The output gap revisited.
Do BTS data on capacity utilisation improve output gap estimates in real time?

Michael Graff and Jan-Egbert Sturm

KOF – Swiss Economic Institute, CH-8092 Zurich

17 November 2009

Extended abstract

Business cycles characteristically manifest themselves in over- or underutilisation of productive resources of an economy. From a theoretical perspective, the output gap, which is defined as the relative deviation of the observed output Y at time t from potential output Y^*, is probably the most comprehensive and convincing concept to describe the cyclical position of an economy. And indeed, it is widely used amongst theorists as well as practitioners. Unfortunately, for practical purposes, the concept depends on the determination of Y^*, which is an inherently unobservable variable. Ideally, a macroeconomic production function should quantify the potential output path. Since this is a formidable task, it is common to refer to statistical procedures designed to isolate the trend of the Y_t series from the cycle and the noise and then to interpret this trend as Y^*_t. This statistical approach impresses through its simplicity. The assumption that trend output corresponds to potential output, however, ignores other information that could lead to a reassessment of the potential. Exogenous shocks or technological developments as well as changes to the capital stock may cause persistent level changes of the potential. Hence, while the output gap might be a useful concept for theoretical thinking about inflationary pressures *ex post*, its practical usefulness is severely impaired or even annihilated by the inherent difficulty to know with sufficient reliability the magnitude of the output gap at the time *when the policy maker needs to know* it, i.e. in *real time*. This to some degree affects all detrending methods used to isolate potential from observed GDP, as in real time, it is virtually impossible to distinguish between cycle and trend.

In this paper, we examine whether the notoriously unreliable real time estimates of the output gap can be improved by referring to measures of physical capital capacity utilisation from business tendency surveys (BTS). These are highly informative data, as they relate directly to the stress on the current capacity to produce goods and services. Moreover, and importantly in our context, these data are usually not revised, so that they are not affected by the end point problem.

To assess this question empirically, we construct a large panel data set, comprising 26 countries and yearly as well as quarterly data from1970 to 2009 (lengths depending on the particular series and country) with qualitative and quantitative information on capacity utilisation from BTS in quarterly frequency and output gap estimates as published by the OECD in real time. We show that the real time output gaps are informationally inefficient in the sense that BTS data available in real time can help to produce estimates that are significantly closer to later releases of output gap estimates. Sensitivity analyses confirm that this finding is robust.

Keywords: Output gap, capacity utilisation, real time analysis, survey data

JEL Classification: D24, E32, E37

0 Introduction

In this paper, we examine whether the notoriously unreliable real time estimates of the output gap can be improved by referring to measures of physical capital capacity utilisation from business tendency surveys (BTS). To assess this question empirically, we construct a large panel data set, comprising 26 countries and yearly as well as quarterly data from1970 to 2009 (lengths depending on the particular series and country) with qualitative and quantitative information on capacity utilisation from BTS in quarterly frequency and output gap estimates as published by the OECD in real time. We show that the real time output gaps are informationally inefficient in the sense that BTS data available in real time can help to produce estimates that are significantly closer to later releases of output gap estimates. Sensitivity analyses confirm that this finding is robust.

The paper is organised as follows: After a discussion of the business cycle and its theoretical and empirical reflection in the output gap (section 1), we discuss and describe our data (section 2). Then, we present our empirical analyses (section 3). Section 4 summarises and concludes.

1 The business cycle and the output gap

Business cycles characteristically manifest themselves in over- or underutilisation of the productive resources of an economy. Consequently, applied business cycle analysis regularly refers to variables reflecting the utilisation of the factors of production – labour and capital –, usually approximated by the unemployment rate and the rate of capital utilisation. Apart from this, estimates of the output gap aim at quantifying aggregate economic capacity utilisation.

Labour and capital utilisation consider just one factor of production and thus may give a distorted picture of the cyclical movement of an economy as a whole, unless labour and capital utilisation are strictly co-moving. This, however, is usually not the case. Labour utilisation is typically *following* the business cycle, and furthermore, is it less cyclically responsive compared to capital utilisation. In addition to this, though unemployment rates at first glance seem to be well-documented and readily available statistics of the business cycle, the data are almost hopeless for international and inter-temporal comparisons.

Capital utilisation appears more convincing, since it typically proceeds simultaneously with the business cycle. Unfortunately, it is not commonly documented statistical information, so that it is not useful for practical purposes.

From a theoretical perspective, the output gap $g = (Y_t - Y^*_t)/Y^*_t$, which is defined as the relative deviation of the observed output Y at time t from potential output Y^*, is probably the most convincing concept to determine the cyclical position of an economy. And indeed, it is widely used amongst practitioners. Moreover, it is a well-established theoretical concept in contemporary economics and plays a crucial role in many structural macro-models. It is frequently (and with some success) referred to in research looking for a scheme to „explain" (reproduce) historical paths of central bank policy settings; and it is probably safe to assume that a substantial number of monetary policy makers as well as fiscal authorities[1] pay close attention to real-time estimates and forecasts of the output gap. As a matter of fact, it would be very hard to understand e.g. the behaviour of the US Federal Reserve System without reference to the output gap or the employment gap.

[1] For example, the recently introduced Swiss „debt break" specifies the budget deficit (surplus) as a function of an HP filtered real-time output gap.

Although the output gap plays such a prominent role in current economic theory and policy, it still is an inherently unmeasurable, equilibrium based construct. It refers to the deviation of realised from potential output, and this in a constantly changing economic environment. Successfully estimating the output gap would require not only reasonably reliable data of current or near future *realisations* of economic activity (which are hard enough to get),[2] but also reasonably reliable estimates or projections of *potential* (*equilibrium*) output Y^*, which – like the business cycle – is an inherently unobservable variable.

Unsurprisingly thus, apart from a general understanding that the output gap g denotes the relative departure of empirical output Y from its „equilibrium" or „potential" Y^*, the current state of the art does not give a conclusive answer to how it should be conceptualised. We can distinguish (at least) three attempts at defining it: a substantive, a statistical and a functional:[3]

(1) The *substantive approach* argues that potential output is a function of the amount of the factors of production voluntarily[4] available at the period under consideration and the technology at hand to combine them to produce goods and services. This is clearly an economically meaningful concept,[5] and – presumably – this is the basis for its popularity. To implement it, however, is a formidable task; and while a number of attempts to estimate full capacity production functions have been conducted, some, if not most, of the results have been rather disappointing. Consequently, the substantive approach, to which the output gap owes much of its credit, is not the one that practical economists usually refer to.

(2) According to the *statistical* approach, potential output is what you get when you send a real GDP series through a low pass filter (frequently the HP filter) and relate it to the unfiltered series.[6] Many of the practical methods nowadays in use to derive estimates of potential output nowadays partly or wholly rely on this statistical approach of extraction of a smooth trend from the historical path of the output series. However, if these methods do not predict a constant growth rate for potential output but allow for some (but not total)[7] adaptation of potential to observed output, real-time gap estimates are imperfect in the sense that they are firstly prone to revisions as new data keep coming in and secondly systematically biased in periods of structural change, since the trend is ultimately identified *ex post* by past *and* future realisations. Regrettably, this is true for linear time-invariant filters and band pass filters alike.[8] Moreover, the only theoretical notion behind this black box approach is that potential GDP is evolving along a path that shows a considerable amount of inertia.

(3) The *functional approach*: Potential output is the level of output at any point in time that results in zero inflationary pressure. This is sometimes labelled NAILO („non-accelerating-

---

[2] Evidently, estimating potential output from unreliable estimates of real output is not likely to produce reliable estimates of output gaps.

[3] See also Chagny and Döpke (2001) and Dergiades and Tsoulfidis (2007).

[4] This needs to be stressed, since potential labour is not just a linear function of a well defined demographic cohort, but intrinsically endogenous, replying to a wide array of economic incentives, regulatory interventions, changing tastes (e.g. for leisure, labour force participation of women, consumption ingredients of prolonged education etc). In addition, effective working hours directly affect the intensity with which the stock of physical (and other) capital is used, so that the effects of fluctuations in effective labour are further amplified.

[5] The neo-Keynesian representation is neatly verbalised by Nelson and Nikolov (2003): „… economic theory suggests…that potential output corresponds to the output level that would prevail in the absence of nominal wage or price rigidity."

[6] Of course, this is not a very useful definition (like saying GDP is what is published by the Statistical Office), but in the end it is not completely without sense, because this is how y^* is frequently computed.

[7] In this trivial case the output gap would vanish (i.e. equal 0 for all observations).

[8] For an elaboration of this point, cf. van Norden (2002).

inflation rate of output")[9] and is conceptually related to – but not identical with – the NAIRU („non-accelerating-inflation rate of unemployment"). The difference between the two is that the first is based on the existence of an equilibrium potential output path, while the latter postulates an equilibrium rate of unemployment, but to the degree that there is a close relation between output and employment (which there certainly is, since the level of employment is the major causal determinant of output, be it factual or potential), the distinction between the two gets academic rather than practical.

Note that this appears like an elegant approach to overcome the practical difficulties with the substantive notion of the output gap. If theory tells you that a positive (negative) output gap creates inflationary (deflationary) pressure and/or over-employment (underemployment) of the factors of production, why not use this theoretical link to identify the output gap inductively by looking at inflationary and/or factor market pressures?[10] To start with, find the points in time when inflationary pressure was zero – e.g. realised inflation (π) equalled expected inflation (π^e) – and/or the points in time when unemployment/capacity utilisation was equal to „equilibrium" – e.g. some longer term average of their past realisations –, and you have identified periods where „functional" potential output equalled observed output. Then, specify functional relationships between the output gap and inflationary pressure and/or unemployment/excess capacity utilisation. Finally, collect data on your indicators and refer to the functional relationships to derive a quantitative measure of the output gap. And indeed, so-called „multivariate filters" that amend the univariate low pass filter approach with additional information, are not uncommon.[11]

However, there are two caveats. Firstly, to incorporate additional indicators for strain on resources into GDP centred estimates of the output gap, they themselves have to be formulated in gaps.[12] In other words, to help gauge the „unobservable" potential output a range of other „unobservables", e.g. the NAIRU and/or „desired" or „equilibrium" capacity utilisation are referred to. Hence, the problem of not being able to measure potential output directly translates into the problem of quantifying the NAIRU[13] and/or „equilibrium" capacity utilisation. The improvement in the augmented output gap measure is therefore subject to the validity of the approaches to estimate the „second order" unobservables. Secondly, potential output is now partly endogenised. Specifically, to the extent that the additional information dominates the output gap estimate, this approach reverses the theoretical relationship

$$\textit{output gap} \rightarrow \textit{inflationary pressure}$$

into an inductive measurement model

$$\textit{inflationary pressure} \rightarrow \textit{output gap},$$

thereby depriving the output gap concept of most of its original substantive content. With potential output being identified contingent on observed inflation and/or inflationary pressure, one can no longer claim that the correlation between such an output gap measure and ob-

9 Cf. Hirose and Kamada (2003).

10 Cf. Laxton and Tetlow (1992) for the seminal contribution for this approach.

11 In particular, a number of central banks, e.g. the Bank of Canada, the Reserve Bank of New Zealand and the Reserve Bank of Australia refer to „multivariate filters".

12 Laxton and Tetlow (1992), Butler (1996). The Reserve Bank of New Zealand's approach follows the same logic, cf. Conway and Hunt (1997).

