Real-time data for Norway: Output gap revisions and challenges for monetary policy

By

TOM BERNHARDSEN, ØYVIND EITRHEIM, ANNE SOFIE JORE

AND ØISTEIN RØISLAND*

Incomplete draft version

Abstract

Monetary policy conducted in real time has to take into account the preliminary nature of recent national accounts data. Not only recent data, but also figures dating many years back are potentially subject to revisions. This means that there is a danger that an important part of the central bank's information set is flawed for a substantial period of time. In this paper we present results based on quarterly vintages of real-time data for Norway from 1993Q1 to 2003Q4. We describe the nature and causes of the data revisions and investigate whether the revisions are true martingale differences or whether they can be forecasted. In the spirit of Orphanides and van Norden (2002), we analyze how data revisions and model uncertainty affect the reliability of output gap estimates. We find that total revisions of output gap estimates are heavily influenced by uncertainty about the trend at the end of the sample and that data revisions are of less importance, i.e., they are of smaller magnitude and show less persistence than other sources of output gap revisions. Finally, we analyze the implications of output gap uncertainty for monetary policy using a small New Keynesian macroeconomic model.

Keywords: Monetary policy, output gap, real-time data, interest rate rules

JEL-Classification: C53, E37, E52, E58

---

* We appreciate comments from the participants at the workshop "Real-time data and monetary policy", organized by Deutsche Bundesbank and North American Journal of Economics and Finance, Eltville, Germany, May 28-29, 2004. The first author is a senior advisor in the monetary policy department, the second author is a director in the research department, the third author is a senior advisor in the monetary policy department, and the fourth author is an assistant director in the monetary policy department in Norges Bank. The views expressed are those of the authors, and do not necessarily represent those of Norges Bank. Please address correspondence to: Øyvind Eitrheim, Research Department, C-51, Norges Bank (the Central Bank of Norway), Box 1179 Sentrum, N-0107 Oslo, Norway. Phone: +4722316161, Fax: +4722424062, Email: oyvind.eitrheim@norges-bank.no

1 Introduction

A central bank processes huge amounts of information when it assesses the state of the economy as part of the monetary policy-making process. Data uncertainty may disturb this assessment, and as data are subsequently are revised this complicates the evaluation of the conduct of monetary policy. Some variables, such as the level of production in a particular period, may be substantially revised over time. Thus, final data - available with a lag of 2-3 years - will typically deviate from real-time data. Hence, final data may - in retrospect - tend to point in the direction of some other path for the policy rate than the one chosen by the central bank on the basis of real-time information. This is not, of course, to say that the central bank decisions were based on bad judgment at the time they were taken. Nevertheless, it is the final data - and not real-time data - that determines what would have been the appropriate monetary policy in the past. In the process of conducting monetary policy, it is therefore important for the central bank to evaluate the consequences of the lack of accuracy of the available real-time information.

While inaccuracy of real-time information may apply to many macroeconomic variables which the central bank assesses when setting the interest rate, the problem is particularly severe for measures of production and economic growth. First, they are crucial input variables for monetary policy decisions: a reliable measure of current production is important for forecasting inflation, and the task of stabilizing the real economy under flexible inflation targeting is dependent on a sound assessment of the current state of the real economy. Second, production data are subject to frequent and sometimes substantial revisions. For other variables important in the monetary policy making process, like consumer price inflation, credit growth and wage growth, real time observations deviate less from final observations and the problems created by data revision are less severe.[1]

Academics and policy makers have recently invested more resources in this area, and there is a growing literature on the properties of real-time data and their consequences for current practices in monetary policy-making. The pioneering work by Croushore and Stark (1999, 2001) (see Croushore and Stark (2000) for a non-technical presentation of the real-time database for the US) has set a standard for the systematic work with real-time data and recent applications include Orphanides (2001), Stark and Croushore (2002) (with comments) and Orphanides and van Norden (2002). Kozicki (2004) provides an overview of this literature in the US and discusses how data revisions may affect the evaluation and conduct of monetary policy.

The output gap is frequently regarded as a basic summary measure of the state of the real economy, and as a theoretical concept, the output gap is a key monetary policy variable. In addition to the real-time data problems mentioned above, there are also important methodological problems associated with finding reliable estimates of the output gap. Orphanides and van Norden (2002) compare a wide range of different models and present an assessment of the reliability of output gap estimates in real-time. They argue that great caution is required and that output gap mismeasurement may pose a serious problem for the correct assessment of the state of the economy. Furthermore, they argue that disregarding this unreliability can lead to flawed policy

[1] The lack of revision in consumer prices does of course not guarantee that observations are accurate and free of biases that may distort policy decisions, cf. the discussion about whether the CPI takes sufficiently into account the quality changes that are typical for many consumer goods.

recommendations. In some situations it may even be advantageous to base monetary policy on information about output growth, and disregard the output gap (Orphanides et al., 2000). In the presence of a high degree of persistence in output gap mismeasurement, growth rates may be more reliable than output gap levels in real-time.

The literature on real-time data and monetary policy has so far been dominated by analysis on US data, typically following up on the seminal work by Dean Croushore and Tom Stark at the Federal Reserve Bank of Philadelphia. This paper provides evidence based on a quarterly real-time database for Norway, which consists of vintages of data from 1993Q1 to 2003Q4. Norway has since 2001 adopted a flexible inflation targeting regime. Details about the Norwegian real-time database with a special focus on mainland GDP are presented in Section 2 along with a descriptive summary of data properties. Section 2 also analyzes whether final data can be predicted on the basis of information available in real time. That is, can variables which are available at the same point in time as the preliminary data, help predict the final data or not? The two polar views are that revisions either represent *news*, hence they are unpredictable on the basis of contemporaneous information, or revisions tend to eliminate *noise* which is present in preliminary data. We test the *news* hypothesis in Section 2.3 using standard efficiency tests. In Section 3 we discuss the problem of estimating output gaps in the face of both model and real-time data uncertainty. We compare results from a subset of the univariate output gap models considered in Orphanides and van Norden (2001) and include output gaps based on a Cobb-Douglas production function. The different output gap estimates obtained from final data are compared with those from real-time data, and total revisions are decomposed into data revisions and other revisions. In Section 4 we discuss implications of output gap uncertainty for monetary policy using a small New Keynesian macroeconomic model. Section 5 provides some conclusions.

2 The real-time database for Norway

2.1 Constructing the database

Since 1993, Norges Bank's macroeconomic model RIMINI has been an important tool for forecasting in Norges Bank.[2] The model is estimated on quarterly data. After each round of forecasting (every third month until 2000 and every four month thereafter) the model's data-base and forecasts have been saved in vintages. This has created an opportunity to construct a real-time database, covering national accounts and other data.

Not all the vintages are complete, however, and additional work has to be done to construct a complete real-time database. So far our main focus has been on GDP for Mainland Norway in market value, measured in fixed prices. For the purpose of estimating output gaps using a production function, vintages of labour market, real capital and GDP data for selected mainland sectors were also constructed.

To construct a quarterly data-base, the starting point was published national accounts figures in Statistics Norway's Economic Bulletin. Figures for 6 to 8 quarters

[2] RIMINI has been used by the Central Bank of Norway for more than a decade to make forecasts for the Norwegian economy 4-8 quarters ahead as part of the Inflation report of the Bank, see Olsen and Wulfsberg (2001).

are published when new national accounts are available. Saved data in Norges Bank were appended to these figures.[3]

Some of the saved vintages were not complete, covering only the last three to four years. To construct a complete vintage, growth rates from other vintages with the same base year were used. In some cases, no vintages with the same base year as the incomplete vintage existed. In those cases, growth rates from some vintage with a different base year had to be used. This means that some of the vintages are not completely accurate. That should not, however, constitute major problems, as the change in historical growth rates associated with a new base year normally is minor.

