Heikki Oksanen

Adjunct Professor, University of Helsinki, hkk.oksanen@gmail.com

This Working Paper version 2018-05-21

Comments welcome.

New Output Gap Estimates for Assessing Fiscal Policy with Lessons for Euro Area Reform

Abstract

The purpose of estimating output gaps (OG) is to distinguish between the cyclical and more persistent changes in macroeconomic development. They are used for identifying the cyclical and structural components in government balances for the purpose of designing fiscal policy.

The OG estimates produced by the leading economic policy institutions have spurred quite heated controversies as they are at the same time vitally important for policy and difficult to define and measure. The debate has intensified, to a great extent due to the difficulty in judging the causes and consequences of the Great Crunch from 2008 onwards.

We present new OG estimates based on the single variable HP filter applied on the GDP, keeping them simple to be easily understood also by non-experts. The new estimates for the HP trend of GDP and the OG for the current year and the next will depend on an explicit assumption of the underlying growth of the GDP in future, allowing this to be varied. In the illustrations the underlying long-term growth for the EA11 countries will be taken from the work of the Ageing Working Group (AWG) in the EU. The new results are based on real time data from the forecasting rounds over 2002-2017. They are compared to the real time OG estimates published by three institutions (EC, IMF and OECD) for the 11 EU member states that formed the euro area in 1999 (EA11), the four largest member states and the US.

The new results give policy-relevant insights to economic development. Their relevance is assessed with regard to their use for policy design, notably to separate the cyclical and structural components of government balances in each conjuncture.

In the euro area, fiscal policy has been procyclical with the only significant exception being the year 2009, when the GDP fell significantly. Procyclicality is caused by short-sightedness stemming from the mistrust that all member states are not willing and capable of assuring the sustainability of their public finances. The flexibility for conducting counter-cyclical policy requires that the long-term sustainability of public finances is assured. This requires reforms of the public pension systems and health-care financing for containing the pressures on public expenditures under population ageing. The various policy issues are discussed with reference to the ongoing discussion on reforming the euro area architecture.

This version 2018-05-21

New Output Gap Estimates for Assessing Fiscal Policy with Lessons for Euro Area Reform

Contents

1. Introduction

Distinguishing the cyclical and more persistent changes in macroeconomic development by estimating output gaps (OGs), together with their projections a few years ahead, has become a dominant feature in designing fiscal policy. In particular, they are used for identifying the cyclical and structural components in government balances, producing crucial data for policy making in the euro area (and in the European Union more generally), as well as in the surveillance work of the International Monetary Fund (IMF) and the OECD also on other counties.

The Hodrick-Prescott filter (HP filter) developed in the 1990s was the first customary method for estimating the trend of the GDP. In general, the HP filter generates a more refined trend than a simple (log)linear regression by comprising a smoothing parameter that synchronises the result with possible persistent changes in the growth rate; the user is allowed to choose the parameter value so that, in the case of the GDP, the computation provides a result where the cyclical fluctuations are distinguished from the more long-term changes in the growth rate as accurately well as possible. The OG was then defined as the difference between the GDP (data for the past and forecast for the future) and the HP trend as a percentage of the latter. It uses only the original statistical series for the GDP, without adding data on other macroeconomic variables.

Subsequently, as this simple method was not considered fully satisfactory, additional data on the stock of capital and the labour market was inserted for estimating the reference level, which was then most often called 'potential GDP' instead of 'trend'. The IMF and the OECD shifted to the new method called the production function approach (PF), and in 2002 the European Commission (EC) introduced its own PF estimates and gave them a dominant role, though continuing to produce also the HP estimates until today (Havik et al., 2014).

The OG estimates take a central role in designing and assessing fiscal policy, notably in determining cyclical and structural components of government balances and indirectly in assessing the sustainability of government debt. They have spurred quite heated controversies as they are at the same time vitally important in designing fiscal policy and difficult to define and measure. Conflicting assessments and policy advice derived from them have not vanished even if the OGs have now been intensively used for 15-20 years; on the contrary the debate has intensified, particularly due to the difficulty in judging the causes and consequences of the Great Crunch of 2008-09 and onwards.

In assessing the various OG estimates it is necessary to emphasise that it is not sufficient to compare the OG estimates as such, but we should remember that the OG is always, by definition, the deviation of the actual GDP (past data or a forecast for the future years) from the estimated HP trend or the potential GDP produced by other methods. Each of these may change and it makes a difference which one.

In this paper we shall mostly use and work on the data on the aggregate of the 11 EU member states that formed the euro area in 1999 (EA11), adding observations on its four largest members (Germany, France, Italy and Spain) and the US.[1]

---

[1] We use a fixed composition of the 11 member states to avoid any effects caused by new members coming in. EA11 accounts for 96% of the GDP in the euro area (EA19) in 2017. As our real time data on the OGs will start from 2002, we could have included Greece, which became a euro member in 2001. However, as it is a special case, it is left out. Its GDP out of EA11 was at its peak around 2.5% and 1.7% in 2017, so including or excluding it from our aggregate hardly affects our results at all. In 2017 the four largest members accounted for 79% of GDP in EA11 and 76% of EA19.

Short survey of previous studies with critical comments

Deutsche Bundesbank (2014) launched a fierce attack on the OG estimates of the IMF and the OECD. It proclaimed that they are biased towards loose fiscal policy and an unintended increase in public debt and that the large retrospective revisions of these estimates disqualify their use for policy recommendations.

Largely to avoid large revisions of the OG estimates afterwards Bundesbank promotes a simple HP filter formula using a (low) smoothing parameter which generates a result where the filtered GDP series follows more closely the actual GDP data than the potential GDP estimates of the other institutions.

Bundesbank raises a most relevant issue. However, emphasising the revisions of the OG estimates provides only a narrow view of the issue. They are not directly comparable to revisions of economic data in general. Essentially, the OG is an indicator of cyclical conditions and it is obvious that the OG for any given year will be revised even several times according to the developments taking place after the year in question. Also the sign of the OG may easily change as the estimates are often not far from zero, and therefore, if the series is shifted in either direction, its sign may change. This may happen especially if a major shock hits the economy.

The sharpest possible counterargument to Bundesbank's view is presented by the experts at the OECD in Turner at al. (2016, 21), an OECD Working Paper (without mentioning the Bundesbank's report explicitly!). They point out that choosing an extreme smoothing parameter of zero in computing the HP-filtered series (applied to the observed GDP data in the past as the Bundesbank is doing) would produce a series that never requires revision as it would be identical to the actual series, and the OG would logically always be zero. More constructively, they consider that the ex post revisions should not be the only or decisive criterion to judge the competing OG estimates.

McMorrow et al. (2015) complement the Bundesbank analysis by assessing also the performance of both PF and HP estimates published regularly by the EC as these were not covered by the Bundesbank. They conclude that the EU's PF method has performed better than the HP filter and the PF estimates by the IMF and the OECD.

Dubious negative assessment of the HP method

McMorrow et al. (2015) discredit the HP method referring especially to its poor performance in the estimates of the EC for the HP trend of the GDP in spring and autumn 2009. Their judgement turns out to be based on the mechanical statistical procedures in dealing with the well-known end-of-the-sample bias in the HP estimates: the HP formula gives a high weight to the latest observations of the original data series, which tends to bend the end of the estimated HP-filtered trend upwards or downwards depending on latest data. As the main interest is usually exactly in those years, i.e. in the current conditions and the short-term forecast, the statisticians have attempted to find operational solutions to deal with it. However, no mechanical solution would work well in all cases. One such failed case turns out to be the HP estimates of the EC in year 2009, when the direness of the economic situation had emerged.

The critical case is in the spring 2009 forecast of the EC. McMorrow et al. (2015, 12) present a graph where the OG for the euro area is zero for both in 2009 and 2010, while all other estimates gave a significant negative value (which became the dominant view and has not been challenged since then). Six months later, in autumn 2009 forecasts the HP estimates were

significantly revised, including a negative value for the OG from the HP filter. This swing in the estimates then led to the denouncement of the HP method in general.

The source of this outcome seems to be an unexplained swing in the HP trend GDP estimates for 2009 and 2010. They were calculated from a statistical series where the GDP data were extended, using a mechanical projection, until 2013 (in the spring 2009 forecast and until 2014 in autumn 2009). In spring 2009, when the economic situation had dramatically changed, the HP trend projection was revised drastically downwards to a persistent and continuous decline in 2009-2013. This absolute decline was projected for several countries, the most significant being Germany and the Netherlands.

This projection was then turned around half a year later in the autumn 2009 forecast. The negative trends had disappeared (except for Ireland and Greece, the latter not being included in our EA11), and the projection for the EA11 GDP trend level for in 2013 had become eight per cent higher than in the spring 2009 forecast (the data are presented in the Technical appendix).

There is no need to try and discover how this remarkable swing in the HP estimated trends by the EC over these two forecasting rounds came about. Most probably, the figures in spring 2009 were produced mechanically based on uncoordinated assumptions. The reasonability and consistency of the results was not examined as the HP estimates are not used in policy analysis, but only produced for reference, hanging over from the negotiations to shift to the PF method in 2002.

The only reason why the swinging HP estimates by the EC in 2009 are referred to here is that those estimates led McMorrow et al. (2015) to discredit the HP method in general as 'unstable' and 'liable to revisions'. This conclusion was inaccurate as those swings were obviously caused by some specific problems in applying the HP method.

The OG estimates are important and need to be assessed carefully

Several recent studies compare the merits of the parallel and often competing estimates for the OG in terms of their stability in the short term and proneness to revision even after several years. For example, Busse (2016) and Kuusi (2017) investigate the impacts of the revisions of the output gaps and cyclically adjusted budget balances under the fiscal framework in the EU. Recent IMF Working papers on improving the estimates for potential output and OG (Blagrave et al., 2015, and Alichi, 2015) develop improvements to their PF methodology.

Practically all studies conclude that the OG estimates are indispensable for policy making – distinguishing trend from the cycle (and both of these from any specific factors) is both important and difficult. First, it is useful to admit that before the Great Crunch of 2008-09, the OG estimates used in policy making failed to guide the policies to dampen the boom. Second, the retrospective revisions of the OG estimates for 2006-2013 is not an adequate reason to abandon them altogether. Instead, the OG estimates should be looked at constructively, though critically, and they can be improved, especially as they are being continuously used for policy design.

The purpose, scope and outline of the present paper

The motivation of the present paper is that there is scope for improvements. Doing this with an open mind, admitting that no perfect method will appear, we shall present new OG estimates produced with a new application of the HP filter. Special emphasis is given to the transparency and simplicity of the method so that the results could be understood also by non-experts and policy makers.

As noted above, the main interest in designing policies is most often in the current situation and in the immediate future. We focus on real time estimates, i.e. the estimates published by the various institutions at the time in the context of their regular forecasts.

In Section 2 we present the new real time estimates for the OGs based on the HP-filtered trend of the GDP. Results for the aggregate of the 11 EU member states that formed the euro area in 1999 (EA11), its four largest members and the US will be reported. The novel idea in the present application is to make it explicit that the estimates for the trend GDP and the OG for the current year and the next will always depend on the assumed path of the GDP in the consecutive years. The new estimates are compared to the real time estimates by the EC, IMF and OECD over the period 2002-2017.[2]

In Section 3 we shall discuss the main use of OG estimates for policy design, which is to separate the cyclical and structural components of government balances. Especially their projections based on planned fiscal policies obtain pivotal importance under the (complex) fiscal policy rules for the euro area (*Vade Mecum on the Stability and Growth Pact,* European Commission, 2017a). In this context we present the data on the retrospective revisions of the OG estimates.

In section 4 we look at the sources of those revisions as the economic developments that cause the revisions may not be determined only by exogenous factors but significantly also by policy decisions. The question is to what extent were the policies responsible for causing the further fall in GDP in 2012-13 and hence the revisions to the OG estimates. We shall discuss the conflicting views on this, highlighting the evidence of the persistent procyclicality of fiscal policy in the euro area.

In Section 5 we discuss the lessons for the ongoing discussion on reforming the euro area architecture more generally. Identifying the problems correctly is indispensable for dealing with them. Procyclical fiscal policies seem to be a central issue and the use of the OG estimates and other features of the Stability and Growth Pact (SGP) need to be properly addressed. Counter-cyclical flexibility in the short term requires that sustainability of public finances in the long term is assured. This is an angle from which to assess the various proposals for reforming the euro.

Sections 6 gives a summary and concludes the paper.

---

[2] The data set required to perform these tasks is quite large: it contains all the forecasting vintages of the EC in 2002-2017 gathered from the original sources and made freely available by the FIRSTRUN (Fiscal Rules and Strategies under Externalities and Uncertainties) project; it contains data by EU member for several macroeconomic indicators, and the historical data goes back to 1960. We do not use all indicators but mainly those for the GDP, estimates of the potential GDP and the trend estimated by their HP application (and by implication for the respective OGs), the various budget balance indicators, etc. Similar data covering the real time estimates by the IMF and the OECD are downloaded for the purposes of the present paper from their respective freely available data banks.

2. New OG estimates generated by the HP method

Real time OG estimates for EA11

The new OG estimates are the deviations of the observed (and forecast) data from a HP-filtered trend of the GDP, which is derived using our new application: the GDP series as an object is based on the real time data on the GDP, including its short-term forecast by the respective institution (for EA11 by the EC; for the US by the IMF) and its extension to future years. The extension is composed as follows: (1) the official forecasts of the respective institution for the current and coming years (t and t+1) are taken as given, (2) an *assumption* for the *underlying* long-term rate of GDP growth from year t+2 onwards is injected, and (3) the GDP from year t+2 onwards is additionally set to adjust so that the OG estimated for year t+1 will fade away in the medium term (an illustration using stylised data is presented in the Technical appendix).

The assumption on the future GDP is vital in this application. It is one source for tracking the revisions of the OG estimates afterwards: the OG will be revised due to the deviation of the GDP from its previously assumed path. Extending the GDP series into the future also allows us to cope with the possible end-point bias in the HP-filter applications. Naturally, the assumption on the underlying future growth is set by the user of our new application, allowing generation of any number of alternative results.

For the base line, the underlying long-term growth rates for the EA11 countries are roughly based on the work of the Ageing Working Group (AWG) in the EU. In the AWG reports 2001, 2006 and 2009 the rough figure for EA11 GDP long-term growth was 1.5 % pa. In the report 2012 it was lowered to 1.3%.[3] As the views of long-term growth obviously started to change earlier, we set the figure at 1.4% for our calculations for 2010 spring and autumn forecasting vintages, and 1.3% from autumn 2011 onwards (the assumptions of the AWG report 2012 were published in 2011). For the four largest countries the growth assumptions are similarly made on the basis of country-specific assumptions in the AWG reports.

The assumption that the OG in year t+1 will trigger a specific adjustment in the medium term, such that the extended GDP series closes the gap by year t+5, follows the conventional practice in the AWG work and elsewhere when long-term projections are constructed; we present below a sensitivity test by extending the adjustment to year t+8 for the autumn 2009 forecasting round, when the OG for EA11 2010 was strongly negative.

The GDP data used starts from 1960 and the extended data runs until 2040, long enough to feed into the calculation the assumed underlying growth rate in future. For estimating the HP-filtered trends we use logarithmic series (as the GDP series normally grow exponentially) and 100 as the smoothing parameter, which is a conventional practice with macroeconomic annual data.[4]

---

[3] This coincides also with the projection by McMorrow et al. (2016, Table 1) who produce a no-policy-change medium-term projection for 2015-2024 of 1.3 % average growth for euro area GDP, based on the negative fallout from the financial crisis and the emerging drag on growth emanating from ageing populations.

[4] As the 10-year average annual growth rates for the EA11 decelerated from five per cent in the 1960s to just above two percent before 2008, the smoothing parameter 100 seems to work well. Higher values would not take into account the deceleration of the long-term growth. In the Technical appendix we show that a small value, like 6.25 favoured by the Bundesbank produces a 'trend' that clearly bends with the cyclical movements in the GDP series and therefore does not measure the amplitude of the cycles in a reasonable way.

The new results are based on real time data from the forecasting rounds over 2002-2017. They are then compared to the real time OG estimates published by the three institutions (EC, IMF and OECD) for EA11, its four largest member states and the US. All autumn forecasting vintages over 2002-2017 are treated, together with the spring forecasting vintages over 2007-2010 (to cover in more detail the developments before and after the outbreak of the economic crisis).

Figure 1 gives a broad picture, the real time OGs for EA11 for the current year (t) generated by the new HP (nHP) method and those of the EC, IMF and OECD in the left panel, and the corresponding results for the following year (t+1) in each case on the right.

The figures show that, with one single exception, the real time OG estimates of the three institutions were clearly negative or zero throughout the whole period 2002-2016. The only exception was the estimate (+0.6 %) by the EC in autumn 2008 for 2008, and even in that case the forecast for the following year was a negative OG.

The OG estimates of the IMF and the OECD until 2009 are quite close to each other, and also the EC estimate is practically identical for 2006-2008, i.e. during the boom that was recognised only afterwards. After 2009 the estimates by the OECD are clearly the most negative ones, while those by the IMF and the EC are close to each other.

Our new HP estimates show slightly positive OGs for the boom period 2006-2008. They are consistently higher than any of the three estimates, but also their average is negative over the whole period since 2002. By construction the HP estimates should be close to zero if the period is sufficiently long and the fluctuations are reasonably regular. This is not the case here as an exceptional crisis occurred (2009), followed by another fall in output (2012-13). Thus, also our new HP method produces a negative average for 2002-2017.

The Great Crunch of 2008-09 is the most important one for interpreting what happened and how it was understood. All real time estimates recognised it in spring 2009, including our new HP-based estimate. The change from earlier estimates was dramatic, the most negative OGs being presented by the IMF and the OECD.

McMorrow et al. (2014, 17) conclude, to the advantage of the estimates by the EC, that the OG estimates in spring 2009 for the euro area of -4.3 % and -5.5 % from the IMF and the OECD respectively turned out to be too low and they were later revised so that they became almost identical to the EU's initial (i.e. real-time) spring 2009 estimate (-2.8 %).

