

Data Uncertainty, the Output Gap and Monetary Policy in the United Kingdom

Graeme Chamberlin¹

¹ Faculty of Business, London South Bank University, London, United Kingdom

Correspondence: Graeme Chamberlin, London South Bank University, London Road, London, SE1 0AA, United Kingdom. E-mail: chambeg4@lsbu.ac.uk

Received: October 13, 2014

Accepted: November 6, 2014

Online Published: December 25, 2014

doi:10.5539/ijef.v7n1p129

URL: <http://dx.doi.org/10.5539/ijef.v7n1p129>

Abstract

It is now widely accepted that early vintages of GDP data are likely to undergo a series of revisions as more information becomes available to National Statistics Institutions. This, however, creates a trade-off between timeliness and accuracy for data users such as monetary policy-makers. This article investigates the importance of data revisions for the operation of monetary policy in the UK. I find data revisions have been smaller in the last decade and early data vintages provide a relatively good signal of data released at a later date. There is though evidence of cyclical and structural breaks in the data revision process which should caution data users against assuming data revisions are predictable. I also look at the role of the output gap in the practical implementation of monetary policy through the Taylor rule. While revisions to GDP data account for some of the mismeasurement of the output gap, I conclude that it is a relatively small culprit. The main cause of output gap revisions is due to the difficulty in estimating the level of potential output towards the end of the sample where statistical filter methods are generally unreliable.

Keywords: GDP, monetary policy, output gap, real-time data, revisions

1. Introduction

The trade-off between timeliness and accuracy in estimates of gross domestic product (GDP) is a well-established phenomenon. In the United Kingdom (UK), the Office for National Statistics (ONS) publishes the *GDP Preliminary Estimate* within 25 days of the reference quarter which is based on the limited information available at the time. Gaps in the data, particularly for the third month of the quarter, need to be plugged with forecasts and imputations for a complete measure of GDP to be recorded.

As time progresses the data content of GDP estimates increases reducing the need to fill gaps and allowing the published estimates to be reconciled with more source data. In the annual *Blue Book*, around 9–12 months after the reference quarter, quarterly data is benchmarked to annual surveys and administrative data as it comes available. When the data passes through its second *Blue Book*, around 18–24 months after the reference quarter, input-output analysis is used to fully reconcile the output, income and expenditure measures of GDP.

The *Blue Book 2* estimate typically marks the end of data-driven revisions. However, methodological changes which lead to the introduction of new sources and techniques to better measure the evolving economy mean the data will continue to be revised. As methodological developments are open-ended it means that GDP could be revised in perpetuity. In a sense there is no such thing as final data.

Even though there is now a general understanding of why revisions occur they still present a challenge to policy-makers in forming their current assessment of the economy. A key question is how much weight should be placed on initial estimates if they continue to be revised, often for several years afterwards. This is especially pertinent when policy is set pre-emptively such as monetary policy. The Bank of England has referred to this issue as ‘dealing with data uncertainty’ (Ashley, Driver, Hayes, & Jeffery, 2005).

In 2004, the then Monetary Policy Committee (MPC) member, Marian Bell (Bell, 2004) commented:

‘The monetary policy committee recognises that revisions are inevitable. Indeed so far as they bring us to a better understanding of reality, we welcome them. There is inevitably a trade-off between timeliness and accuracy, but in general, we would prefer early imperfect data to late perfect series – it gives us something to work with.’

Later that year, Rachel Lomax (Lomax, 2004), the then Deputy Governor of the Bank of England said:

‘If the ONS waited two or three years before publishing their first estimates they would have reasonably complete information. But it would be of rather limited use for policy purposes. We need more timely indicators of economic activity, even though these will tend to be less accurate than later estimates.’

This article considers the implications of data revisions for monetary policy in the UK. The first half of the paper looks at the effect of data revisions on GDP estimates. In section 2, I outline the main features of a real-time database for UK GDP and analyze average revisions over time. I find revisions to GDP exhibit both a number of structural breaks and a general cyclical pattern. This should warn statisticians and policy-makers that care must be taken if using past information on revisions to try and improve preliminary GDP estimates, for example by applying bias adjustments. In section 3 I look at how clearly early data vintages signal more mature or revised data. I find preliminary estimates of GDP, especially over the last decade, have provided a high signal to noise ratio.

The second part of the paper looks more closely at the effect of data revisions on the operation of monetary policy through the measurement of the output gap. In section 4, I discuss the literature on how the output gap feeds into monetary policy rules and the problems arising from output gap mismeasurement. In section 5 I then investigate the relative importance of data revisions in contributing to output gap mismeasurement. I find, in accordance with most of the published literature, data revisions actually play a relatively small role in accounting for output gap mismeasurement. The final section offers some concluding thoughts including on how policy-makers may respond to data uncertainty.

2. UK GDP Data in Real-Time

Real-time data is a collection of different vintages of GDP which shows the actual data available to users at specific points in time. I use the real-time data set referenced in Brown, Buccellato, Chamberlin, Dey-Chowdhury and Youll (2009) which shows a snapshot of UK GDP taken from monthly ONS publications from 1961 to the current day. The dataset allows us to see how revisions have affected the measured path of the economy over time.

2.1 Features of the Real-Time GDP Data

Figure 1 presents eight different vintages of GDP growth (quarter on same quarter a year ago) between 1961 Q2 and 2013 Q4. These are the first published estimate (T) for each quarter, and then the respective data three months (T+3), six months (T+6), 12 months/one year (T+12), 24 months/two years (T+24), 36 months/three years (T+36), 48 months/four years (T+48) and 60 months/five years (T+60) later. This is a sample of the data held in the real-time database although any data vintage published since 1961 could have been presented.

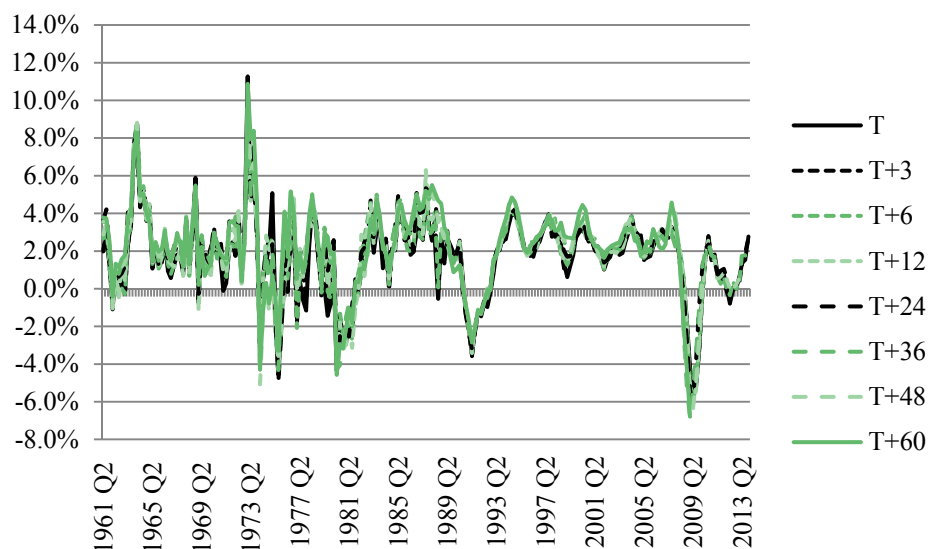


Figure 1. Real time GDP data (% growth, quarter on same quarter 1 year ago)

Plotting several different data vintages in the same chart gives a visual of real-time data. The myriad of estimates for each quarter is what Bell (2004) referred to as the 'statistical fog' which policy-makers had to see through.

To make things a bit clearer Figure 2 directly compares the first data vintage with that published five years (60 months) later. This data suggests that, over a long period of time, first estimates have provided a good guide to the state of the economy as described by more mature data. So, although early estimates are likely to be revised which creates uncertainty for users, it seems rare that the data is changed in such a way that it fundamentally alters the economic story being told.