2.2 Description of the database

There are three main sources for national accounts data to change over time. First, the earliest estimates are based on preliminary and incomplete information. Second, the base year is changed each year, and third, the national accounts are occasionally subject to major revisions.

In Norway quarterly, unadjusted national accounts data are published in the third month after the end of the quarter. Each time data for a new quarter are published, earlier data are also revised. In the second quarter of year (t+3), national accounts data for year (t) are final, and year (t) is the new base year.[4]

The first published figures are of course based on the most incomplete information, and naturally the figures will change when new information becomes available. The revisions can be quite substantial in the first few quarters after the initial publication. After 10 to 13 quarters, quarterly account figures are final, but year-on-year growth rates are in general fairly close to the final growth rates after 3 to 4 quarters. Changes of the base year obviously entail levels changes in historical national accounts data. To some extent, annual growth is also affected, but the effect is typically minor due to the frequency of the base year switch.

In the last ten years, there have been two major revisions of the national accounts:

- From 1995, the guidelines of the System of National Accounts SNA 1993 and European System of Accounts ESA 1995 were gradually implemented. In the third quarter of 1995, revised quarterly national account figures from 1993 to the second quarter of 1995 were published. Historical national accounts were revised stepwise. In 1997 data going back to 1978 were published. For research purposes, data from 1970 to 1978 have been recalculated by the Research Department in Statistics Norway. These are not official National Account statistics, but were made available for research purposes in the first quarter of 2001.

---

3 Labour market, real capital and GDP data for selected mainland sectors were constructed solely on the basis of saved data in Norges Bank.

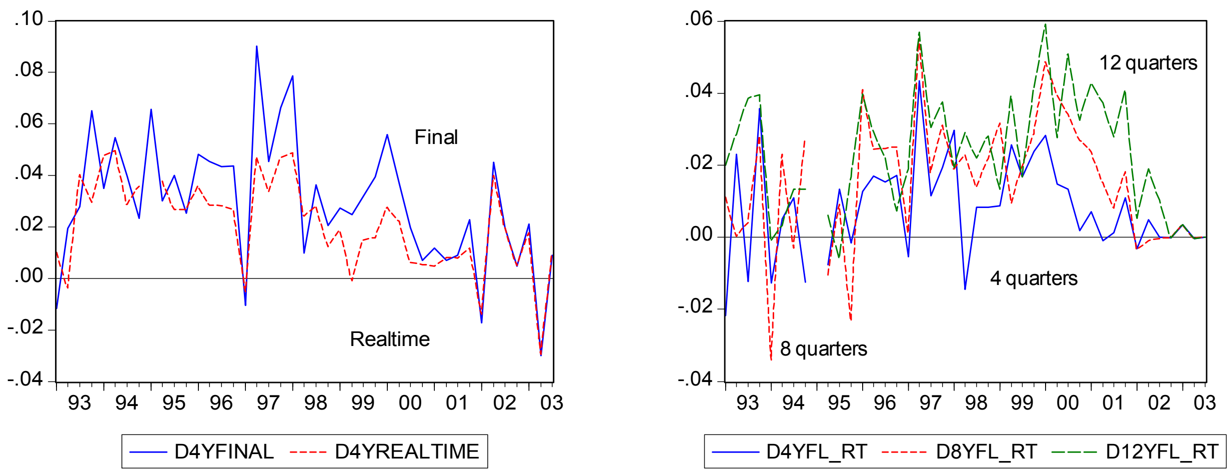
4 From 2004, the publishing of final national accounts figures and the base year change are moved forward. Final national accounts figures for year (t) are now published in the last quarter of year (t+2). In principle, the base year is changed every year. In connection with major revisions of the national accounts, however, the base year was kept unchanged for a longer period. The revisions are described below.

- Statistics Norway has collected new structural statistics for many industries over the last years. For some service industries the new statistics entailed changes that were deemed too substantial to be included on a continuing basis. Statistics Norway therefore decided to undertake a new revision of the national accounts, incorporating the new structural statistics in a coordinated way from 1991. In connection with this revision, some recalculations and corrections also affected national accounts in the previous decades. The revised figures were published in 2002.

The complete national accounts are based on unadjusted figures. Seasonally adjusted figures are presented for some main aggregates, but until recently these series have not been published as historical time series. Accordingly, the vintages that are saved in Norges Bank are all unadjusted. Both the major revisions led to substantial changes in the seasonal pattern.

Figure 1(a) depicts the year-on-year growth (indicated by D4Y) in the unadjusted real-time[5] and final data. The final data is the vintage published in 2003Q4. Growth rates are generally higher in the final data. The change in the level of GDP after several years of revisions is illustrated in Figure 1(b), showing accumulated revisions over 4, 8 and 12 quarters. For example, in the fourth quarter of 2001, the accumulated growth over the last 12 quarters turned out to be 4 percentage points higher than initially measured at the end of that year.

Figure 1: Final and real-time output growth, accumulated revisions over 4, 8 and 12 quarters. Vintages 1993Q1 - 2002Q1



(a) Final and real-time output growth

(b) Accumulated revisions

Looking at changes in average growth rates is another way of illustrating the overall effects of the data revisions. The main impact of the two major revisions is illustrated in Table 1, showing average annual growth rates of Mainland Norway GDP

[5] In connection with the major revision in 1995, national account figures were not published in 1995Q1.

over five-year periods. Base year changes also contribute to revised average growth rates. But compared to the major revisions, these effects are minor. The vintage 1995Q1 is the last published vintage prior to the main revision of 1995, the vintage 2002Q1 is the last published vintage prior to the revision in 2002 and the vintage 2003Q4 is the latest available national accounts at the time of writing.[6]

Table 1: Annual growth rate of Mainland Norway GDP over five-year periods

Vintage Period	1995q1	2002q1	2003q4
1974 to 1979	3,6	3,9	3,8
1979 to 1984	2,0	2,3	2,2
1984 to 1989	1,5	1,8	1,9
1989 to 1994	1,7	2,3	2,3
1994 to 1999		3,1	4,1
1991 to 1996		3,1	3,5
1996 to 2001		2,3	3,2

The 1995 revision raised the level of GDP, increasing average annual growth rates in the 1970's and 1980's by 0.3 percentage point. The assessment of the upturn in the first part of the 1990's also changed. Prior to the main revision, the national account figures indicated that the turning point after the pronounced downturn in the last half of the 1980's occurred in 1992. According to the revised account figures, the upswing started already in 1990/91.