13 For a fundamental critique of the NAIRU cf. Hagger and Groenewold (2003). Evidence for the practical usefulness of a Phillips curve relationship to forecast inflation is mixed. For example, Gruen et al. (2002) report encouraging evidence from Australia, whereas Robinson et al. (2003) point to difficulties with real-time estimates, and Lansing (2002) argues that it is of little or no use for the USA.

served inflation represents a structural relationship. It is there by construction.[14] Hence, with a functional measurement approach, the output gap loses some of its original sense and should properly be regarded as an econometric indicator of inflationary pressure.[15]

To summarise, the potential output path Y^*_t should ideally be quantified referring to a full-blown production function. Since this is a formidable task, however, it is common to refer to either to univariate statistical procedures – filters – that are designed to isolate the trend of the Y_t series from the cycle (and the noise) or to eclectic approaches such as „multivariate filters" and then to interpret this trend as Y^*_t. Various filters are doing the job fairly well, and the statistical approach impresses through its simplicity. The assumption that the univariate output trend corresponds to potential output, however, suffers from the fact that it ignores all other information that could lead to a reassessment of the potential. Exogenous shocks or technological developments which may lead to persistent level changes of the potential are ignored, as are changes to the stock of accumulated factors of production (physical and human capital) due to changes to net investment ratios. The last point is particularly critical: while shocks to observed output – which are filtered out by a low pass filter – rightly appear as deviations from potential, technical change or evolution of the economy's capital stock are not duly considered when determining potential output with a low pass filter, which would identify them as cyclical.

It is hence not a surprise that serious doubts have been expressed as to whether the output gap is a *practically useful* concept. In two seminal papers Orphanides and van Norden (2002, 2003) argue and illustrate empirically that while the output gap might be a useful concept for theoretical thinking about inflationary pressures, and while in addition to this, this usefulness is empirically well-established *ex post*, its practical usefulness is severely impaired or even annihilated by the inherent difficulty to know with sufficient reliability the magnitude of the output gap at the time *when the policy maker needs to know* it, i.e. in *real time*. Specifically, we are confronted with the „end point problem", which reflects the fact that without knowledge of the future, it is impossible to distinguish between cycle and trend, so that when shifts of the latter are eventually discovered, prior estimates of potential output have to be revised. This view is supported by a large body of empirical evidence, which also suggests that the end point problem associated with the output gap may already have led to severe misjudgements and resulting policy mistakes that only become clear in hindsight. Notably, Orphanides (2003, p. 997) compares a reconstructed real-time output gap series for the US going back to 1951 with today's view and finds persistent underestimation through most of the period until the mid-eighties. In the mid-seventies, the misperception amounted to an incredible *ten percentage points* of potential output, which in a simulated real-time Taylor rule framework would suggest that the Fed's monetary policy during the „Great inflation" was by no means meant to be permissive. Similarly, Nelson and Nikolov (2003) reconstruct a real-time output gap series for the UK going back to 1965 and plug this into a standard monetary policy framework. They find that the Bank of England's failure to lean against inflation in the early 1970s can be attributed to a real-time perception of the output gap that was *seven percentage points* lower than what one would quantify it nowadays. Cayen and von Norden (2002) conduct similar analysis for Canada since 1981 and find revisions of up to *six percentage points*

14 This circularity can be traced to the very origins of multivariate filtering; cf. Laxton and Tetlow (1992: i): „… if movements of potential output have a different effect on inflation than do cyclical movements in output, then information on inflation may be useful in identifying potential output."

15 An algebraic representation may help to clarify this point: When potential output Y^* is made endogenous on inflation π and a vector of indicators for strain on resources s, we get $Y^* = Y - f(\pi, s)$. Recall the substantive definition of the output gap (in absolute terms): $g^{abs} = Y - Y^*$. Accordingly, in the limiting case that potential output is derived exclusively by correcting observed output by the deviation of potential from observed output, the magnitude of which is estimated through the indicator model f, both potential and observed output are cancelled from the „gap": $Y^* = Y - f(\pi, s) \Rightarrow g^{abs} = Y - Y^* = f(\pi, s)$.

of potential GDP.[16] For Japan, Hirose and Kamada (2003) find that since 1995 an output gap which is derived by an HP filter augmented with a Phillips curve relationship would have suffered revisions of the same magnitude. An analysis for Finland (Billmeier 2004) finds that out of nine output gap measures none would add significantly to a univariate autoregressive explanation of annual CPI inflation from 1980–2002 and attributes this to the fact that a „statistically satisfying measure of potential output" might not be feasible for a high volatility observed (yearly) output series like the Finnish one (p. 27). The empirical literature thus casts serious doubt on the practical usefulness of the prevailing output gap measures in *real time.* When information on the state of the economy is most important, estimates of the output gap are to be uncomfortably unreliable; and this includes gaps resulting from multivariate filtering.[17]

Obviously, real-time uncertainty about the magnitude of the output gap is not merely a theoretical concern. There is evidence that reliance on the output gap might be responsible for some of the gravest central bank mistakes of the last decades, when real-time output gap measures failed to take account of changes in the growth rate of potential output. More recently, the discussion about the retarded effects of the IT revolution and the „jobless" recovery in the US points to the possibility of another major change in the growth rate of potential output,[18] and the question in how much the global recession that was triggered by the subprime mortgage turmoil in the US will affect potential output in the years to come is now on the agenda.

What lessons can we learn from this? In our view, apart from a serious warning to take real time estimates of the output gap with more than just a grain of salt, the importance of the output gap merits devoting more effort to improve the lamentable real time characteristics of its prevailing empirical implementations. In particular, we shall now turn to some information that – to the best of our knowledge – has so far not been systematically exploited to improve output gap estimates in real time, which is the assessment of the degree of capacity utilisation by firms as reflected in business tendency survey (BTS).

BTS are nowadays conducted in a considerable and increasing number of countries. For our purposes, they are invaluable, as they reflect unique information on technical capacity. In particular, firms are asked to make a judgement on the degree of utilisation of their „technical capacities" in qualitative terms (normal, above or below normal); and many surveys also to give a quantitative estimate of the firm's rate of capacity utilisation in per cent.

The rate of capacity utilisation that can be inferred from these surveys is an important business cycle indicator, as it relates directly to the stress on the current capacity to produce goods and services. From a policy perspective, technical bottlenecks indicate a positive output gap, whereas idle capacity above normal would have it negative. Yet, it is not obvious which rate of capacity utilisation should be regarded as normal. Moreover, the level of normal capacity utilisation can change over time. When the substitutability of physical capital declines, firms will tend to keep more idle reserves to make sure they can cope with unexpected orders. On the other hand, with technical and organisational progress making production more flexible,

---

16 While Cayen and van Norden evaluate a wide range of output gap estimation methodologies, they lamentably do not include the Bank of Canada's multivariate filter. They note (p. 58) that this would be „interesting". The reason for this omission is probably that the Bank of Canada's multivariate filter was only installed in the mid-nineties, and whereas the other methodologies allow for „backcasts" to the beginning of the 1980s, the multivariate filter cannot easily be simulated. Note that the same is true for the Reserve Bank of New Zealand's MV filter.

17 See e.g. Gruen at al. (2002), Orphanides and van Norden (2002), Rünstler (2002) and Graff (2004).

18 As Kahn and Rich (2003) have recently pointed out, accepting the „new economy" story and assuming a sustainable acceleration of potential output growth would significantly lower our present real-time estimates of the output gap which tend to attribute fast growth to cycle rather than trend.

the normal rate of capacity utilisation could increase. In a similar fashion, a move on to just-in-time production could lift the normal rate of capacity utilisation. These reflections imply that the rate of capacity utilisation is not, or not always, a stationary variable. Moreover, due to the ambiguity of the theoretical predictions, it is not clear whether we should expect an increase or a decline in the level that is considered normal.[19]

Accordingly, although BTS deliver highly relevant and timely information on firms' self assessment of the stress on their technical capacities, they cannot be used directly to compute economy-wide measures of capacity utilisation. However, this information could be extremely useful to add confidence to timely estimates of the output gap, as BTS data are usually not revised; they are final as soon as a survey is completed.

In what follows, we shall demonstrate that the above conjecture is reflected in the real time data. In particular, we show that some important and widely circulated output gap estimates – those published bi-annually by the OECD – are indeed informationally inefficient in the sense that BTS data available in real time can contribute to produce estimates that are closer to the final values.

2 Data

The sample

We refer to large panel data set, comprising 26 countries from1970 to 2009 (lengths depending on the particular series and country) with qualitative and quantitative information on capacity utilisation from BTS in quarterly frequency and output gap estimates as published by the OECD in real time.

The BTS data on capacity utilisation are all available in quarterly frequency, and they also relate to quarters.[20] The OECD output gap estimates are released bi-annually and relate to years as well as to quarters. The natural choice thus appears to perform our analyses with quarterly data. However, as in most empirical work, we are faced with a *trade off* between the time span covered and the depth of the information, as the OECD started to release output gap estimates referring to years in December 1995, whereas the quarterly data were first published in December 2003. Although the quarterly frequency is certainly more appropriate to trace the business cycle, and at this stage (2009) it also offers more degrees of freedom than yearly date that go eight years further back into the past, the yearly data cover more cycles and turning points. Since revisions to real time estimates of the output gap are especially pronounced at turning points, a look at the yearly frequency may add confidence to the results obtained with quarterly data or offer additional insights. Moreover, quarterly GDP data – which are the basis on which output gap estimates are built – are often of questionable quality.[21] They are commonly obtained from a quarterly breakdown of yearly aggregates from national accounts with help of quarterly indicators series. The quarterly pattern, which is to some degree arbitrarily imposed on the yearly aggregates then accounts for a major share of the variance of a series. Yearly series do not suffer from these problems. We shall hence conduct our analyses both with quarterly and yearly data.

---

19 See e.g. Shapiro et. al. (1989), Bansak et al. (2007) and Etter et al. (2008).

20 Some of our sources report monthly series; we transform these by taking quarterly averages.

21 See Agénor et al. (2000).

Output gap

Considering the fact that there is a variety of techniques to estimate output gaps, the choice of how to specify our key variable is based on economic as well as pragmatic reasons.

Starting with the latter, for the purpose of this paper, the preferred output gap estimates should either have a sufficient long and documented history or else be computed in a way to enable us to reconstruct vintages of real time data that are comparable across time and countries. For this reason, we cannot resort to collect the various output gap estimates published by central banks, statistical offices or other institutions. Moreover, despite increasing efforts to provide researchers with real time data, the information required to simulate sophisticated real time estimates of the output gap for a panel comprising a broad sample of countries as well as time series that are long enough for a meaningful revision analysis are not available at this stage, and it will take more effort and time until this may eventually become an option.

Regarding the information available, the two feasible options are either to find reasonably sophisticated estimates referring that have a history which is well enough documented, or to refer to real time vintages of GDP only and compute output gap vintages based on univariate methods, e.g. with a low pass filter. This latter however, would result in output gap vintages that consider nothing apart from GDP and admittedly suffer from a well-known end point problem, so that it should not come as a big surprise that one could have done better in real time. We shall hence refer to published data that are based on a unifying framework.

The output gap data corresponding closest to our requirements are those of the OECD. According to the documentation given in the OECD Economic Outlook, these estimates are usually based on multivariate techniques with reference to economic theory:

> "The output gap is measured as the percentage difference between actual GDP in constant prices, and estimated potential GDP. The latter is estimated using a production function approach for all countries except Portugal, taking into account the capital stock, changes in labour supply, factor productivity and underlying non-accelerating wage rates of unemployment or the NAWRU for each Member country. Potential output for Portugal is calculated using a Hodrick-Prescott filter of actual output." [22]

Vintages of the OECD output gap estimates are documented since 1995, and the cross-sectional coverage corresponds roughly to the OECD member countries, so that reconstruction of a reasonably large real time panel is possible. The estimates are released bi-annually at the occasion of the publication of the OECD Economic Outlook in June and September. They relate to years as well as to quarters. Output gap estimates referring to years started to be included in the OECD Economic Outlook data set from No. 57 onward in December 1995; quarterly data were first published in December 2003 (Economic Outlook No. 74).

Our data are exclusively obtained online via „Source OECD". In August 2008, the OECD released a real time data base of its output gap estimates,[23] but the online data sets referring to the OECD Economic Outlook issues No. 57 to 85 (June 2009) cover considerably more data points than the real time data base, so that we construct a broader and longer real time panel. The time reference of our vintages is shown in table 1.