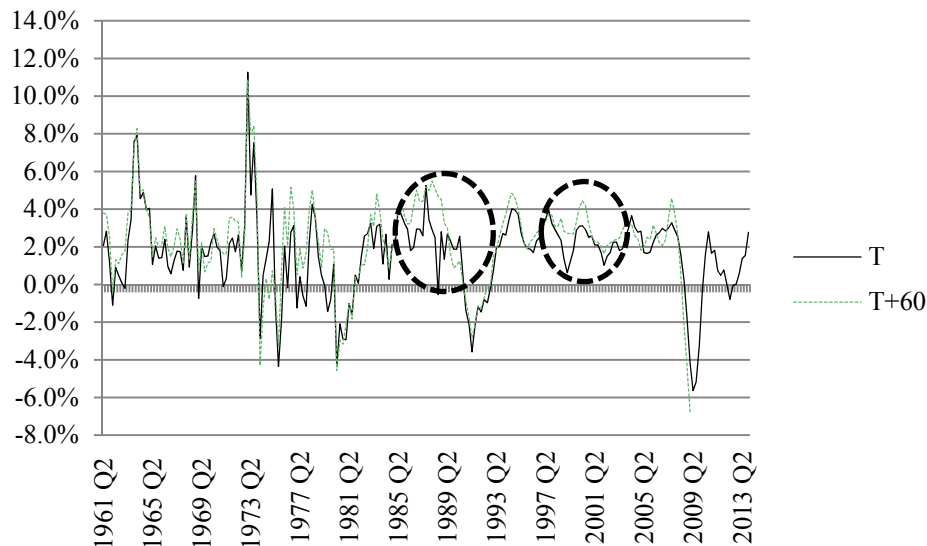


Figure 2. Real time GDP data, T and T+60 (% growth, quarter on same quarter 1 year ago)

That is not to say there haven't been cases where early data have been revised in a significant way. In Figure 2, I have circled two periods where revisions have had an important effect on our reading of the economy. Firstly, preliminary GDP data underestimated the strength of the economy in the late 1980s. Early data vintages suggested a sharp slowdown in 1987 and 1988, but later data showed growth to be more resilient.

Secondly, there was a period of significant upward revisions during 1998 and 1999. Preliminary estimates of GDP had pointed to a slowdown in growth, a view widely backed by other business survey data. The Asian financial crisis of 1997-8 and the Russian crisis of 1998 which was then propagated to Latin America, certainly led to uncertainty and instability in the global economy at this time, but there was also an easing in UK monetary policy that supported domestic demand. This was picked up in later data vintages which showed the economy, especially aggregate demand, to have been stronger over these two years.

2.2 Average Revisions

Average (arithmetic mean) revisions between the first published estimate and those published 24 and 60 months later for the quarter on quarter growth rate for GDP are presented in Figure 3. I separate the analysis into sub-periods corresponding to four distinct periods in the UK's economic history. These are the period of relatively successful stabilization policy (1961 Q2 to 1972 Q4), the period of stagflation (1973 Q1 to 1984 Q4), the boom and bust of the late 1980s and early 1990s (1985 Q1 to 1996 Q4) and finally the Great Moderation and Great Crash (1997 Q1 to 2009 Q1). This final period was also marked by significantly lower inflation compared to previous decades.

Figure 3 shows that revisions to GDP have generally been positive. Although interesting in their own right, mean revisions are not a generally accepted measure for judging the size and significance of revisions. This is because average revisions can be relatively low simply because, in a given period, large upward revisions have been offset by large downward revisions.

The cyclical nature of revisions is clearly evident from Figure 4. In order to see how revisions vary over the economic cycle, a centered five-year rolling average of revisions between the first published and a number of subsequent data vintages are presented. Downward revisions were reported in both the recessions of the early

1980s, early 1990s and 2009, and very significant downward revisions in the 1973 recession. Upward revisions have usually occurred in times when the UK economy is expanding.

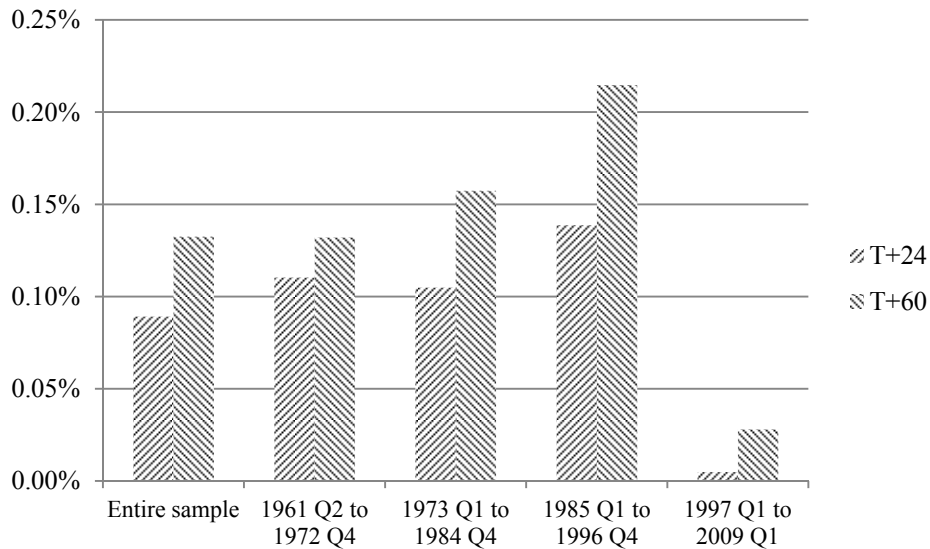


Figure 3. Mean revisions between T and T+24 and T+60 GDP vintages (percentage points change to quarter on quarter growth rates)

This is a typical feature of National Accounts data around the world. Because GDP estimates are based on surveys of firms, usually incorporating all the large firms and a randomized sample of medium and small enterprises, the business register may lag changes in the economy. For instance, smaller but fast growing new firms and new sectors may be under-represented in samples during periods of stronger economic growth.

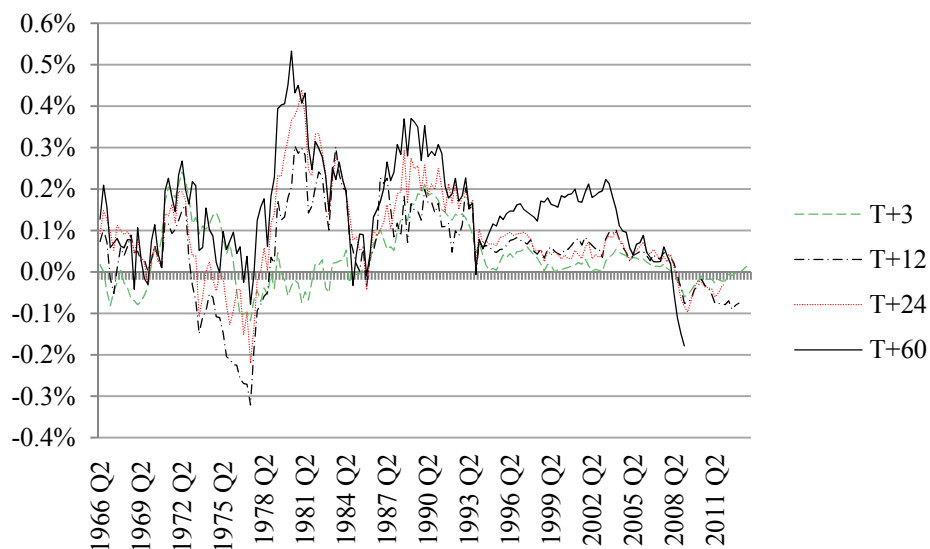


Figure 4. Rolling centered five-year average of GDP revisions (percentage points change to quarter on quarter growth rates)

The cyclical nature of revisions should caution policymakers and statisticians against relying on past information on revisions to predict revisions in the future, for instance by applying bias adjustments to preliminary estimates of GDP. During the Great Moderation period (1993–2008) there was a long period of upward GDP revisions

which led to calls for a bias adjustment to be applied to GDP data. However, that would have led to large errors as the economy entered recession in the second half of 2008 when preliminary data were subsequently revised downwards.

Further proof that the profile of GDP revisions is unstable over time and the past may not necessarily be a good guide to the future is shown in Figure 5. Mean absolute revisions do not allow positive and negative revisions to cancel out so are a better indicator of the ONS's revisions performance over time. Using the same sub-periods it is clear that over the last decade revisions from the preliminary to more mature (T+24 and T+60) estimates have improved. Figure 5 suggests in addition to a cyclical influence the profile of revisions is also subject to structural breaks. This could be related to improvements in the collection and publication of data, or changes in the structure and volatility of the economy.