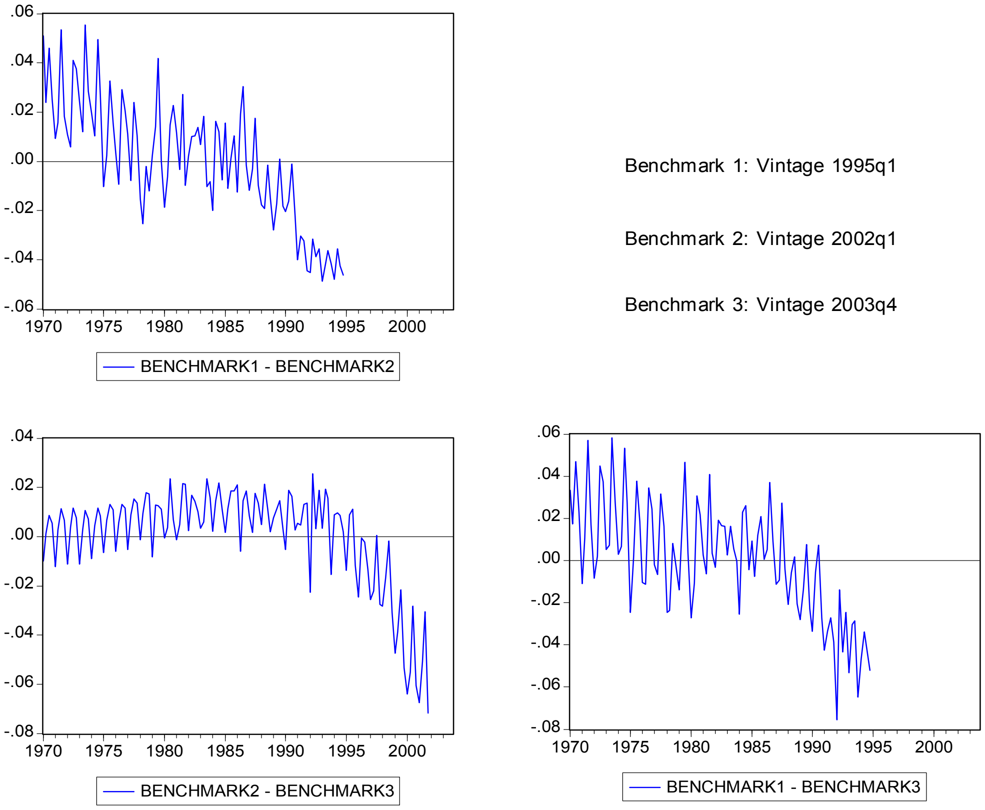
In the 2002 revision, new structural statistics were incorporated in the account figures from 1991. The impact of the 2002 revision is highlighted in the last two lines in Table 1. New structural statistics raised the average annual growth rate by 0.4 per cent in the first half of the decade and by 0.9 per cent in the second half. Some additional, partly methodological, changes also affected national accounts figures in the two previous decades.

Following Croushore and Stark (1999, 2001), the effect of the two main revisions on the level of GDP is also illustrated in Figure 2, depicting the differences between the log levels of the benchmark vintages used in Table 1, adjusted for mean differences between the vintages. The upper left panel illustrates the effect of the main revision in 1995. The lower right panel illustrates the combined impact of the main revision in 1995 and the revision in 2002. The left panel isolates the effect of the 2002 revision.

The log level ratio is decreasing in all three charts, further illustrating the upward shifts of GDP in the major revisions. The main revision in 1995 increased the GDP level for the whole period, but the upward revision was particularly sharp in 1991, at the beginning of the long upswing in the 1990's. The revision in 2002 further increased the GDP level. The changes in the seasonal pattern are pronounced in all the charts, creating of noise in the ratios.

[6] Each vintage contains national accounts figures up to and including the previous quarter.

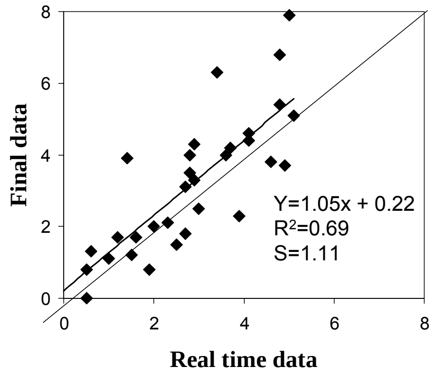
Figure 2: Log output ratios for three different vintages of real-time data



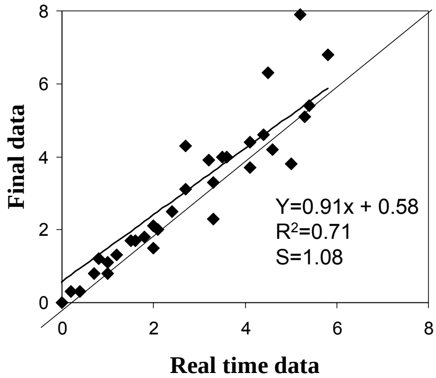
2.3 Can final growth data be predicted?

Three main reasons for revisions of national accounts data are described in the previous subsection. Changes due to incomplete information and base-year changes are occurring regularly, while major revisions normally are undertaken more seldom. In this subsection we aim to assess how reliable early estimates of national accounts data are in general, given available information. We investigate whether final data are predictable, given available information at the time of the first estimate. As major revisions are infrequent, and may lead to changing growth rates several decades after the earliest estimates, it does not seem reasonable to define final data across major revisions. For that reason, we leave out changes due to those revisions. The changes due to major revisions are taken into account partly by leaving out some observations, and partly by considering when growth rates can be judged to be final.

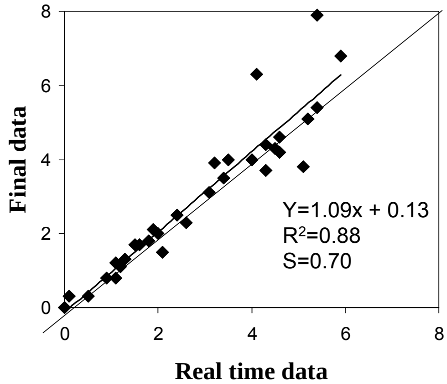
Figure 3: Real-time observations versus final data. Vintages 1993Q1 - 2002Q3



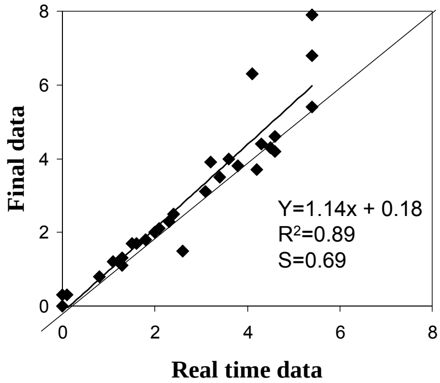
(a) 1. observation versus final data



(b) 2. observation versus final data



(c) 3. observation versus final data



(d) 4. observation versus final data

To obtain a broad picture of how growth data evolves over time, Figure 3(a) to 3(d) show real-time year-on-year growth at the time of publication and the revised growth rates in the three subsequent quarters, all four series in scatterplots along with the final revised growth. For example, Figure 3(c) shows revised growth two quarters after the time of publication and final revised growth. In the absence of real-time measurement problems, all data would lie on the 45-degrees line, meaning that real time data would not be revised. As time evolves, it is expected that the revised data will coincide with the final data, meaning that the numbers will approach the 45-degrees line over time. Note that in the figures the regression line and the standard deviation of the regression have also been included. For all the four regression lines, the hypothesis of a zero intercept and a coefficient equal to one turns out not to be rejected. This supports the view that real time growth rate mismeasurements do not contain a systematic pattern.

Figure 3(a) shows that real-time growth data deviate considerably from final data for many of the observations (the standard deviation of the regression is equal to 1.1 percentage point). For some observations the difference between real-time and final growth is substantial, indicating that in the worst case, incorrect real-time information may have serious consequences for monetary policy.

Turning to the evolution of growth data in the subsequent three quarters, the figures reveal that there are no substantial revisions in the second observation, as the standard deviation of the regression remains broadly unchanged. However, this picture changes with the third and fourth observations, as the standard deviations then take values of around 0.70 percentage point. While measurement errors are considerably reduced half a year after the data was published for the first time, some errors remains. After three quarters, an error of twice the standard deviation would be around 1.5 percentage point, not trivial in terms of assessing the consequences for monetary policy.

We now turn to the question of whether final growth can be predicted on the basis of real-time information. The revision process can be characterized in terms of two polar alternatives dubbed *news* and *noise* by Mankiw et al. (1984) and Mankiw and Shapiro (1986). In the *news* view real-time information contain all available information at the time of announcement. Faust et al. (2000) test this hypothesis on real-time and final quarterly GDP growth rates for the G-7 countries using a forecast efficiency test. This resembles the well-known efficiency test in finance, where the crucial question is whether future asset prices can be predicted on the basis of all kind of information publicly available at the time the test is performed. In our context, the question is to what extent the difference between final data and real-time data contains a systematic pattern. The test can be conducted on the basis of the regression

$$\Delta y_t^f - \Delta y_t^r = \alpha + \beta \Delta y_t^r + \gamma Z_t^r + \varepsilon_t \qquad (1)$$

where Δy_t^r is the growth rate published in real time (first time publication), Δy_t^f is the final growth rate, Z_t^r is a vector of real time controlling variables, which could possibly have explanatory power on the left-hand side variable and α, β and γ are coefficients to be estimated. Furthermore, ε_t is a disturbance term.