For a broader picture we need to remember that the differences in the OG estimates always come both from the differences in the estimates of the potential GDP and of the forecast for the GDP itself. It turns out that the more negative OGs by the IMF and the OECD partly stemmed from their more pessimistic GDP forecasts for 2009, which were subsequently revised upwards. Also, the not-so-negative OG estimate by the EC for 2009 partly stems from a comparable low estimate of the potential GDP (in the middle of the crisis), which was later revised upwards.

Figure 1. Output gap estimates for EA11 in real time forecasting vintages 2002-2017. For the current year (t) on the left and for the next year (t+1) on the right.
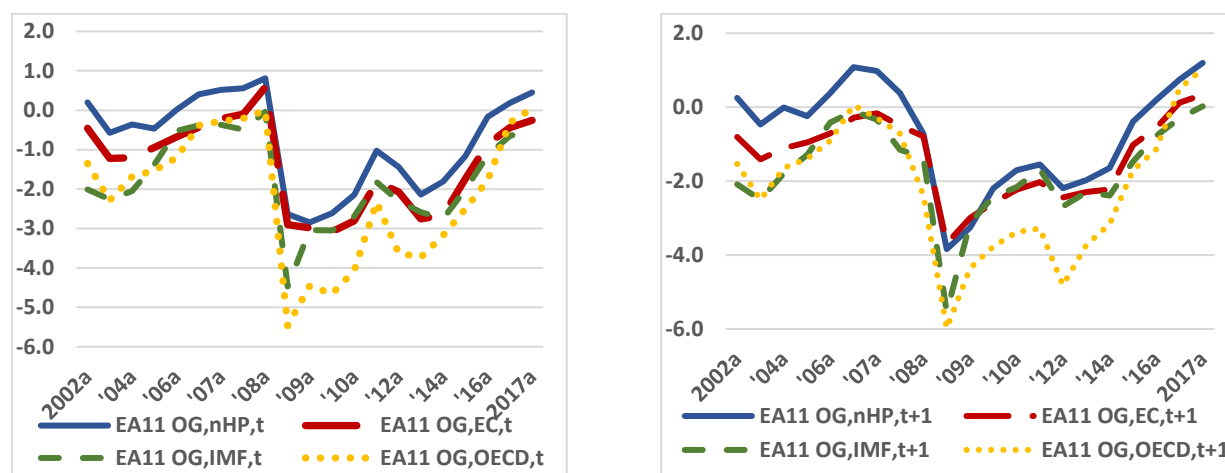

Legend: OG = output gap; a = autumn forecast, s = spring forecast. nHP = new HP-based estimate based on real time data from the EC including forecasts for t+1; EC = European Commission.

Reconsideration of the nHP trend and potential output 2007-2013

The Great Crunch 2008-09 led to revised estimates for the HP trend and potential GDP by the three institutions. Our new HP estimates automatically generate the HP projection for the GDP so that the levels projected at each forecasting vintage can be easily compared at any chosen future year.

Figure 2 shows the results for EA11 from selected vintages: before the Great Crunch (autumn 2007), after it was recognised (autumn 2009), when a revival of growth was projected (autumn 2011) though failed to happen, when the negative growth 2012-2013 was about to be over (autumn 2013), and for hindsight, according to the recent retroactive estimates (autumn 2017). The ex post GDP data is presented for reference.

Figure 2. New HP estimates for EA11 GDP in 2007-2013 according to selected forecasting vintages.
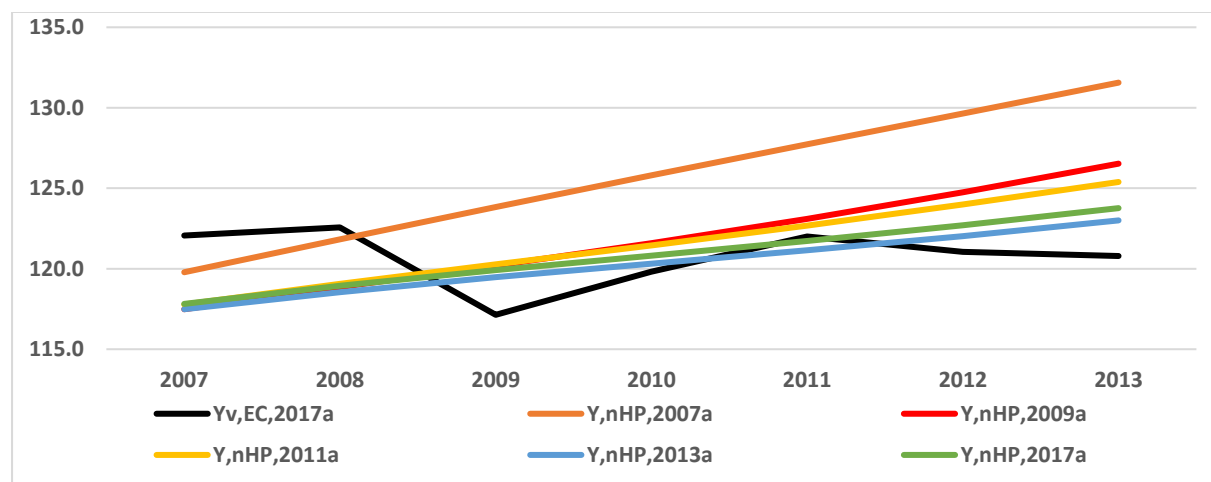
Legend: all series ind 1998=100; Yv,EC,2017a = EA11 GDP volume based on EC autumn 2017 data; Y,nHP = new HP estimate for the trend of GDP; a = autumn forecasting vintage of the year; for other details see the text.

The nHP trend estimate in autumn 2009 for 2009 was 2.9% lower than projected in autumn 2007, and as the projected growth rate (the slope of the projection) also declined the relative difference was on the increase. After a marginal revision in the autumn 2011 vintage the second recession in 2012-13 led to a further decrease in the nHP trend estimate: the estimate in the autumn 2013 vintage for 2013 was 1.9% lower than projected in autumn 2011. The revival of growth then led to a small upward shift shown in the autumn 2017 vintage; the projected level in autumn 2017 for 2017 is 7% lower than projected 10 years earlier (not shown in the graph).

The projections for potential GDP by the three institutions also show that the prospects were significantly revised downwards, although some interesting differences also appear. Their time spans vary and impede straightforward comparisons (some graphs and data are shown in the Technical appendix), but the most interesting observations are the following.

The downward revisions of the EC projections were larger than in our nHP estimates in autumn 2009 and remained so throughout. In autumn 2007 the IMF did not publish an estimate for the potential GDP for year t+2, but the level and slope of for 2007-08 was buoyant. Later, its projections do not deviate significantly from those by the EC. The OECD was revising its projections down more slowly than the others, and still in autumn 2013 its projection for potential GDP was higher than those of the other institutions. By the latest projections in autumn 2017 its view has converged to those of the others. Our nHP autumn 2017 estimate for 2017 is a good one percentage point lower than the other three, consistently with its regular pattern.

Real time OG estimates for the US

Figure 3 presents the new HP OG estimates and those of the three institutions for the US from the forecasting vintages from autumn 2002 to autumn 2017. Again, the forecast for the current year (t) is on the left and for the following year (t+1) on the right.

Figure 3. Output gap estimates for the US in real time forecasting vintages 2002-2017. For the current year (t) on the left and for the next year (t+1) on the right.
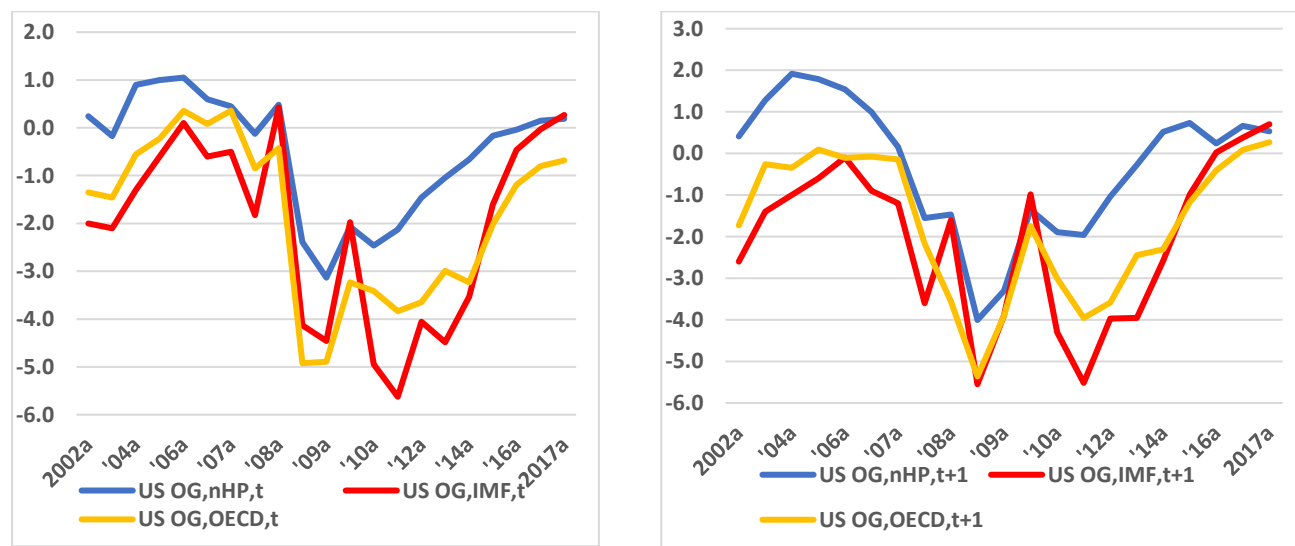

Legend: OG = output gap; a = autumn forecast, s = spring forecast. nHP = new HP-based estimate based on real time data from the IMF, including forecasts for t+1.

The nHP estimates are based on the IMF data and forecasts for the current and following year. The assumption on the underlying GDP growth rates is set at a round number 2 % throughout, i.e. higher than for the EA11. This is based roughly on the differences in the projections for the rate of growth of the population in the long term.

Like for EA11 above the OG,nHP estimates are consistently higher than the other two. In this case the OECD estimates are slightly positive in 2006-07, while the IMF estimates just hit zero in autumn 2006 and 2008 forecasts. The averages over 2002-2015 reveal a negative tendency in both IMF and OECD estimates, while it is just over half a percentage point (pp.) in the nHP. Apart from this, the fall in the estimates from autumn 2007 to autumn 2009 is interesting. The largest decline is in the OECD estimate, while it is 4 pp. for IMF and 3.5 pp. for the nHP.

The nHP trend estimates for the US in Figure 4 since autumn 2007 and the corresponding estimates for the potential GDP by the IMF and the OECD (in the Technical appendix) show a roughly similar pattern as for EA11 above. The largest negative correction in the vintage data shown is in autumn 2011, being revised slightly upwards since 2013. The graph also shows the GDP path ex post depicting that in the US GDP was growing steadily, unlike the second recession occurring in EA11.

Figure 4. New HP estimates for the US GDP in 2007-2013 according to selected forecasting vintages.
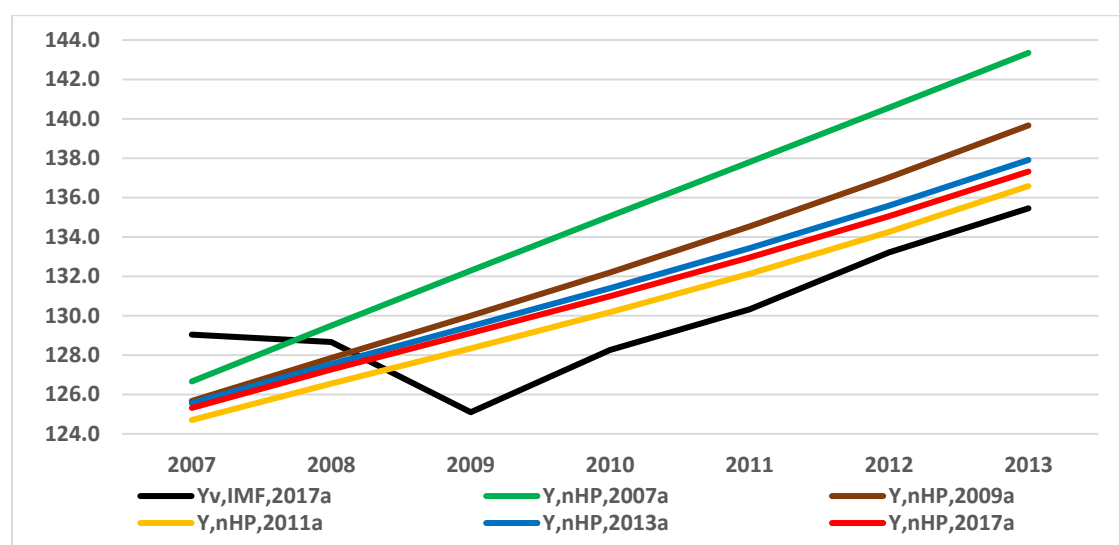
Legend: all series ind 1998=100; Yv,IMF,2017a = US GDP based on IMF autumn 2017 forecast; Y,nHP = new HP estimate for the trend of GDP; a = autumn forecasting vintage of the year; for other details see the text.

Real time OG estimates for the four largest euro area countries

Figure 5 presents the new HP OG estimates and those of the three institutions for the four largest euro area economies from autumn 2002 to autumn 2017, for the current year (t) on the left and for the following year (t+1) on the right.

Figure 5. Output gap estimates in real time for DE, FR, IT and ES, forecasting vintages 2002-2017. For the current year (t) on the left and for the next year (t+1) on the right.
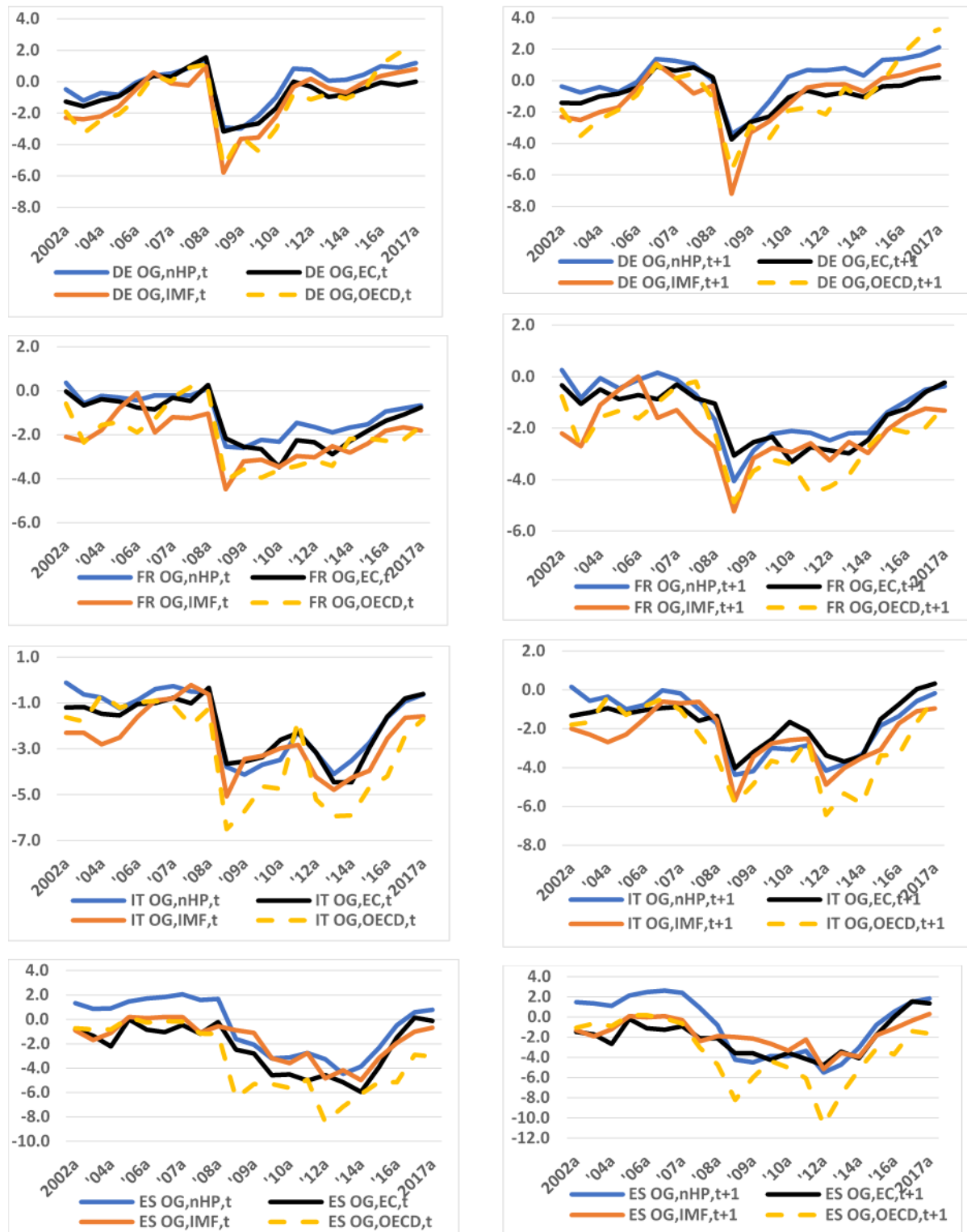
Legend: OG = output gap; a = autumn forecast, s = spring forecast. nHP = new HP filter estimate based on real time data from the EC, including forecasts for t+1. DE = Germany, FR = France, IT = Italy, ES = Spain.

The patterns are broadly the same as for EA11. For Germany the estimates of the institutions were mostly positive in 2007-08, but for all the other three countries almost consistently negative or zero, with the exception of the EC estimate in autumn 2008 for 2008. The estimates for Spain were practically zero or negative throughout the boom 2005-07 as we know it in retrospect (the issues in not identifying it in real time is left out of the scope here).

A preliminary assessment of the merits of the new HP estimates

In the present paper we do not engage in high powered statistical tests for comparing the various HP trend and potential output estimates and the corresponding OG estimates. Instead, we prefer to discuss the various substantive problems in such comparisons, like for example that the differences in the OGs between those produced by the various institutions and our nHP estimates stem not only from the method to estimate the potential GDP or its HP trend, but also from the short-term forecasts for the current and following year. There can be many reasons for those differences, including the assumption on the fiscal policy and the size of the fiscal multiplier in each occasion, making it difficult to reveal them by simple statistical tests.