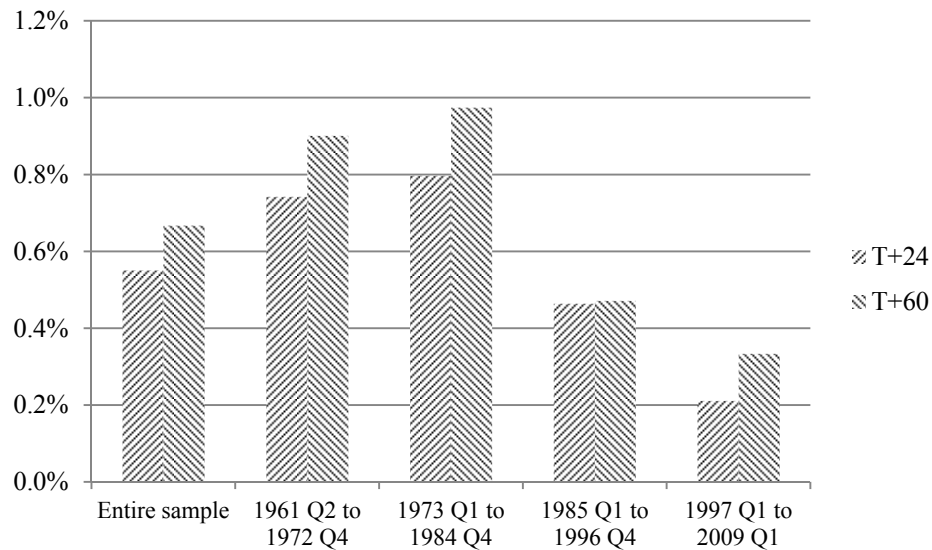


Figure 5. Mean absolute revisions (percentage points to quarter on quarter growth rates)

3. GDP Quality and the Signal to Noise Ratio

Judging the quality of the National Accounts is a subjective exercise as there is no agreed definition or metric for doing so, and most National Statistics Institutions refrain from trying to attach a value to the quality of GDP data. Instead quality is expressed verbally in terms like accuracy, timeliness, punctuality, accessibility, clarity, comparability and coherence.

Oller and Teterukovsky (2006) suggest the key quality measure for preliminary or early estimates of GDP is how early and how accurately the value of a macroeconomic variable can be assessed with some certainty. This relates to the 'accuracy' benchmark and 'timeliness' in their availability for use in policy-making. This approach puts information on revisions at the centre of the quality assessment. The more early data has to be revised the less predictable and trustworthy these vintages become.

In this section I look at the signal to noise quality measure referred to by Oller and Teterukovsky (2006). This is based on standardized mean squared revisions and essentially tells us how predictable mature GDP data is given preliminary estimates.

3.1 Mean Squared Revisions

Mean squared revisions (MSR) are a commonly used diagnostic which penalizes large revisions at the expense of smaller ones. From a user perspective, a few large revisions to GDP are more likely to disrupt the reading of the economy than a preponderance of small ones, so producers of economic statistics are more sensitive to avoiding large revisions.

The mean squared revision of a particular data vintage (i) with respect to the T+60 vintage is given by

$$MSR_{T+i:T+60} = \frac{\sum_{n=1}^N (y_{n,T+i} - y_{n,T+60})^2}{N} \quad (1)$$

For all vintages $i = 0, \dots, 60$.

Where N is the number of quarters in the sample or sub-sample and $y_{n,T+i}$ is the i -th vintage of GDP data pertaining to the n -th quarter. Obviously, when $i = 60$ there is no revision between the $T+60$ vintage and its self so the MSR will equal zero.

Using the same sub-periods as before the mean squared revisions with respect to the $T+60$ vintage are shown in Figure 6. It is clear that in the last two sub-periods, and particular the more recent sub-period, the revisions performance of early GDP vintages has improved markedly.

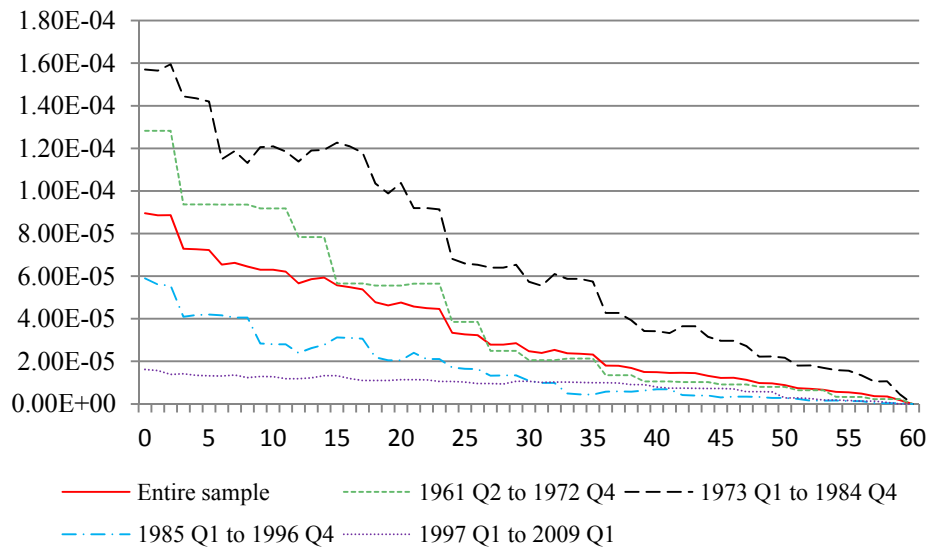


Figure 6. Mean squared revisions over different sub-periods (quarter on quarter growth rates)

3.2 Standardized MSR and the Signal to Noise Ratio

Oller and Teterukovsky (2006) have developed a quality measure for national accounts data based on standardizing mean squared revisions. This is simply the MSR of each vintage divided by the variance of the final data vintage ($T+60$ vintage in this case).

$$\overline{MSR}_{T+i:T+60} = \frac{MSR_{T+i:T+60}}{var(y_{T+60})} \quad (2)$$

For all vintages $i = 0, \dots, 60$.

Figure 7 gives an indication why standardized MSR is an interesting quality measure. It shows the MSR of the first data vintage in each sub-period alongside the variance of the $T+60$ vintage. It is clear that the MSR in the final two sub-periods is relatively low, but at the same time the variance of the underlying mature data is also significantly lower. This suggests one of the reasons why revisions performance has improved is because the economy has become more stable. MSR in the stagflation decade (1973–1984) was relatively high, but this was also a period of high nominal and real instability in the UK economy which hampered the efforts of statisticians in recording economic activity. The standardized measure therefore takes into account the volatility of the economy in judging revisions performance.

Figure 8 shows how standardized MSR varies over 60 data vintages and across different sub-periods. Once again early vintages in the latest sub-period have the lowest standardized MSR suggesting a recent improvement in the quality of early GDP estimates. An interesting observation is that, taking into account the underlying volatility of the economy, revisions performance in the stagflation decade no longer looks out of step with the entire sample. On the other hand, revisions performance during the boom and bust of the late 1980s and early 1990s no longer looks so good.

Oller and Teterukovsky (2006) define the signal to noise ratio equal to 1 minus the standardized MSR. Clearly the lower the standardized MSR, the higher the signal to noise ratio for any given data vintage. A standardized MSR above unity implies a ‘worthless’ GDP vintage, in that the signal to noise ratio would be negative and in this case statisticians would have been better off publishing an average of the past mature data rather than that particular data vintage.

The Oller and Teterukovsky (2006) signal to noise ratio for a particular data vintage is synonymous with the degree of determination (the R^2 statistic) in a regression of that data vintage against the mature data. A high signal to noise ratio means that data vintage is a good predictor of the mature data. This analysis has again shown that, even taking into account the lower volatility of the UK economy, revisions performance during the last decade has improved. Based on the recent history of revisions data users should therefore be reasonably confident that preliminary measures of GDP published by the Office for National Statistics are a good reflection on the current state of the UK economy.

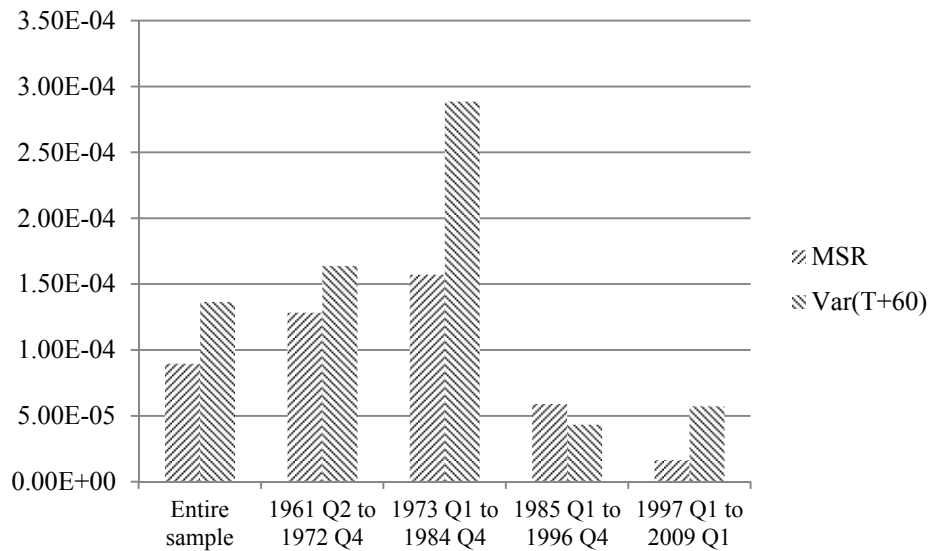


Figure 7. Mean squared revisions between T and T+60 (MSR) and variance of T+60 (Var(T+60)) (quarter on quarter growth rates)