Under the null hypothesis $\alpha = 0$, $\beta = 0$, $\gamma = 0$ and ε is white noise. Then final growth data do not deviate systematically from real-time data, i.e., $\Delta y_t^f = \Delta y_t^r + \varepsilon_t$. If the null hypothesis does not hold, i.e., if at least one of the coefficients differs from zero, or if the disturbance term contains a systematic pattern, deviations from real-time data can be predicted. In that case we have some information, not embedded in the real-time growth data, which could help us predicting the final revised growth data. This again, could be valuable information in the monetary policy making process.

To test for this, a large set of macroeconomic variables, which could have an explanatory power on the left-hand side variable, were included in the model (one by one). Table 2 shows the F-statistic (and the corresponding probability value) for the joint null hypothesis that the coefficients and the constant term in regression (1) are zero. For all regressions the null hypothesis could not be rejected. Hence available real-time macroeconomic information indicate that final revised growth cannot be predicted beyond the information contained in the data published in real-time.[7]

[7] One should note that the appropriate testing procedure would normally be to include all, or at least some explanatory variables in the regression simultaneously. However, in our case, due to lack of degrees of freedom the additional explanatory variables were included only one by one in the regression.

Table 2: Omitted variable tests for additional effects on revisions from macroeconomic variables

Labour market variables	
New Jobs	$F_{OVT}(3, 30) = 0.1872$ [0.8303]
Vacancies	$F_{OVT}(3, 30) = 0.2616$ [0.7716]
Employment and vacancies	$F_{OVT}(3, 30) = 0.1814$ [0.8350]
Unemployment	$F_{OVT}(3, 30) = 0.2298$ [0.7961]
Δ(Unemployment)	$F_{OVT}(3, 30) = 1.5212$ [0.2354]
Hours worked	$F_{OVT}(3, 30) = 0.2616$ [0.7716]
Goods market variables	
Industrial production	$F_{OVT}(3, 30) = 0.1144$ [0.8923]
Δ(Industrial production)	$F_{OVT}(3, 30) = 0.3211$ [0.7279]
Retail sales	$F_{OVT}(3, 30) = 0.069$ [0.9338]
Δ(Retail sales)	$F_{OVT}(3, 30) = 0.2422$ [0.7864]
New orders	$F_{OVT}(3, 30) = 0.0671$ [0.9352]
Δ(New orders)	$F_{OVT}(3, 30) = 0.3681$ [0.6952]
Industrial investment	$F_{OVT}(3, 30) = 0.2616$ [0.7716]
Δ(Industrial investment)	$F_{OVT}(3, 30) = 0.4538$ [0.6397]
Bankruptcies	$F_{OVT}(3, 30) = 0.3716$ [0.6928]
Financial market variables	
Credit growth, C1	$F_{OVT}(3, 30) = 0.0700$ [0.9324]
Δ(Credit growth, C1)	$F_{OVT}(3, 30) = 0.6087$ [0.5509]
Credit growth, C2	$F_{OVT}(3, 30) = 0.0698$ [0.9327]
Δ(Credit growth, C2)	$F_{OVT}(3, 30) = 0.7627$ [0.4755]
Credit growth, C3	$F_{OVT}(3, 30) = 0.0682$ [0.9342]
Δ(Credit growth, C3)	$F_{OVT}(3, 30) = 1.1621$ [0.3269]
Nominal effective exchange rate	$F_{OVT}(3, 30) = 0.3213$ [0.7277]
Δ(Nominal effective exchange rate)	$F_{OVT}(3, 30) = 0.3036$ [0.7405]
Slope of the yield curve	$F_{OVT}(3, 30) = 0.8261$ [0.1922]

Similar hypotheses have also been analyzed by Faust et al. (2000) using data for the G-7 countries. The results reported for the period 1988Q1 to 1997Q4 indicate that the *news* hypothesis can be firmly rejected for countries like Germany, Italy, Japan and UK, where GDP-revisions seem to be highly predictable[8]. For Canada, France and US the results in Faust et al. (2000) indicated that revisions were unpredictable at the 10 per cent confidence level over the period 1988Q1 to 1997Q4. This is supportive evidence for the *news* view and holds for the case with no additional variables Z_t , i.e., assuming $\gamma = 0$. When additional variables - available at the same time as the preliminary GDP-figures - were included in the regression, it turned out that one or more of them were significant in the analysis of data for G-7 countries.

[8] In fact the results in Faust et al. (2000) yield some support to the *noise* view, i.e., that subsequent revisions tend to remove measurement errors.

3 Output gap estimates in real-time

3.1 Models

The output gap is usually defined as the deviation of actual output from potential output. Traditionally, potential output is measured by some sort of trend. From an operational point of view, the output gap is thus the deviation of output from its trend.[9]

A detrending method decomposes (the log of) real output, y_t, into a trend component, μ_t, and a cyclical component, z_t, see Orphanides and van Norden (2002) for an overview of the relative merits of different detrending methods using final and real-time data. We can accordingly write

$$y_t = \mu_t + z_t, \tag{2}$$

where the cyclical component, z_t, may be used as a measure of the output gap, defined as $ygap_t = y_t - \mu_t$. There is considerable uncertainty with respect to the measurement of potential output, and in this paper we will use estimates of the trend, μ_t, as our estimate of the potential output. In the following we consider four models of the output gap:[10]

QT	Quadratic trend	$\mu_t = \alpha + \beta t + \gamma t^2 + \varepsilon_t$
UC	Harvey (1985), Clark (1987) (local trend model)	$\mu_t = \delta_{t-1} + \mu_{t-1} + \eta_t$ $\delta_t = \delta_{t-1} + \upsilon_t$ $z_t = \rho_1 z_{t-1} + \rho_2 z_{t-2} + \varepsilon_t$
HP	Hodrick-Prescott (λ=1600)	$\mu_t = \arg\min \sum_{t=1}^{T} \left\{ (y_t - \mu_t)^2 + \lambda \left(\Delta^2 \mu_{t+1} \right) \right\}$
PF	Production Function model	$\mu_t = \hat{\alpha} + 0.67 l_t^* + 0.33 k_t^* + tfp_t^*$

Before proceeding, we will take a closer look at the production function method. In contrast to univariate models, the production function model takes account of the underlying structure of the economy. We follow the approach in Nymoen and Frøyland (2000), basing the calculations on a production function for manufacturing industries,

[9] This measure is, however, not necessarily consistent with the definition from the more recent New Keynesian theoretical framework. The sort of gap the central bank should attempt to close according to theory, is the deviation of output from the output level that would have occurred if all prices and wages were fully flexible. Despite the theoretical attractiveness of the flex-price output gap, it is extremely difficult to measure, and most central banks use more traditional methods of detrending.

[10] Four additional univariate models are considered in Bernardsen et al. (2004), essentially matching the set of univariate models presented in Orphanides and van Norden (2002). The estimation results for the unobserved component (UC) models in Orphanides and van Norden (1999, 2002) are based on Kalman filter algorithms in the TSM-module in GAUSS. The only UC model reported here is the Harvey-Clark model with local trend. We are grateful to Simon van Norden for providing access to his procedures written in RATS and Gauss for estimating the univariate models.

building and construction and retail trade, which accounts for about 75 per cent of GDP in the mainland Norwegian economy. Potential output may be interpreted as representing the supply side of the economy, and output gaps accordingly represent excess demand or supply.