We noted the general tendency in the real time OG estimates of the three institutions that they have been almost always negative. There can be justified reasons for this. For example, as the OECD puts it, one of the factors behind their estimate for the unemployment gap is that estimated equilibrium level of unemployment depends on the inflation rate (and expectations thereof). It is plausible that most of the time since 2002 unemployment has exceeded this reference level and therefore this factor tends to keep their estimate for the OG below zero.

These types of possible substantive issues mean that the different estimates may serve different purposes. None is superior in all respects and no simple statistical comparisons incorporate all relevant aspects. The question nevertheless arises how an OG estimate that is almost always negative can be interpreted as cyclical as in ordinary language a cycle should fluctuate on both sides of the reference level. Perhaps, then, one should rather pay attention to the deviations from the long-term average of the estimated series or, alternatively, use the annual changes, which is another way to eliminate a possible bias in the level.

Looking at the graphs above, we can conclude that in broad terms the various estimates give a similar picture apart from the level over the whole period on average. Our new HP estimates have several merits. First, they are simple to be explained and they are not based on any refined (and therefore always contestable) economic theory. Second, the meaning of the relatively small number of assumptions put on top of the GDP data can be easily understood and the sensitivity of the results using alternative assumptions can be easily worked out.

The assumption on the underlying GDP growth rate is the key and it should be tested. The result for the EA11 OG estimate in 2009 when we have the largest deviations from zero, is that if the assumed underlying growth rate was already in autumn 2009 lowered to 1.3% instead of 1.5% in the base line, the OG would be 0.2 pp. higher (-2.65 instead of -2.84 in autumn 2009 for 2009). Thus, limited variation of this assumption hardly affects the result at all. In section 5 below we present a more dramatic test by assuming zero growth from 2018 onwards.

Our nHP method also offers the possibility to make tests the other way round: we could ask which assumption for the underlying long-term growth rate is required in each situation to attain the OG estimates of the three institutions. The result EA11 is that growth rates should have even accelerated from the well above 2% rate prevailing in 2006-2007.

Another assumption to be tested is the speed of adjustment of the GDP from t+1 to the trend level. In the base line we used the conventional assumption that the gap is closed by year t+5.

There is hardly any serious question about the reasonability in normal fluctuations, but for the extreme situations a longer adjustment path should be tested. The result is that if we allow three more years in the autumn 2009 estimate for 2010, the OG estimate is -2.9% instead of -3.3% in the base case. Thus, reasonable variation in this assumption does not affect the results significantly but noting that there is no definitely correct result is a healthy reminder that what we can achieve by any method is a range of (hopefully) useful estimates.

The smoothing parameter of 100 assumed for our estimates is not based on any hard facts or analysis but it is merely a conventionally used value in filtering annual GDP data. In the Technical appendix we present the results derived from a log-linear trend and the HP filter with a smoothing parameter of 6.25 favoured by the Deutsche Bundesbank (2014) mentioned above. Based on them, 100 as the value seems reasonable and limited variation around it would not affect the results significantly.

Our new HP application could be modified in the respect of which forecast is used as the basis. For each real time estimate we used the forecast of the respective institution (EC for the EA11 and IMF for US) for the following year (t+1) and allowed our application to produce the projection from then onwards. The method would easily allow us to use also the existing forecast for year t+2 (regularly available in the autumn forecasts and apply our method from there onwards). However, looking at our results and the EC forecasts for t+2 it seems that such an alternative computation would give almost the same results as here.

In the previous studies the relative merits of the OG estimates have been judged according to their retroactive revisions. The adequacy of those assessments should be discussed in the context of the use of the OG estimates in policy design. This is where we turn next.

3. Using the OGs and structural balances for policy: a critical view

Distinguishing the cyclical and structural components of government balances

The OGs are used to design policies in real time and to assess them ex post, mainly for the short term, but there are also connections to fiscal targets in the medium term and sustainability of government finances in the long term.

The primary use of the OGs is to separate the cyclical and structural components of government balances in each conjuncture. This is a centre piece of fiscal policy making, originating from the work at the OECD since the early 1990s and fully rooted in the regular reports on all countries by all the three institutions, including extensive data available in their data banks (see Mourre et al., 2014, for references).

The standard procedure is to estimate the cyclical component of government budget balance by multiplying the estimated OG by the semi-elasticity of the budget balance (defined as the effect of movements in the GDP on the budget balance as a percentage of GDP). The latest estimate for the semi-elasticity for the EU countries is 0.50 for the EU, ranging from 0.31 to 0.65 across member states (Mourre et al., 2014, 6). Removing the cyclical component gives the cyclically adjusted balance, and subtracting one-off and temporary measures results in the structural balance, which has recently gained a central role in designing fiscal policy.

Applying the standard procedure on real time data gives the real time estimates for the cyclical and structural components. In monitoring and designing policies the three most relevant OG estimates are those for the previous, current and next year, although all estimation methods

always produce also the whole history backwards (and, depending on the exercise, also the forecast for year t+2 and even a mechanical extrapolation until t+5).

Use of structural balances in euro area policy making

The structural balances have gained a pivotal role in the fiscal policy rules for the euro area, explained in the most recent edition of the *Vade Mecum on the Stability and Growth Pact* (European Commission, 2017a, where presentation of these complex rules requires 224 pages).

Probably the best interpretation why this has happened is that budgetary surveillance had to be improved, after the experience of the Great Crunch of 2008-09 at the latest. The original SGP of 1997 focussed on the headline budget balances, but it became obvious that setting sensible fiscal policy targets the cyclical component of the balances had to be acknowledged. As well-working indicators for measuring policy efforts are not available, compromised solutions had to be accepted even if, again in this case as in the euro area (and in the EU) in general, the results were simple and open to criticism.

As the new OG real time estimates for the period 2002-2017 are on average one percentage point higher than the ones by the three institutions, the cyclical components would be roughly 0.5 percentage points higher and the estimates for the structural balances correspondingly lower. These are the rough numbers based on the real time OG estimates for the current year and the next, while the corresponding figures for the differences in the OGs for the subperiods and institutions vary roughly between 0.3 and 1.8 percentage points.

Here, we do not ponder further details along these lines, by country or otherwise. Instead, we debate the conceptual basis of these budget balance estimates, questioning their usefulness and accuracy for policy design.

How to define a cyclical component in government balances?

What should we mean by the cyclical component of the budget balance? Consider a thought experiment: assume an economy in equilibrium (or steady state growth) that is hit by an exogenous factor which moves cyclically, i.e. it affects the economy first, say, negatively and then turns positive in a symmetric and regular manner; assume then that the government pursues fiscal policy that completely eliminates these effects on the GDP; this means that the OG remains at zero through the whole cycle; obviously, the budget balance first deteriorates due to expansive fiscal policy and then improves, and at the end of the cycle the equilibrium returns; as the OG is zero over the whole cycle, the cyclical component of the budget balance according to the standard definition is zero throughout. Does this make sense? Not necessarily as it would be reasonable to say that all the variation in the budget balance was cyclical as it stemmed from the specific joint effect of the exogenous cyclical factor and the countercyclical policy pursued.

This was naturally an extreme example to express an analytical conundrum. The point here is that the standard definitions of cyclical and structural components in government budgets ignore the effects of changes in fiscal policy on the GDP and via that to the measured OGs that are used to define and measure the cyclical components. The loop from fiscal policy to the GDP and the OG is missing, even though the existence of this loop is a core part of conventional economics (ignored only by the extremists who rely on the simplest possible interpretation of the 'Ricardian equivalence' and declare that fiscal policy does not matter; this was not the view

of David Ricardo, who formulated the equivalence theorem but considered that neither the assumptions nor the implications were valid, see Elmendorf and Mankiw, 1999).

Ex post revisions of the OG estimates obscuring policy

If the preceding argument were too abstract, we should recognise various other reasons why the structural balance is not a reliable indicator of discrete policy actions. One of them is that the OG estimates are regularly changed afterwards as new data comes in, and thereby the estimates for the cyclical and structural components in government balances are also changed.

The significant revisions seriously hamper the use of structural balances (or their change from one year to the next) as indicators or even proxies of discretionary policy actions. Logically, a measure of a discrete policy action at any given time should not depend, even in retrospect, on what happened in the economy afterwards. Sure, its combined effects may depend on other factors, including their future developments, but the original identification and definition should not. Due to its ex post revisions the structural balance (or its change) does not fulfil this logic.

As the revisions of the estimates for the structural balances are directly derived from the revisions of the OG estimates (multiplying them by the semi-elasticity mentioned above) we can here be confined to observing the revisions only in the latter.

The revisions of all these OG estimates over the period 2002-2017 studied here were indeed significant. Figure 6 shows the OG estimates in real time for EA11 for year t and ex post, the latter being the data of the latest forecasting vintage, autumn 2017. Our new HP estimates and those by all three institutions are shown in separate graphs by their producer.

The revisions are significant and otherwise roughly of the same magnitude (though on average greatest for the OECD). The largest revisions concern the years 2006-2008 in all of them. Busse (2016) provides a detailed analysis of the revisions in the EC estimates for the OGs and the related cyclical and cyclically adjusted balances.

We do not enter into a more extensive tracking of the sources of the revisions, but as we are focussing on the OG estimates for the current and next year we should remember that the revisions come partly from replacing the forecasts for the GDP and other economic variables by their outcomes. This source of revisions is conceptually the same as any updates of statistical data.

Instead, the revisions in the OG estimates mainly come from their intrinsic nature as measures of the cyclical phase and this will, ex post, always depend on what happened to the cycle after any given year. This is not a matter of updating the data but what happened in the economy.

Figure 6. Output gap estimates for EA11 in real time for the current year (t) and corresponding estimates in autumn 2017; new HP estimates and by institutions (European Commission, IMF and OECD).
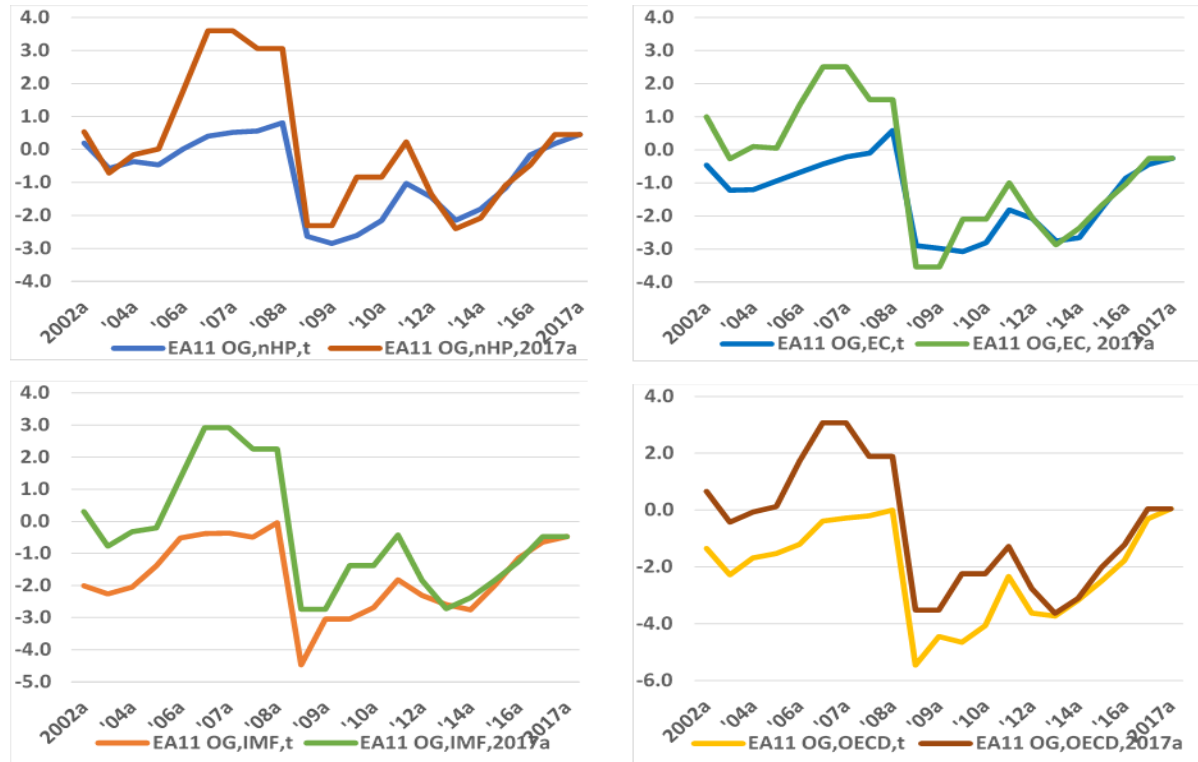
Legend: nHP OG = new HP output gap estimate; EC = European Commission, a = autumn forecast, s = spring forecast.

Our new HP estimates have the advantage that, by their construction, the main factors behind the revisions can be relatively easily tracked to the deviation of what happened to GDP growth since 2006 compared to the assumed underlying growth rate assumed in 2006-07 when the real time nHP estimates were calculated. The assumption on future growth in estimating our new HP trend for EA11 was 1.5 % until 2009, including for example in the real time estimate in 2007 (the peak of the boom). As we now observe ex post that since then the average 10-year growth fell to 0.6 % by 2017, a significant revision of the ex post OG estimates followed.[5]

Tracking the sources of revisions of the PF-based OG estimates of the three institutions must be left beyond the scope of the present paper as it would require going into the detailed construction of each of the estimation methods and their parameters, including changes in them during the period investigated (as the estimates have been modified to improve their accuracy). Also, almost all previous literature avoids this as it would be a major task.[6]

---

[5] Had we used for each HP trend estimate, for example, the average growth over the preceding 10 years, which was still well over two per cent until 2007, the real time nHP trend GDP estimate in 2007 would have been even higher and the nHP OG estimate correspondingly lower, and hence the retroactive revision even larger than shown in Figure 6. The presentation in the Technical appendix with the help of Figure TA.2 based on stylised data illuminates the revision in the nHP trend and the derived OGs caused by a shock to the economy.

[6] Turner et al. (2014) study the various sources of the changes in the OECD estimates for potential output. Recognising the importance of these estimates in policy assessment in all the three institutions, the experts should possibly compare the differences in their methods for tracking the deviations in their results in detail.

Revisions of the OG estimates for the US and the four largest euro area countries are shown in the Technical appendix.

4. Did policies cause the further fall in GDP in 2012-13 and hence the revisions to the OG estimates?

Going back to the basics: estimating the OGs is done to appraise the phase of the economic cycle and distinguish the cyclical component from a possibly more permanent change in economic conditions. Naturally, the assessment will change as time goes by and fresh information of the direction and persistence of the changes become more apparent. For any estimation method, the estimates for the OGs at any given time will always be revised retroactively. This happens irrespective of the causes of the new developments. Most importantly, these causes may include the policies pursued and normally do so.

The revisions of the OGs retroactively has received a lot of attention in the literature and the propensity to revisions has been used as an important criterion for comparing their reliability and usefulness. The underlying causes of the revisions have nevertheless been overlooked in most studies. This possibly indicates an implicit assumption that only exogenous factors were driving the developments. This is surprising as any study on fiscal policy should allow the possibility that fiscal policy matters: why study it if does not?

Busse (2016, 18), in his comprehensive study of the revisions of the OGs, discusses briefly the possibility that the 'errors' in real time estimate may lead to a recommendation to tighten fiscal policy that might unnecessarily jeopardise economic growth; and, vice versa, if the real time OG estimate turns out to be too pessimistic, fiscal stance might become unnecessarily easy. He nevertheless does not pursue the questions of the frequency and effects of such possible occasions of unintended fiscal stance.

Kuusi (2017) studies the OG revisions in the EU in 2002-2014 noting that they indicate significant uncertainty for policy making. He finds that in the median EU country the optimal fiscal policy, as specified with the help of his DSGE model, is mildly countercyclical under the output gap uncertainty, as opposed to appearing strongly countercyclical under the retrospective output gap estimates. Any sensible analysis can be expected to lead to the same result: as distinguishing between cyclical and trend shocks in real time is always difficult, the optimal policy tends to be more cautious than what it would have been if the amplitude of the cycle were perceived to be larger. Additionally, using the retrospective OG estimates for his comparison does not answer the more interesting question of how the policy as pursued in reality diverged from optimal policy as specified under his model.

Other indicators for discretionary policy measures

Although the structural balances are given a central place in the euro area (and EU) fiscal rules (*Vade Mecum on the Stability and Growth Pact,* European Commission, 2017a), there is considerable uncertainty about them both in real time and retrospectively. This has given rise to development of complementary indicators of the fiscal stance that help to analyse the developments before and after the Great Crunch.

Carnot and de Castro (2015a and b) present a new indicator named discretionary fiscal effort (DFE) that combines data on changes in policy on public revenues and refinements to measuring changes in policy on the expenditure side. The latter, defining discretionary

measures on public expenditures, is never clear-cut. For example, unchanged policies combined with moving underlying factors may lead to an increase in expenditures as a percentage of GDP (or other scale factors). This happens in welfare states as population ageing-related public expenditures tend to increase if deliberate policy changes are not made. Despite this, for practical purposes, it is common to define a constant share of GDP as the no-policy-change scenario.

The same applies on the revenue side as taxes are not strictly proportional but a constant share of GDP is only an approximation of neutral policy. Such compromises in defining the concepts are unavoidable and acceptable as there is no way to make a complete mapping between the matrix of the policy parameters and their effects on public revenues and expenditures, not to mention covering all the EU countries. Yet, the national experts working for their ministers of finance regularly need to make estimates of how the proposed changes in various policy parameters affect public finances.

Carnot and de Castro (2015a and b) compare their DFE indicator to *changes* in structural primary balances based on the OG estimates of the EC. Comparing to the primary balance is justified as it removes the effect of changes in the interest rate. Focussing on *changes* in structural balances is appropriate as the issue pertains to *changes* in policy; focussing on changes also eliminates the possible bias in their levels (due to the possible negative bias in the underlying OGs). They cover all EU member states over the period 2004-2015.