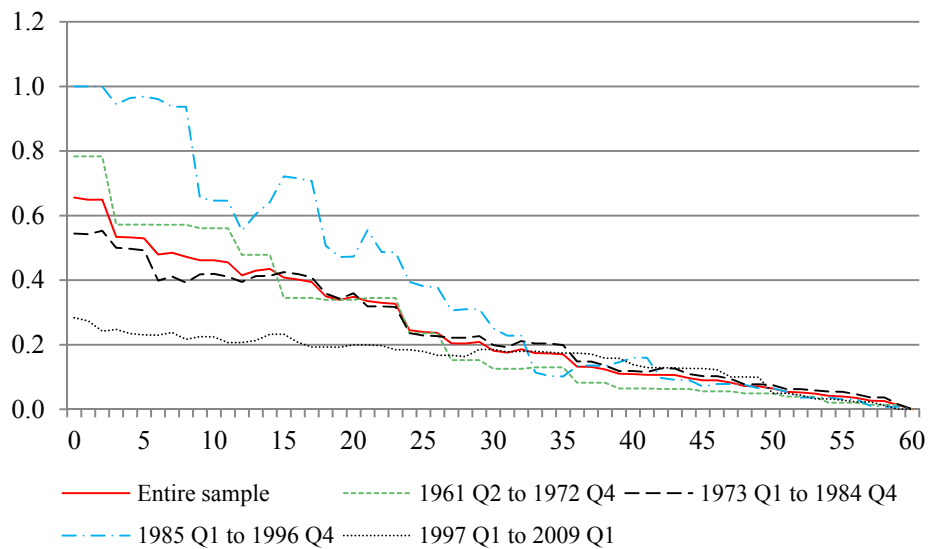


Figure 8. Standardized mean squared revisions over different sub-periods (quarter on quarter growth rates)

4. Taylor Rules, the Output Gap and Monetary Policy in Real-Time

This section presents some of the literature on how revisions to GDP might affect the operation of monetary policy through the measurement of the output gap and the Taylor rule for setting nominal interest rates.

4.1 The Taylor Rule and Monetary Policy

The Taylor rule (Taylor, 1993) is a simple description of how monetary policy is set in practice. It suggests that central banks set their main lending rate according to the following equation:

$$i_t = \pi_t + r_t^* + \alpha_\pi(\pi_t - \pi^*) + \alpha_y(y_t - y_t^*) \quad (3)$$

Where i_t is the nominal interest rate, π_t the current rate of inflation, r_t^* the real equilibrium rate of interest, π^* the target rate of inflation and $(y_t - y_t^*)$ the difference between actual and potential output or the output gap. The respective weights α_π, α_y determine how interest rates are changed in relation to deviation of inflation from target and the output gap.

Using $\alpha_\pi = \alpha_y = 0.5$ and $r_t^* = \pi^* = 2\%$ Taylor (1993) finds that the rule fits the movement in the US Federal Funds rate between 1987 and 1992.

Since then the Taylor rule and its forward-looking variants based on inflation expectations and forecasts of output growth (Clarida, Gali, & Gertler, 1998) have become a staple in the macroeconomic models used by policy-makers. This especially applies to monetary policy where the move to independent central banks in recent decades has emphasized rules over discretion.

The Taylor rule though has a number of limitations. First it relies on the equilibrium real rate of interest and the level of potential output both of which are unobserved variables. Clark and Kozicki (2005) show estimates of the equilibrium real interest rate are very unstable. Difficulties in measuring the level of potential output and the output gap is the main focus of the next section.

Second, the Taylor rule is often viewed as an excessive simplification of policy setting. Gerdesmeier and Roffia (2005) debate whether the output gap and the current rate of inflation actually capture all the information required for monetary policy setting. They point to other relevant variables such as credit aggregates, stock prices, fiscal aggregates, wage agreements, exchange rates and commodity prices. They also ask whether the Taylor rule can adequately pick up the optimal reaction to demand and supply or permanent and temporary shocks as monetary policy may and should react differently in each case.

Gerberding, Seitz and Worms (2005) conclude the Taylor rule is not a good description for German monetary policy between 1979 and 1998. The Bundesbank more closely followed a target for long-term growth in the broad money supply. When this monetary target rule is included in real-time estimates of the Taylor rule the output gap measure drops out.

4.2 Real-Time Estimates of the Output Gap and Monetary Policy Mistakes

Key to the practical operation of monetary policy in the Taylor rule is an accurate assessment of the output gap. In a series of papers, Orphanides compares the operation of the Taylor rule for the US Federal Funds rate using both real-time and the latest available data. He finds that in many situations, the interest rate prescription differs depending on the data the policymakers were looking at.

Orphanides (2001) finds the Taylor rule fits the evolution of the Federal Funds Rate fairly well between 1987 and 1992 concurring with Taylor (1993). However, this finding is based on using revised data. Using real-time or unrevised estimates of the output gap available to policymakers at that point in time would have given a different path to interest rates. In real-time the recovery from recession in 1992 looked worse than it did in the subsequently revised data. As a result the Federal funds rate was higher by 200 basis points through to 1992 than it would have been if the weaker real-time data was applied to the Taylor rule.

Orphanides (2002) identifies the great inflation of the 1970s as a severe macroeconomic policy failure after the relative success of stabilization policy of the 1960s. During the 1970s the natural rate of unemployment was underestimated so the rise in unemployment was treated as cyclical rather than structural, and associated with a negative rather than a positive output gap. Monetary policy was misconceived in that the interest rate was too low to combat rising inflation, but not set in a way completely at odds with the Taylor rule given the real-time view on the output gap.

Orphanides (2003a) argues that misperceptions on the economy's productive potential have been a recurrent source of errors in fitting Taylor rules in real-time for the US. In 1955, 1965 and 1978 later data vintages would have called for tighter monetary policy than the prevailing Federal funds rate at the time. Past minutes of the Fed

meetings showed that in each case the extent of overheating was sorely underestimated. In fact the output gap typically had the wrong sign as policy-makers were slow to pick up on persistent slowdowns in productivity growth.

This issue was analyzed by Lansing (2002) using a small macroeconomic model of the US economy where the Federal Reserve sets monetary policy according to its perceived output gap in a Taylor rule while it is the actual output gap based on later data vintages which feeds into the model affecting aggregate demand and inflation. Simulations show that important inflation dynamics in the US can be explained by abrupt changes in potential output which the Fed initially mistakes for cyclical factors. For example the peak in inflation in 1979 Q2 was accounted for by the productivity slowdown in the mid 1970s. The author also reports that the downward movement in inflation after 1995 is associated with the productivity boom in the mid 1990s.

Nelson and Nicolov (2003) provide the most comprehensive analysis for the UK looking at the role of output gap mismeasurement in explaining UK inflation in the 1970s and 1980s. They build a database of real-time output gap measures from real-time datasets and collections of policy statements and reports on the potential level of output. They find significant revisions to the output gap series contributed to the inflation spikes in the 1970s and particularly the 1980s. In 1972 the slowdown in growth was originally perceived as cyclical rather than structural so the output gap was tighter than initially thought and monetary policy subsequently too loose. During the late 1980s they found economic growth was weaker in preliminary data and as a result the budget of 1989 was, in hindsight, too expansionary.

A number of authors have estimated the direct empirical relationship between the output gap and the rate of inflation. Kamada (2005) for Japan, Marcellino and Musso (2010) and Guerin, Maurin and Mohr (2011) for the Euro Zone, and Orphanides and van Norden (2005) for the US all find that while output gap measures based on later data vintages are statistically significant as predictors of inflation, output gap measures based on real-time data generally fall out of the estimated equation.

In summary, mistakes in the operation of monetary policy can often be traced to mismeasurement of the output gap in a real-time setting. The key issue, and the crux of this paper, is the extent to which this mismeasurement is as a result of revisions to the underlying data or due to the difficulty in measuring an economy's potential level of output. In the next section I try to answer this for the UK output gap.

5. Real-Time and Latest Measures of the UK Output Gap

Constructing an output gap series requires an estimate of the economy's potential level of output which is an unobserved variable. Ladiray, Mazzi and Sartori (2003) and Murray (2014) both provide good surveys of the various approaches.

These primarily fall into four main categories. First, univariate non-structural methods are basic statistical approaches using filters. They do not rely on any economic theory and as they require little information are easy to apply. Secondly, multivariate approaches apply the same techniques but supplement the statistical approach with information from economic variables and relationships such as the Phillips curve.