As the publication rhythm is bi-annual, whereas the frequency of the series is quarterly and annual, we define as "real time" those quarterly data that are published in the corresponding semester (i.e. two data points), and those annual estimates of g that are published in December

[22] This text is now found in every issue of the OECD Economic Outlook, along with reference to a paper by Giorno et al. (1996).

[23] http://www.oecd.org/document/1/0,3343,en_2649_33715_41054465_1_1_1_1,00.html (OECD Quarterly output gap revisions database, August 2008); for a documentation, see Tosetto (2008).

of that year (one data point). For the annual data, the June release of the same year is considered to be a two quarter ahead forecast, the June release of the following year is first revision after two quarters, and so on. Equivalently, for the quarterly series, the values reported for g referring to the two quarters preceding the publication quarter are regarded as the first revision with a lag of two quarters, and the number referring to the two quarters following the publication quarter are regarded as forecasts with a lead of two quarters, and so on.

The OECD output gap data release and revision sequence for quarterly and annual frequency is illustrated in figures 1 and 2.

Figure 1: OECD output gap release and revision sequence for quarterly frequency

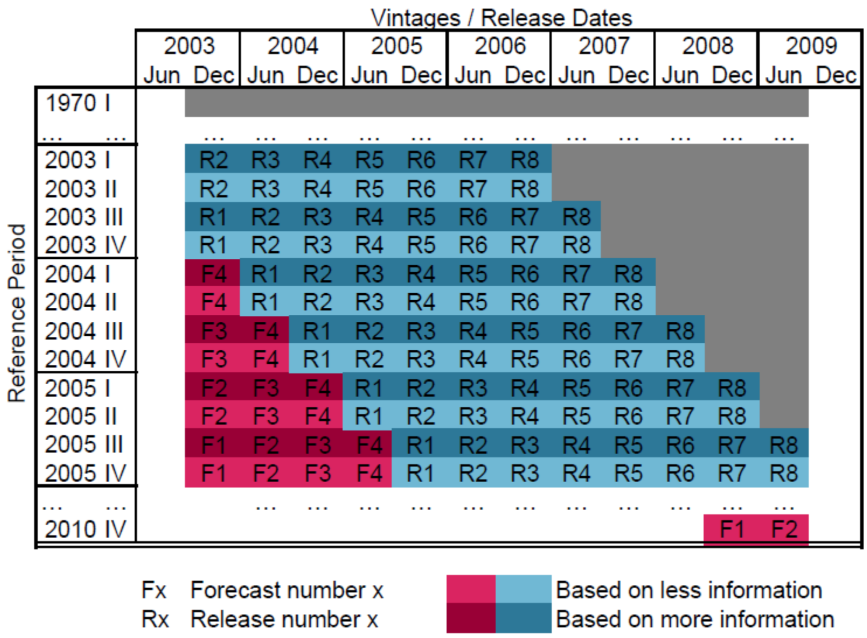
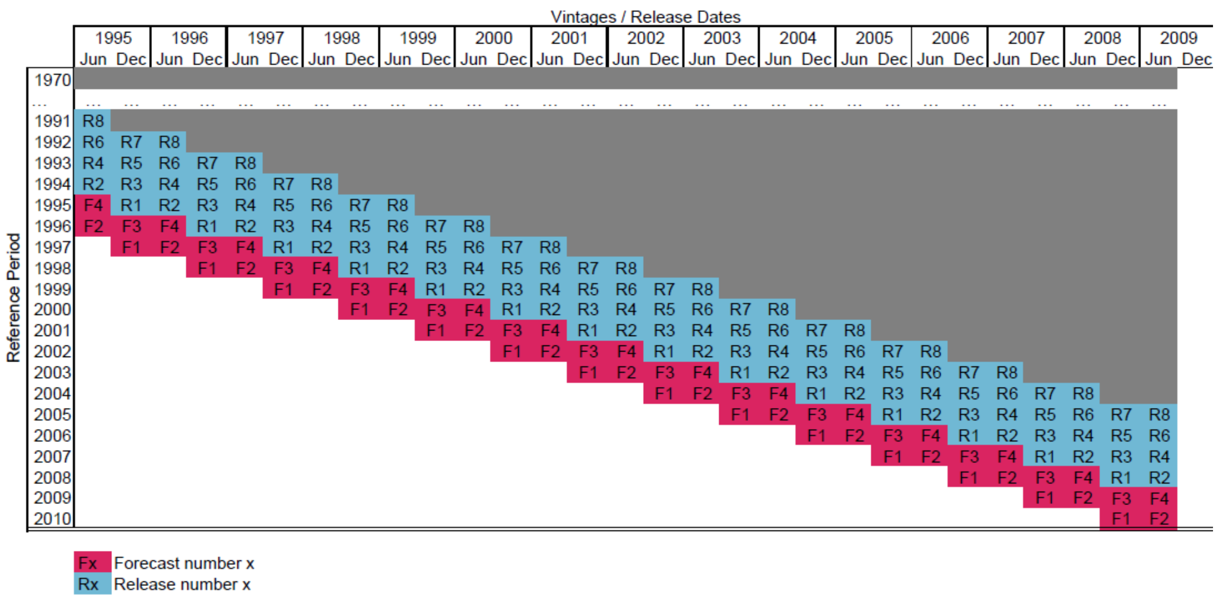


Figure 2: OECD output gap release and revision sequence for annual frequency



After a check for consistency and completeness of the output gap vintages, we carefully amended the data for two countries:

- The OECD has to this date not published output gap estimates for Slovakia. However, in the IMF World Economic Outlook Database, the April 2009 vintage for the first time includes annual output gap estimates for Slovakia.[24] A quick check demonstrates that this series has a correlation of $r = 0.73$ with the OECD December 2008 estimates, but only $r = 0.48$ with the June 2009 vintage. Hence, leaving methodical differences aside, it appears that the information reflected by the IMF April 2009 data are closer to the OECD December 2008 estimates than to the OECD June 2009 data. Accordingly, we refer to the IMF data to replace the missing values for Slovakia in the OECD December 2008. As there is only one vintage, this country does not enter into the revision analysis, but we keep it in the panel for general statistical observations, e.g. descriptive statistics.
- Apart from this, we refer to the IMF data to backcast the output gap for Germany prior to 1991, i.e. when the German Democratic Republic ceased to exist and joined the Federal Republic of Germany. We spline the IMF series to the OECD series in 1991.

The resulting output gap vintages comprise a maximum of 26 countries for which we could also find data on capacity utilisation from surveys. These countries are shown in table 2, along with the length of the quarterly and annual real time output gap series.

[24] The IMF publishes only yearly data, starting in 1980. Also, the country coverage of the data base is less than that of the OECD, and the IMF output gap series obviously suffer from severe end point problems at the left margin of the series (e.g. for Belgium, the April 2009 IMF output gaps are consistently above 10% from 1980–1995, which does not make any sense; similar problems are encountered regarding Finland, Italy and New Zealand). Accordingly, we did not seriously consider basing our analyses on the IMF rather than the OECD data.

Table 1: OECD output gap estimates in real time: reporting and reference periods

OECD Economic Outlook, online data	Reporting time	Latest reference year	Latest reference quarter	
			OECD real time data base	Online data, June 2009
No. 57	1995q2	1996	–	–
No. 58	1995q4	1997	–	–
No. 59	1996q2	1997	–	–
No. 60	1996q4	1998	–	–
No. 61	1997q2	1998	–	–
No. 62	1997q4	1999	–	–
No. 63	1998q2	1999	–	–
No. 64	1998q4	2000	–	–
No. 65	1999q2	2000	–	–
No. 66	1999q4	2001	–	–
No. 67	2000q2	2001	–	–
No. 68	2000q4	2002	–	–
No. 69	2001q2	2002	–	–
No. 70	2001q4	2003	–	–
No. 71	2002q2	2003	–	–
No. 72	2002q4	2004	–	–
No. 73	2003q2	2004	–	–
No. 74	2003q4	2005	2003q2	2005q4
No. 75	2004q2	2005	2003q4	2005q4
No. 76	2004q4	2006	2004q2	2006q4
No. 77	2005q2	2006	2004q4	2006q4
No. 78	2005q4	2007	2005q2	2007q4
No. 79	2006q2	2007	2005q4	2007q4
No. 80	2006q4	2008	2006q2	2008q4
No. 81	2007q2	2008	2006q4	2008q4
No. 82	2007q4	2009	2007q2	2009q4
No. 83	2008q2	2009	2007q4	2009q4
No. 84	2008q4	2010	–	2010q4
No. 85	2009q2	2010	–	2010q4

Table 2: Real time output gap data, reference period 1970q1–2009q2

Country	Output gap in % of potential output	
	Quarterly	Annual
Australia	2003q3–2009q2	1995–2008
Austria	2003q3–2004q2	1995–2008
Belgium		1995–2008
Canada	2003q3–2009q2	1995–2008
Czech Republic	2005q3–2009q2	2005–2008
Denmark	2007q1–2009q2	1995–2008
Finland	2003q3–2009q2	1995–2008
France	2003q3–2009q2	1995–2008
Germany	2003q3–2009q2	1995–2008
Greece	2004q1–2004q2	1995–2008
Hungary		2005–2008
Ireland	2003q3–2009q2	1995–2008
Italy	2003q3–2009q2	1995–2008
Japan	2003q3–2009q2	1995–2008
Luxemburg		2005–2008
Netherlands	2003q3–2009q2	1995–2008
New Zealand	2003q3–2009q2	1997–2008
Norway	2003q3–2009q2	1995–2008
Poland	2007q3–2009q2	2006–2008
Portugal		1995–2008
Slovakia		2008
Spain		1995–2008
Sweden	2003q3–2009q2	1995–2008
Switzerland		1995–2008
United Kingdom	2003q3–2009q2	1995–2008
USA	2003q3–2009q2	1995–2008

Capacity utilisation

Our second variable – capacity utilisation as reflected in business tendency surveys – nowadays belongs to the core items of the EU harmonised BTS, but most OECD countries as well as a few others are conducting surveys including similar questions. One is quantitative and asks respondents to assess the level of capacity utilisation of their firm in per cent; the other is qualitative and asks for an assessment of the present rate of capacity utilisation on a three point scale (low, normal, high).

The BTS data on capacity utilisation were obtained from various sources. For EU-members, the Economics and Financial Affairs division of the European Commission publishes quarterly BTS data on capacity utilisation in per cent as well as the balance indicator for the assessment of the present level on capacity utilisation, both as original values and seasonally adjusted series.[25] The time series generally start in 1985q1, or with the beginning of EU membership, if the latter took effect later. Data for non EU-members countries as well as for some member countries going back beyond 1985q1 were taken from the OECD Main Economic Indicators online data base.[26] A limited amount of data was taken from other sources than those mentioned above, as more variables or data points were available than from the OECD and ECFIN databases:

- National Bank of Belgium: not seasonally adjusted rate of capacity utilisation in the Belgian manufacturing industry;
- Board of Governors of the Federal Reserve System, US industrial capacity utilisation, seasonally adjusted, quarterly average of monthly data;[27]
- KOF Swiss Economic Institute: judgement of capacity utilisation in the Swiss manufacturing industry, original and seasonally adjusted series;
- METI Ministry of Economics, Trade and Industry, operating ratio in the Japanese industry, index (set to 100 for 2005) seasonally adjusted and unadjusted series, quarterly average of monthly data;
- NZ Institute of Economic Research (NZIER), Wellington, Quarterly Survey of Business Opinion: not seasonally adjusted rate of capacity utilisation in the NZ economy (manufacturers and builders), quarterly average of monthly data;[28]
- Statistics Canada: seasonally adjusted rate of capacity utilisation in the Canadian manufacturing industry.

