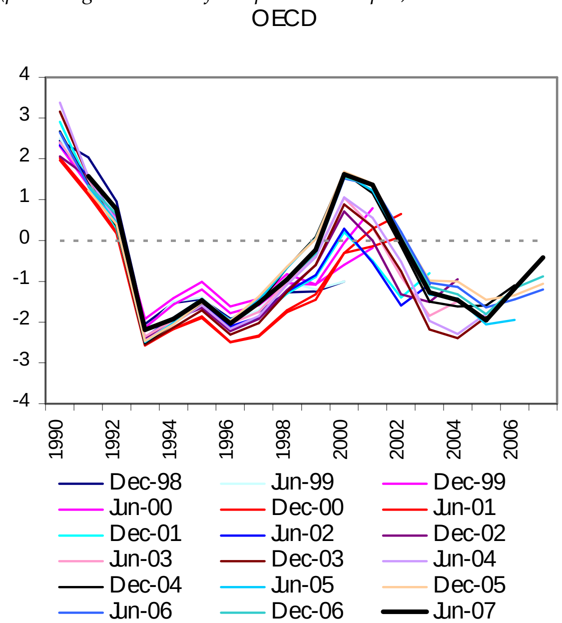
(percentage deviations from potential output)



Sources: IMF.

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

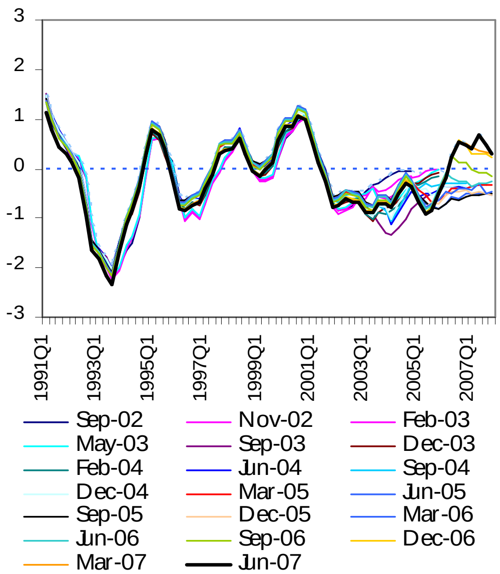
(percentage deviations from potential output)



Sources: OECD.

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

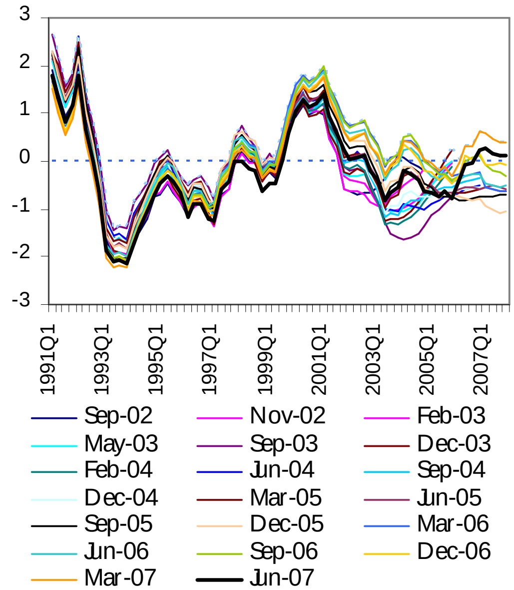
(percentage deviations from potential output)



Sources: Own calculations.

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

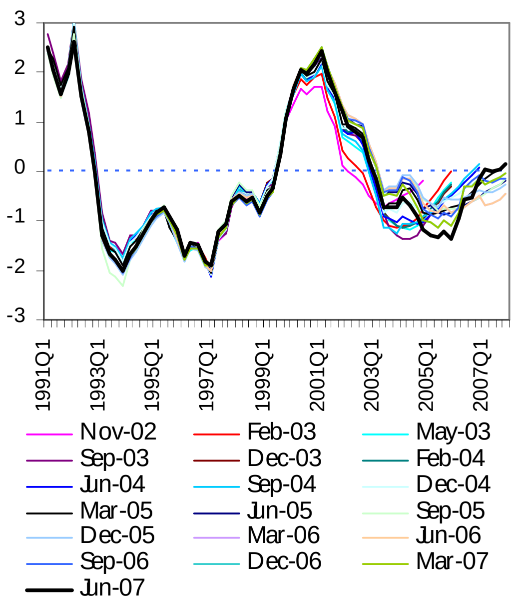
(percentage deviations from potential output)



Sources: Own calculations.