Third, direct approaches to estimating the output gap rely on using survey measures such as capacity utilization and recruitment difficulties providing a measure of spare capacity in the economy. Finally structural approaches are based on relationships grounded in economic theory relationships. This may involve the identification of shocks in an SVAR model or by directly estimating an aggregate production function for the economy.

Because the focus of this paper is to investigate the effects of data revisions on the practice of monetary policy I concentrate on simple univariate models which are easy to apply in real-time and because they are less subject to parameter instability it is easier to identify the effects of data revisions on the measured output gap.

5.1 Sources of Output Gap Uncertainty in Real-Time

There are three main sources of output gap uncertainty. First there are data revisions so the underlying GDP data may look different in real-time to revised data leading to different output gap estimates. Second, univariate models based on filters are unreliable at the end of the sample and are likely to be revised as more data in terms of new observations become available. I call this the benefit of hindsight. Third there are a number of ways of estimating the output gap and no definitive test to determine the most accurate which I call model uncertainty.

Orphanides and van Noorden (2002) investigate the relative importance of these sources of output gap uncertainty by comparing three different measures. First an estimate based on the latest or revised data. Second an estimate based on real-time data which gives the estimate of the output gap at each point in time using the actual data that would have faced the policymaker. Finally, a quasi real-time estimate of the output gap is a

recursive measure on the latest fully revised data.

The difference between the latest and quasi real-time output gap measures reflects the benefit of hindsight in being able to more accurately apply the filter over the entire sample rather than up to a particular point in time. The difference between real-time and quasi real-time measures reflects the contribution of data revisions. The effect of model uncertainty can be gauged by using a number of univariate approaches and seeing how the measures and the relative importance of hindsight and data revisions in each case.

They test the reliability of US output gap measures in real-time using 12 different models. They show revisions to the output gap are often as large as the output gap itself but data revisions only account for a relatively small part of the revision. Revisions to the level of potential output towards the end of the sample are the main factor.

These findings have been corroborated in many studies. Ince and Papell (2012) update the original Orphanides and van Norden (2002) paper to a sample covering the period 1973–2012 and extend the analysis to 9 other OECD countries. They confirm the original findings by reporting a low correlation between real-time and revised output gap estimates, but high correlation between real-time and quasi real-time estimates. They report their results robust across all countries and various de-trending methods.

Cayen and van Norden (2005) produce similar analysis to Orphanides and van Norden (2002) for Canada. Using 12 different methods for de-trending GDP data to construct the output gap they find data revisions are a more important source of output gap mismeasurement relative to the US study, but still less important than end of sample instability. Similar results for the Euro Zone are reported by Marcellino and Musso (2010) and Brazil by Cusinato, Minella and Porto Junior (2010).

5.2 Measures of UK Output Gaps in Real-Time

Using the real-time data set for the UK I use four basic univariate models to produce real-time, quasi real-time and latest estimates (data published in May 2014) of the output gap. These are the Harvey-Clark unobserved components model, the Harvey-Jaeger variant with a periodic cycle, the Hodrick-Prescott filter and the Christiano-Fitzgerald band pass filter.

There are a number of de-trending methods used in these other studies that I have chosen not to use here. This includes those based on linear or deterministic trends. Following the work of Nelson and Plosser (1982), which finds it hard to reject the presence of a unit root in most macroeconomic time series, it seems plausible to reject the idea of a deterministic trend in GDP in favor of a stochastic trend.

I also do not apply the Beveridge-Nelson decomposition which is also a frequently used method for splitting a macroeconomic time series into trend and cycle components. Here most of the variance in the time series is allocated to the trend component and cycles tend to be small and noisy.

An interesting paper by Morley, Nelson and Zivot (2003) explains why Beveridge-Nelson and unobserved-components models produce such different cyclical series. At the heart of the matter is the amount of variance in GDP that should be allocated to the stochastic trend and the cycle. Stock and Watson (1988) summarize this debate by ascribing that one economist's trend can be another's cycle. However, if we accept the output gap is an important measure describing the economy it makes sense to use methods that produce a relatively clear cyclical series.

In the estimates that follow I have worked in logarithms and expressed the output gap as the deviation of the level of output from its potential level. The output gap measures are therefore a proxy for the percentage deviation from potential output.

5.2.1 Harvey-Clarke

The unobserved components model of Harvey (1985) and Clark (1987, 1989) is a general framework for the estimation of trends using the Kalman filter.

$$\begin{aligned}
 y_t &= y_t^* + c_t \\
 y_t^* &= y_{t-1}^* + \mu_{t-1} + \tau_t \quad \tau_t \sim (0, \sigma_\tau^2) \\
 \mu_t &= \mu_{t-1} + \eta_t \quad \eta_t \sim (0, \sigma_\eta^2) \\
 \Gamma(L)c_t &= v_t \quad v_t \sim (0, \sigma_v^2)
 \end{aligned}
 \tag{4}$$

The first equation is the measurement equation and splits the GDP data (y_t) into trend (y_t^*) and cycle (c_t) series. The next three state equations describe the motion of the trend and cycle components. In this case the trend follows a random walk with drift process and the cycle a stationary auto-regressive model. Following Watson (1986) the cyclical components follow an AR(2) process.

The variances play a critical role in allocating the variance between trend and cycle and therefore the relative smoothness of potential output and the size of the output gap. For instance as σ_{τ}^2 and σ_{η}^2 fall the stochastic trend becomes smoother and more of the variance in output is accounted for by the cycle variance σ_v^2 .

I take the following parameters from Clark (1989) for the US: AR parameters are $\phi_1 = 1.57$, $\phi_2 = -0.63$ and standard deviations are $\sigma_{\tau} = 0.0072$, $\sigma_{\eta} = 0$ and $\sigma_v = 0.0061$. The three calculated output gap series are presented in Figure 9.

5.2.2 Harvey-Jaeger

Harvey and Jaeger (1993) use a very similar model to the Harvey-Clarke model but in their case the cyclical component is determined by a periodic function which is similar to a standard AR model.

$$\begin{aligned}
 c_t &= \rho \cos \lambda_c c_{t-1} + \rho \sin \lambda_c c_{t-1}^* + \xi_t \quad \xi_t \sim (0, \sigma_{\xi}^2) \\
 c_t^* &= -\rho \sin \lambda_c c_{t-1} + \rho \cos \lambda_c c_{t-1}^* + \xi_t^* \quad \xi_t^* \sim (0, \sigma_{\xi^*}^2)
 \end{aligned}
 \tag{5}$$

Where $0 \leq \rho \leq 1$ is a dampening factor and λ_c is the frequency of the cycle in radians.

I apply the following parameters estimated for US GDP to my real-time dataset: $\sigma_{\eta}^2 = 8 * 10^{-7}$, $\sigma_{\tau}^2 = 0$, $\sigma_{\xi}^2 = \sigma_{\xi^*}^2 = 625 * 10^{-7}$, $\rho = 0.92$ and $2\pi/\lambda_c = 22.2$ quarters. The three respective measures of the output gap are plotted in Figure 10.