The timeliness of the data is remarkable, as they are already available at the end of the first month of quarter that they are referring to. The numbers for 2009q4 e.g. were available in the end of October 2009.

The data that we were able to find are shown in table 3.

[25] See http://ec.europa.eu/economy_finance/db_indicators/surveys9185_en.htm.

[26] See „Leading Indicators and Tendency Surveys“, http://www.oecd.org.

[27] The Federal Reserve's capacity utilisation data are estimates that are not only derived from BTS, hence there are more revisions than those due to seasonal filter.

[28] We herewith thankfully acknowledge.

Table 3: BTS data on capacity utilisation, as reported, 1970q1–2009q4

	BTS data on capacity utilisation			
Country	*CapU* in %, original data	*CapU* in %, seasonally adjusted	*CapU*, judgement, balance, original data	*CapU*, judgement, balance, seasonally adjusted
Australia	1996q1–2007q2	2007q1–2009q2	1998q2–2009q2	
Austria		1996q1–2009q4		1996q1–2009q4
Belgium	1980q1–2009q4	1978q2–2009q4	1985q1–2009q4	1985q1–2009q4
Canada	1970q1–2008q4	2006q4–2009q1	1980q1–2007q3	1980q1–2007q3
Czech Republic		1991q4–2009q4		1995q1–2009q4
Denmark	1987q1–2009q4	1987q1–2009q4	1985q1–2009q4	1985q1–2009q4
Finland		1993q1–2009q4		1985q1–2009q4
France	1985q1–2009q4	1976q1–2009q4	1985q1–2009q4	1985q1–2009q4
Germany	1985q1–2009q4	1970q1–2009q4	1985q1–2009q4	1985q1–2009q4
Greece		1985q1–2009q4		1985q1–2009q4
Hungary	1996q1–2009q4	1996q1–2009q4	1999q2–2009q4	1999q2–2009q4
Ireland	1985q1–2008q2	1985q1–2008q2	1985q1–2008q2	1985q1–2008q2
Italy		1970q1–2009q4		1985q1–2009q4
Japan	1978q1–2009q1♣	1978q1–2009q1♣	1974q2–2009q2	1974q2–2009q2
Luxemburg	1985q1–2009q4	1985q1–2009q4	1985q1–2009q4	1985q1–2009q4
Netherlands	1985q1–2009q4	1971q4–2009q4	1985q1–2009q4	1985q1–2009q4
New Zealand	1970q1–2009q2			
Norway	1987q1–2009q1	1987q1–2009q1		
Poland	1992q2–2009q4	1992q2–2009q4	1992q2–2009q4	1992q2–2009q4
Portugal	1987q1–2009q4	1977q1–2009q4	1987q1–2009q4	1987q1–2009q4
Slovakia		1993q4–2009q4	1998q2–2007q3	1993q4–2009q4
Spain	1987q2–2009q4	1970q1–2009q4	1987q2–2009q4	1987q2–2009q4
Sweden	1996q1–2009q4	1996q1–2009q4	1996q2–2009q4	1996q2–2009q4
Switzerland	1970q1–2009q4	1970q1–2009q4	1970q1–2009q4	1970q1–2009q4
United Kingdom	1985q1–2009q4	1985q1–2009q4	1985q1–2009q4	1985q1–2009q4
USA	1985q1–2007q2	1972q1–2009q2		

♣ Index series, set equal to 100 for 2005

As can be seen from table 3, the coverage is better for the quantitative question (*CapU*) than for the qualitative (*CapU_bal*). As the quantitative indicator is also easier to interpret[29] and conceptionally closer related to the numerical estimates of the output gap, we choose $CapU_{it}$ as our preferred indicator, i.e. capacity utilisation from survey data on in per cent, where *i* denotes the country that the data refer to and *t* the quarter.

When we could refer to BTS data on capacity utilisation that did not have our target format *CapU* in per cent without seasonally adjustment, we devoted some effort to derive reasonable estimates of the target series. In particular, we performed the following steps:

- For Australia, the *CapU* series ends at 2007q2. However, the National Australia Bank's (NAB) quarterly economic surveys, published on its website, list seasonally adjusted capacity utilisation until 2009q2. As the seasonally unadjusted data for Australia do not exhibit significant seasonality, seasonal adjustment does not significantly alter the seasonal pattern ($F = 0.055$). Accordingly, we fill the eight missing quarters of *CapU* with a spline to the seasonally adjusted NAB series at the 2007q2 quarter.
- For Canada, *CapU* is reported until 2008q4, whereas seasonally adjusted data are available until 2001q1. As the seasonally unadjusted data for Canada do not exhibit significant seasonality as well ($F = 0.031$), we fill the first quarter of 2009 referring to the seasonally adjusted series with a spline performed at the 2008q4 quarter.
- For Hungary, there are three missing values in a row within the target series (1996q2–1996q4). We fill these by linear interpolation.
- For Japan, our sources show *CapU* as an index series only (set to 100 for 2005). We adjust this series by shifting it down so that its maximum (113.6 for the index) equals the average of the maxima of our target series across all other countries (86.9%).
- For the United States, *CapU* is reported until 2007q2, whereas seasonally adjusted data are available until 2009q2. As the seasonally unadjusted data for the USA also do not show significant seasonality ($F = 0.006$), we fill the eight missing quarters of $CapU_t$ with a spline to the seasonally adjusted series performed at the 2007q2 quarter.

For two countries where the seasonally adjusted series *CapU_sa* have missing values or do not reflect capacity utilisation as a percentage, we could compute reasonable estimates:

- For Japan, our sources show *CapU_sa* as an index series only (set to 100 for 2005). We adjust this series by shifting it down so that its maximum (114.4 for the index) equals the average of the maxima of the *CapU_sa* series across all other countries (86.6%).
- For the Netherlands, the seasonally adjusted series reported by the OECD goes back to 1971q4, whereas the unadjusted data start in 1985q1 only. There are, however, six missing values within the *CapU_sa* series (1972q3, 1973q3, 1974q3, 1975q3, 1976q3, and 1977q3). We fill these by linear interpolation.

For our purposes, an important characteristic of the BTS data is that they are usually not revised, but final as soon as a survey is completed.[30] Yet, even if the data are not revised, we have to be aware of an end point problems due to seasonal filtering, as this may lead to revisions as new data points are added. While seasonal adjustment can in principle be performed with constant seasonal factors, thus avoiding subsequent revisions, this is the exception rather

---

[29] The interpretation of qualitative question is plagued by the fact that „normal“ capacity utilisation is not well defined; see above (section 1).

[30] Revisions may occur when early information is taken from a sub-sample of respondents while the survey is not yet completed. Completion from there on, however, is usually only a question of a few days, so that published BTS data is usually final.

than the rule. Most standard procedures to eliminate season rely on recursive or rolling estimates of seasonal factors, and noise is also addressed, e.g. via outlier detection or moving averages. Accordingly, what is called „seasonal adjustment" to some degree amounts to outright low pass filtering, where the end point problem of symmetric filters can lead to massive subsequent revisions. Our analysis would hence ideally initially refer to unfiltered data. For a number of countries, this is feasible. Unfortunately, some countries publish seasonally adjusted data only.

For Austria, the Czech Republic; Finland, Greece, Italy and Slovakia, capacity utilisation is available only as seasonally adjusted time series. If we want to resort only to BTS data that would have been available in real time, the sample reduces from 26 to 20 countries.

The unadjusted data are not affected by the end point problem. On the other hand, as some survey data are heavily affected by season and noise, it may be necessary to eliminate seasonal patterns and to increase the signal-to-noise ratio before they reveal useful information on the cyclical position of an economy. In other words, filtering may indeed be mandatory, though – for our purposes – we need to make sure not to refer to information that would not have been available in real time. Thus, for all countries with original data, we performed recursive seasonal adjustments. As the data are by construction stationary, we referred to the simple additive seasonal components model. In addition, we also computed a simple symmetrical low pass filtered series of *CapU*, where the trend-cycle is the nine quarters moving average of the seasonally adjusted series. Accordingly, for those countries that report original series *CapU*, we constructed four types of filtered series:

1. An ex post seasonally adjusted series *CapU_sa_xp*, computed from all available data points of *CapU*,
2. A real time seasonally adjusted series *CapU_sa_rt*, computed by recursive filtering of *CapU*,
3. An ex post low pass filtered series *CapU_lp_xp*, computed from all available data points of *CapU*,
4. A real time low pass filtered series *CapU_lp_rt*, computed by recursive filtering of *CapU*.

The resulting series and their coverage across countries and time are shown in table 4.

Table 4: Data on CapU after estimation of missing values and filtering, 1970q1–2009q4

Country	BTS data on capacity utilisation			
	CapU	*CapU_sa*	*CapU_sa_xp* *CapU_lp_xp*	*CapU_sa_rt* *CapU_lp_rt*
Australia	1996q1–2009q2	2007q1–2009q2	1996q1–2009q2	1999q4–2009q2
Austria		1996q1–2009q4		
Belgium	1980q1–2009q4	1978q2–2009q4	1980q1–2009q4	1983q4–2009q4
Canada	1970q1–2009q1	2007q3–2009q1	1970q1–2009q1	1980q1–2009q1
Czech Republic		1991q4–2009q4		
Denmark	1987q1–2009q4	1987q1–2009q4	1987q1–2009q4	1990q4–2009q4
Finland		1993q1–2009q4		
France	1985q1–2009q4	1976q1–2009q4	1985q1–2009q4	1988q4–2009q4
Germany	1985q1–2009q4	1970q1–2009q4	1985q1–2009q4	1988q4–2009q4
Greece		1985q1–2009q4		
Hungary	1996q1–2009q4	1996q1–2009q4	1996q1–2009q4	1999q4–2009q4
Ireland	1985q1–2008q2	1985q1–2008q2	1985q1–2008q2	1988q4–2008q4
Italy		1970q1–2009q4		
Japan	1978q1–2009q1	1978q1–2009q1	1978q1–2009q1	1981q1–2009q1
Luxemburg	1985q1–2009q4	1985q1–2009q4	1985q1–2009q4	1988q4–2009q4
Netherlands	1985q1–2009q4	1971q4–2009q4	1985q1–2009q4	1988q4–2009q4
New Zealand	1970q1–2009q2		1970q1–2009q2	1980q1–2009q2
Norway	1987q1–2009q1	1987q1–2009q1	1987q1–2009q1	1990q4–2009q1
Poland	1992q2–2009q4	1992q2–2009q4	1992q2–2009q4	1996q1–2009q4
Portugal	1987q1–2009q4	1977q1–2009q4	1987q1–2009q4	1990q4–2009q4
Slovakia		1993q4–2009q4		
Spain	1987q2–2009q4	1970q1–2009q4	1987q2–2009q4	1991q1–2009q4
Sweden	1996q1–2009q4	1996q1–2009q4	1996q1–2009q4	1999q4–2009q4
Switzerland	1970q1–2009q4	1970q1–2009q4	1970q1–2009q4	1980q1–2009q4
United Kingdom	1985q1–2009q4	1985q1–2009q4	1985q1–2009q4	1988q4–2009q4
United States	1985q1–2009q2	1972q1–2009q2	1985q1–2009q2	1988q4–2009q2

3 Empirical Analyses

The empirical analyses aim at showing whether the OECD output gap estimates could have been improved in real time resorting to then available BTS data on capacity utilisation. By improvement, we mean that the modified estimates are closer to later releases of the same series. In other words, since the true reference series is unobservable, we have to make the assumption that the output gap estimates are getting more reliable as time passes and more information is getting available.