Since spring 2014 the EC has made available (in AMECO) a parallel set of four discretionary fiscal measures: general government current expenditure, current revenue, capital expenditure and capital transfers received. This data represents the effects of changes to these variables stemming from changes in policy parameters. It starts from 2010 and is published twice a year in connection with the regular forecasts.

Figure 7 shows a comparison of these data for EA11. The two variables presented by Carnot and de Castro are there, but as their data is from the spring 2014 forecasting round we also show them corrected by using our new OG real time estimates. In addition, discretionary fiscal measures, the total of the four components, from the AMECO data are shown.

Differences between these variables occur, notably regarding the AMECO data for 2010-11 and 2013. We confine ourselves here to noting that fiscal policy was procyclical almost always during 2004-2015 and for most of the euro area counties (and in the EU more generally) with one exception, the year 2009 (Carnot and de Castro, 2015a, Annex 2, shows this pattern by country). Specifically, the contractionary procyclicality in 2011-2013 is confirmed by the new discretionary measures data in AMECO, showing a tightened fiscal policy, which was then followed by the renewed negative growth in EA11.[7]

This observation is relevant for our study of the OGs as it is at least a partial explanation of their revisions extending backwards to the pre-crisis period.

---

[7] It would be useful to test and develop the Carnot and de Castro (2015b) DFE indicator by basing it on real time data. Also the AMECO discretionary measures data could be investigated. Unfortunately, it only starts from 2010, i.e. the most interesting year 2009 does not get new light. Yet, inspecting the real time data available since spring 2014 could be interesting, including the information content of the forecasts for these series produced by the EC as part of its regular forecasts. As the structural balances (and their changes) are suffering from several problems there are strong reasons to develop this type of bottom-up measures for analysing fiscal policy. Regrettably, the explanations of how the published indicators are constructed and collected are currently missing from the AMECO website.

Figure 7. Discretionary fiscal effort estimates for EA11 2004-15 with comparisons.
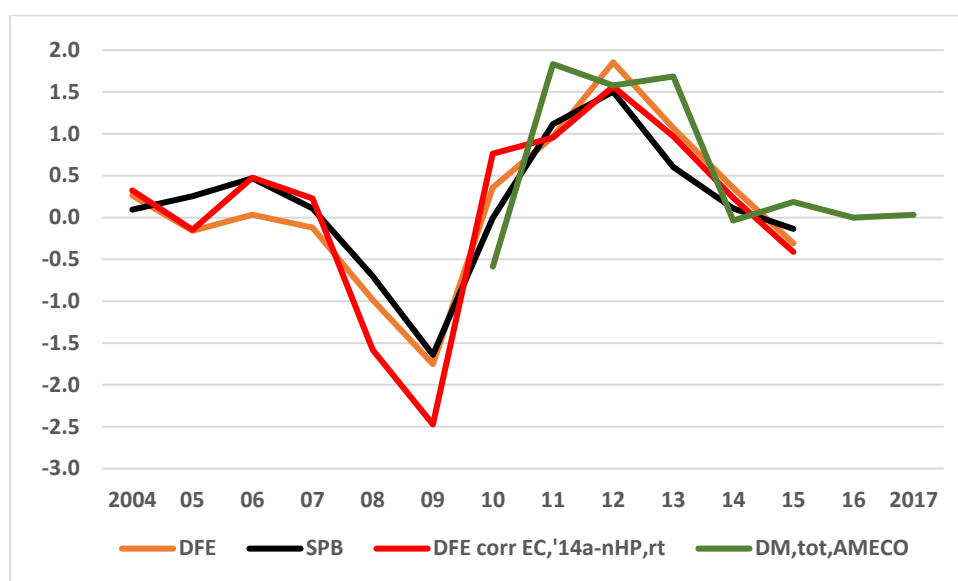
Legend: DFE estimates by Carnot and de Castro (2014), SPB = change in primary structural balance by them, EA11 calculated by the author from their data by member states; DFE corr EC,14a-nHP,rt = DFE corrected by the difference between EC autumn 2014 forecasting vintage data and the new HP OG real time estimates by the author, using semi-elasticity of 0.5; DM,tot,AMECO = Discretionary measures data from AMECO, total of the four series for EA11; the latter covers 2010-2017; its source is EC autumn 2017 forecast in AMECO .

Tight fiscal policy and the further fall in GDP in 2012-13 in the euro area

The crisis in 2009 was first followed by counter-cyclical fiscal policy via the automatic stabilisers and otherwise, but from 2011 onwards policy was tightened (Figure 7 above and Carnot and de Castro, 2015a and b). A number of leading economists writing in Baldwin et al. (2015, 10-11) observe that *'[T]he Eurozone as a whole saw its 2010 primary deficit move from about minus €350 billion in 2010 to €10 billion in 2014. This was a massive contractionary shock – equal to 4 percentage points of the monetary union's economy'*. They consider that this triggered the further recession, the fall in the euro area GDP in 2012-13. Despite negative GDP growth, aggregate government deficits in EA11 were cut over this period, while the US was growing at a rate of more than two per cent in 2010-15, helped by running twice as large fiscal deficits over 2011-14 (and in fact already in 2008-10).

The wide consensus that tight fiscal policy in the euro area in 2011-13 hampered growth was later adopted also by the European Commission (2016, 2), although it added that this was at least partly inevitable as *'many Member States engaged in fiscal consolidation to preserve their access to the markets at the height of the sovereign debt crisis'*. This latter view is still controversial. Baldwin and Giavazzi (2015, 47-48) showed that policy was strongly tightened also in counties that had access to the markets, Germany's share of tightening being more than its relative share in the euro area (see also Baldwin, 2017, 233, 244-245). In spring 2010 Germany's government deficit was forecast at 5.0% of GDP, and it was pressed to zero by 2012 and subsequently to a surplus.

More generally, Fatas and Summers (2018) contest the advisability of fiscal consolidation in the circumstances after the Great Crunch of 2008-09. According to them, consolidation was self-defeating, i.e. it did not decrease the deficits at all in most countries but caused a prolonged recession. For the clarity of assessing their result, we should note the observation that the deficit did decrease over the period 2011-2015 in the euro area (and elsewhere) does not disprove their conclusion. Instead, a plausible explanation is that other factors gradually turned supportive to growth and compensated the negative effects of the consolidations as performed.

Fatas and Summers (2018) reach their result by combining their estimates for the fiscal multipliers and the effect of fiscal contraction on permanent output (hysteresis) and estimates of long-term growth and real interest rates. They find that under a wide range of alternative values for these parameters fiscal consolidation is self-defeating (for a survey of estimating the fiscal multipliers see also Carreras et al., 2016).

They point out that this does not mean that the government should always spend more (or reduce taxes) to restore the sustainability of their debts. The conditions under which their proposition is valid are not universal, but according to them there is no reasonable doubt that they were at place in the European economies in 2008-2014.

Are more accurate estimates a solution?

Even if we consider that the retroactive revisions of the OGs do not necessarily disprove the use of the real time OG estimates for policy, we should not underestimate the problem with the uncertainty of their level and changes from one year to the next. This uncertainty is demonstrated by the differences between the estimates from our four different sources (our nHP-based and those from the three institutions), and this uncertainty feeds into the structural balances via the semi-elasticity of 0.5 for EA11 on average. Broadly expressed, the results often vary in the range of +/- half a percentage point; and here we ignore the much larger difference in the forecasting vintages of the year 2009.

Due to this we should conclude that whatever improvements one could try and achieve, the accuracy of the type of rules set and monitoring of their implementation foreseen in the *Vade Mecum on the Stability and Growth Pact* (European Commission, 2017a) is an illusion. Also Busse (23016, 30-31) considers that the errors may lead to ill-fitting policy advice and unwarranted sanctions, but then carefully judges that the SGP, thanks to its political flexibility, is faring quite well in dealing with the uncertainty and revisions of the OGs and structural balances, yet calling for continuous monitoring and further work.

His conclusion might be fairly balanced but we should still have a critical eye on the incomprehensible complexity of the fiscal rules and compacts developed partly before and especially after the Great Crunch, and ask whether they fit the real economic policy needs and whether they, due to their complexity, pose a challenge to democratic decision making. This takes us to discuss the reform of the euro.

5. Lessons for reforming euro area architecture

Assessing the OGs

The fiscal tightness in 2011-14 implies for the OG estimates that their retrospective revisions do not necessarily indicate problems with their quality by rather with the policies that depressed the economies to the extent that the retrospective OG estimates changed significantly: the preceding boom looks even more elevated than at the time as the prolonged recession is added to the picture. The strongly positive OGs for the years prior to 2008 in retrospect do not only arise from learning what the situation was at the time but the revisions upwards are also mirror images of the prolonged recession from 2009 onwards.

This inference is valid for all OG estimates, including those produced by our new HP-based method. We should humbly admit that our new estimates would not have revealed the strength of the boom much better than the others: the new OGs were slightly less than one percentage point above zero. It is prudent to say that the factors behind the strength of unsustainable boom that then turned into the Great Crunch in 2008-09 were such that no improvements to the OG estimation methods would prevent the repetition of the same ignorance in future.

Also the situation and forecasts in 2011 give an important lesson. The view that it was appropriate to move to a tight fiscal policy was at least partly based on the relatively positive forecast for GDP growth in 2012-13. This lifted up the OG estimates (including our new HP-based ones as they are based on the short-term forecast of the EC at each forecasting round), but this was probably not the most important source of the tight fiscal stance as pursued. A more plausible explanation is that the short-term forecast itself was overly optimistic, and this probably followed from an erroneously low estimate of the fiscal multiplier, i.e. the effect of the planned and proposed fiscal tightness was not correctly fed into the forecast. The consequence of all this was then the fall in the growth rate below zero for 2012-13 in the EA11.

These short comments imply that even if the OGs are produced for discovering the phase of the economic cycle and help designing policies to deal with them, we should not consider that the procyclicality of fiscal policies in the euro area resulted to any significant extent from unreliable OG estimates. The reasons behind procyclicality are probably much deeper.

Procyclicality stems from mistrust that led to short-sighted rules

The purpose of the fiscal rules originally stipulated in the Maastricht Treaty was to ensure smooth functioning of the single currency. The famous reference values for deficit and debt were set *'in view of identifying gross errors'* (TFEU 126) and the purpose was to give recommendations to correct them, but originally this would not necessarily have caused procyclicality. Instead, it was suspicion and mistrust among the prospective members in the currency union that led to the short-sighted rules set in the SGP of 1997.

The SGP became to alleviate the suspicion in Germany that the euro about to be started in 1999 was not backed by sufficient fiscal discipline in all prospective member states. This suspicion goes back to the political nature of the process towards a European monetary union. The roots go further back, but autumn 1989 was decisive. President Francois Mitterrand connected German reunification to Germany becoming a member of the currency union as outlined in the

Delors-report in April 1989. Chancellor Helmut Kohl committed to this in Strasbourg European Council in December 1989. He told later that the hard negotiations were one of his darkest moments (Der Spiegel, 1998), and that joining the euro would not have won support in a referendum (The Telegraph 2013).

The German constitutional court ruled in October 1993 that ratification of the Maastricht Treaty in Germany was acceptable provided that the currency union respects stringent fiscal and monetary policies (Tuori and Tuori, 2014, 200).

This is, in short, the historical background for the strict time schedule for correcting excessive deficits set in the SGP in 1997. It then led to the procyclicality that still prevails, and further developments have not eliminated it.

Short-sighted procyclicality was most strongly demonstrated in the so-called SGP crisis in 2003. This crisis did not happen due to joint violation of the rules by Germany and France as is commonly argued. True, France did not want to follow the rules, but Germany's case was different. Germany made the argument that it had implemented the previously given recommendations for correcting its excessive deficit, but new adverse factors impeded attaining the deficit target. It wanted the previous recommendations to be revised and be given more time to consolidate. However, the European Commission and some member states maintained that postponing the procedure was illegal. This interpretation turned out to be erroneous as the EU Court of Justice (ECJ) later ruled that the European Commission could indeed have tabled a proposal for revising the previously given recommendations (paragraph 92 in Court of Justice of the EU, 2004). This careful reading of the ECJ ruling has been presented in the literature (Beetsma and Oksanen, 2008, 565; Heipertz and Verdun, 2010, 2, 128-162), but the other narrative that Germany and France jointly violated the rules has been politically more suitable to be kept alive.

However, a positive effect of this episode was that the SGP was revised in 2005, notably making it explicit that the stages under the SGP can be repeated, the recommendations revised and deadlines for correcting excessive deficits extended, if unexpected adverse economic events occur. When the Great Crunch erupted, the revision of the rules probably helped to facilitate significant increases in deficits. This flexibility then remained the only significant exception to the observation that fiscal policy has been procyclical (Carnot and de Castro, 2015a, Annex 2). In this regard all the new rules and fiscal compacts since the crisis have not led to fundamental improvements. Instead, the diversity of circumstances and differences among member states has led to an overkill complexity of the fiscal rules, so that also the European Commission (2017b, 12) EMU roadmap includes their substantial simplification as one of the important tasks by 2025 (!).

Apart from complexity of the rules the substantive content of policies in the recent past guide the views for the future. As noted above, several experts (Baldwin and Giavazzi, 2015, Baldwin, 2017, Bayoumi, 2017) question whether fiscal tightness during the sovereign debt crisis 2011-13 was reasonable in terms of benefits and costs, and Fatas and Summers (2018) go further by concluding that it was not effective at all: fiscal consolidation was self-defeating, and it did not help to reduce the debt burden; so, the pursued benefits were not attained. This contrasts strongly with the official view that consolidation was both necessary and useful for preserving the euro.

Over the coming decades economists and economic historians which will certainly work on the most serious economic crisis since the 1930s in general and specifically on the role of the euro to complement the picture. But we do not have the luxury to only wait for more elaborate judgements as the reform of the euro is the topic of today and the responses will determine the future.

Proposals to improve financial infrastructure

Almost all proposed euro area reforms are linked to fiscal policy and institutions and have close links to measuring GDP fluctuations, hence the OGs. Even the reforms that are farther away are relevant here as a failure in one policy area spills over to fiscal policy, for example via costs of rescuing banks.

A group of 14 French and German economists (Bénassy-Quéré et al., 2018) make several proposals for improving financial and fiscal stability. One of them includes provisions for economic, legal and institutional underpinnings for orderly sovereign-debt restructuring. The purpose is to make the no-bail out rule credible again, after the leeway that had to be improvised as a consequence of the Great Crunch. According to the authors such rules should be phased in gradually and announced when the member states have normal access to the markets (Bénassy-Quéré et al., 2018, 12-13).

They also make a proposal to create a synthetic euro area safe asset backed by a standardised diversified portfolio of sovereign bonds. Carefully constructed this could contribute to financial stability.

De Grauwe and Ji (2018) are sceptical and warn that 'financial engineering' proposed by the group may create a false sense of stability despite being far from sufficient.

Bénassy-Quéré et al. (2018) proposal for an expenditure rule

The group of 14 (Bénassy-Quéré et al., 2018, 9-11) includes also a proposal to eliminate procyclicality of the fiscal rules by replacing the focus on the (structural) deficit by a simple expenditure rule guided by a long-term debt reduction target. The new rule would stipulate that government expenditure must not grow faster than long-term nominal output and below that rate in countries where debt-to-GDP ratios must be reduced. According to the authors, this rule would help stabilise economic cycles, since cyclical changes in revenue do not need to be offset by changes in expenditure.

Although this proposal at first sight seems well-founded, it may encounter difficulties. *Goodhart's law* says that when a measure becomes a target, it ceases to be a good measure (Strathern, 1997). This law originally appeared in the context of responses of economic agents to policy targets, but it obviously also applies to the actions of the member states under the EU rules.

As public expenditure schemes are large and complex the opportunities to manipulate the indicators to be monitored are many. Public transfers to households can be taxable or tax exempt; in many instances tax rebates given to enterprises may replace direct subsidies. The relevant parameters in the various tax and transfer schemes may change over time for many reasons, affecting what is recorded as gross public expenditures. Even intentional manipulation

cannot be ruled out if a target is binding. If the target is set by an EU rule, eliminating manipulation and correcting for changes in the policy parameters would require detailed rules and monitoring. So, a door again opens for complex rules.

The current rules regarding the fully funded second pillar pension systems are an example of Goodhart's law under the current accounting standards. The second pillars are classified as part of the private sector in the national accounts even if they are mandatory, and their surplus and assets are not included in the government accounts. Also, if a second pillar is created by shifting pension contributions to the newly created pension funds but the government remains liable for the pension rights accumulated so far, the effect on government finances is negative for a transition period that lasts for decades. Despite renewed attempts, no proper solution to this issue has been agreed, despite the fact that an appropriate consolidation of the second pillar accounts to government finances for the purposes of the EDP would remove this obvious flaw (Beetsma and Oksanen 2008, 568-9). Under these circumstances, reversals of pension reforms in some countries have taken place as this has improved the government accounts, especially when the deficit targets became hard to attain (Bielawska, Chłoń-Domińczak and Stańko 2015, 9, 85-91). This effect of Goodhart's law certainly was not the purpose of the SGP rules.

A further example is that public expenditures can be shifted to the private sector accounts by various type of public-private partnerships. This has required special attention under the rules on deficits and debt.

These same problems would be repeated under the expenditure rule.

Asymmetric shocks and financial flows

From the beginning of preparing for the euro one of the key issues was the challenge of asymmetric shocks in the monetary union where the exchange rate is not any more a tool of the national authorities. The competitive views were revolving around the size of the common budget and, in case it were not foreseen, possible specific arrangements for dampening the effects of such shocks. The US was always used as the reference with an eye of different mechanisms dampening the asymmetric developments across the US states (Oksanen, 2016a). Recent papers by Alcidi, D'Imperio and Thirion (2017) and Alcidi and Thirion (2017) give a carefully gathered summary of the studies, benefitting from the recent data covering the changes in the dynamics after 2008.

One key finding is that in the US smoothing is larger than in the euro area mainly due to capital market integration, meaning that cross-ownership of capital distributes the effects of asymmetric shocks on capital income across the federation. The other is that in the euro area smoothing effect via the government budgets is normally larger than inter-state fiscal transfers in the US federal budget, and the state budgets do not contribute to it due to the common balanced budget requirement).