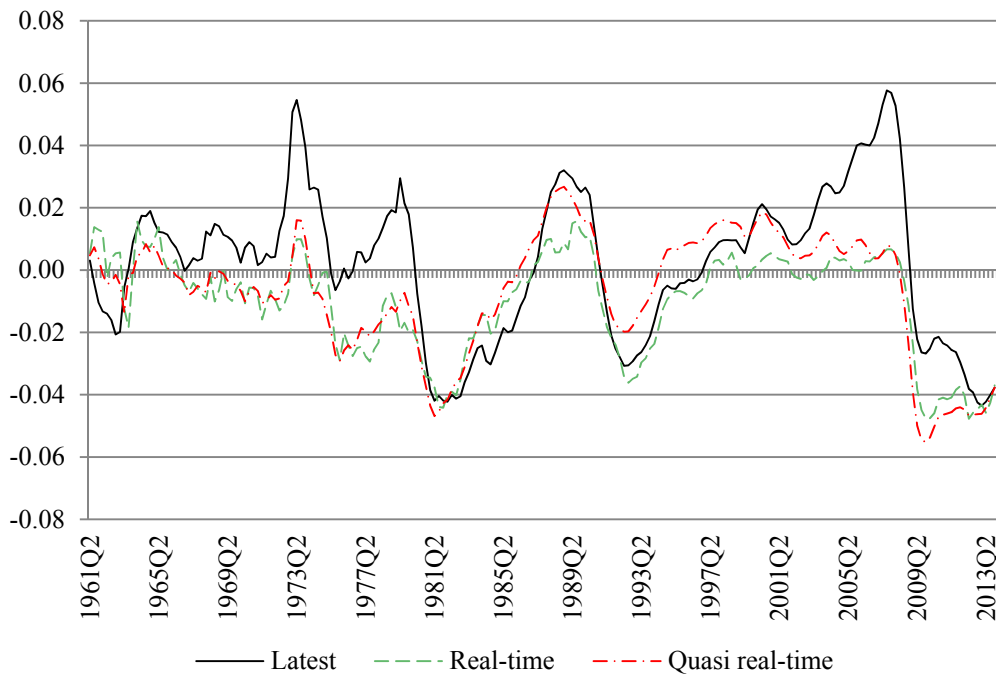


Figure 9. Harvey-Clarke output gap

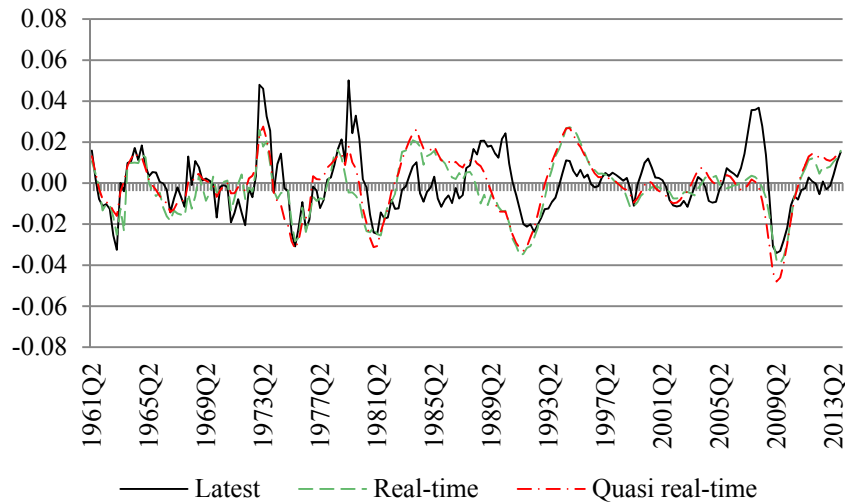


Figure 10. Harvey-Jaeger output gap

5.2.3 The Hodrick-Prescott Filter

The Hodrick and Prescott (1997) or H-P filter is commonly used in macroeconomics. The trend and cycle components of a time series are determined by minimizing the following loss function:

$$\sum_{t=1}^T \left[\frac{1}{\sigma_1^2} (c_t)^2 + \frac{1}{\sigma_2^2} (\Delta y_{t+1}^* - \Delta y_t^*)^2 \right] \tag{6}$$

The variances σ_1^2 and σ_2^2 are respectively those of the output gap and trend growth respectively and are chosen so that cycles are not too large and the growth rate of potential output is smooth. Of course there is a trade-off between the two which is reflected in the loss function.

These variances can be estimated by maximum likelihood, but I follow Hodrick and Prescott (1997) who applied the smoothing parameter $\lambda = \sigma_1^2 / \sigma_2^2 = 1600$ based on US quarterly data. The three output gap measures are presented in Figure 11.

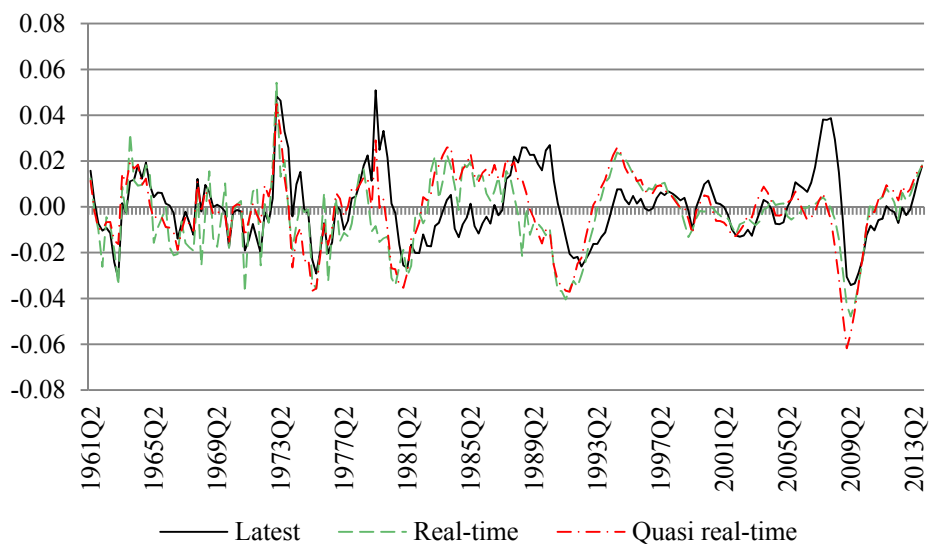


Figure 11. Hodrick-Prescott output gap

5.2.4 Christiano-Fitzgerald Band Pass Filter

Band pass filters aim to isolate the part of a time series that lies within a particular band of frequencies. They

make use of spectral representation theory which implies any series can be split into different frequency components. For business cycles this is typically thought to be in the 6–32 quarters (1.5–8 years) part of the spectrum. The HP filter with a smoothing parameter $\lambda = 1600$ typically takes out anything over 8 years leaving both the cycle and the noise/irregular components intact. The band pass filter derived by Christiano and Fitzgerald (1999) will also remove the short-term irregular component of the data series, so will provide a similar but smoother cycle. The three output gap measures are shown in Figure 12.

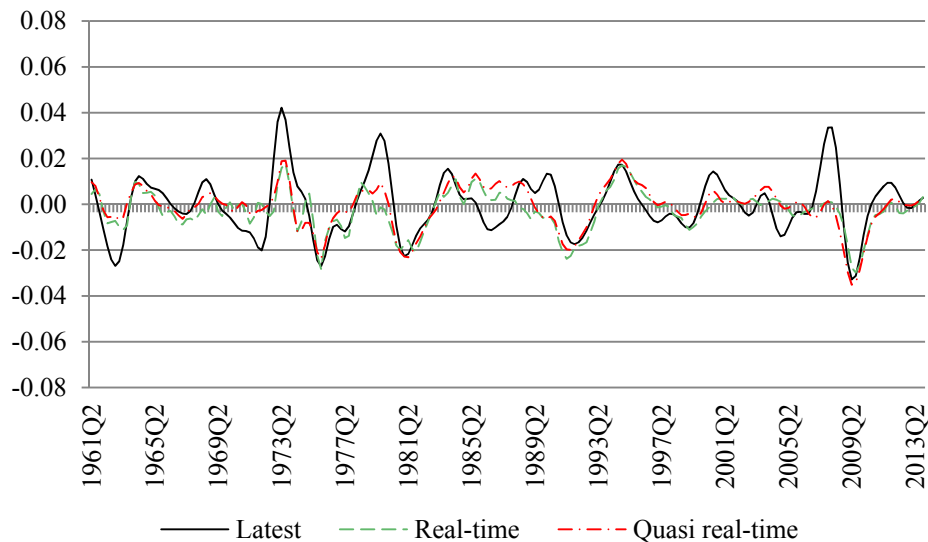


Figure 12. Christiano-Fitzgerald output gap

5.3 Comparing Real-Time, Quasi Real-Time and Latest Estimates of the Output Gap

Looking through figures 9–12 it is fairly clear that the real-time measures of the output gap track the quasi real-time measures fairly closely while the estimates based on the latest data are more different. This suggests data revisions are less important than new data observations at the end of the sample in accounting for the overall difference between real-time and the latest output gap estimates. Following Cusinato et al (2010) and Cayen and van Norden (2005) I present a number of simple metrics comparing the various output gap estimates.

5.3.1 Correlations

Table 1 records the correlations between different output gap measures. It is clear for each of the four estimation approaches the correlation between latest and real-time measures is very similar to that of the correlation between the latest and quasi real-time estimates. In comparison the correlations between real-time and quasi real-time estimates are much higher.

Table 1. Correlations

| | Harvey-Clark | Harvey-Jaeger | Hodrick-Prescott | Christiano-Fitzgerald |
|----------------------------|--------------|---------------|------------------|-----------------------|
| Latest: Real-time | 0.77 | 0.53 | 0.47 | 0.62 |
| Latest: Quasi real-time | 0.73 | 0.56 | 0.50 | 0.64 |
| Quasi real-time: Real-time | 0.92 | 0.92 | 0.85 | 0.91 |

5.3.2 Average Revisions

Table 2 shows the average absolute revisions between the three measures of the output gap. For each estimation approach the mean absolute revision between the latest and quasi real-time estimates is around twice as high as that between quasi real-time and real-time estimates. Table 3 presents the same information but for root mean squared revisions and shows the same results. This suggests that output gap revisions stemming from new data at the end of the sample are twice as large as those resulting from data revisions.