The aggregated production function is assumed to be of a Cobb-Douglas type with constant returns to scale[11]:

$$y_t = \alpha_0 + \alpha_1 l_t + (1-\alpha_1)k_t + e_t \qquad (3)$$

where the variables y_t, (production, i.e. value added), l_t, (person-hours) and k_t, (capital stock) are measured as logarithms. e_t, represents total factor productivity, α_1 and (1- α_1) are elasticities and α_0 is a constant. The elasticities are given by the income factor shares of the two production factors (see Nymoen and Frøyland (2000) for details).

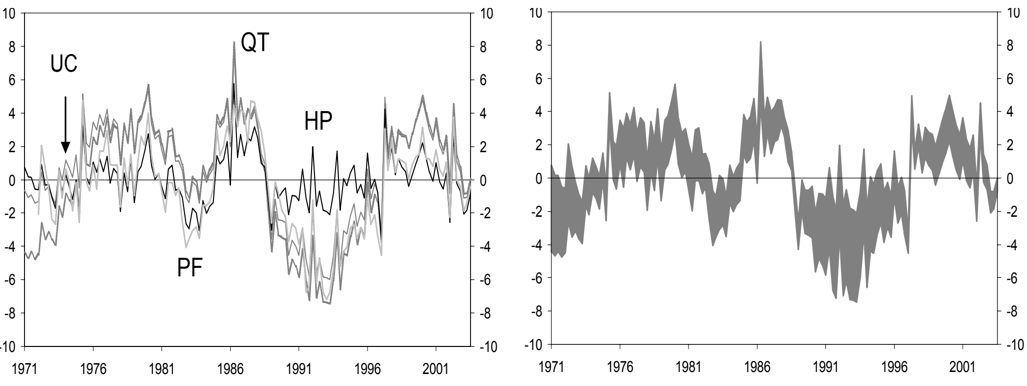
Potential output is the hypothetical level of output when person-hours, capital stock and total factor productivity are at their equilibrium (potential) levels. These levels are not observable and must be estimated. Potential person-hours is a function of the potential levels of the labor force, unemployment and average working hours per employee. These levels are estimated using a Hodrick-Prescott filter. Total factor productivity is calculated as the residual from estimating equation (3) using the least squares method. The residuals are then smoothed by means of a Hodrick-Prescott filter. The capital stock is assumed to be fully utilized at any time (cf. Bernhardsen et al., 2004).

3.2 Output gap revisions

Figure 4(a) shows estimated output gaps from final data (i.e., vintage 2003Q4) over the last three decades for each of the four models. In Figure 4(b) we illustrate the output gap uncertainty in the sense of "thick modelling" (Granger and Jeon (2004)) by means of the envelope of the trajectories in Figure 4(a). We observe that the envelope is quite wide in the early 1970s and early 1990s, but one should be aware of the possibility that the envelope is delineated by "outlier models", cf. e.g., the HP model in Figure 4(a) which shows a smaller negative output gap in the early 1990s than the other models. Within this period, we have observed three pronounced periods of expansion. There are no authoritatively determined business cycles, or dating of recessions or output gaps in Norway. There is, however, general agreement that the upswing in the 1990's started in 1991 and peaked in 1998 with the international financial turmoil. Activity stayed at a high level until 2000. Since then growth has decreased and output is generally viewed to have been below potential in 2003. The second quarter of 2003 seems to be the starting point of a new expansionary period.

[11] As in Nymoen and Frøyland (2000), we use the calculation method from the OECD, described in Giorno et al. (1995).

Figure 4: Final output gaps. Vintage 2003Q4



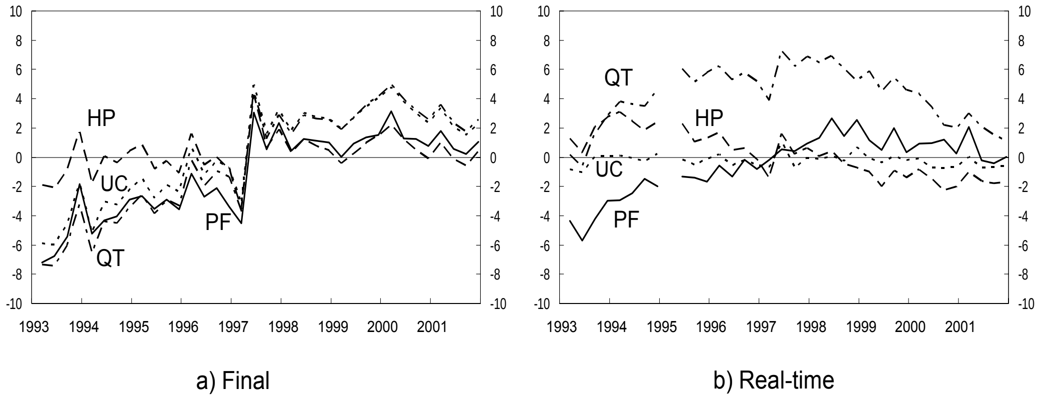
a) Four output gaps b) Thick modelling

The period covered by the real-time data-base thus starts close to the beginning of a cycle that peaks in the late 1990's and ends with a trough in 2003. Output gaps estimated on final and real-time data for each of the models described in the previous section are presented in Figures 5(a) and 5(b). Real-time output gaps are calculated for vintages 1993Q1 to 2002Q1.[12] The figures reveal that the size of the output gaps covers a wide range, particularly measured in real-time. Output gaps measured by the alternative models generally move in the same direction, both as real-time and as final gaps. Two of the real-time output gaps are negative in the last half of the 1990s, while the final output gaps are generally positive since 1997.

We follow Orphanides and van Norden (1999), analysing the revisions of output gaps for each model. Total revisions, calculated as the difference between final and real-time estimates of the output gaps, have two main sources: revisions of national accounts data and effects stemming from new observations as time passes. To decompose total revisions, we calculate quasi-real output gaps. These are constructed in the same way as real-time output gaps, i.e., using only data up to and including the period in question, but instead of using preliminary vintages of real-time data, we use final data truncated at the relevant period.

[12] Real-time estimates are calculated in the following way, described in Orphanides and van Norden (1999): For each vintage, output gaps are estimated. The last value in each of the output gap series is the first available estimate of the output gap for that particular period. These estimates are picked from each output gap vintage, forming a time series of real-time output gaps.

Figure 5: Final and real-time output gaps. Vintages 1993Q1 – 2002Q1



The differences between quasi-real and real-time output gaps are entirely due to revisions of data being used over the sample range. We define these as data revisions. The differences between final and quasi-real output gaps are associated with the addition of new observations, i.e., for the periods *after* the relevant period, and we have referred to these differences simply as *other revisions*. Other revisions have different sources in different models. In models with two-sided filters, these revisions are associated with end-point instability problems.[13] In UC models, other revisions are mainly caused by parameter instability.

Descriptive statistics for estimates of real-time, quasi-real and final output gaps and for total revisions are shown in Table 3. Correlations between real-time and final output gaps are generally low, except for the PF model. The means and standard deviations of total revisions are generally quite large. Typically, we observe that for all models except for the PF-model, the magnitude of the mean or the standard deviation of total revisions is larger than the corresponding statistic of the final output gap. Total revisions exhibit a particularly high degree of persistence in the QT model, with an autocorrelation coefficient equal to 0.94. Persistence is substantial also in the UC and HP models. Again, the PF model is an exception with an autocorrelation coefficient of 0.38.

[13] See Bernhardsen et al. (2004) for a brief discussion of end-point problems in the HP model.