Thus, the importance of flexibility of government budgets in the euro area under both common and asymmetric shocks come in again here. However, what was said above holds in normal circumstance. In the periphery of the euro area all smoothing practically ceased since 2010 (Alcidi and Thirion, 2017, 15) due to fiscal austerity, whether it was necessary and optimal or excessive.

The above picture is otherwise quite comprehensive, but one important element should be added. The ECB has an important role in distributing liquidity in the euro area (partly also beyond it) though its Target2 mechanism ('Target' stands for 'Trans-European Automated Real-time Gross Settlement Express Transfer system'; Target2 balances are the claims and liabilities of euro area national central banks vis-à-vis the ECB that result from cross-border payments settled in central bank money (European Central Bank, 2017).

When private financing from the core of the euro area was withdrawn from the troubled periphery, Target2 gained utmost importance. As private financing flew to the core, the banks placed this liquidity in the ECB which then channelled it via Target2 to in the southern members of the central banking system allowing them to provide liquidity and financing to their financial institutions and via them to their governments (Minenna, 2017). At the end of 2012, as a result of the first peak of Target2 balances, the positive balance of Germany was 24% of its GDP, relatively even more for Finland (35%) and highest for Luxembourg as a banking centre (240%). The negative balance for Spain was -32% of its GDP and relatively more for the smaller troubled members (Ireland, Portugal and Greece). After some decline, the numbers are again in the beginning of 2018 in the same order of magnitude (except for Ireland that has done away with its negative balance; data is available at Statistical Data Warehouse/ecb.europa.eu).

These stocks have been large, and they have fluctuated so that, noting the largest provider, Germany increased its balance by 5-7% of its GDP in 2010-12, then reduced it by 5% in 2013, and has increased it again since 2014. The negative balance of Spain reflects these movements.

The main point here is that the Target2 flows are far greater than financing via the special arrangements and the newly created special institution, European Stability Mechanism (ESM). Complemented by various other new measures, the ECB has performed its responsibilities in providing liquidity, the total amount for the euro area as a whole, and distributing it though Target2 to places where it is needed. Without the ECB undertaking these tasks during the crisis the euro would have already collapsed.

For the analytical work on smoothing the shocks the lesson is that it would be useful to add these flows as an important channel. They are stemming from a central euro area institution (the ECB) and have obviously smoothened the effects of the shocks. Their existence is as important as any other items in the balance of payments and other statistics of the euro member states, and the stocks and flows are so large that they should not be ignored. Putting it the other way round: if this function of the ECB to channel financing back to the south had not happened during the crisis, the results from the analysis of the various mechanism for smoothing would have looked quite different. Smoothing would not only have ceased in the periphery but it would have turned negative, i.e. those channels would have aggravated the asymmetric shocks.

A proposal for smoothing asymmetric shocks and the OGs

Proposals for smoothing asymmetric shocks among members of the monetary union remained in the background for the first decade of the euro, but came back to the agenda especially as the Great Crunch and its consequences revealed significant asymmetries.

Creating a stabilisation function under the next multiannual financial framework to deal with large asymmetric shocks is listed in the European Commission (2017b, 8, 12) roadmap of December 2017. Also the group of 14 French and German economists (Bénassy-Quéré et al., 2018, 14-16) propose a fund providing payouts mainly based on unemployment indicators, the system being financed by the member states according to their economic size. Andor et al. (2018), another recent group of 14 economists, propose an unemployment insurance scheme for the euro area to complement the national systems.

These proposals encounter several problems: tackling only large asymmetric shocks limits their effectiveness; unemployment is a lagged indicator of exogenous shocks; conditionality to compliance with other schemes and rules (which can be controversial) easily leads to disputes; even though it is declared that they should not lead to permanent redistribution, this is far from being assured; the contributions paid to the scheme should not necessarily be proportional to GDP but they could also be geared to dampening the asymmetric booms. All in all, creating a system that would not be effective for smoothing but aggravate mistrust and disputes is most probably not a reform that the euro area needs.

Oksanen (2016a) has proposed an alternative scheme where payments in and paybacks depend on the relative OGs of the member states, being, for example, proportional to half of the percentage deviation of each member state's OG from the euro area average. Importantly, his proposal contains a rule that after an agreed period, say seven years, the net balances are measured and netted out in constant instalments over the subsequent seven-year period. This would make sure that permanent redistribution would not emerge (but all desired redistribution would be arranged under the funds specifically designed for that purpose).

Oksanen (2016a) considers that, for the proposed mechanism, improving the real time OG estimates would be advisable, though not be indispensable for starting it. The details could also be improved afterwards as no member state would be winning or losing from such changes, thanks to the rule of netting out the balances over time in any event.

Figure 8 presents how our nHP real time OG estimates would have worked for the proposed mechanism. It shows the differences between the OGs for each of the four largest euro area countries to the EA11 average. The estimates for the current year are presented, while similar graphs the forecast for the next year could be added (and those for the previous year if deemed interesting). The parallel data based on the OGs by the EC are shown for comparison.

This first glance at the use of the nHP OG estimates is encouraging. It shows that for the two opposite countries, Germany and Spain, the new OG estimates would have worked better than the EC estimates before the crisis. The relative figure for Spain is close to two pp. in 2005-07, while in the EC estimates it was positive (below one) only in autumn 2005 forecast, and otherwise negative. Not identifying the unsustainability of the boom in Spain was one fault that had to be recognised in the hard way.

This is not a place to go much further in illustrating and testing the proposed mechanism. It suffices to note that it has the advantage that it only requires a light administration and avoids disputes that aggravate mistrust. Perhaps most importantly, it would serve the need for smoothing asymmetric shocks effectively enough and remove this issue from the reform agenda and leave space for more demanding tasks.

Figure 8. Output gap estimates, real time, differences to EA11, four largest EA11 member states 2002-2017; new HP and European Commission estimates.
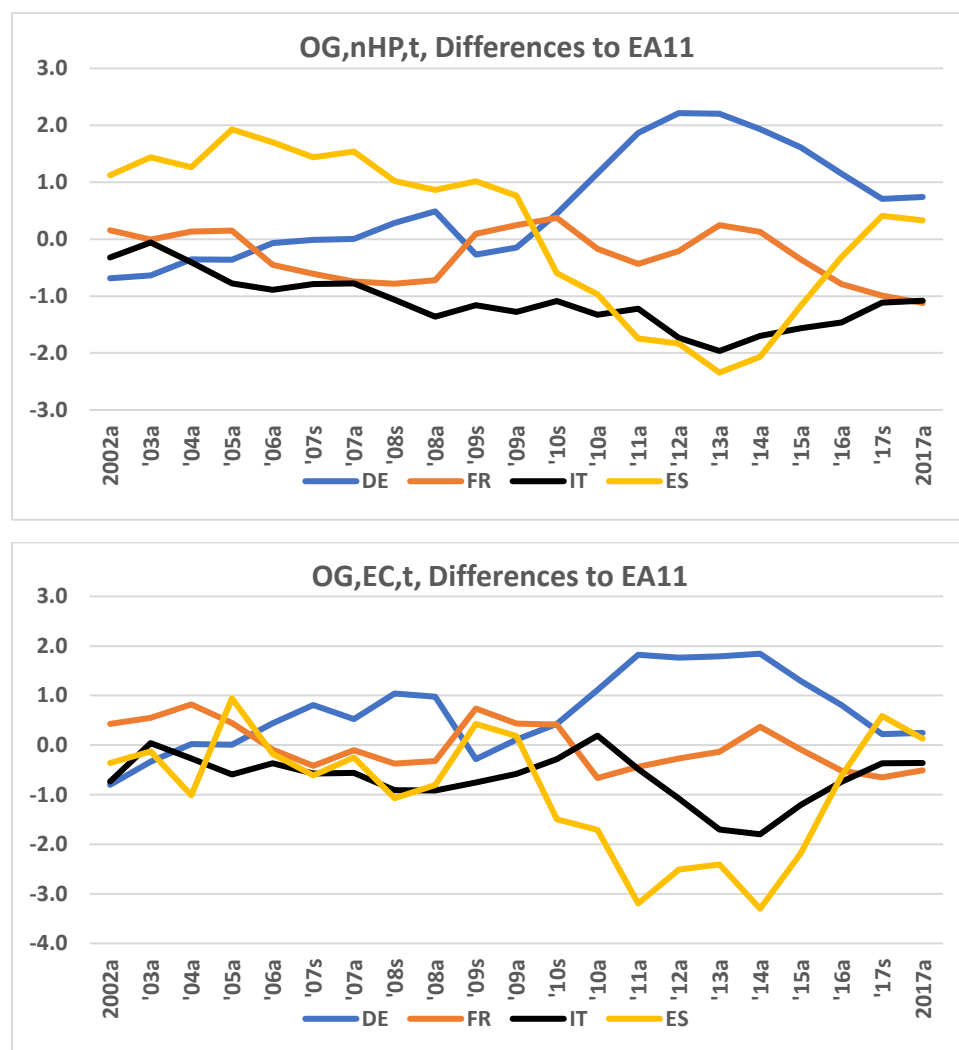
Legend: OG,nHP,t = new HP OG estimates in real time estimates for the current year in each forecasting vintage; OG,EC,t = the corresponding European Commission real time estimates; DE = Germany, FR = France, IT = Italy, ES = Spain.

A large budget for the euro?

A limited and practical smoothing mechanism could also take heat from the view that without a large common budget a monetary union would not survive. This view goes in the EU back to the MacDougall Report in 1977, where it was argued that it is needed both for permanent redistribution and smoothing short-term asymmetries (Oksanen, 2016a, 332, 336-337). This view did not prevail in Maastricht in December 1991 when establishing the monetary union was written into the EU Treaty. Like the proposals for specific smoothing mechanisms, proposals for a significantly larger budget for the euro area have come up again, triggered by the crisis, for example by Andor et al. (2018).

The main institutional difference between the euro area (and EU) and the US is that in the euro area national governments run 98% of the public budgets (2 % at the EU level), while in the

US the share of the federal budget of general government is 60% (relative to the GDP the EU-US difference is even more pronounced). We have above referred to the studies comparing the euro to the US in several other ways, showing that there are various parallel mechanisms working in each, and specific limited mechanisms could be considered. For these reasons there are no strong economic arguments for a significant increase of the EU/euro area budget.

In addition, it is a commonly held view is that multiplying the size of the EU/euro area budget is currently politically completely unrealistic so that it would probably take a few generations of Europeans before it could become a serious option. This view is based on the solid argument that a significant increase in the budget requires unanimity. Under these conditions a view that a large budget is indispensable means de facto, perhaps ironically to its proponents, that the monetary union should be dismantled.

Short-term flexibility in government budgets requires long-term sustainability

An overarching theme in the present paper is the need to eliminate procyclical policies and maintain the capacity to react in the event of a serious crisis, be they common or asymmetric shocks in the monetary union. Economically optimal speed of adjustment derived from empirically tested analysis should replace the rigid and short-sighted budget discipline rules.

This principle sounds like common wisdom to solve the issues with procyclicality. However, for various political and economic reasons flexibility in fiscal policy in the short term cannot be allowed without assuring sustainability of public finances in the long term. This holds for any economy in isolation and all member states jointly in the monetary union. Those who suspect that some of the members are not willing and/or capable to pursue sound public finances have reasons for their mistrust. Unfortunately, the Greek case proved that they had a point.

The fundamental question is what should be done to improve credibility that all members conduct fiscal policies that are sound and sustainable in the long term. We referred above to the proposal by Bénassy-Quéré et al. (2018, 12-13) for creating the provisions for orderly sovereign-debt restructuring in order to force the governments and their creditors to set targets that are sound in the long term. This may help, but it is safe to accept that most probably no institutional reform can do this alone, and they may even not be the main resolution to lasting credibility and confidence.

It should be understood and accepted that the policies required for long-term sustainability are mostly in the competence of the member states. They are measures that deal with the pressures of population ageing on public expenditures, including reforms of the public pension systems and health-care financing. They are needed in all countries, but policies and politics tend to be too short-sighted to produce effective results. The dilemma is that the changes to population ageing-related expenditures can only be gradual, stretched over several electoral cycles. This is economically necessary for treating all generations fairly, and indispensable politically as too swift reform proposals do not pass though the political process. The detriment of this is that there is no single institution that can make the required firm commitments for the long term.

Yet, even if there are no straightforward responses, assuring sustainability of public finances has been on the agenda in the European welfare states for, say 30 years, after starting to gain

attention in the late 1980s. Keeping up the pressure for further reforms is probably necessary not only for the sake of sustainability of public finances but also for maintaining the high level of social protection in the EU as a further crisis could easily trigger disorderly adjustments to it.

The need for reforms that respond to the needs in each member state and can be orderly legislated and implemented by them (it was necessary that the outsiders forced Greece to make the inevitable pension reforms as the previous Greek governments had overlooked them, but what has been going on should not be repeated). But as sustainability of public finances is a common concern in the euro, useful work on ageing-related public expenditures has been undertaken by the member states and the EC since establishing the euro (reported at regular intervals by the Ageing Working Group of the Economic Policy Committee).

Unfortunately, for various understandable reasons, the analysis is often too narrow. An example is that the medium-term objects (MTOs) for budget balances are derived from the long-term expenditure projections in a way that does not give sufficient attention to the factors that determine those projections. Under the current approach in the *Vade Mecum* (European Commission, 2017a) the MTOs are built on the expenditure projections based on current policies, and then made less demanding by ad hoc adjustments; this tends to diminish the pressure on containing expenditure increases. In many cases, perhaps quite generally, it would be more reasonable to move in several steps: as the requirements for medium-term adjustment of government finances seem hard, the determinants of expenditure in the long term should be reviewed and proposals to revise them should be made, and the medium-term object should be derived only from a revised long-term expenditure projection (Oksanen, 2016b, 383-387).

The main reasons why the EU and its member states have not been able to establish adequate and accurate assessments and procedures for containing the ageing-related expenditures and failed to derive well-working MTOs from them is that these expenditures are under national competence and the institutions and political circumstances vary greatly. Therefore, it is difficult to establish approaches and procedures that could fit them all.

Yet, it is important to be aware of the potential biases and narrowness of the various indicators and the guidelines based on them.

Finally, as an example of challenging the quasi conventional wisdom: there can be firm grounds for criticising the common view that increasing the old age retirement age so that that it would keep up with the increase in life expectance (by maintaining the current share of participating in working life of expected adult life time). This may not be sufficient, for example because health care expenditures may still increase more than proportionally. Yet, attempts to increase the old age retirement age more significantly hardly appear anywhere.

The contribution of the OG estimates, old and new, to improve the euro

The preceding discussion covered several issues that go beyond the narrowly defined task of estimating the OGs. The OGs are nevertheless an important input for designing policies, primarily but not only fiscal policy. They are of pivotal importance in the rule book of the euro area and underlie important procedures in the joint decision making in the euro area (and the EU).

The problem with the OGs used in all the three institutions, the EC, IMF and OECD, is that the estimation methods are complex and the results have been challenged on various grounds. Some criticism is well-founded, like pointing out that the real time OGs have been negative almost without exception. So, there is at least a terminological problem as in ordinary language a cycle refers to fluctuations on both sides of something that is understood as normal. A prudent conclusion is that parallel estimates are useful as they measure slightly different aspects of the economies. An OG estimate that is consistently negative may indicate in a useful way that the economy produces persistently less than its potential.

In the present paper we showed that the HP-based estimates have been criticised, at least partially, on dubious grounds. Careful use of the method can distinguish the cycle from changes in the trend in a way that compares well with the results from other methods. A clear advantage of the HP method is that it is simple, based only on one single time series, the GDP. This makes it clear that the results do not depend on complex views of the dynamics of the economy that can always be contested.

The contribution here is to present a new application of the HP method, where we insert an explicit link to the view about underlying growth of the economy in future. This does not complicate the method unduly, but it is useful in two ways. First, it makes explicit the question of what might be happening to long-term growth prospects, allowing us to make alternative assumptions on it; this question is encountered continually, especially when a major shock hits the economy. Second, it is obvious under all OG estimation methods that the phase of the cycle always depends not only on the current situation (supplemented by the short-term forecast) but also on the changes further on in the future. Our new method makes this explicit. This helps to understand that the results are uncertain, and that they should be treated with care and caution.

As a final illustration of what can be done with the new method: imagine that currently, based on the latest forecast in autumn 2017, it is perceived that the economic growth will soon go persistently to zero. This could be argued on several grounds: even if technological progress may continue, more resources are required for repairing the environment and containing climate change and a smaller share is available for private and public consumption, so that the GDP measures the material resources available for consumption and well-being less accurately than hitherto; for a long time the income level of large segments of the population has not grown in the US, the leading major economy, as the richest are taking an increasing share; and globalisation in general shifts production outside Europe. Whatever one might think of these arguments, an experiment with the new HP method would be interesting: assume zero for the underlying growth from 2018 onwards. The result is shown in Figure 9 together with the base line result assuming 1.3% growth and the OG estimate of the EC in autumn 2017.

The interesting result is that for 2018 the nHP OG estimate is +3.3%, almost exactly equal to the retroactive estimate for 2007, the peak of the overheated boom (this estimate does not change much with the reduced growth assumption from 2018 onwards).

The high OG for 2018 does not mean that fiscal policy should be immediately tightened – we remember the warning by Fatas and Summers (2018) that fiscal consolidation can be self-defeating. However, the new dramatic result could serve as a wake-up call for thinking seriously about long-term challenges. More modestly, it is an illustration of a possible use of the new HP method.