Table 2. Absolute revisions

| | Harvey-Clark | Harvey-Jaeger | Hodrick-Prescott | Christiano-Fitzgerald |
|----------------------------|--------------|---------------|------------------|-----------------------|
| Latest: Real-time | 0.0156 | 0.0105 | 0.0124 | 0.0077 |
| Latest: Quasi real-time | 0.0148 | 0.0105 | 0.0120 | 0.0075 |
| Quasi real-time: Real-time | 0.0069 | 0.0044 | 0.0066 | 0.0033 |

Table 3. Root mean squared revisions

| | Harvey-Clark | Harvey-Jaeger | Hodrick-Prescott | Christiano-Fitzgerald |
|----------------------------|--------------|---------------|------------------|-----------------------|
| Latest: Real-time | 0.020 | 0.014 | 0.016 | 0.010 |
| Latest: Quasi real-time | 0.018 | 0.014 | 0.016 | 0.010 |
| Quasi real-time: Real-time | 0.009 | 0.006 | 0.009 | 0.004 |

5.3.3 Noise to Signal Ratio

The noise to signal ratio provides an indication of how clearer signal of the latest output gap estimate is provided by respectively the real-time and quasi real-time estimates. If data revisions are a large and relevant source of noise then we'd expect the noise to signal ratio between latest and real-time estimates to be much higher than between latest and quasi real-time estimates.

In Table 4 the noise to signal ratio estimates have been calculated by dividing the standard deviation of the revisions between the latest output gap and the real-time and quasi real-time estimates by the standard deviation of the final output gap series. In Table 5 the noise to signal ratio was calculated by dividing the root mean squared revisions by the standard deviation of the latest output gap series. It is evident in both cases that when it comes to providing a signal for the latest measures of the output gap quasi real-time estimates do not conclusively outperform real-time estimates.

Table 4. Noise to signal ratio 1

| | Harvey-Clark | Harvey-Jaeger | Hodrick-Prescott | Christiano-Fitzgerald |
|-------------------------|--------------|---------------|------------------|-----------------------|
| Latest: Real-time | 0.64 | 0.94 | 1.05 | 0.78 |
| Latest: Quasi real-time | 0.69 | 0.93 | 1.03 | 0.77 |

Table 5. Noise to signal ratio 2

| | Harvey-Clark | Harvey-Jaeger | Hodrick-Prescott | Christiano-Fitzgerald |
|-------------------------|--------------|---------------|------------------|-----------------------|
| Latest: Real-time | 0.83 | 0.95 | 1.07 | 0.81 |
| Latest: Quasi real-time | 0.79 | 0.93 | 1.03 | 0.77 |

5.3.4 Opposite Sign Tests

In Table 6 I show the proportion of observations where real-time and quasi real-time estimates have the opposite sign to the latest output gap estimate. Once again if data revisions were to have an important effect on the direction of policy we'd expect the quasi real-time estimates to match the sign of the latest estimates much more closely than the real-time estimates. While the results in Table 6 shows the proportion of opposite signs is lower across the various estimation techniques, it is not distinctly so. It should also be considered that many recordings of the wrong sign are when the estimated output gap is small and close to zero.

Table 6. Opposite sign test

| | Harvey-Clark | Harvey-Jaeger | Hodrick-Prescott | Christiano-Fitzgerald |
|-------------------------|--------------|---------------|------------------|-----------------------|
| Latest: Real-time | 33.2 | 39.8 | 40.3 | 35.5 |
| Latest: Quasi real-time | 31.3 | 33.6 | 37.4 | 28.9 |

5.3.5 Large Revisions

Table 7 shows the proportion of observations where revisions to the latest estimates from quasi real-time and real-time estimates are larger than the latest output gap measure itself. Again there is not conclusive evidence

that quasi real-time estimates significantly outperform real-time estimates of the output gap. Again, many of these instances of large revisions are recorded when the estimate of the output gap is close to zero.

Table 7. Large revisions

| | Harvey-Clark | Harvey-Jaeger | Hodrick-Prescott | Christiano-Fitzgerald |
|-------------------------|--------------|---------------|------------------|-----------------------|
| Latest: Real-time | 39.8 | 54.5 | 58.8 | 40.8 |
| Latest: Quasi real-time | 40.3 | 50.2 | 55.5 | 35.5 |

6. Concluding Comments

Looking at a real-time data set for UK GDP it is rare, but not impossible, for revisions to significantly alter the story being told by the data. The evidence suggests early estimates of GDP, especially over the last decade, have provided a relatively good signal of later data vintages. This should be good news for policy-makers who require both timely and accurate measures of economic activity.

This is particularly important in monetary policy where the independent Bank of England takes pre-emptive action to hit an inflation target in the medium-term. The Taylor rule identifies the importance of the output gap as a source of inflationary pressure in the economy and many researchers have concluded output gap mismeasurement has resulted in past mistakes in the operation of monetary policy.

Data revisions can feed directly into output gap measures. However, the analysis in this paper suggests it accounts for a relatively small part of output gap uncertainty. The difficulty in using statistical filters towards the end of the sample in order to place the potential level of output is the main cause of output gap revisions. This issue has been identified by Watson (2007) among others, and suggests hindsight is more valuable than smaller data revisions in estimating the output gap.

Despite this, the Bank of England and other monetary policy authorities around the world have heavily researched approaches to deal with what they describe as the ‘data uncertainty’ problem. Recent attempts have focused on the predictability of revisions by using bias adjustments or other data sources such as business survey data and sentiment indicators. Ashley et al (2005) and Cunningham and Jeffrey (2007) describe how the Bank of England use other data along with official estimates of GDP in forming their view on the strength of the economy.

Alternatively, monetary policy-makers may choose to respond less aggressively to uncertain data. Naturally this would imply a smaller coefficient on the output gap in the Taylor rule. Kozicki (2004) and Orthanides (2003b) note that this results in a more gradual movement in interest rates to output gap movements or an interest rate smoothing outcome. In a similar vein Jaaskela and Yates (2005) suggest the Taylor rule should put more weight on data that is not revised such as consumer price inflation or use lagged measures of the output gap which have already been revised.

Finally, monetary policy-makers might act to deal directly with measurement errors in GDP data. Walsh (2003) advocates speed limit rules in monetary policy based on the change in the output gap where errors will be less persistent than in the level of the output gap. Lomax (2004) suggests targeting nominal GDP rather than real GDP growth rates, as price deflators are often found to be a major source of revisions to real GDP growth.

In all cases though it is worth remembering that the pattern of GDP revisions over time has itself been unstable, prone to both cyclical influences and structural breaks as a result of the way the data is constructed or due changes in the structure and volatility of the economy. Therefore efforts to reduce data uncertainty may be successful only temporarily. It also suggests that for policy-makers general economic uncertainty in forecasting economic activity is still the bigger problem.

References

- Ashley, J., Driver, R., Hayes, S., & Jeffery, C. (2005). Dealing with data uncertainty. *Bank of England Quarterly Bulletin*, 23–29. Retrieved from <http://www.bankofengland.co.uk/publications/Documents/quarterlybulletin/qb050101.pdf>
- Bell, M. (2004). *Monetary policy, data uncertainty and the supply side: Living with the statistical fog*. Speech to the Society of Business Economists at the Chartered Institute of Public Finance and Accountancy (London), Wednesday 15th December 2004. Retrieved from <http://www.bankofengland.co.uk/archive/Documents/historicpubs/speeches/2004/speech228.pdf>