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

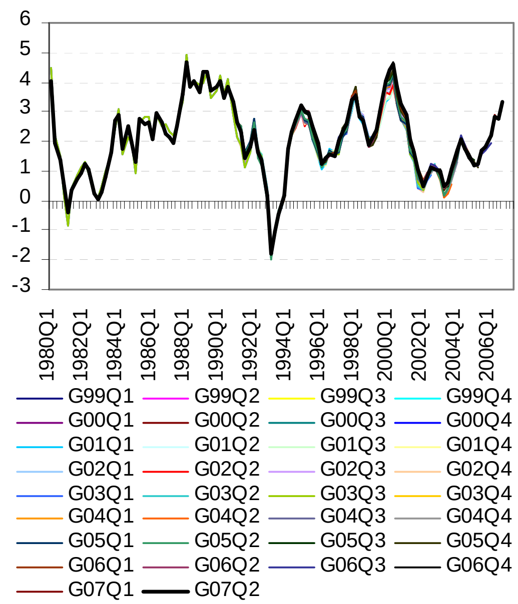
(percentage deviations from potential output)



Sources: Own calculations.

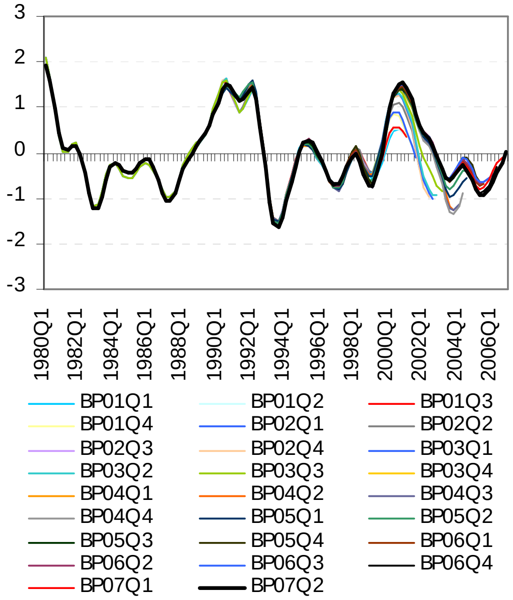
Chart H: Vintages of quarterly estimates of euro area real GDP growth

(annual percentage changes)



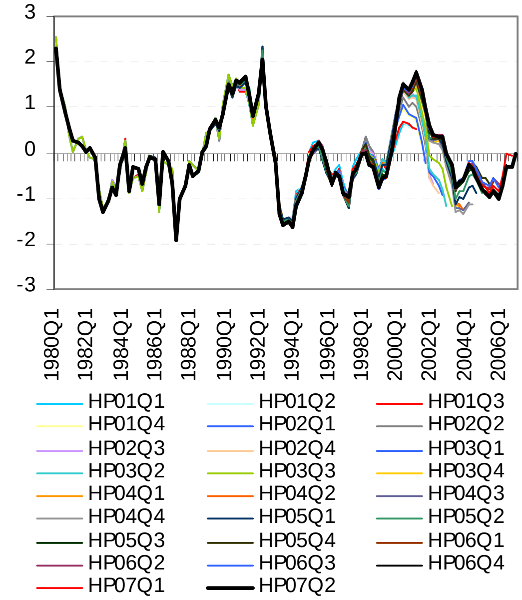
Sources: EABCN.

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



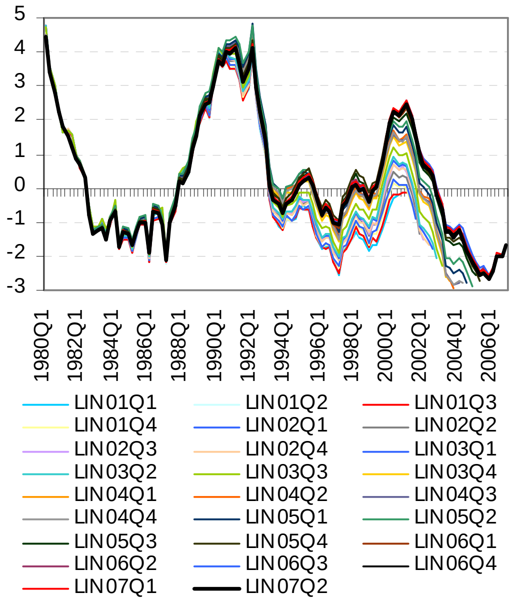
Sources: EABCN and own calculations.

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.