Figure 9. A new OG estimate for EA11 assuming zero growth from 2018 onwards compared.
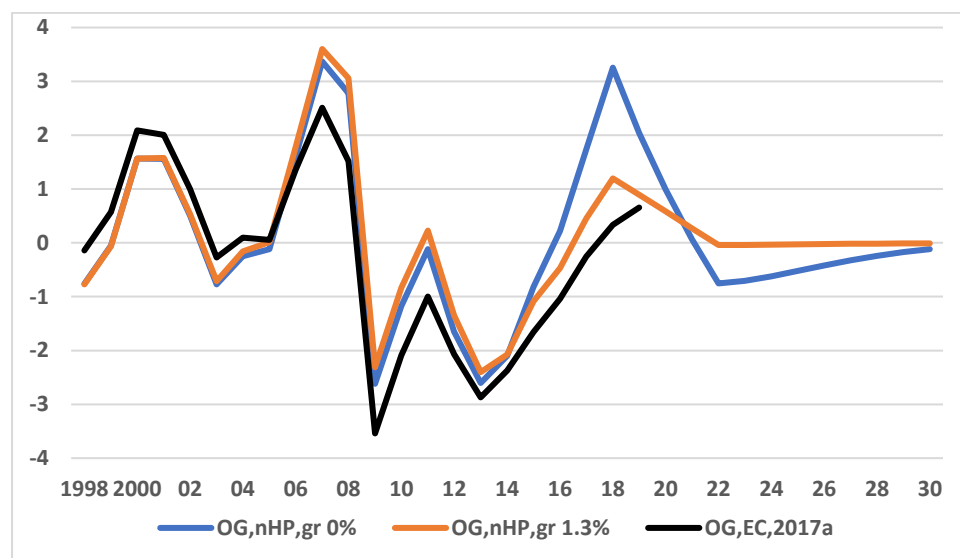
Legend: nHP OG estimates in real time based on EC autumn 2017 forecast; gr 0% = zero underlying GDP growth assumed from 2018 onwards; gr 1.3% = the base line growth assumption; OG,EC,2007a = EC OG estimate in autumn 2017.

6. Summary and Conclusions

We have presented here new output gap (OG) estimates based on a simple statistical method, the Hodrick-Prescott (HP) filter, to distinguish the cycle from a possible change in the trend of the GDP.

The HP method has been criticised, but partly on inaccurate grounds. Careful use of the method can give results that compare well with the results from other methods. A clear advantage of the HP method is that it is simple, based only on one single time series, so that the results do not depend on complex and contestable views and theories of the dynamics of the economy.

The novel feature of our new HP application is that we insert a link to the perception of underlying long-term growth prospects of the economy. The application allows the user to make alternative assumptions on it. It also makes explicit that the phase of the cycle always depends not only on the current situation and short-term forecast but crucially also on the changes in the future. This helps to give a healthy warning that all estimates of the phase of the cycle are uncertain and they are to be used with prudence.

Comparing the new OG estimates to those regularly produced and published by the three main institutions, the European Commission (EC), IMF and OECD, shows that they are not worse in any respect, but they have the advantage of being more easily understood also by non-experts.

As the retrospective revisions in the OG estimates have been widely used as a criterion of their reliability, we emphasised that the revisions do not necessarily indicate weaknesses in their construction. The recent history after the Great Crunch of 2008-09 is a case in point. The revisions of the OG estimates, extending backwards were the consequence of the tight fiscal

policies that caused the fall in the GDP in 2012-13. This was the main factor behind the revisions of the OG estimates produced by all three institutions and also of our new estimates.

With the exception of year 2009 when the significant fall in GDP struck, there is overwhelming evidence that fiscal policy has been persistently procyclical in the euro area. As the OGs are produced for guiding towards counter-cyclical policies, it is necessary to analyse the factors that are behind short-sightedness and procyclicality. The most plausible narrative is that mistrust prevailed and still prevails between the member states: they do not trust that all members are willing and capable of targeting fiscal policy at sustainability of public finances so that they would credibly meet their debt obligations in all circumstances. This led to short-sighted fiscal rules as there is no way to absolutely convincingly commit governments to sound policies in the long term.

The purpose of estimating the OGs and giving them a pivotal role in the fiscal rules for the euro area should have helped to pursue counter-cyclical policy, or at least to avoid procyclicality. In this regard the results have not been successful. Instead, the mistrust first led to detailed rules on budget discipline which did not work in the unexpected circumstances. This then led to even more complicated rules especially after the Great Crunch. Complexity became so evident that there is now a wide agreement that simplification is needed, but how to do this is a major open question.

A successful reform must contain flexibility for counter-cyclical policy, which means that the national budgets are working as stabilisers for both common and asymmetric shocks.

In addition, in the debate on the reform of the euro a stabilisation fund for smoothing large asymmetric shocks has been proposed. The proposed unemployment-based schemes seem both ineffective and administratively cumbersome, while an effective quasi-automatic transfer mechanism could be more efficient and help smoothing both large and small asymmetric shocks. It could be based on the deviation of the real time OGs of the member states from the euro area average and contain clear provisions to assure that no permanent transfers would take place. A first assessment indicates that our new OG estimates might be more useful for such a mechanism than those produced by the EC. This said, the details of the new mechanism could be improved even after having started it as the principle that no member state would permanently gain or lose applies also when changes in the details of the mechanism were possibly made.

Flexibility in fiscal policy in the short term cannot in practice be allowed without assuring the sustainability of public finances. This should be in the forefront of designing reforms for the euro area.

One useful proposal aims at making the no-bail out rule credible again by creating the necessary rules for orderly sovereign-debt restructuring in case of insolvency. This would eliminate bail-outs and thereby strengthen the responsibility of both governments and their creditors to behave prudently, assuring long-term sustainability of public debts.

However, no institutional reform alone is sufficient for assuring long-term sustainability of public finances. Policies to contain the increase of population ageing-related public expenditures are required in all countries, and this is under the competence of the member states.

As sustainability of public finances is nevertheless a common concern in the euro, joint work on ageing-related public expenditures is undertaken by the member states and the EC to develop analysis and provide useful experience to each other in designing the reforms. In some respects, the analysis can nevertheless, for understandable reasons, be too narrow and exclude sufficiently effective reform options. There is a risk that the proposals for increasing the old age retirement age have been too complacent and in practice hindered serious discussion on significantly larger increases.

In reforming the euro it is equally important to eliminate the tendency of procyclical fiscal policy and to assure long-term sustainability. Simple and easily understandable estimates for distinguishing the phase of the economic cycle from possible changes in the trend of the GDP can be conducive elements for a comprehensive analysis covering both the near future and more distant trends.

Heikki Oksanen

Technical appendix

to paper 'New Output Gap Estimates for Assessing Fiscal Policy with Lessons for Euro Area Reform'

This appendix provides additional data on EA11, the US and the four largest euro area countries. Most of it comes from public sources on the internet made available by the three institutions. Most of the EA11 data is downloaded from the website of the FIRSTRUN project, http://www.firstrun.eu/.

The sections in this appendix refer to those in the main paper.

Section 1 Introduction

Dubious negative assessment of the HP method

Figure TA.1. HP trend estimates for GDP by the European Commission (EC) for Germany, the Netherlands and EA11 in autumn 2008, spring 2009 and autumn 2009 forecasts.

Legend: 2008a, 2009s and 2009a are autumn 2008, spring 2009 and autumn 2009 forecasts, respectively. The first two until 2013 and the last one until 2014. Note: the scales differ.
Source: European Commission AMECO and CIRCABC data bases.
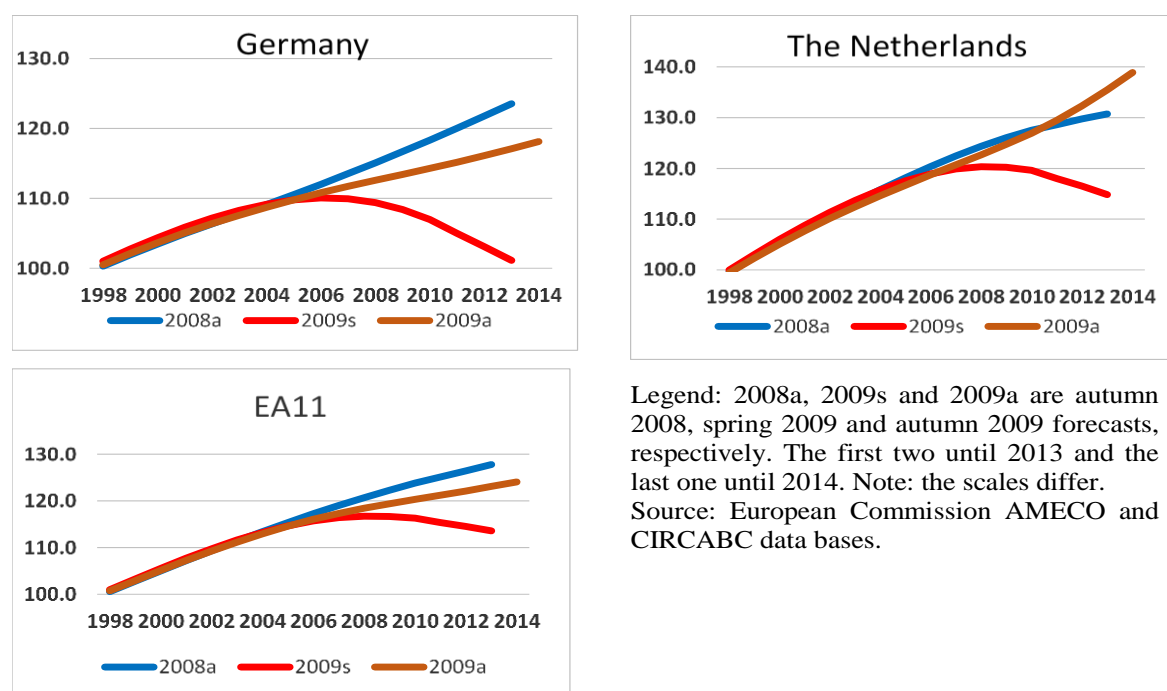
McMorrow et al. (2015, 12) present a graph where the OG for the euro area in the spring 2009 forecast of the EC is zero for both in 2009 and 2010 (while all other estimates gave a significant negative value). Six months later, in autumn 2009 forecasts the HP estimates of the EC for the OG were significantly revised, including a negative value for the OG from the HP filter. This swing in the estimates then led the authors to denounce the HP method in general.

Figure TA.1 shows a swing in the HP trend GDP estimates of the EC for Germany and the Netherlands published in autumn 2008, and in spring and autumn 2009, and the calculations by the author for EA11 (from the country-wise EC estimates). In spring 2009 their HP trend projections were revised drastically downwards from autumn 2008 to a persistent and continuous decline in 2009-2013, especially dramatically for Germany: its trend level would have declined almost to the 1998 level by 2013. An absolute decline was projected for several countries, including the Netherlands, so that the similar pattern appears for EA11.

This projection was then turned around half a year later in the autumn 2009 forecast. The negative trends had disappeared (except for Ireland and Greece, the latter being not included in our EA11), and the projection for the EA11 GDP trend level for in 2013 had become eight per cent higher than estimated in the spring 2009 forecast. This swing in the estimates is not noted by McMorrow et al. (2015).

Section 2. New OG estimates generated by the HP method

Graphical illustration of the method

We refer here to the change from 2008 to 2009 using stylised numbers which resemble the data and perceptions at the time for EA11.

In Figure TA.2 GDP has been growing 1.5% pa. until 2008 and is expected to continue. In this steady state the trend (or potential) of the GDP is identical with the data, and thereby the OG is equal to zero throughout. Then, suddenly, in 2009 GDP falls by 5% and is expected to grow by 1% in 2010 (in the official forecasts almost without exceptions GDP is forecast to grow in year t+1, and, for that matter, also in t+2 when made available).

The GDP series is extended beyond t+1 based on an assumption of a constant underlying rate of GDP growth, and assuming that in years from t+2 to t+5 it additionally grows by one quarter of the OG in year t+1. This latter variable is estimated by our application. From t+6 onwards the GDP is assumed to grow by its assumed underlying rate.

A freely available HP filter device available at Federal Reserve Bank St. Louis Economic Research website https://dge.repec.org/cgi-bin/hpfilter.cgi is used.

The HP filter is applied in logs and with smoothing parameter 100 on the extended GDP series explained above (as in most of the estimates presented in this paper). In this stylised example the data starts from the year 1980 and extends to 2040.

Figure TA.2. Stylised illustration of the new HP (nHP) method
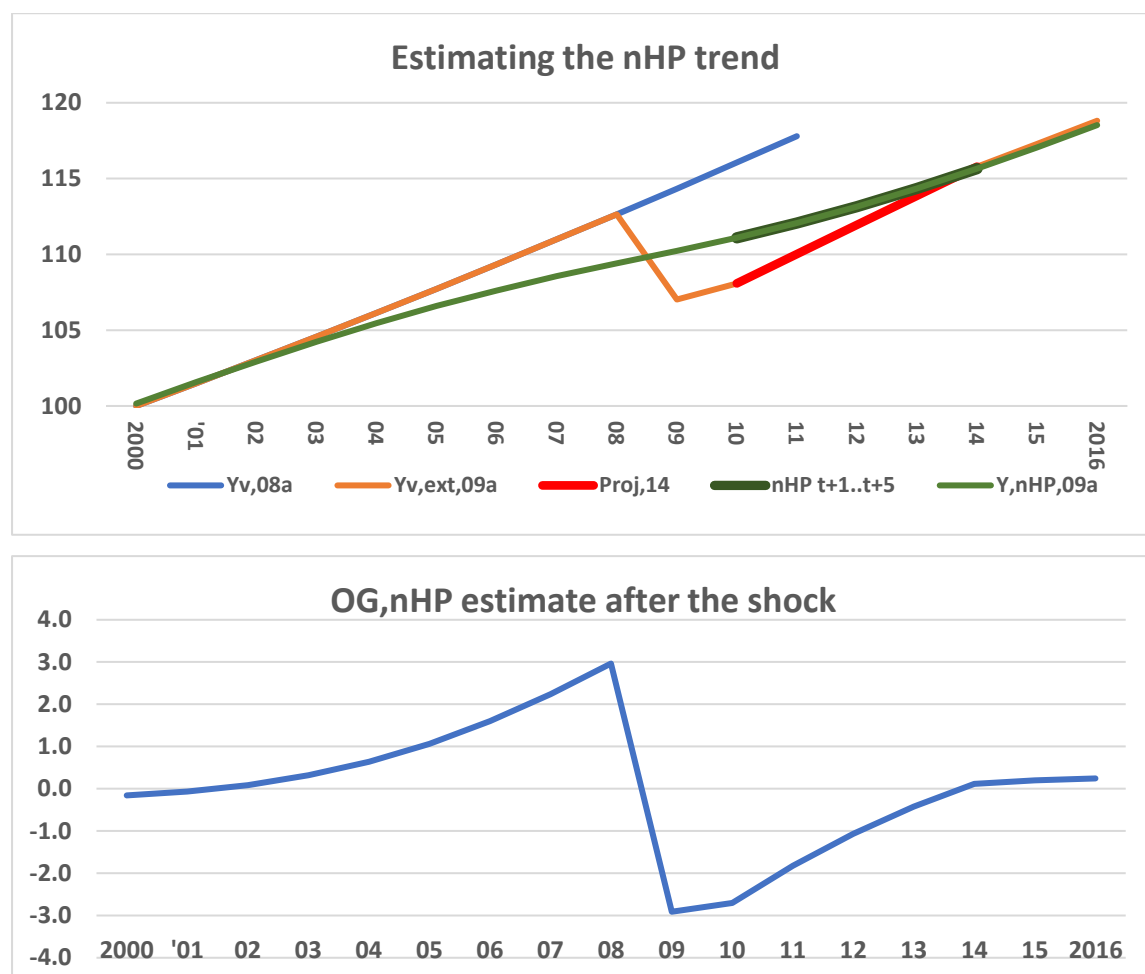
Legend: Yv,08a = GDP and its forecast in 'autumn 2008'; Yv,ext,09a = GDP and its forecast after the shock ('autumn 2009'); Proj,14 = GDP projection until 2014; nHP t+1..t+5 = the resulting nHP estimate for those years; Y,nHP,09a = the nHP estimate for the entire data series. nHP OG = estimate for the output gap after the shock.

As the OG for year t+1 enters in the formula for the extended series and it affects the estimated nHP trend, an iterative procedure is needed: a value for OG in t+1 is set, the nHP trend is calculated and the result for OG in t+1 noted. This is repeated a few times to arrive at a result where the value set for OG,t+1 and the result from the nHP estimate coincide.

The result then gives the time series for the nHP trend and OG. One interesting observation from this stylised example is how the OGs for the previous years change when the shock hits the economy. This has implications for using the OG for estimating the cyclical and structural balances in the government budgets. Not only the OGs in the past change but also changes in the current and next year take place regardless of possible changes in fiscal policy parameters. Also, it should be noted that in the newly estimated series the annual change of the OG from 2008 to 2009 is 5.9%, while the GDP is changing by 5%.

The basically same features apply also to the OG estimates of the institution based on more complex production function methods.

Illustration with long series

Figures TA.3 illustrate the HP method. In the first graph we have the GDP data for EA11 1960-2017, its log-linear trend and the nHP trend calculated with smoothing parameter 100 (used for the base line results in this paper). The log-linear trend obviously does not give a useful estimate for measuring the cyclical movements. The reason is that it does not take into account the decline in the rate of growth which took place over this period (from 5% to 2% by 2008).

The HP method is developed exactly to take into account such changes in the long-term trend and to provide the deviations of the observed (or forecasted) data series from such an adjusted trend.

The second graph covers the years 1998-2017 showing the EA11 observed GDP and the HP estimates using alternatively the smoothing parameter 100 or 6.25, the latter being preferred by the Deutsche Bundesbank (2014). It shows that with the value 6.25 the HP result follows the movements in the original series in such a way that part of the cyclical movements remains in the estimated adjusted HP trend.

Complementing the second graph, the third one shows that for 1960-2017 the amplitude of the OGs based on smoothing parameter 6.25 is about half of what it is with value 100. Also the OG by the EC from the autumn 2017 forecast (for 1998-2017) is shown, indicating that its amplitude is roughly the same as for the nHP estimate.