Formally, this means to resort to show that early (provisional) estimates P of the output gap are informationally inefficient in the sense that BTS data on capacity utilisation C available in real time can help to produce estimates that are significantly closer to later ("final") releases F of the output gap. Alternatively, informational efficiency implies that revisions $R = F - P$ are not predictable, so that the regression

$$R = \alpha_0 + \alpha_1 C$$

should result in parameters α_0 und α_1 that are not (significantly) different from zero. As

$$R = F - P \Leftrightarrow F = P + R,$$

we cannot reject the null hypothesis "informational efficiency" when the regression of later releases on earlier releases of the output gap and BTS data that were available no later than the other right hand variable

$$F = \beta_0 + \beta_1 P + \beta_2 C$$

yields parameters β_0 und β_2 that are not (significantly) different from zero, and β_1 not (significantly) different from unity.

Throughout the analyses, we shall carefully simulate real time conditions, making sure that the information set represented by C is referred to only in as much as would have been possible in real time. Accordingly, we have to distinguish the reference point in time or period r from the point in time or period s when an estimation is performed, so that all output gap releases will be denoted as $OG^s{}_r$. The equation to be estimated is hence

$$OG^{t+u}{}_t = \beta_0 + \beta_1\ OG^{t+v}{}_t + \beta_2\ C^{t+w}{}_t,$$

where $v \geq 0$, $v \geq w$ and $u > v$. When $v = 0$, we are dealing with a real time estimation, and $v > 0$ implies a "postcast", which is performed after the end of the reference period. The inequality $u > v$ specifies that the later (revised) values that serve as the reference series are published later that the right hand side variables, which means that they can refer to a larger set of information.

A crucial issue is to determine F, i.e. the "final" value of the output gap that we need as a benchmark against which to assess earlier releases. This is not a trivial task, as in practice, output gap estimates tend to be revised for extended periods without settling down or converging to a readily identifiable final value. As examples, we plot the 2003q1, 2004q1, 2005q1, 2006q1 and 2007q1quarterly output gap for Germany through time, starting with the (close to) real time release in June 2004 and then successively up to the latest release in June 2009 (figure 3) as well as the annual German output gap for 1995, 2000 and 2005, starting with the real time release in December 1995 and then successively up to the latest release in June 2009 (figure 4). Obviously, after five years, we still cannot be sure about the magnitude the quarterly output gaps, and the same holds for the yearly output gaps after some 15 years. Some of the series do not at all appear to settle down or converge to any stationary value.[31]

[31] See Graff (2004) for plenty of similar evidence from New Zealand.

Figure 3: Quarterly OECD output gaps for Germany, 2003q1, 2004q1, 2005q1, 2006q1 and 2007q1, releases June 2004 to June 2009

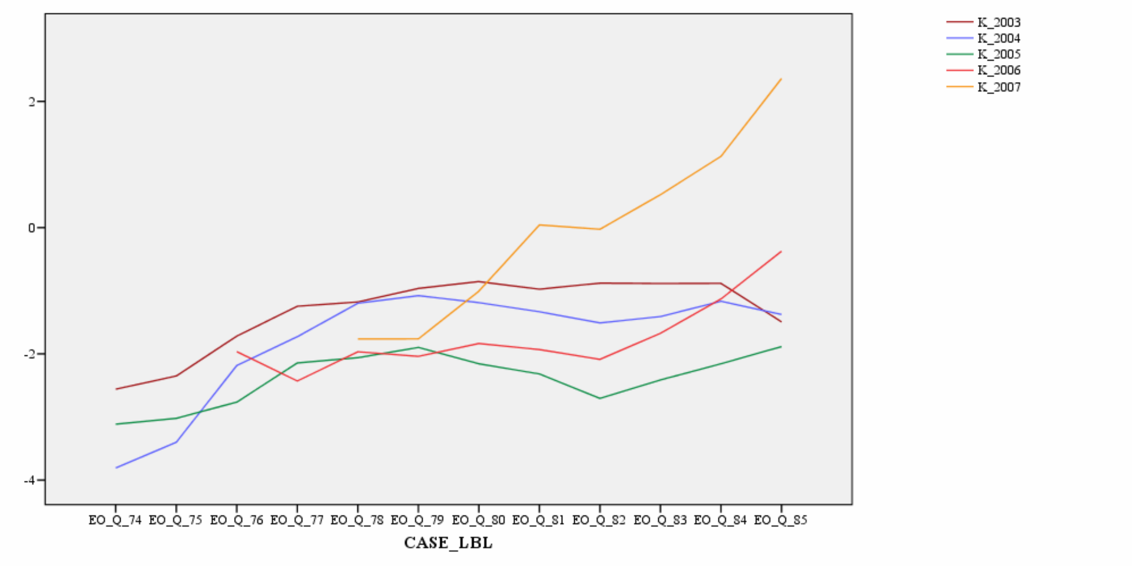
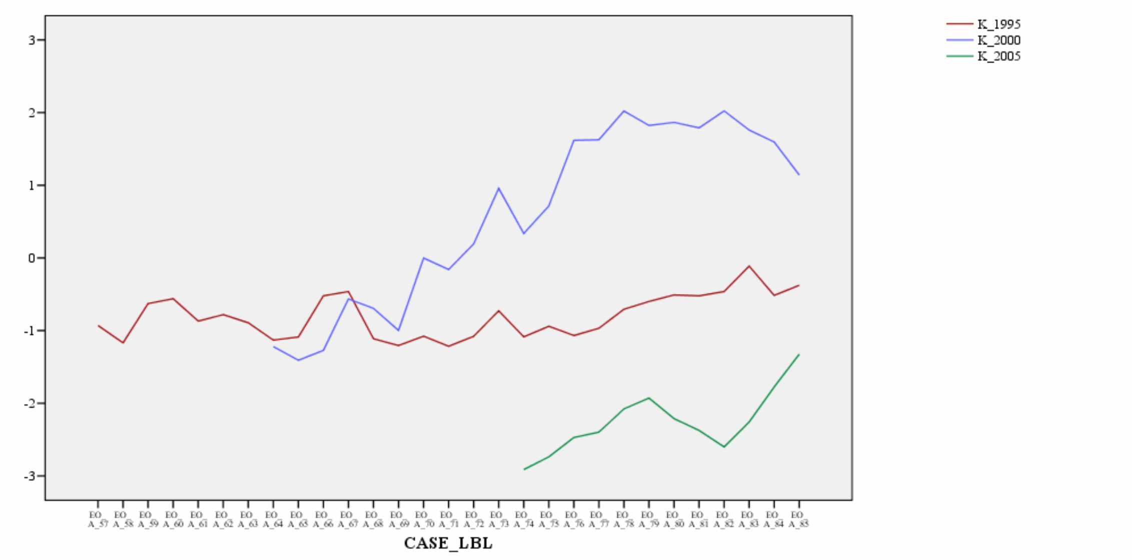


Figure 4: Annual OECD output gaps for Germany, 1995, 2000, and 2005, releases December 1995 to June 2009



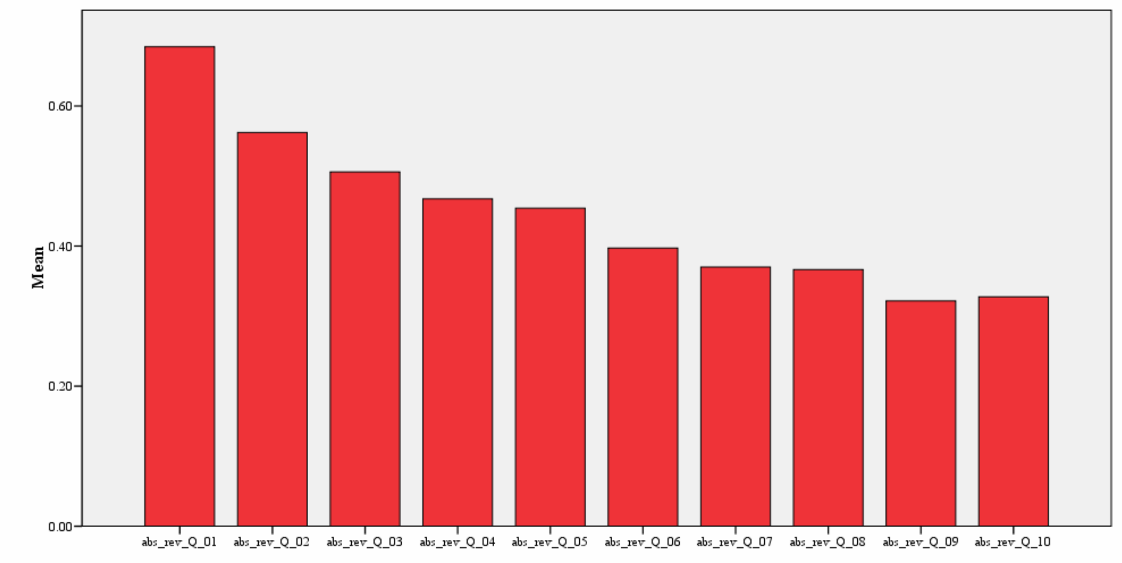
As the choice of Germany is somewhat arbitrary (although inspection of the data confirms many instances where revisions as happening as permanently – if not more pronounced – as in the examples shown above), we now compute some descriptive statistics of the revision process, both for the quarterly and the annual OECD output gaps, relating to the whole sample. First, let us look at the subsequent revisions of the quarterly data. Table 5 summarises the average revisions of all quarterly output gaps starting with the Economic Outlook No. 74 (December 2003) data base. *Q_rev_no_01* denotes the first revision of the real time release (after two quarters), *Q_rev_no_02* denotes the following revision of the second release (after another two quarters, i.e. four quarters after real time), and so forth.

As the revisions are not cumulated, the evolution of the mean of the absolute revisions should indicate when the revision process on average tends to ebb off. With the data at hand, however, though the revisions decline from close to 0.7 percentage points in magnitude to roughly 3.3 percentage points, they do not appear to converge to zero. Rather, 3.3 percentage points seem to mark a bottom line for a continuing revisions. The bar chart of the tabled values (figure 5) visualises this finding.

Table 5: Average revisions of quarterly OECD output gaps, 2004q2 to 2009q2

	Minimum	Maximum	Mean	Mean absolute	N
Q_rev_no_01	–3.60	7.14	0.14	0.68	338
Q_rev_no_02	–5.63	4.29	0.24	0.56	338
Q_rev_no_03	–3.27	2.78	0.21	0.51	338
Q_rev_no_04	–4.06	2.44	0.18	0.47	338
Q_rev_no_05	–4.85	2.66	0.17	0.45	338
Q_rev_no_06	–2.94	2.57	0.15	0.40	338
Q_rev_no_07	–3.45	2.38	0.09	0.37	338
Q_rev_no_08	–3.23	2.15	0.09	0.37	338
Q_rev_no_09	–2.10	1.85	0.06	0.32	338
Q_rev_no_10	–1.85	1.45	0.05	0.33	338

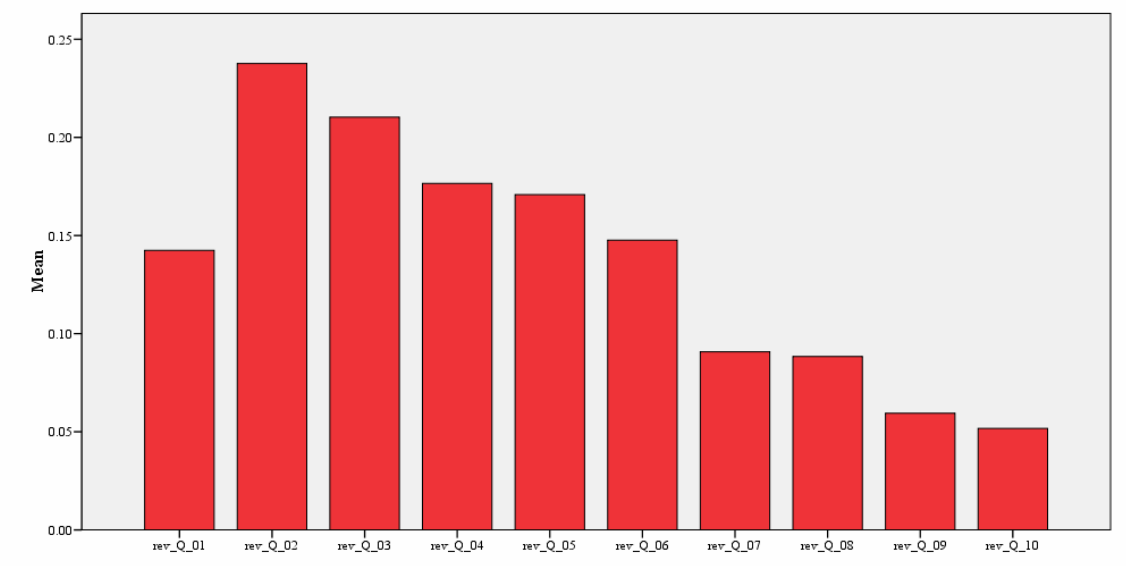
Figure 5: Average absolute revisions of quarterly OECD output gaps, 2004q2 to 2009q2



Looking at the average revisions rather than at the average absolute revisions (figure 6), we see that during the given period, all revisions have been upwards. As the sample is short, this may reflect a tendency to underestimate the strength of the upswing that dominated those years.