Table 3: Summary statistics for output gaps and total revisions (levels). Vintages 1993Q1 – 2002Q1

Method	MEAN	SD	MIN	MAX	CORR	AR
Quadratic Trend (QT)						
RTGAP	4.23	2.00	-0.67	7.30	0.33	
QRGAP	5.22	2.26	0.22	11.10	0.64	
FLGAP	-0.22	3.70	-7.43	4.97	1.00	
Total revisions	-4.39	3.57	-10.35	1.51		0.94
Harvey-Clark (UC)						
RTGAP	-0.25	0.43	-1.06	1.02	0.22	
QRGAP	-0.17	0.72	-1.41	2.81	0.30	
FLGAP	0.28	3.14	-5.98	4.95	1.00	
Total revisions	0.58	3.10	-5.28	4.91		0.83
HP1600 (HP)						
RTGAP	0.16	1.56	-2.25	3.09	-0.01	
QRGAP	0.41	1.70	-2.35	4.83	0.39	
FLGAP	0.20	1.39	-3.59	4.23	1.00	
Total revisions	0.02	2.13	-4.92	3.06		0.73
Production Function (PF)						
RTGAP	-0.43	1.96	-5.69	2.68	0.87	
QRGAP	0.06	2.20	-5.10	5.05	0.95	
FLGAP	-1.08	2.83	-7.21	3.12	1.00	
Total revisions	-0.65	1.51	-4.34	2.55		0.38

RTGAP is real-time output gaps, QRGAP is quasi-real output gaps and FLGAP is final output gaps. MEAN is mean value, SD is standard deviation, MIN and MAX are minimum and maximum values, respectively. CORR is the correlation between final output gaps and, respectively, real-time and quasi-real output gaps. AR is the first order autocorrelation coefficient.

To facilitate comparisons between different models, some measures independent of the size of the gaps are presented in Table 4. The statistics are indications of the reliability of the real-time output gaps compared to the final output gaps, in the sense that they are measures of how different final output gaps are from real-time output gaps. The statistics do not say anything about how reliable the different models are as tools for measuring of the "true" output gap.

Table 4: Measures of reliability of output gaps (levels). Vintages 1993Q1 – 2002Q1

Method	CORR	N/S	OPSIGN	XSIZE
Quadratic Trend (QT)	0.33	1.530	0.44	0.64
Harvey-Clark (UC)	0.22	1.005	0.53	0.53
HP 1600	-0.01	1.532	0.53	0.75
Production Function (PF)	0.87	0.580	0.06	0.17

CORR is the correlation between final and real-time output gaps. N/S is a noise-to-signal ratio, the ratio of the root-mean-squared revisions divided by the standard deviation of the final estimates of the output gap. OPSIGN measures the frequency with which real-time and final estimates have opposite signs. XSIZE shows the frequency with which the absolute value of the total revision is larger than the absolute value of the Final output gaps.

The noise-to-signal ratio reported in Table 4 indicates a high degree of noise in the real-time output gaps calculated by the QT, UC and HP models. These models also produce opposite signs in real-time and final output gaps with a frequency of around 0.50, and in 50 to 75 per cent of the time the absolute value of the total revision is larger than the absolute value of the final output gaps. In contrast the PF model seems to be an exception in the sense that the different measures of reliability indicate that this model is more reliable in real-time than the other models.

The statistics in Tables 3 and 4 confirm the visual impression from Figures 5(a) and 5(b). The reliability of the various models is in general poor. Total revisions are large and persistent, and correlations between real time and final estimates are low. Output gaps estimated by the PF model exhibits somewhat more favorable statistics than output gaps produced by univariate filtering models. Compared to the US data analyzed by Orphanides and van Norden, the real-time univariate estimates of the output gap based on Norwegian data seem in general to be even less reliable than similar estimates for the US. This points in the direction of applying some caution and a fair amount of judgment when assessing the level of the output gap in real-time.

In addition, as pointed out in Orphanides and van Norden (1999), the approach taken here probably overestimates the precision and accuracy in all the detrending models. The size of the revision errors measured here must be interpreted as the lower limit of the real revision errors.

3.3 Decomposition of Output Gap Revisions

Figures 6(a) – 6(d) show detailed pictures of total revisions and their decomposition into data revisions and other revisions for each of the four models. The figures indicate that for all models, the addition of new observations is more important than data revisions in explaining total revisions. Table 5 provides summary statistics.

Total revisions in the QT model are large, with a mean value of -4.39 per cent. Data revisions increase the output gap by approximately 1 per cent in average, while the mean of other revisions is -5.38 per cent. At the other end of the spectrum, mean revisions in the HP model is only 0.02 per cent, as data and other revisions are of the same size but of opposite signs.

The persistence of total revisions measured by their first order autocorrelation coefficient varies between 0.94 for the QT model and 0.38 for the PF model. The persistence of other revisions is large for all models, typically showing autocorrelation coefficients above 0.90. We find that other revisions account for the bulk of the persistence in total revisions of the output gap, and this tendency is more pronounced in Norwegian than in US real-time data, cf. Orphanides and van Norden (2002, Table 4). The degree of persistence in data revisions is low compared with other revisions. This is consistent with the lack of predictability of future revisions of output growth data reported in section 2.

Figure 6: Total revisions, data revisions and other revisions

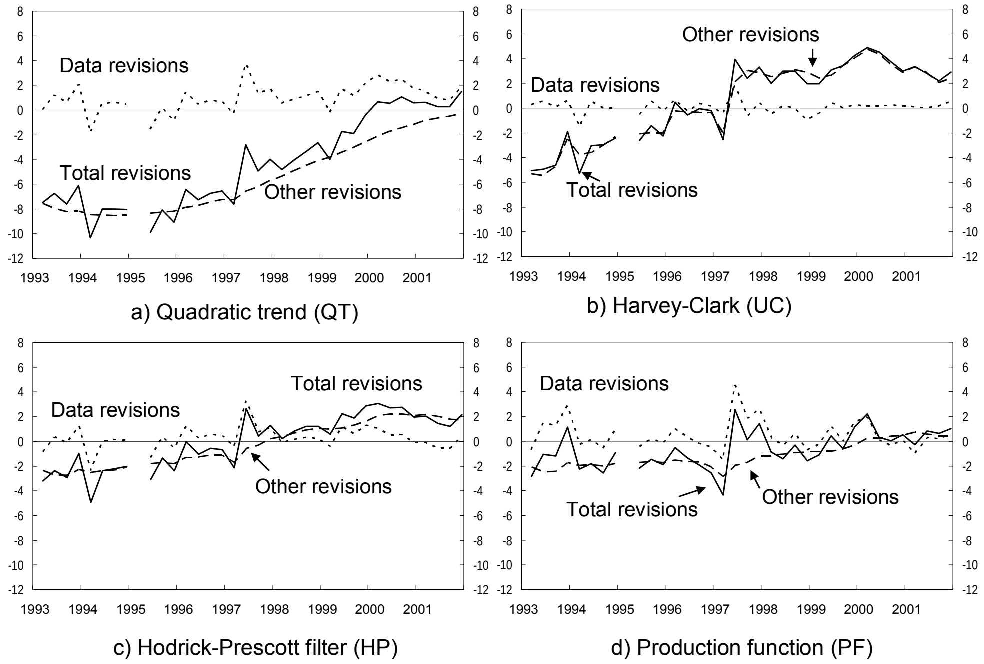


Table 5: Summary statistics for output gap revisions (levels): Vintages 1993Q1 – 2002Q1

Method	MEAN	SD	MIN	MAX	AR	N/S
Quadratic Trend (QT)						
Total Revision	-4.39	3.57	-10.35	1.51	0.94	1.530
Data Revision	0.99	1.14	-1.87	3.79	0.53	0.407
Other Revision	-5.38	2.87	-8.58	-0.33	0.98	1.648
Harvey-Clark (UC)						
Total Revision	0.58	3.10	-5.28	4.91	0.83	1.005
Data Revision	0.08	0.55	-1.49	1.79	-0.41	0.176
Other Revision	0.50	3.03	-5.45	4.75	0.91	0.979
HP1600 (HP)						
Total Revision	0.02	2.13	-4.92	3.06	0.73	1.532
Data Revision	0.25	0.91	-2.37	3.25	0.04	0.679
Other Revision	-0.23	1.75	-2.77	2.21	0.96	1.270
Production Function (PF)						
Total Revision	-0.65	1.51	-4.34	2.55	0.38	0.580
Data Revision	0.49	1.20	-1.50	4.53	0.23	0.458
Other Revision	-1.14	1.00	-2.84	0.70	0.95	0.536

For explanations, see explanations to Tables 3 and 4.