Figure TA.3. Performance of the HP method with different parameters for smoothing applied to EA11 GDP.
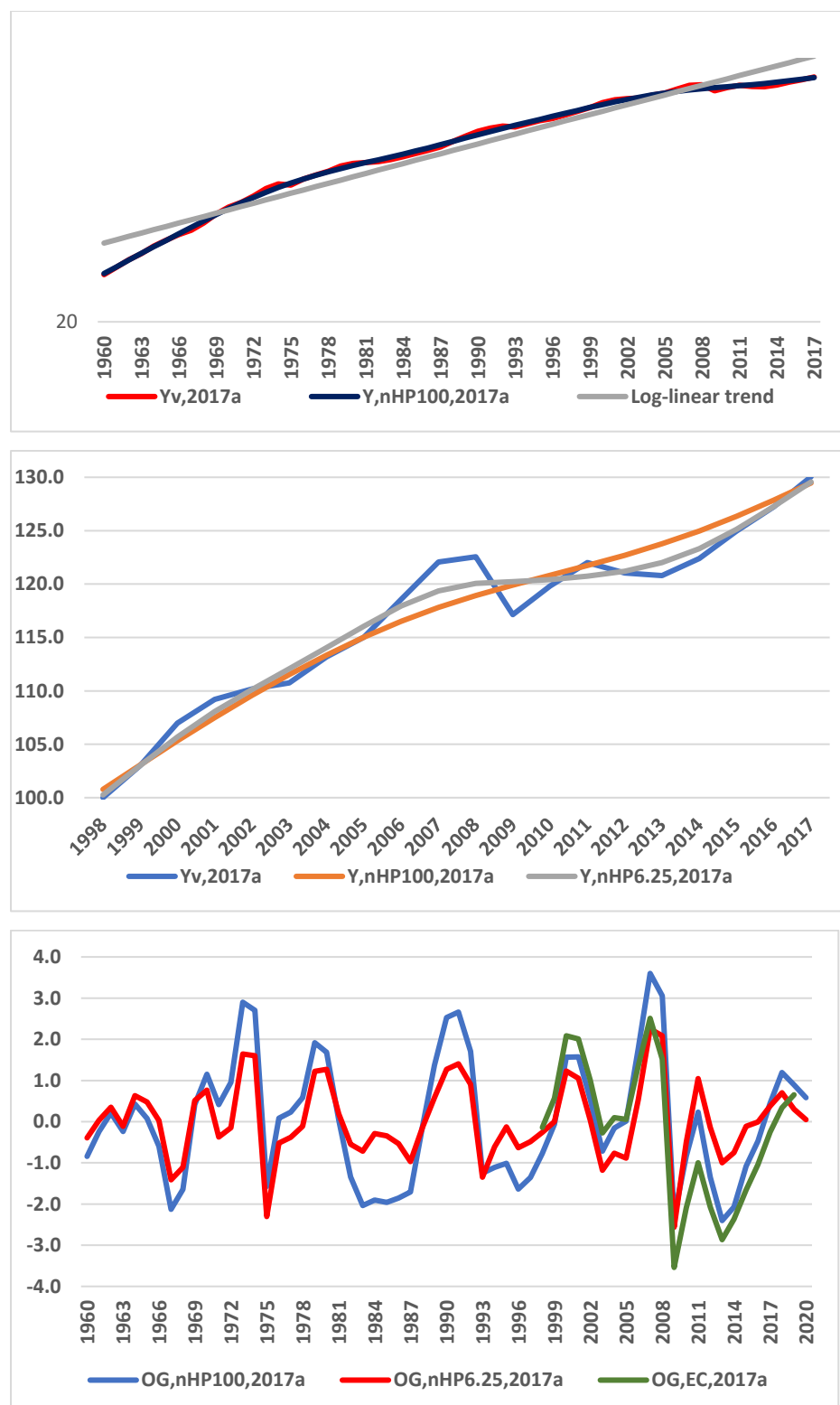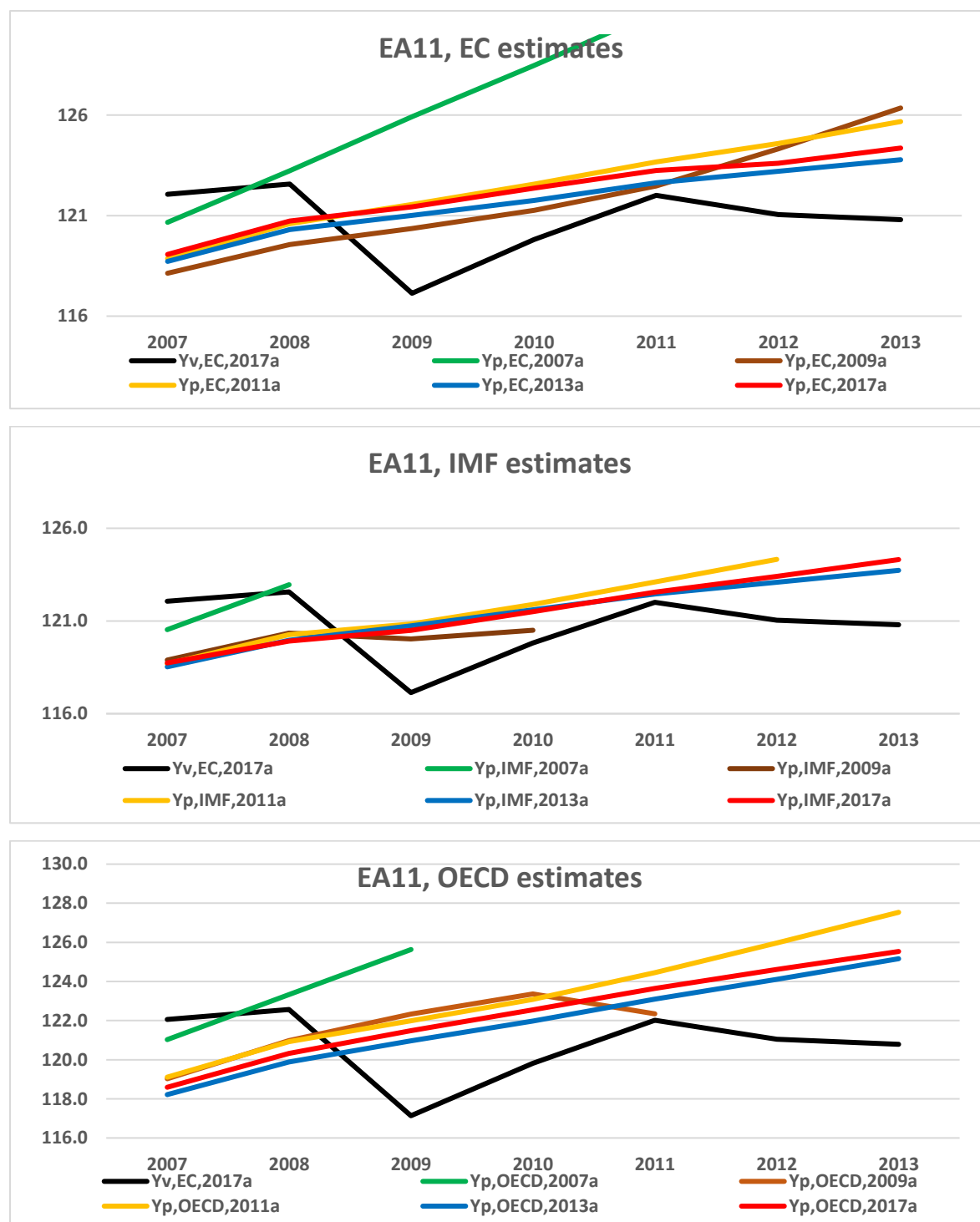

Legend: Yv = GDP for EA11; OG = output gap; nHP100 and 6.25 = HP trend estimate using smoothing parameter 100 and 6.25. 2017a = AMECO data autumn 2017; in the upper panel log scale.
Source: EC, AMECO.

Revisions of potential GDP estimates for EA11 in 2007-2013 of the three institutions

In the main text, Figure 2, we show the new HP estimates for EA11 GDP in 2007-2013 according to selected forecasting vintages. Below are the corresponding estimates of the three institutions.

Figure TA.4. Potential GDP estimates for EA11 in 2007-2013 according to selected forecasting vintages of the three institutions.

Legend: Yp = potential GDP, a = autumn forecasting vintage of the year indicated; Yv,EC,2017a = the ex post GDP data as reported by EC/AMECO in autumn 2017.

Revisions of potential GDP estimates for the US in 2007-2013 of the IMF and OECD

In the main text, Figure 4, we show the new HP estimates for the US GDP in 2007-2013 according to selected forecasting vintages. Below are the corresponding estimates of the IMF and OECD.

Figure TA.5. Potential GDP estimates for the US in 2007-2013 according to selected forecasting vintages of the IMF and OECD.
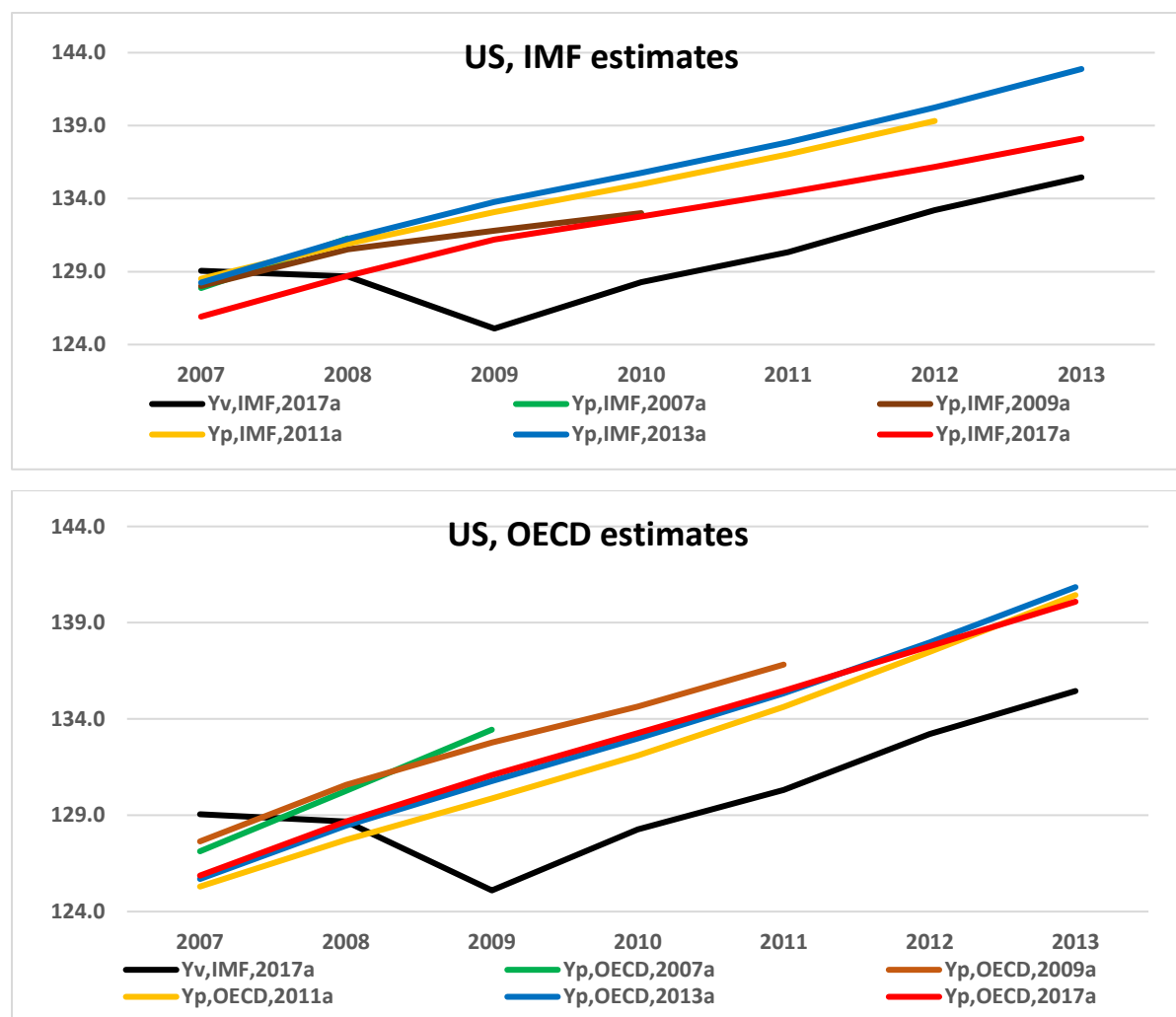
Legend: Y,p = potential GDP, a = autumn forecasting vintage of the year indicated; Yv,IMF,2017a = the ex post GDP data as reported by IMF in autumn 2017.

Section 3. Using the OGs and structural balances for policy: a critical view

Figure TA.6. Output gap estimates for the US in real time for the current year (t) and corresponding estimates in autumn 2017; new HP estimates and by IMF and OECD.

US, OG,nHP
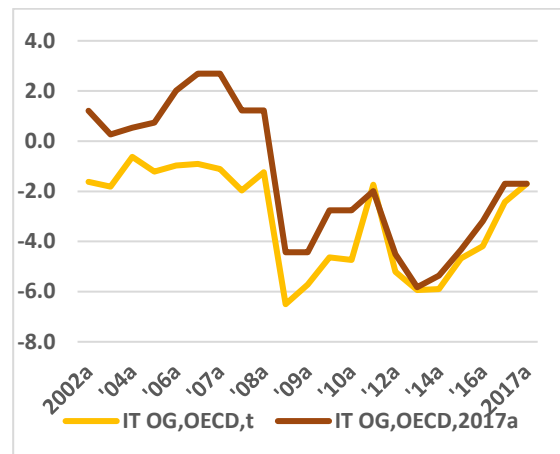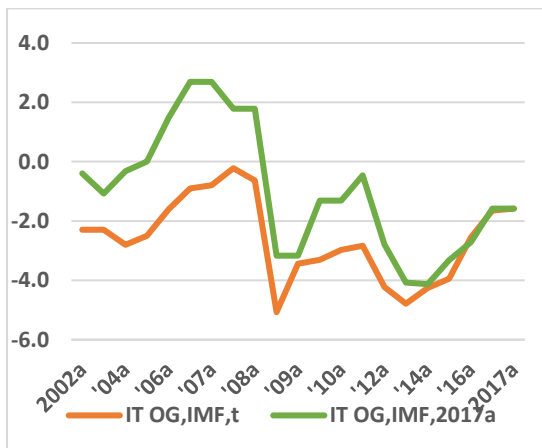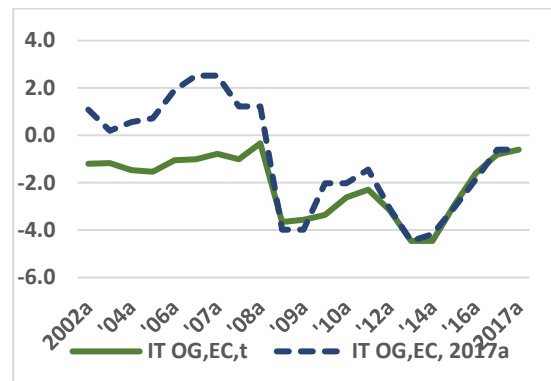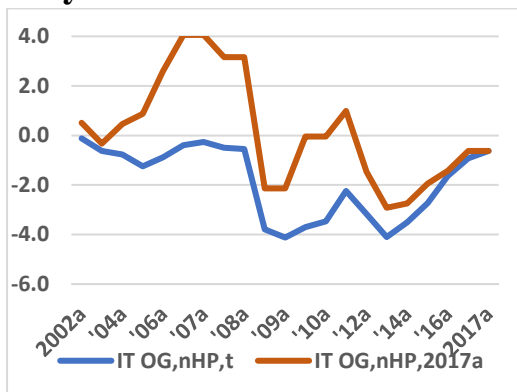