As more data will be getting available, we should expect to encounter a fair number of downward revisions; otherwise, we would have to conclude that there is some systematic downward bias in the OECD releases that reduces as time passes, which is not a reasonable assumption.

Figure 6: Average revisions of quarterly OECD output gaps, 2004q2 to 2009q2



Turning to the annual data allows us to trace the revision process back to 1995. Table 6 summarises the average revisions of all annual output gaps starting with the Economic Outlook No. 57 (December 1997) data base. *A_rev_no_01* denotes the first revision of the real time release (after two quarters), *A_rev_no_02* denotes the following revision of the second release (after another two quarters, i.e. four quarters after real time), and so forth.

As with the quarterly output gap estimates, the evolution of the mean of the absolute revisions should indicate when the revision process on average tends to ebb off. In particular, we are now dealing with some 15 rather than 5 years, and we thus might expect to see that means of the absolute converging towards zero. Yet, this is clearly not the case. During the first few years of the revision process, the absolute revisions decline from about 0.5 percentage points to about 0.4 percentage points, but then they keep lingering at that level. Figure 7 visualises this finding.

The average revisions to the annual output gaps (table 6 column 4, visualised in figure 8) show a cyclical pattern, which is what we would expect if observes of the business cycle are systematically under-and overestimating the position of the economy in different phases of the cycle. As with the longer historical time covered, we now also encounter a fair number of downward revisions, we are happy to conclude that there is no systematic downward bias in the OECD releases that reduces as time passes. Nevertheless, we observe some negative autocorrelation, particularly in the later years. Suffice to say here that we assume that the reason for this lies in the fact that the difference in the periodicity of the annual reference series and the semi-annual release pattern implies that the second release during a given reference year is based on a larger information set than the preceding one. This, however, certainly merits further investigation.

Table 6: Average revisions of annual OECD output gaps, 1995q4 to 2009q2

	Minimum	Maximum	Mean	Mean absolute	N
A_rev_no_01	–2.90	3.06	0.15	0.47	307
A_rev_no_02	–2.53	3.22	0.16	0.48	303
A_rev_no_03	–2.62	3.67	0.15	0.44	307
A_rev_no_04	–4.15	3.31	0.10	0.49	303
A_rev_no_05	–2.21	2.51	0.12	0.40	307
A_rev_no_06	–3.62	2.90	0.05	0.46	303
A_rev_no_07	–1.77	2.25	0.07	0.36	307
A_rev_no_08	–3.89	2.21	0.02	0.42	303
A_rev_no_09	–2.23	2.30	0.04	0.31	307
A_rev_no_10	–3.53	1.89	–0.02	0.38	303
A_rev_no_11	–2.39	2.50	0.03	0.31	307
A_rev_no_12	–3.12	2.13	–0.03	0.38	303
A_rev_no_13	–2.86	3.85	–0.00	0.34	307
A_rev_no_14	–2.83	2.58	–0.04	0.35	303
A_rev_no_15	–3.01	2.86	–0.03	0.35	307
A_rev_no_16	–2.48	2.88	–0.03	0.35	303
A_rev_no_17	–2.77	2.70	–0.03	0.33	307
A_rev_no_18	–2.24	2.09	–0.04	0.34	302
A_rev_no_19	–2.33	3.02	0.04	0.31	306
A_rev_no_20	–2.36	2.46	–0.04	0.34	301
A_rev_no_21	–2.41	3.01	0.06	0.33	305
A_rev_no_22	–2.65	3.98	–0.03	0.35	301
A_rev_no_23	–2.79	2.62	0.07	0.33	304
A_rev_no_24	–3.17	3.34	–0.03	0.35	301
A_rev_no_25	–2.85	2.73	0.09	0.34	303
A_rev_no_26	–3.64	2.76	–0.03	0.37	300
A_rev_no_27	–4.41	5.20	0.08	0.35	301
A_rev_no_28	–4.54	2.58	–0.06	0.40	298
A_rev_no_29	–6.55	5.61	0.07	0.36	298
A_rev_no_30	–5.05	1.98	–0.09	0.41	295

Figure 7: Average absolute revisions of annual OECD output gaps, 1995q2 to 2009q2

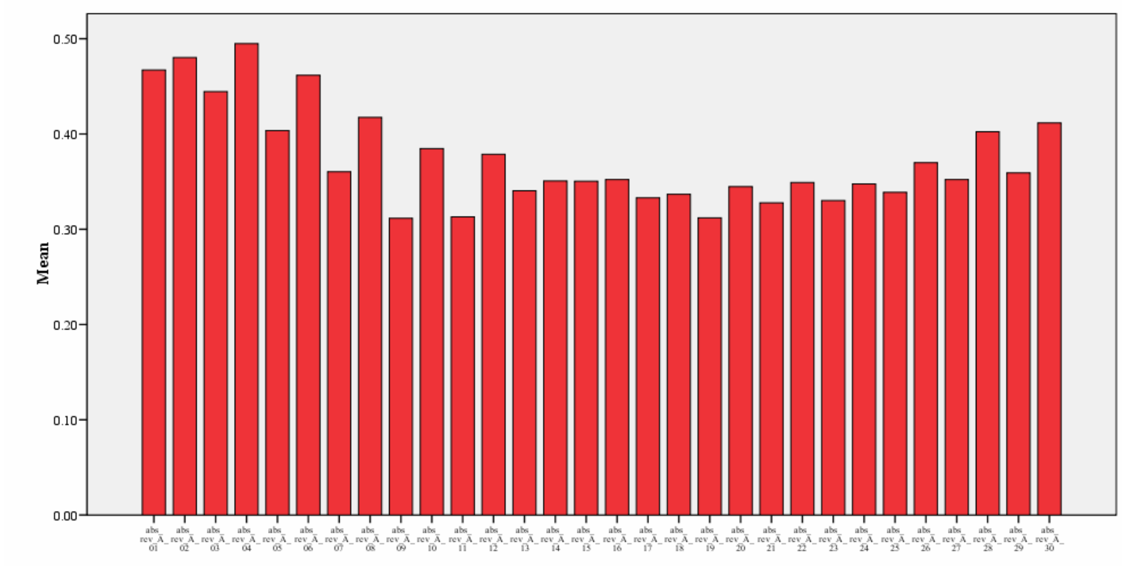
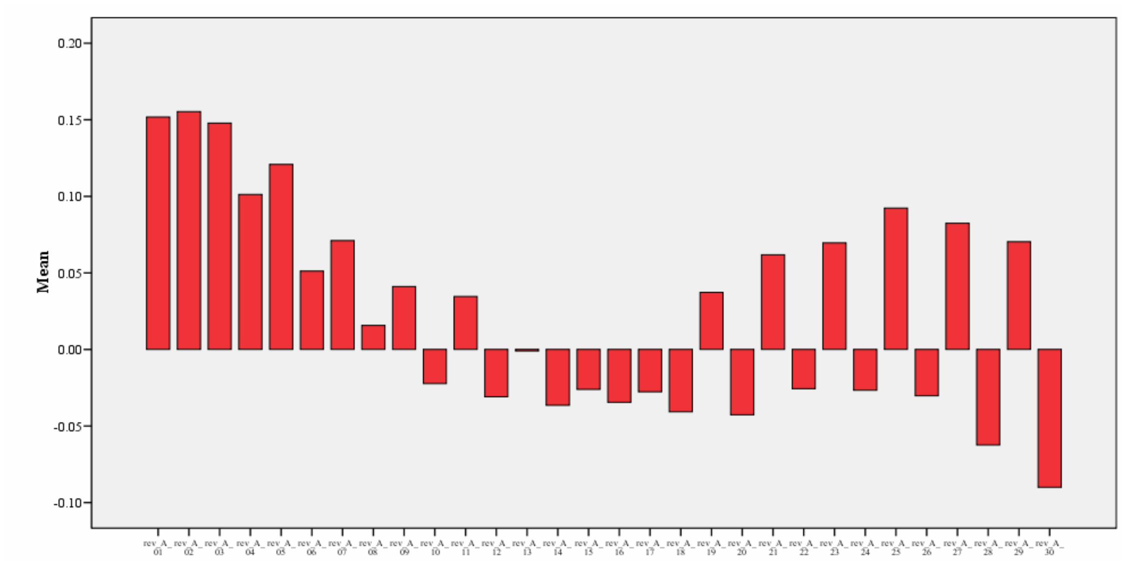


Figure 8: Average revisions of annual OECD output gaps, 1995q4 to 2009q2



It now may appear as a natural choice to resort to the last available release of g (OG^l)as the benchmark. However, $OG^l_t \neq OG^{t+u}_t$ unless $l = t + u$, which is true for one point of the reference series only. The time series OG^l_t is characterised by decreasing distance to the reference period r as we move to the right, eventually collapsing to zero. Contrary to this, in OG^{t+u}_t, the distance to r is held constant, so that this series imposes a constant revision period for the reference series, which is econometrically sounder. However, as we increase u to allow for a longer revision period, we are loosing observations at the right margin of the OG^{t+u}_t series. Given the shortness of our empirical real time series OG^{t+v}_t, we then face a trade-off between the length of the revision period and the number of observations that we can refer to in order to compare earlier releases against a benchmark.

Given this, we are not going to fix a constant integer for u; instead, we shall address the trade-off and vary u from 2 to up to 20 quarters for the quarterly data and to up to 32 quarters for the annual data, e.g. from the first release two quarters after real time to the tenth (sixteenth) revision five (eight) years after the reference period.

We now analyse the quarterly data. After this, we shall refer to the yearly data, which are *per se* less informative than the quarterly data, but which are available for more countries and for some additional earlier years.

Quarterly data

First, let us look at the correlation matrix of the real time and the subsequent quarterly output gap releases (table 7), where v, according to the notation introduced above, denotes the time lag between reference period and release.