The noise-to-signal ratios for total and other revisions are 1 or higher for the QT, UC and the HP models, indicating large and biased revisions. The ratios are markedly smaller for data revisions. The PF model exhibits less biased revisions, with a noise-to-signals ratio of around 0.50 for all revisions.

The sharp increase in GDP in the second quarter of 1997[14] leads to a shift in other revisions for the UC model, indicating parameter instability. In the HP model other revisions are small in the middle of the period, but increase both in the beginning and in the end of the period – indicating end-point instability.

Our main conclusions are that revisions to real-time output gaps are large and persistent, and that the addition of new data is the main source of revisions for all the models. We do not find all output gap estimates equally reasonable. In real time, the HP and UC models produce output gaps that are close to zero or positive in the first half of the 1990s and negative from 1998. Such a development is contrary to the general consensus about developments in the Norwegian economy. Output gaps calculated by the QT model are positive for virtually the whole period, which also seems unreasonable.

Of the models considered here, the PF model exhibits somewhat more favourable properties than the univariate models, including more reasonable output gaps in real-time. It remains to be tested, however, how well real-time output gaps behave in the context of a larger model.

3.4 Revisions of levels versus revisions of changes

Figures 5(a) and 5(b) indicate that output gaps calculated on the basis of final and real-time data generally move in the same direction. Revisions of the *change* in the output gap may therefore be better behaved than revisions of the *level* of the output gaps. Orphanides et al. (2000), Orphanides (2003) and Walsh (2003a,b) have investigated this hypothesis for US data, and Cayen and van Norden (2004) study revisions in the change in output gaps for Canadian data.

In Table 6 we present some statistics comparing revisions of changes and levels of output gaps for Norwegian data. For most of the models, mean values for the revisions of the change in the output gap are smaller than the corresponding figures for revisions in output gap levels. The exception is the HP model, where the mean value of the revisions is higher for the change in the output gap than for its level. Standard deviations of the revisions are also generally smaller for changes than for levels, except for the PF model. Compared with the findings for US and Canadian data, the gains of converting real time output gap levels to changes in the output gaps seem small. The mean values and standard deviations of revisions in the changes in the output gaps are still substantial. This is confirmed by the statistics in the four last columns of Table 6. The maximum revisions are higher for changes than for levels. The auto-correlation coefficients lie in the area -0.5 to -0.7 for all models. Improvements in the noise-to-signal ratio are small, and for the PF model the noise-to-signal ratio increases.

One reason for the substantial remaining volatility in the revisions of the change in the output gaps can be traced to the development in 1997. Figure 7 illustrates this for the PF model, which is extreme in the sense that the noise-to-signal ratio increases by more than 60 per cent moving from levels to changes. In the final data, the *level* of the output gap increases from -4.5 per cent in the first quarter to 3.1 per cent in the second quarter. The *change* in the output gap moves from -1.1 per cent to 7.6 per cent – around 1 percentage point more than for the level. The increase in the real-time output gap is less than 1 percentage point measured both in levels and changes, meaning that revisions are greater for the changes than for levels.

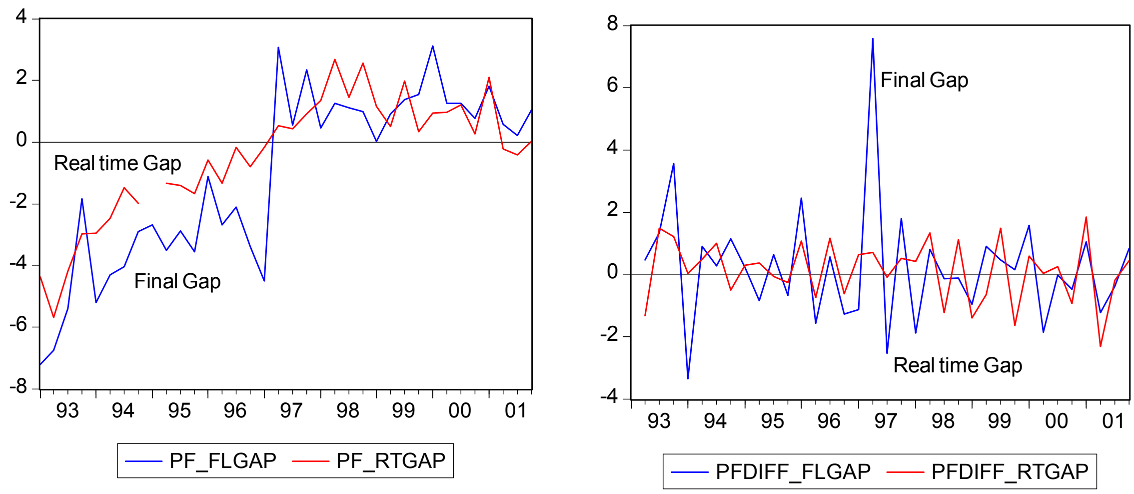
[14] The development in 1997 is a puzzle, showing up after the last major revision of the National accounts.

Table 6: Summary statistics for total revisions (levels versus changes). Vintages 1993Q1 – 2002Q1

FORM	MEAN	SD	MIN	MAX	AR	N/S
Quadratic Trend (QT)						
Level	-4.39	3.57	-10.35	1.51	0.94	1.530
Difference	0.26	1.59	-4.02	4.80	-0.59	0.806
Harvey-Clark (UC)						
Level	0.58	3.10	-5.28	4.91	0.83	1.005
Difference	0.21	1.67	-3.38	6.47	-0.54	0.844
HP1600 (HP)						
Level	0.02	2.13	-4.92	3.06	0.73	1.532
Difference	0.14	1.51	-3.91	3.69	-0.67	0.767
Production Function (PF)						
Level	-0.65	1.51	-4.34	2.55	0.38	0.580
Difference	0.12	1.79	-3.40	6.89	-0.48	0.945

For explanations, see explanations to Tables 3 and 4.

Figure 7: Real-time and final output gaps calculated by the production function model. Levels and changes. Vintages 1993Q1 - 2002Q1



(a) Levels

(b) Percentage change

In conclusion, improvements in reliability from using revisions of the *changes* in the output gaps instead of revisions of their *levels* are generally smaller for Norwegian than Canadian and US data. The most reliable model, the PF model, exhibited less favourable properties when moving from levels to changes. In the next section, we consider the impact of output gap uncertainty for monetary policy.

4 Monetary policy in real-time - output gap uncertainty

The output gap is a key variable in interest rate decisions. First, the output gap affects inflation, and the central bank must respond to the gap in order to stabilize inflation. Second, the output gap is an independent term in the loss function under flexible inflation targeting. In the short and medium run there may be a tradeoff between stabilizing inflation around the target and stabilizing the output gap, and the central bank must strike a balance between the two objectives.

An important question is whether output gap uncertainty has implications for how central banks should respond to their estimate of the gap. Svensson and Woodford (2003) show that in models with forward-looking variables, the optimal response to an optimal estimate of potential output displays certainty equivalence. They find, however, that the optimal response to an imperfect observation of output depends on the noise in this observation.

Monetary policy is often described by simple instrument rules, like the Taylor rule, rather than the complex optimal rules considered by Svensson and Woodford. It is well known that certainty equivalence does not hold if the central bank follows simple rules, and Smets (2002) finds that the optimal coefficients in the Taylor rule are smaller under output gap uncertainty. There is an ongoing debate on whether monetary policy should be described by minimizing a loss function (optimal policy) or whether it is better described by simpler Taylor-type instrument rules, and we will not follow up this debate here. Because of the intuitive appeal of simple instrument rules like the Taylor rule, we will discuss the challenges of output gap uncertainty for monetary policy using such rules. We will apply a simple version of an aggregated New Keynesian macroeconomic model for this purpose.