US, OG,OECD

US, OG,IMF

Legend: see the main text.
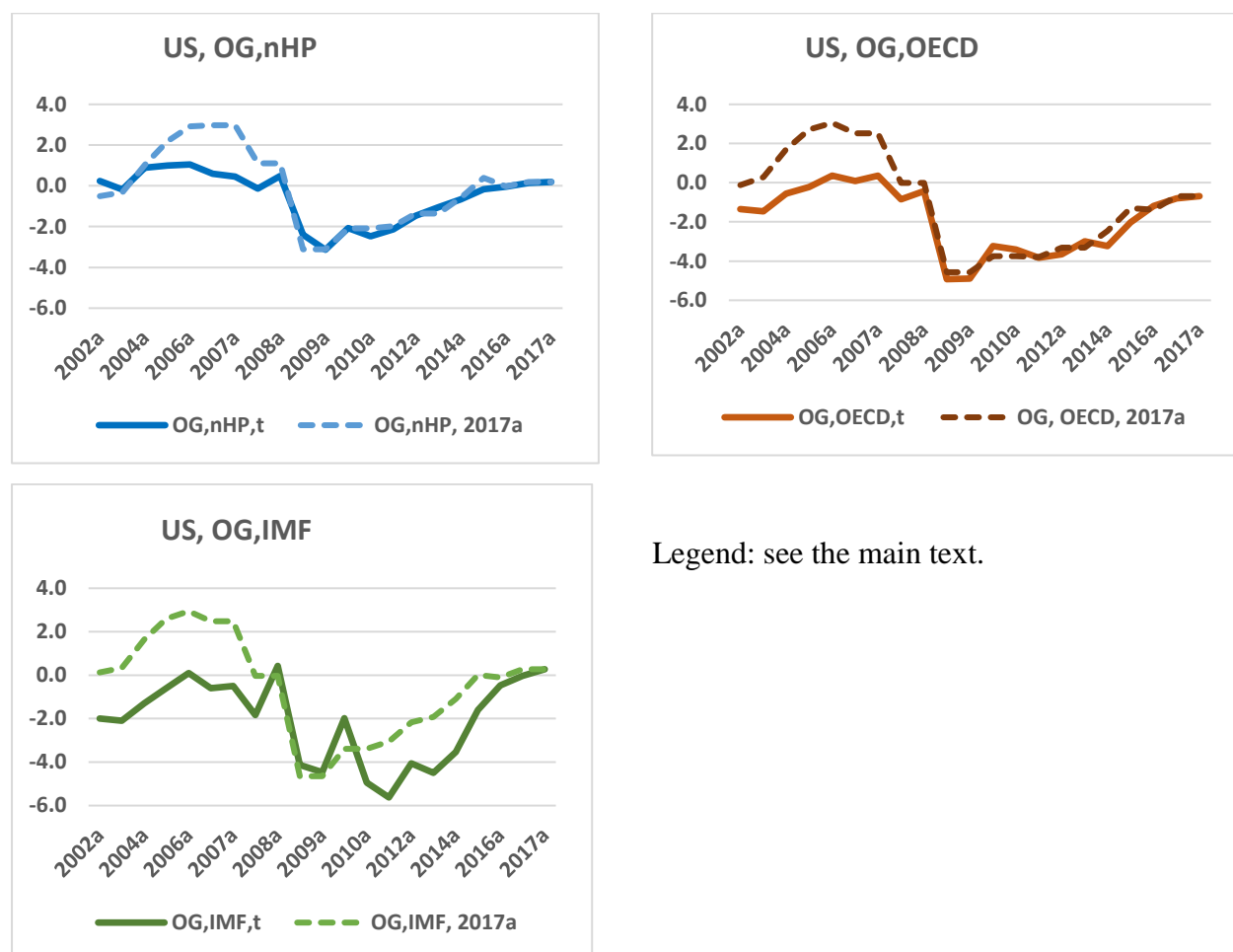
Figure TA.7. Output gap estimates for the four largest euro area counties in real time for the current year (t) and corresponding estimates in autumn 2017; new HP estimates and by EC, IMF and OECD.
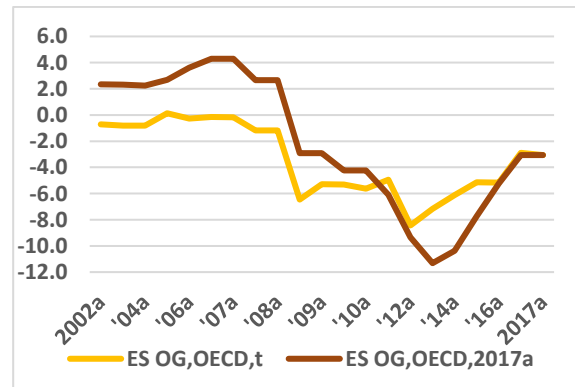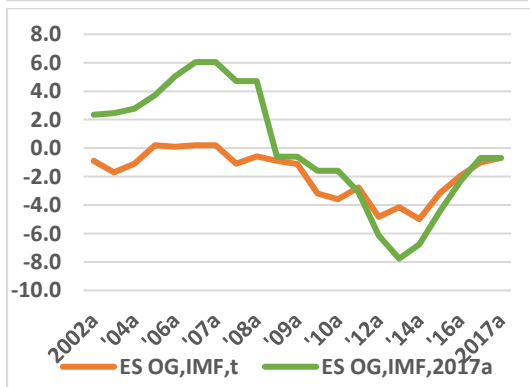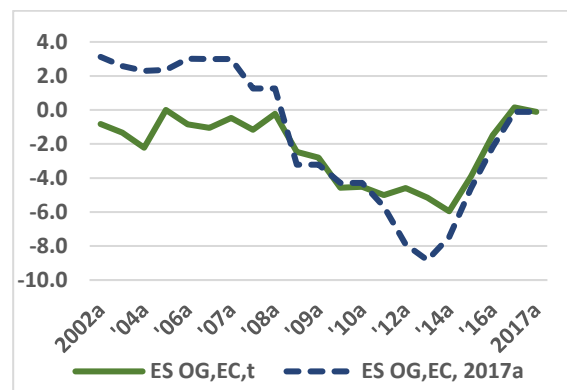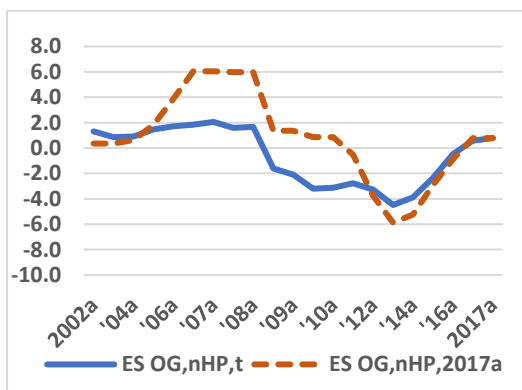
Germany
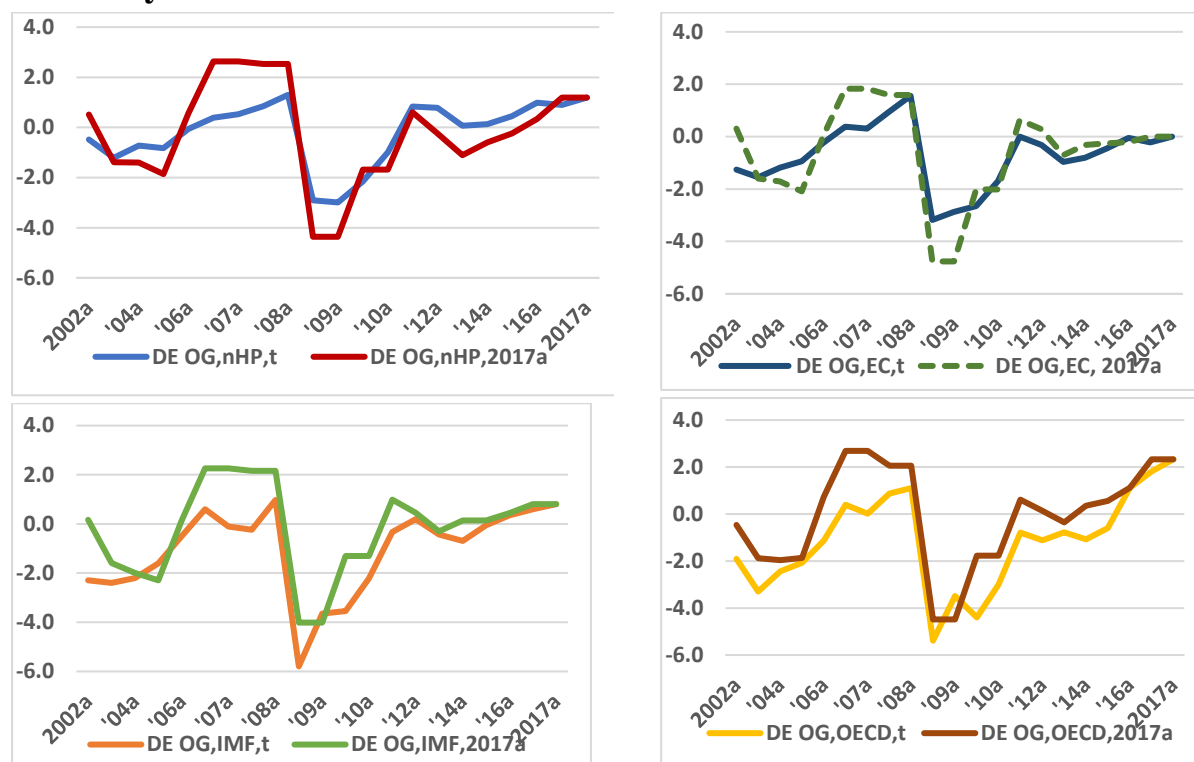
France
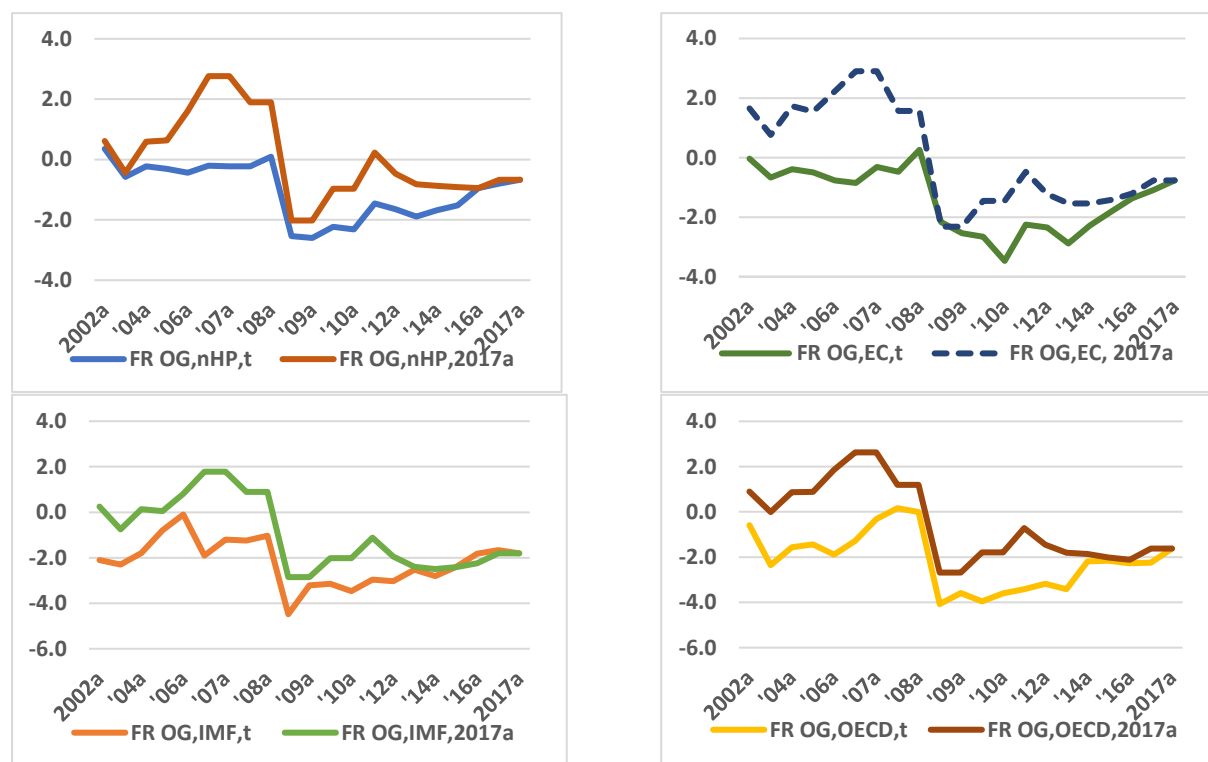
Italy

Spain

References

Alichi, Ali (2015) 'A New Methodology for Estimating the Output Gap in the United States', IMF WP/15/144, July 2015.

Alcidi, Cinzia and Gilles Thirion, (2017), 'Fiscal Risk Sharing and Resilience to Shocks: Lessons for the euro area from the US', CEPS Working Document No 2017/07, May 2017.

Andor, Laszlo, Pervenche Berès, Lorenzo Bini Smaghi, Laurence Boone, Sebastian Dullien, Guillaume Duval, Luis Garicano, Michael A. Landesmann, George Papaconstantinou, Antonio Roldan, Gerhard Schick, Xavier Timbeau, Achim Truger and Shahin Vallée (2018), 'Blueprint for a democratic renewal of the eurozone, Strengthening the monetary union requires a significant leap in governance', Opinion, *Politico*, 28 February 2018, updated 20 April 2018, https://www.politico.eu/article/opinion-blueprint-for-a-democratic-renewal-of-the-eurozone/.

Bayoumi, Tamim (2017), *Unfinished business, The unexplored causes of the financial crisis and the lessons yet to be learned,* Yale University Press, New Haven and London.

Beetsma, R. and H. Oksanen (2008), "Pensions under Ageing Populations and the EU Stability and Growth Pact", CESifo Economic Studies 54, 4/2008, 563–592.

Bénassy-Quéré, Agnès, Markus K Brunnermeier, Henrik Enderlein, Emmanuel Farhi, Marcel Fratzscher, Clemens Fuest, Pierre-Olivier Gourinchas, Philippe Martin, Jean Pisani-Ferry, Hélène Rey, Isabel Schnabel, Nicolas Véron, Beatrice Weder di Mauro, Jeromin Zettelmeyer (2018), 'Reconciling risk sharing with market discipline: A constructive approach to euro area reform', *CEPR, Policy Insight* No. 91, 17 January 2018.

Bielawska, K., A. Chłoń-Domińczak and D. Stańko (2015), "Retreat from mandatory pension funds in countries of the Eastern and Central Europe in result of financial and fiscal crisis: Causes, effects and recommendations for fiscal rules", Warsaw, June 2015.

Blagrave, P., R. Garcia-Saltos, D. Laxton, and F. Zhang (2015) 'A Simple Multivariate Filter for Estimating Potential Output', IMF Working Paper WP/15/79, April 2015.

Busse, Matthias (2016), 'Revisions of the cyclically adjusted budget balance: How large are they and how do they impact the European fiscal framework?', FIRSTRUN (Fiscal Rules and Strategies under Externalities and Uncertainties) Deliverable 2.4, 2016-11-15.

Carnot, Nicolas and Francisco de Castro (2015a), 'The Discretionary Fiscal Effort: An Assessment of Fiscal Policy and its Output Effect', European Commission, Directorate-General for Economic and Financial Affairs, Economic Papers 543, February 2015.

Carnot, Nicolas and Francisco de Castro (2015b), 'The Discretionary Fiscal Effort: An Assessment of Fiscal Policy and its Output Effect', *Hacienda Pública Española / Review of Public Economics,* 215-(4/2015): 63-94.

Carreras, Oriol, Simon Kirby, Iana Liadze and Rebecca Piggott (2016), 'Quantifying Fiscal Multipliers', NIESR Discussion Paper No. 469, 20 October 2016.

Cassidy, John (2010), *How markets fail - the logic of economic calamities*, Penguin edition.

Court of Justice of the European Union (2004) "Case C-27/04, Judgment of the Court", 13 July 2004.

Deutsche Bundesbank (2014), 'On the reliability of international organisations' estimates of the output gap', *Deutsche Bundesbank Monthly Report (April),* 13-35.

De Grauwe, Paul and Yuemei Ji (2018), 'Financial engineering will not stabilise an unstable euro area', VOX, CEPR's Policy Portal, 19 March 2018, https://voxeu.org/article/financial-engineering-will-not-stabilise-unstable-euro-area.

Der Spiegel (1998), "'Dunkelste Stunden' - Der Kanzler öffnet die Akten über die deutsche Einheit. Die Dokumente zeigen: Frankreich hat das schnelle Ende der Mark erzwungen", *Der Spiegel* 18/1998, 27.4.1998, http://www.spiegel.de/spiegel/print/d-7870401.html.

Elmendorf, Douglas W. and N. Gregory Mankiw, (1999), "Government debt", in *Handbook of Macroeconomics, Volume 1,* in J.B. Taylor and M. M. Woodford, eds., Elsevier, Amsterdam etc., 1615-1669.

European Central Bank (2017), 'The ECB's asset purchase programme and TARGET balances: monetary policy implementation and beyond', ECB Economic Bulletin, Issue 3 / 2017, 21-26.

European Commission (2013a), 'Paper on automatic stabilisers', DG Employment, Social Affairs and Inclusion. Main authors Olivier Bontout and Guy Lejeune with Magdalena Ciesielska and Roberta Di Girolamo. Brussels, 04 October 2013.

European Commission (2016), 'Towards a positive fiscal stance for the euro area', 16.11.2016, COM(2016) 727 final.

European Commission (2017a), *Vade Mecum on the Stability and Growth Pact, 2017 Edition,* Directorate-General for Economic and Financial Affairs, European Economy, Institutional Paper 052.

European Commission (2017b), 'Further steps towards completing Europe's Economic and Monetary Union: a roadmap', 6.12.2017, COM(2017) 821 final.

Fatás, Antonio and Lawrence H. Summers (2018), 'The permanent effects of fiscal consolidations', *Journal of International Economics* 112, May 2018, 238-250 (online 2 December 2017).

Havik, Karel, Kieran Mc Morrow, Fabrice Orlandi, Christophe Planas, Rafal Raciborski, Werner Röger, Alessandro Rossi, Anna Thum-Thysen, Valerie Vandermeulen (2014), 'The production function methodology for calculating potential growth rates and output gaps', European Commission, Directorate-General for Economic and Financial Affairs, Economic Papers 535, November 2014.

Heipertz, M. and A. Verdun (2010), *Ruling Europe: The Politics of the Stability and Growth Pact,* Cambridge University Press, Cambridge.

Kuusi (2017), 'Output gap uncertainty and the optimal fiscal policy in the EU', FIRSTRUN (Fiscal Rules and Strategies under Externalities and Uncertainties) Deliverable 2.7, 2017-09-04.

Mc Morrow, Kieran, Werner. Roeger, Valerie Vandermeulen and Karel Havik (2015), 'An assessment of the relative quality of the Output Gap estimates produced by the EU's Production Function Methodology', European Commission, Directorate-General for Economic and Financial Affairs, Discussion Paper 020, December 2015.

Mc Morrow, K., Orlandi, F., Raciborski, R. et al. (2016), 'Medium term economic dynamics of the Euro Area', *International Economics and Economic Policy* 13(1), 27-43.

Minenna, Marcello (2017), 'The ECB's story on Target2 doesn't add up', Financial Times 14 September, 2017, http://marcello.minenna.it/wp-content/uploads/2017/09/20170914_FT_Alphaville.pdf.

Mourre G., Astarita C. and S. Princen (2014), 'Adjusting the budget balance for the business cycle: the EU methodology', European Commission, Directorate-General for Economic and Financial Affairs, Economic Papers 536, November 2014.

Oksanen, Heikki (2016a), 'Smoothing Asymmetric Shocks vs. Redistribution in the Euro Area: A Simple Proposal for Dealing with Mistrust', *CESifo Economic Studies* **62(2)**: 332–375.

Oksanen, Heikki (2016b): 'Coping with the effects of population ageing on public finances in the European Union and China', *Economic and Political Studies* 4(4): 368-396.

The Telegraph (2013), "Helmut Kohl: I acted like a dictator to bring in the euro", *The Telegraph* 9.4.2013, report on an interview with Helmut Kohl made in 2002, http://www.telegraph.co.uk/news/worldnews/europe/germany/9981932/Helmut-Kohl-I-acted-like-a-dictator-to-bring-in-the-euro.html.

Strathern, Marilyn (1997), ''Improving ratings': audit in the British University system', European Review, 5(3), 305-321.

Tuori, Kaarlo and Klaus Tuori (2014), *The Eurozone Crisis, A Constitutional Analysis*, Cambridge, Studies in European Law and Policy, Cambridge University Press.

Turner, David, Maria Chiara Cavalleri, Yvan Guillemette, Alexandre Kopoin, Patrice Ollivaud and Elena Rusticelli (2016), 'An investigation into improving the real-time reliability of OECD output gap estimates', OECD Economics Department Working Papers, No. 1294, 14 April 2016, OECD Publishing, Paris.