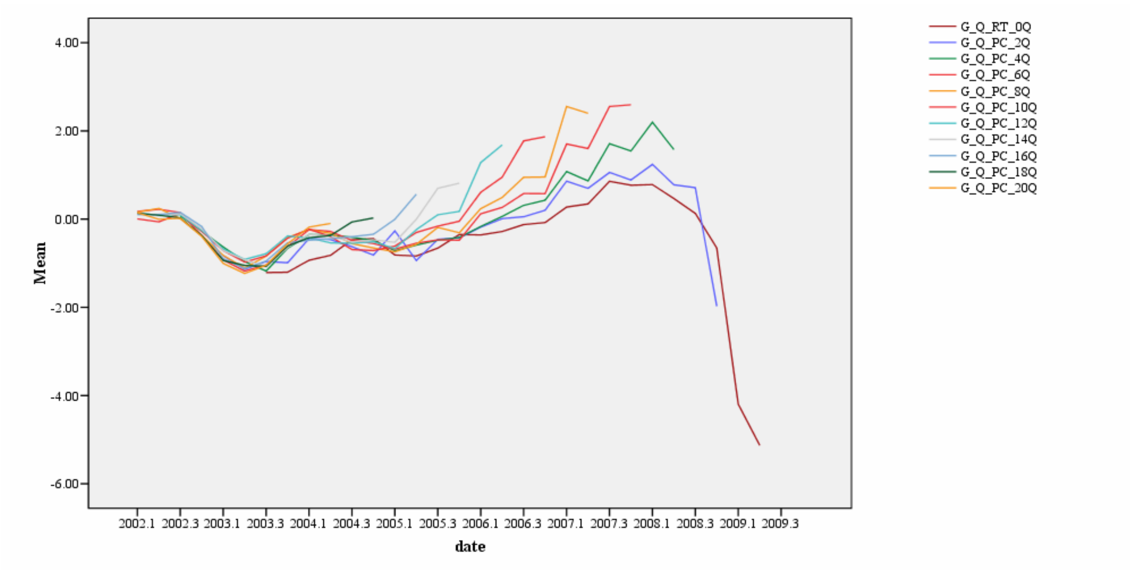
Table 7: Correlation matrix OECD Output gap, n = 56

	$v = 2$	$v = 4$	$v = 6$	$v = 8$	$v = 10$	$v = 12$	$v = 14$	$v = 16$	$v = 18$	$v = 20$
$v = 0$	0.75	0.63	0.57	0.65	0.66	0.72	0.70	0.70	0.64	0.67
$v = 2$		0.85	0.87	0.82	0.80	0.77	0.78	0.74	0.76	0.78
$v = 4$			0.95	0.92	0.89	0.87	0.86	0.80	0.83	0.83
$v = 6$				0.93	0.90	0.87	0.87	0.77	0.81	0.81
$v = 8$					0.98	0.95	0.93	0.88	0.85	0.83
$v = 10$						0.97	0.93	0.89	0.88	0.82
$v = 12$							0.95	0.92	0.91	0.84
$v = 14$								0.95	0.94	0.91
$v = 16$									0.95	0.90
$v = 18$										0.92

The table shows that the correlation generally decreases with the distance between releases, as can be expected. Moreover, the real time release ($v = 0$) and the first revision ($v = 2$) show lower overall correlation with subsequent releases than releases a year or longer after the reference period, indicating relatively pronounced revisions during the first year, which may partly be due to the fact that the real time estimation is performed before the national statistical offices publish their first provisional quarterly GDP estimates, thus requiring the OECD to refer to their own GDP estimates, adding difficulty to an already tricky task. Nevertheless, the correlation is far from perfect for pairs with $v \geq 4$, i.e. when "official" GDP data should usually be available. Finally, notice that the correlation has been computed for those observations only, where all pairs are available. This makes the correlation comparable, however at the price of reducing the sample to $n = 56$. Yet, an inspection of the correlation matrix where each cell refers to the pairwise available observations (not reproduced), in qualitative term confirms these findings.

Figure 9 plots the 2003q4 to 2009q2 vintages of the average of the OECD quarterly output gaps and hence visualises the revision process quantified above. Apart from showing the magnitude of revisions, one can readily detect the dominant tendency to revise upwards as well as the dramatic drop of the output gap caused by the 2008/2009 recession.

Figure 9: 2003q4 to 2009q2 vintages of average OECD quarterly output gaps



Next, let us inspect the correlation between the quarterly output gap (last vintage, published in June 2009) and the survey data on capacity utilisation.

Table 6: Correlation OECD output gap(last vintage) and capacity utilisation, quarterly data

	CapU	*CapU_sa*	*CapU_sa_xp*	*CapU_sa_rt*	*CapU_lp_ex*	*CapU_lp_rt*
Bivariate r	0.33	0.34	0.34	0.35	0.33	0.34
Partial β_1	0.33	0.35	0.33	0.34	0.38	0.38
N	1319	1446	1316	1125	1316	1125

The table shows that the last vintage of the OECD output gap (June 2009) is significantly correlated with the survey data on capacity utilisation. Moreover, though the correlation is significantly positive, it is only modest, and filtering the BTS data does not change this finding.

Could differences in levels be responsible for the relatively low correlation? If yes, the inclusion of country fixed effects β_i should increase the (partial) correlation β_1 in the following regression:

$$OG_{it} = \beta_i + \beta_1 C_{it}.$$

The country fixed effects are indeed jointly highly significant, which implies that we have to stick to them throughout our analyses; nevertheless the partial correlation β_1 is hardly higher than the bivariate correlation r, unless we refer to the low-pass filtered *CapU* series. Accordingly, the improvement of the signal-to noise ratio is reflected by the increase in the correlation, once we control for country specific level effects. Notice further that our low-pass filter is asymmetric, so that *CapU_lp_xp* is affected by the end point problem. However, if we refer to the recursively filtered series *CapU_lp_rt*, which consists only of values that would have been available in real time, we still arrive at a correlation that appears better than that with the unfiltered or only seasonally adjusted series. The improvement of the signal to noise ratio hence dominates the phase shift that is caused by the fact that recursive filtering renders the filter asymmetric.

After these preliminary steps, we are now ready to run our information inefficiency regressions for the quarterly frequency

$$OG^{t+u}{}_t = \beta_i + \beta_1\ OG^t{}_t + \beta_2\ C^t{}_t,$$

where $20 \geq u > 2$. When $\beta_2 \neq 0$, we can reject the null hypothesis of informational efficiency of the real time output gap with respect to BTS data on capacity utilisation that one would have been able to refer to in real time. As we face a trade-off between an expected increase in the reliability of the left hand variable – e.g. the benchmark – as u is increased, and a reduction in the degrees of freedom, which makes it more difficult to reject the null, we step by step increase u from 2 to 20 quarters. For space reasons, we report a selection of our findings ($u = 2$, $u = 12$ and $u = 20$) which allows to infer the structural relationship. The results are presented in table 7.

Table 7: Informational inefficiency regressions, with country fixed effects, $C = C^t{}_t$

u	$u = 2$ quarters			$u = 12$ quarters			$u = 20$ quarters		
Filter	*LP_rt*	*Sa_rt*	none	*LP_rt*	*Sa_rt*	none	*LP_rt*	*Sa_rt*	none
β_1	0.72	0.73	0.75	0.54	0.55	0.56	0.18	0.23	0.22
$p(\beta_1)$♣	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**	0.17	0.11	0.12
β_2	0.15	0.14	0.11	0.16	0.14	0.13	0.25	0.18	0.14
$p(\beta_2)$♣	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**	**≤ 0.05**	**≤ 0.05**	**≤ 0.10**
n	276	276	276	144	144	144	48	48	48
R^2	0.83	0.83	0.83	0.80	0.80	0.80	0.88	0.88	0.88

♣ One-tailed significance tests

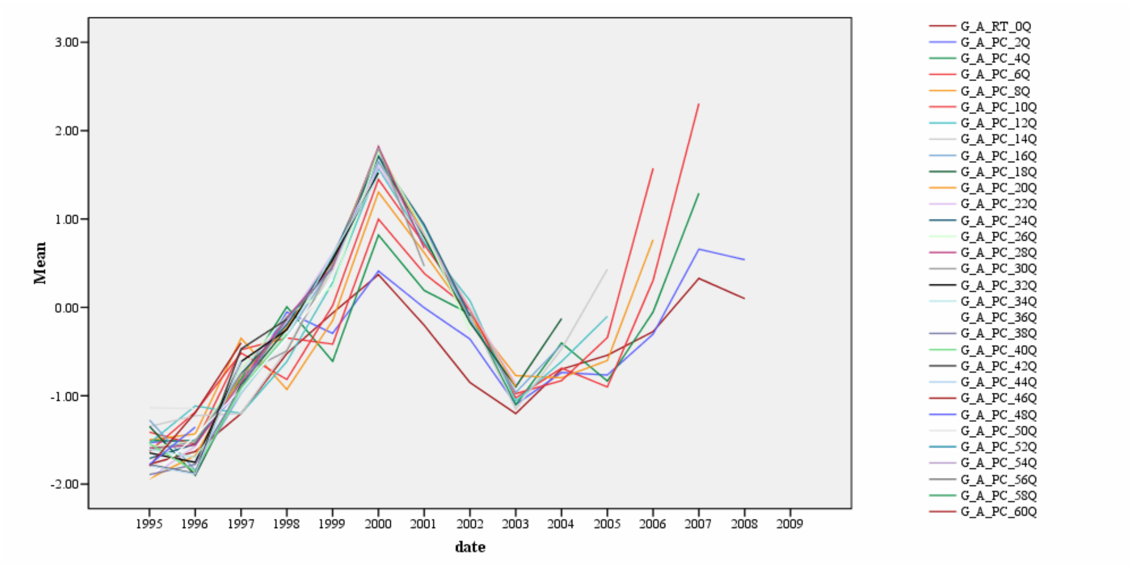
The table shows that the null hypothesis of informational efficiency of the OECD output gap released in real time can clearly be rejected. For the first revision ($u = 2$) as well as revisions up to some four years, rejection is at $p \leq 1\%$. If we take the OECD estimate five years after the reference period as a benchmark (which is as far as we can go with the data at hand), the significance fails to pass the 1% level, however, the rejection of the null remains significant when we allow for a 5% or 10%, which still is impressive, given that with increasing u the sample size is markedly going down. Moreover, at $u = 20$, in the informational inefficiency regression, the real time release of the output gap fails to be a significant predictor of the later release even at the 10% level. As this regression can be interpreted as an encompassing test of non-nested models, this implies the astonishing conclusion that after five years, the original estimates of the output gap have less predictive power for the current output gap estimates than the BTS data collected five years ago.

Annual data

Finally, we look at the annual data, which increases the sample size from 14 to 25 countries and lets us go back in time until 1995 rather than 2003.

Figure 10 plots the 1995q2 to 2009q2 vintages of the average of the OECD annual output gaps. Apart from showing the magnitude of revisions – especially around the upper turning point around 2000 –, one can again detect the dominant tendency to revise upwards. As the annual real time data and in 2008, the 2008/2009 recession is not as apparent as in the quarterly data shown above.

Figure 10: 1995q2 to 2009q2 vintages of average OECD annual output gaps



The BTS data are quarterly and hence have to be aggregated into annual frequency. We consider four alternatives:

- The last quarters of year *t* ($CapU^{tq4}{}_t$)
- The mean of the last two quarters of year *t* ($CapU^{tq3-4}{}_t$)
- The mean of the last two quarters of year *t* ($CapU^{tq2-4}{}_t$)
- The mean of the four quarters of year *t* ($CapU^{tq1-4}{}_t$)

Next, let us inspect the correlation between the annual output gap (last vintage, published in June 2009) and the annualised survey data on capacity utilisation. Table 8 shows that the last vintage of the OECD output gap is significantly correlated with the annualised BTS data. Moreover, though the correlation is most pronounced when we refer to the annual average to annualise the data. The country fixed effects are again jointly highly significant, but the partial correlations are practically the same as the bivariate ones. As the annual average eliminates any seasonality, the information inefficiency regressions for the annual frequency will refer to original and recursively low-pass filtered BTS series only.

Table 8: Correlation OECD output gap and capacity utilisation, annual data, n = 401

	$CapU^{tq4}$	$CapU^{tq3-4}$	$CapU^{tq2-4}$	$CapU^{tq1-4}$	$CapU^{tq1-4}sa\_rt$	$CapU^{tq1-4}lp\_rt$
Bivariate r	0.36	0.38	0.34	0.40	0.40	0.40
Partial β_1	0.36	0.40	0.40	0.40	0.41	0.39

The information inefficiency regressions for the annual frequency, referring to $CapU^{tq1-4}$, are summarised in table 9. The revision period is specified by $32 \geq u > 2$, measured in quarters.