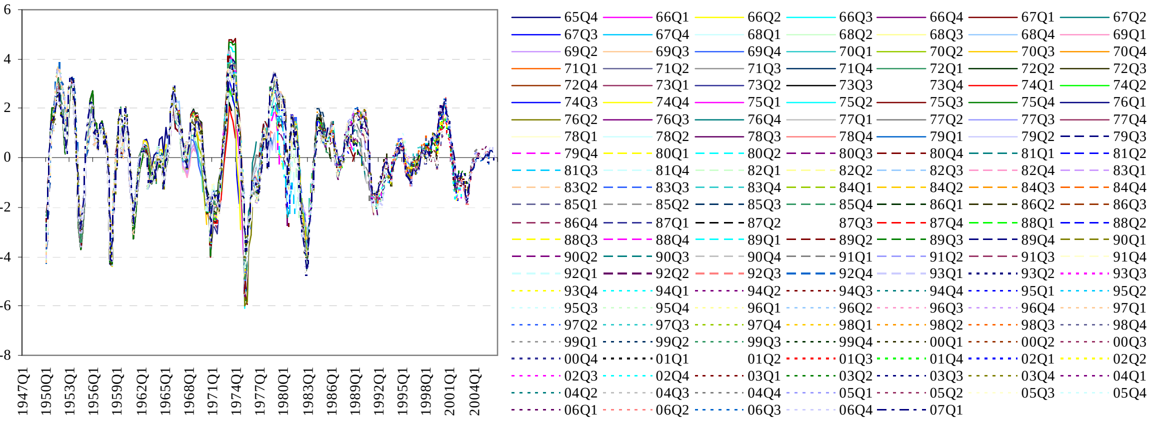
"sign" refers to the percentage of times the real time estimate has the same sign as the final estimate

"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)

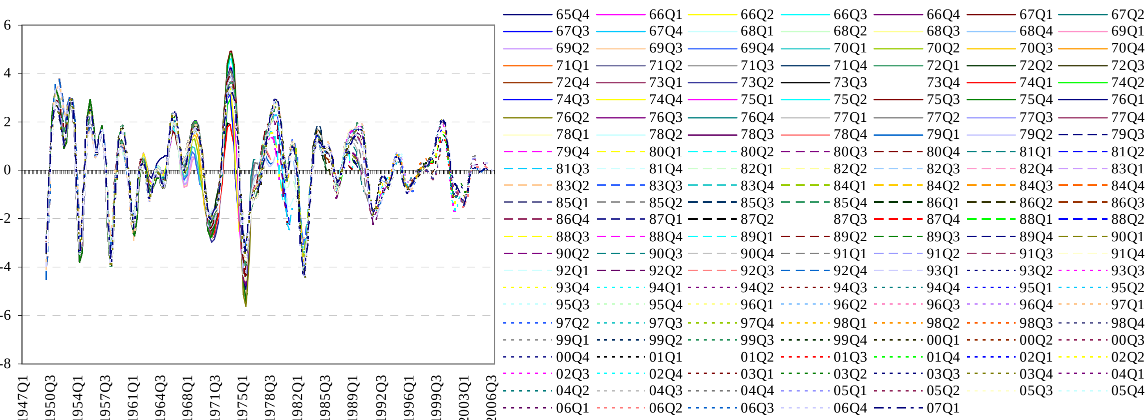
(percentage deviations from trend output)



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)

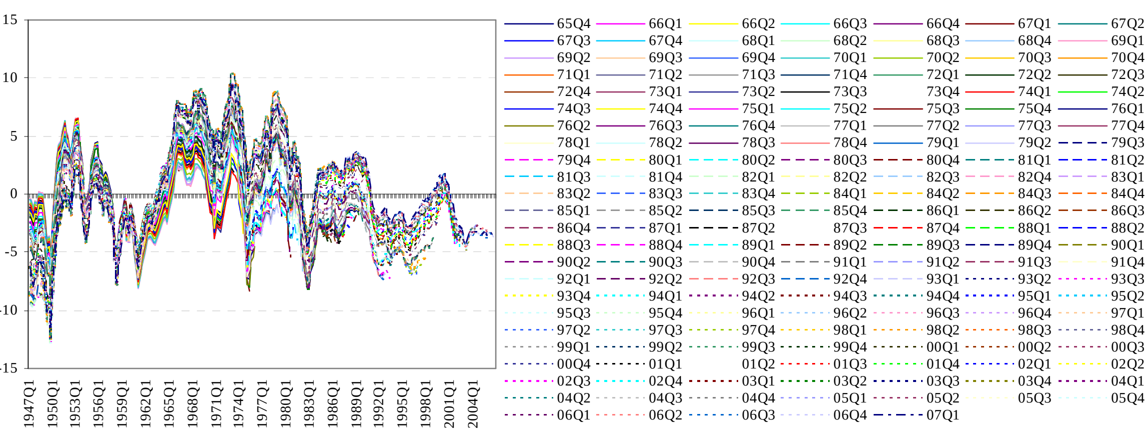
(percentage deviations from trend output)



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.