- Brown, G., Buccellato, T., Chamberlin, G., Dey-Chowdhury, S., & Youll, R. (2009). Understanding the quality of early estimates of Gross Domestic Product. Office for National Statistics. Retrieved from <http://www.ons.gov.uk/ons/rel/gva/gross-domestic-product--preliminary-estimate/understanding-the-quality-of-early-estimates-of-gross-domestic-product/understanding-the-quality-of-early-estimates-of-gross-domestic-product-.pdf>
- Cayen, J. P., & Van Norden, S. (2005). The reliability of Canadian output gap estimates. *North American Journal of Economics and Finance*, 16, 373–393. <http://dx.doi.org/10.1016/j.najef.2005.06.003>
- Christiano, L., & Fitzgerald, T. (1999). *The Band Pass Filter*. Federal Reserve Bank of Cleveland WP 99/06. Retrieved from <http://www.clevelandfed.org/research/workpaper/1999/wp9906.pdf>
- Clarida, R., Gali, J., & Gertler, M. (1998). Monetary policy rules in practice: Some international evidence. *European Economic Review*, 42, 1033–1067. [http://dx.doi.org/10.1016/S0014-2921\(98\)00016-6](http://dx.doi.org/10.1016/S0014-2921(98)00016-6)
- Clark, P. (1987). The cyclical component of US economic activity. *Quarterly Journal of Economics*, 797–814. <http://dx.doi.org/10.2307/1884282>
- Clark, P. (1989). Trend revision in real output and unemployment. *Journal of Econometrics*, 40, 15–32. [http://dx.doi.org/10.1016/0304-4076\(89\)90027-4](http://dx.doi.org/10.1016/0304-4076(89)90027-4)
- Clark, T., & Kozicki, S. (2005). Estimating equilibrium real interest rates in real time. *North American Journal of Economics and Finance*, 16, 395–413. <http://dx.doi.org/10.1016/j.najef.2005.04.002>
- Cunningham, A., & Jeffery, C. (2007). Extracting a better signal from uncertain data. *Bank of England Quarterly Bulletin*, 364–375. Retrieved from <http://www.bankofengland.co.uk/publications/Documents/quarterlybulletin/qb070301.pdf>
- Cusinato, R., Minella, A., & Porto, J. S. (2010). *Output gap and GDP in Brazil: a real-time data analysis*. Banco Central do Brasil Working Paper 203. Retrieved from <http://www.bcb.gov.br/pec/wps/ingl/wps203.pdf>
- Garratt, A., Lee, K., Mise, E., & Shields, K. (2008). Real-time representations of the output gap. *The Review of Economics and Statistics*, 90(4), 792–804. <http://dx.doi.org/10.1162/rest.90.4.792>
- Geberding, C., Seitz, F., & Worms, A. (2005). How the Bundesbank really conducted monetary policy. *North American Journal of Economics and Finance*, 16, 277–292. <http://dx.doi.org/10.1016/j.najef.2005.05.003>
- Gerdesmeier, D., & Roffia, B. (2005). The relevance of real-time data in estimating reaction functions for the euro area. *North American Journal of Economics and Finance*, 16, 293–307. <http://dx.doi.org/10.1016/j.najef.2005.04.001>
- Guering, P., Maurin, L., & Mohr, M. (2011). *Trend-cycle decomposition of output and euro area inflation forecasts*. European Central Bank Working Paper series no. 1384/October 2011. Retrieved from <http://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1384.pdf>
- Harvey, A. (1985). Trends and cycles in macroeconomic time series. *Journal of Business & Economic Statistics*, 3(3), 216–227. <http://dx.doi.org/10.1080/07350015.1985.10509453>
- Harvey, A., & Jaeger, A. (1993). Detrending stylized facts and the business cycle. *Journal of Applied Econometrics*, 8, 231–247. <http://dx.doi.org/10.1002/jae.3950080302>
- Herrmann, H., Orphanides, A., & Siklos, P. (2005). Real time data and monetary policy. *North American Journal of Economics and Finance*, 16, 271–276. <http://dx.doi.org/10.1016/j.najef.2005.06.002>
- Hodrick, R., & Prescott, E. (1997). Postwar US business cycles: an empirical investigation. *Journal of Money, Credit and Banking*, 29(1), 1–16. <http://dx.doi.org/10.2307/2953682>
- Ince, O., & Papell, D. (2012). The (un)reliability of real time output gap estimates with revised data. Retrieved from <http://www.uh.edu/~dpapell/Unreliability%20Real%20Time.pdf>
- Jaaskela, J., & Yates, T. (2005). *Monetary policy and data uncertainty*. Bank of England working paper no. 281. Retrieved from <http://www.bankofengland.co.uk/research/Documents/workingpapers/2005/WP281.pdf>
- Kamada, K. (2005). Real time estimation of the output gap in Japan and its usefulness for inflation forecasting and policymaking. *North American Journal of Economics and Finance*, 16, 309–332. <http://dx.doi.org/10.1016/j.najef.2005.05.004>
- Kozicki, S. (2004). How do data revisions affect the evaluation and conduct of monetary policy? *Federal Reserve Bank Kansas City Economic Review*, 5–38. Retrieved from <https://www.kansascityfed.org/publicat/econrev/PDF/1q04kozi.pdf>

- Ladiray, D., Mazzi, G., & Sartori, F. (2003). *Statistical methods for potential output estimation and cycle extraction*. Eurostat working paper, European Commission. Retrieved from http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-AN-03-015/EN/KS-AN-03-015-EN.PDF
- Lansing, K. (2002). *Learning about a shift in trend output: implications for monetary policy and inflation*. Federal Reserve Bank of San Francisco Working paper. Retrieved from <http://www.frbsf.org/economic-research/files/wp00-16bk1.pdf>
- Lomax, R. (2004). *Stability and statistics*. Speech to the North Wales Business Club dinner, Llandudno, Wales, 23rd. Retrieved from <http://www.bankofengland.co.uk/archive/Documents/historicpubs/speeches/2004/speech232.pdf>
- Marrcellino, M., & Musso, A. (2010). *Real time estimates of the euro area output gap: reliability and forecast performance*. European Central Bank working paper 1157. Retrieved from <http://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1157.pdf>
- Morley, J., Nelson, C., & Zivot, E. (2003). Why are Beveridge-Nelson and unobserved components of GDP so different? *The Review of Economics and Statistics*, 85(2), 235–243. <http://dx.doi.org/10.1162/003465303765299765>
- Murray, J. (2014). *Output gap measurement: judgement and uncertainty*. Office for Budget Responsibility Working Paper 5. Retrieved from <http://budgetresponsibility.org.uk/wordpress/docs/WorkingPaperNo5.pdf>
- Nelson, C., & Plosser, C. (1982). Trends and random walks in macroeconomic time series. *Journal of Monetary Economics*, 10, 139–162. [http://dx.doi.org/10.1016/0304-3932\(82\)90012-5](http://dx.doi.org/10.1016/0304-3932(82)90012-5)
- Nelson, E., & Nicolov, K. (2003). UK inflation in the 1970s and 1980s: the role of output gap mismeasurement. *Journal of Economics and Business*, 55, 353–370. [http://dx.doi.org/10.1016/S0148-6195\(03\)00030-4](http://dx.doi.org/10.1016/S0148-6195(03)00030-4)
- Oller, L. E., & Teterukovsky, A. (2006). *Quantifying the quality of macroeconomic variables*. Methodology reports from Statistics Sweden 2006:1. Retrieved from http://www.scb.se/statistik/_publikationer/OV9999_2006A01_BR_X103OP0601.pdf
- Orphanides, A. (2001). Monetary policies based on real-time data. *American Economic Review*, 91(4), 964–985. <http://dx.doi.org/10.1257/aer.91.4.964>
- Orphanides, A. (2002). Monetary policy rules and the great inflation. *American Economic Review Papers and Proceedings*, 92(2), 115–120. <http://dx.doi.org/10.1257/000282802320189104>
- Orphanides, A. (2003a). Historical monetary policy analysis and the Taylor Rule. *Journal of Monetary Economics*, 50, 983–1022. [http://dx.doi.org/10.1016/S0304-3932\(03\)00065-5](http://dx.doi.org/10.1016/S0304-3932(03)00065-5)
- Orphanides, A. (2003b). Monetary policy evaluation with noisy information. *Journal of Monetary Economics*, 50, 605–631. [http://dx.doi.org/10.1016/S0304-3932\(03\)00027-8](http://dx.doi.org/10.1016/S0304-3932(03)00027-8)
- Orphanides, A., & Van Noorden, S. (2002). The unreliability of output-gap estimates in real time. *The Review of Economics and Statistics*, 84(4), 569–583. <http://dx.doi.org/10.1162/003465302760556422>
- Orphanides, A., & Van Norden, S. (2005). The reliability of inflation forecasts based on output gap estimates in real time. *Journal of Money Credit and Banking*, 37(4), 583–601. <http://dx.doi.org/10.1353/mcb.2005.0033>
- Stock, J., & Watson, M. (1988). Variable trends in economic time series. *Journal of Economic Perspectives*, 2, 147–174. <http://dx.doi.org/10.1257/jep.2.3.147>
- Taylor, J. (1993). Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214. [http://dx.doi.org/10.1016/0167-2231\(93\)90009-L](http://dx.doi.org/10.1016/0167-2231(93)90009-L)
- Walsh, C. (2003). Speed limit policies: the output gap and optimal monetary policy. *American Economic Review*, 93(1), 265–278. <http://dx.doi.org/10.1257/000282803321455278>
- Watson, M. (1986). Univariate detrending methods with stochastic trends. *Journal of Monetary Economics*, 18, 49–75. [http://dx.doi.org/10.1016/0304-3932\(86\)90054-1](http://dx.doi.org/10.1016/0304-3932(86)90054-1)
- Watson, M. (2007). How accurate are real time estimates of output trends and gaps? *Federal Reserve Bank of Richmond Economic Quarterly*, 93(2), 143–161. Retrieved from https://richmondfed.org/publications/research/economic_quarterly/2007/spring/pdf/watson.pdf

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).