Table 9: Informational inefficiency regressions, with country fixed effects, $C = C^t_t$

	u = 2 quarters		u = 3 years		u = 5 years		u = 8 years	
Filter	*LP_rt*	none	*LP_rt*	none	*LP_rt*	none	*LP_rt*	none
β_1	0.80	0.82	0.80	0.81	0.95	0.98	1.08	1.09
$p(\beta_1)$♣	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**	**≤ 0.01**
β_2	0.07	0.06	0.08	0.06	0.08	0.07	0.18	0.21
$p(\beta_2)$♣	**≤ 0.01**	**≤ 0.05**	**≤ 0.10**	**≤ 0.10**	**≤ 0.10**	**≤ 0.10**	**≤ 0.05**	**≤ 0.01**
n	236	244	177	185	145	140	90	98
R^2	0.92	0.92	0.82	0.82	0.83	0.82	0.88	0.87

♣ One-tailed significance tests

The table shows that the null hypothesis of informational efficiency of the OECD output gap released in real time can also clearly be rejected when we refer to annual data going back to 1995. For the earlier revisions (we show $u = 2$), rejection is at $p \leq 1\%$, referring to the recursively low-pass filtered BTS data. For longer revision intervals, he have to lift the hurdle to 5% or 10% in order to be able to call our results significant at conventional levels, but the overall picture is clear. Interestingly, at $u = 8$ years, β_2 again passes the 1% hurdle, so that even after eight years of revision, BTS data available in real time add information to the real time release of the output gap in terms of predicting the current output gap vintage.

4 Summary and conclusions

The output gap might be a useful concept for theoretical thinking about inflationary pressures *ex post*, its practical usefulness is severely impaired or even annihilated by the inherent difficulty to know with sufficient reliability the magnitude of the output gap at the time *when the policy maker needs to know* it, i.e. in *real time*.

We show that this verdict holds for the OECD output gap estimates, which are on average massively revised. Moreover, as revisions tend to continue, it remains hard to reliably quantify the output gap for a particular period, even with the benefit of hindsight.

In this paper, we examine whether the notoriously unreliable real time estimates of the output gap can be improved by referring to measures of physical capital capacity utilisation from business tendency surveys (BTS). These are highly informative data, as they relate directly to the stress on the current capacity to produce goods and services. Moreover, and importantly in our context, these data are usually not revised, so that they are not affected by the end point problem.

To assess this question empirically, we construct a large panel data set, comprising 26 countries and yearly as well as quarterly data from1970 to 2009 (lengths depending on the particular series and country) with qualitative and quantitative information on capacity utilisation from BTS in quarterly frequency and output gap estimates as published by the OECD in real time. We show that the real time output gaps are informationally inefficient in the sense that BTS data available in real time can help to produce estimates that are significantly closer to later releases of output gap estimates. Sensitivity analyses confirm that this finding is robust.

Starting from this, future research will have to show how these findings can be used to improve output gap estimates in real time.

References

Bansak, C. A., Morin, N. J. and Starr, M. A. (2007), Technology, Capital Spending, and Capacity Utilization, Economic Inquiry, 45(3), 631–645.

Billmeier, A. (2004), Measuring a Roller Coaster: Evidence on the Finnish Output Gap, IMF Working Paper WP/04/57.

Butler, L. (1996), A Semi-Structural Method to Estimate Potential Output: Combining Economic Theory with a Time-Series Filter, Bank of Canada Technical Report No. 77.

Cayen, J.-P. and S. van Norden (2002), La fiabilité des estimations de l'écart de production au Canada, Bank of Canada Working Paper 2002-10.

Chagny O. and Döpke, J. (2001), Measures of the Output Gap in the Euro-Zone: An Empirical Assessment of Selected Methods. Kiel Working Paper No. 1053.

Conway, P. and B. Hunt (1997), Estimating Potential Output: a Semi-structural Approach, Reserve Bank of New Zealand Discussion Paper G97/9.

Dergiades T. und Tsoulfidis, L. (2007), A New Method for the Estimation of Capacity Utilisation: Theory and Empirical Evidence from 14 EU Countries, Bulletin of Economic Research, 59(4), 361–381.

Eidgenössisches Finanzdepartement (2001), Die Schuldenbremse. Dokumentation, 2nd ed., Bern (www.efd.admin.ch/dokumentation/gesetzgebung/00573/00869/index.html?lang=de).

Etter, R. Graff, M. and Müller, J. (2008), Is „Normal“ Capacity Utilisation Constant over Time? Analyses with Macro and Micro Data from Business Tendency Surveys, 29th CIRET Conference 2008, Santaigo de Chile, 8–10 October.

Giorno, C., Richardson, P., Roseveare, D. and van den Noord, P. (1996), Potential Output, Output Gaps and Structural Budget Balances, OECD Economic Studies No. 24.

Graff, M. (2004), Estimates of the output gap in real time: how well have we been doing? Reserve Bank of New Zealand Discussion Paper DP2004/04, Wellington, May 2004.

Gruen, D., , T. and Stone, A. (2002), Output Gaps in Real Time: Are They Reliable Enough to Use for Monetary Policy? Reserve Bank of Australia Research Discussion Paper 2002-06.

Hagger, A. J. and N. Groenewold (2003), Time to Ditch the Natural Rate? *Economic Record*, 79, 324–335.

Hirose, Y. and K. Kamada (2003), A New Technique for Simultaneous Estimation of Potential Output and the Phillips Curve, Bank of Japan Monetary and Economic Studies, August 2003.

Kahn, J. A. and R. Rich (2003), Tracking the New Economy: Using Growth Theory to Detect Changes in Trend Productivity, Federal Reserve Bank of New York Staff Reports No. 159.

Lansing, K. (2002), Can the Phillip Curve Help Forecast Inflation? Federal Reserve Bank of San Francisco Economic Letter No. 2002-29.

Laxton, D. and R. Tetlow (1992), A Simple Multivariate Filter for the Measurement of Potential Output, Bank of Canada Technical Report No. 59.

Nelson, E. and Nikolov, K. (2003), UK Inflation in the 1970s and 1980s: the Role of Output Gap Mismeasurement, Journal of Economics and Business, 55 (4), 353–370.

Orphanides, A. (2003), Historical Monetary Policy Analysis and the Taylor Rule, Journal of Monetary Economics, 50 (5), 983–1022.

Orphanides, A. and van Norden, S. (2002), The Unreliability of Output-gap Estimates in Real Time, Review of Economics and Statistics, 84, 569–583.

Orphanides, A. and van Norden, S. (2003), The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time, www.hec.ca/pages/simon.van-norden/wps/RT2JMCB2.pdf.

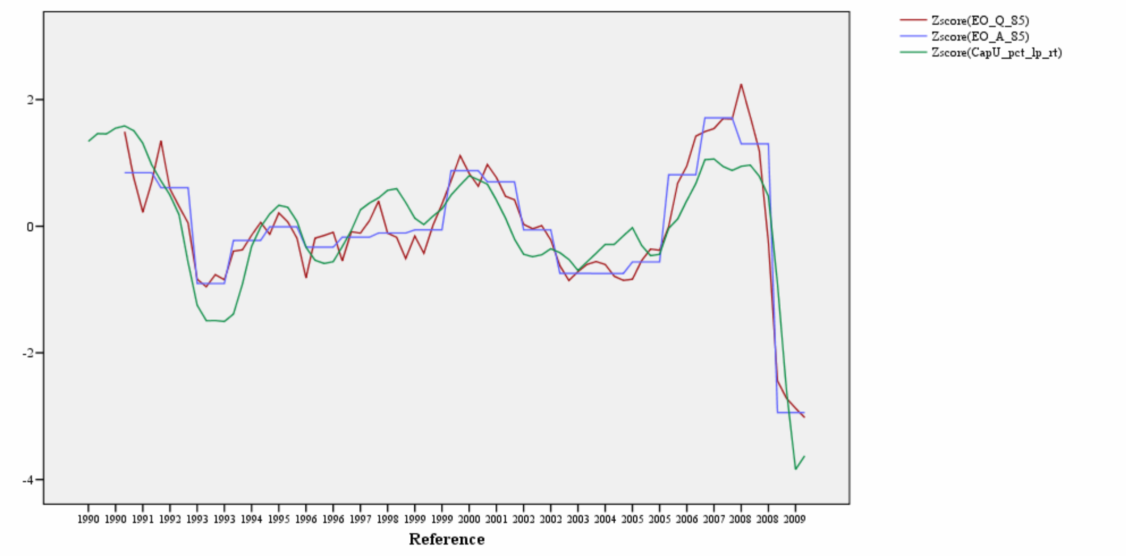
Robinson, T., A. Stone and M. van Zyl (2003), The Real-Time Forecasting Performance of Phillips Curves Reserve Bank of Australia Research Discussion Paper 2003-12.

Rünstler, G. (2002), The Information Content of Real-Time Output Gap Estimates: An Application to the Euro Area, ECB Working Paper No. 182.

Tosetto, E. (2008), Revisions of Quarterly Output Gap Estimates for 15 OECD Member Countries, OECD, Statistics Directorate, 26 September 2008.

Appendix

Quarterly and annual output gap, capacity utilisation (recursively low-pass filtered), Germany 1990–2009, standardised over the sample period



2003q4 to 2009q2 vintages of German OECD quarterly output gaps

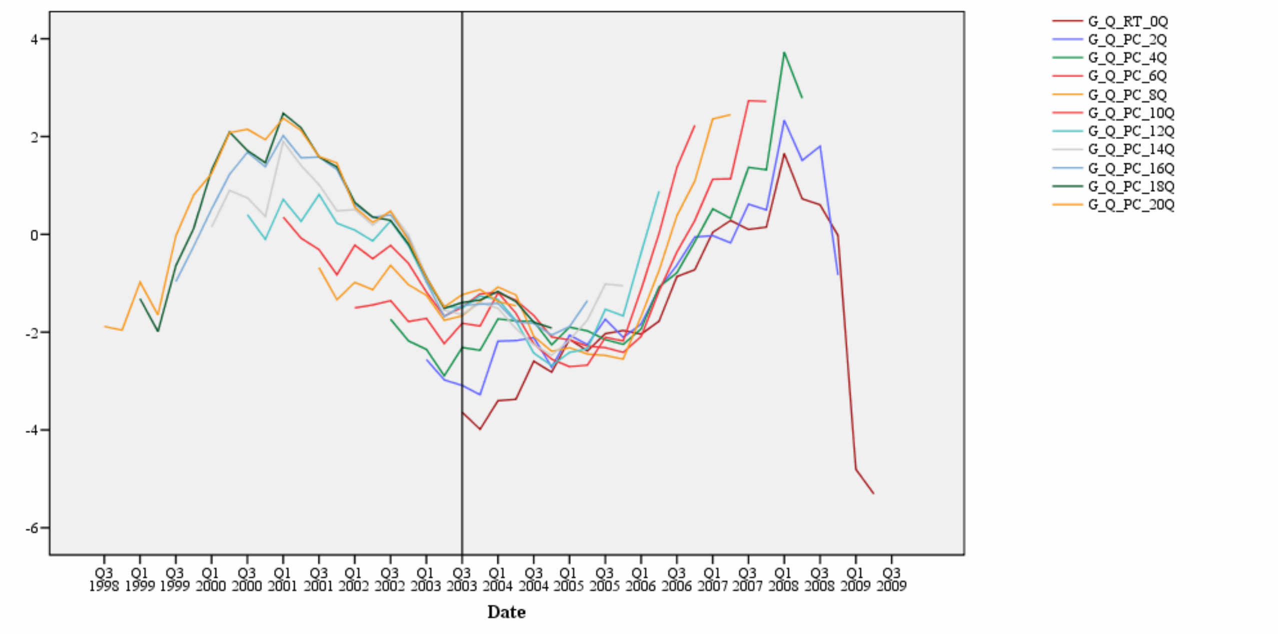


Figure #: 1995q2 to 2009q2 vintages of German OECD annual output gaps