The model is given by

$$\pi_t = 0.8\pi_{t-1} + 0.2E_t\pi_{t+1} + \gamma y_t + 0.1z_t + \varepsilon_t^{\pi} \tag{5}$$

$$y_t = 0.85y_{t-1} + 0.1E_t y_{t+1} - 0.1(i_{t-1} - E_{t-1}\pi_t) + 0.05z_{t-1} + \varepsilon_t^{y} \tag{6}$$

$$z_t = 0.4z_{t-1} + 0.6E_t z_{t+1} - 0.2\{(i_t - E_t\pi_{t+1}) - (i_t^f - E_t\pi_t^f)\} + \varepsilon_t^{z} \tag{7}$$

$$i_t = \alpha_i i_{t-1} + \alpha_\pi \pi_t + \alpha_y y_t^0 + \alpha_{\Delta y}\Delta y_t^0 \tag{8}$$

$$y_t^0 = y_t + \varepsilon_t^0 \tag{9}$$

$$\varepsilon_t^0 = \rho\varepsilon_{t-1}^0 + \eta_t^0 \tag{10}$$

The parameters in the model have been calibrated to map properties of the Norwegian economy.[15] The first equation is a "hybrid" open-economy New Keynesian Phillips curve, where π_t is the rate of (CPI) inflation, y_t is the "true" (but unobservable) output gap, z_t is the (log of) the real exchange rate, measured as deviation from the equilibrium real exchange rate, and ε_t^{π} is a cost-push shock. The second equation represents aggregate demand, where i_t is the nominal short-term interest rate. $i_t - E_t\pi_{t+1}$ is then the real interest rate, and the neutral real interest rate is for simplicity normalized to zero. Equation (7) is the UIP condition, where a lag is introduced to better

[15] See Husebø, McCaw, Olsen and Røisland (2004) for a discussion of a version of the model.

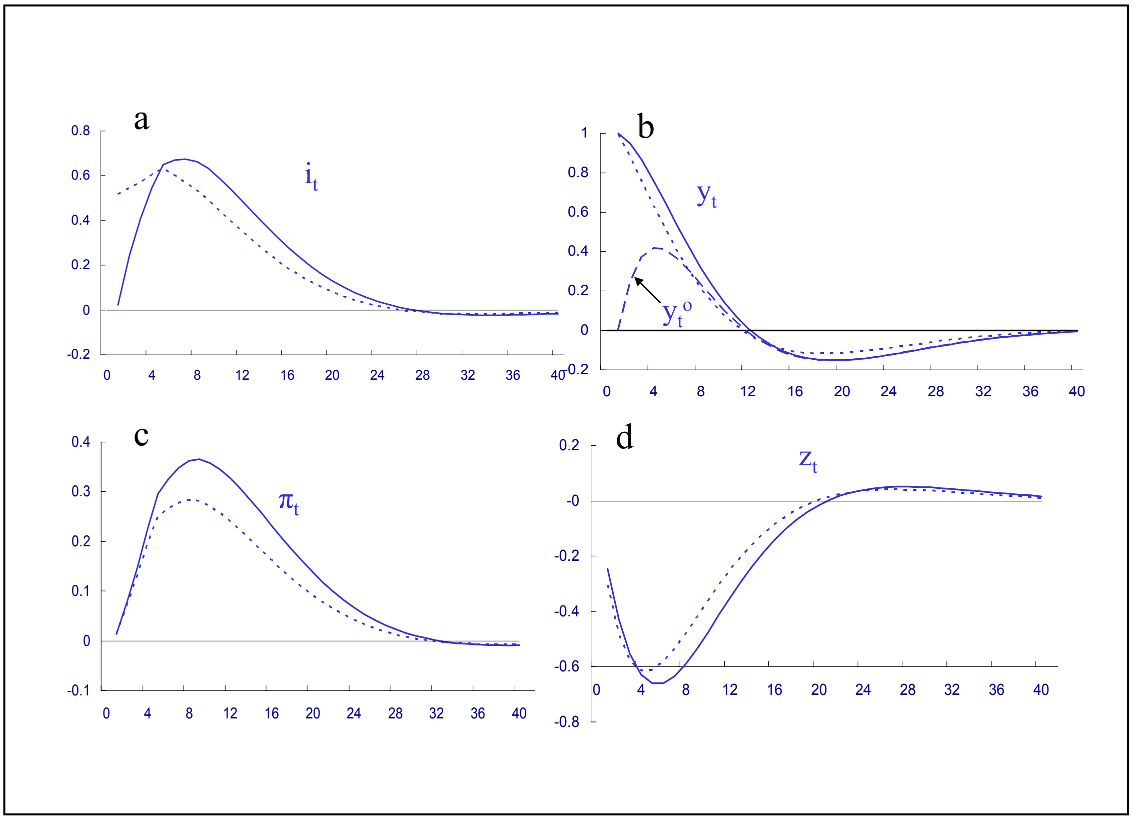
capture observed dynamics in the real exchange rate (short-run deviations from UIP). The coefficient of 0.2 on the real interest rate differential ensures that UIP holds in steady state. Equation (8) represents the monetary policy rule, which is a generalized Taylor rule. Since the central bank cannot observe the true output gap, it responds to its real time estimate of the gap, y_t^0. The most notable difference between equation (8) and the classic Taylor rule is the inclusion of the change in the output gap. The motivation for this is twofold. First, as shown by Woodford (1999), optimal monetary policy under commitment in forward-looking models is characterized by history-dependence (inertia). By responding to the change in the gap, one achieves history-dependence in monetary policy. This contributes to more stable inflation expectations and thereby stabilizes actual inflation, as expectations of future inflation affect current inflation. Second, if there is persistence in the output gap mismeasurement, Orphanides et al. (2000) show that there is a case for responding to the change in the gap, since, with a high degree of persistence, the errors in the estimate of the change are less severe than the errors in the level of the gap. Equation (9) states that the central bank's estimate of the output gap is subject to errors, represented by ε_t^0. For simplicity, we assume that these errors are exogenous to the rest of the model and follow an AR(1) process given by equation (10).

Since we cannot observe the true output gap, the process for the true measurement errors ε_t^0 cannot be estimated. This is an important and unavoidable problem for monetary policy; there is no way of knowing the true degree of output gap mismeasurement. The best one can do is to use the difference between the real time estimate and the estimate based on final data as a proxy for the true measurement error. This is the approach taken in this paper. Since the gap estimates vary depending on which model is used, we must choose a model (or a combination of estimates) when estimating equation (5). Since the output gap estimates presented in Norges Bank's Inflation Reports are based on the HP-filter, we will use this as our benchmark. When estimating the process for the difference between the real-time estimate and the final estimate, we find

$$\varepsilon_t^0 = 0.7\varepsilon_{t-1}^0 + \eta_t^0 , \quad \hat{\sigma}_{\eta_t^0} = 1.3 \tag{11}$$

The degree of persistence is somewhat less than what Orphanides et al. (2000) find for the US data, while the standard error of η_t^0 is somewhat higher. To illustrate the challenges for monetary policy in practice resulting from output gap mismeasurements, we will first consider a specific case where the true output gap increases due to a temporary decrease in potential output, while the central bank does not immediately observe this drop in potential output and estimates the gap to be unchanged. The central bank is here assumed to set the interest rate according to a classic Taylor rule which is nested in (8) as a special case. The impulse-responses following such a shock are illustrated in Figure (8).

Figure 8: Unobserved increase in the true output gap (productivity slowdown). Number of periods on the horizontal axis.



Since the central bank does not observe the increase in the output gap, it does not raise the interest rate as a direct response to this.The increase in the true output gap is followed, however, by an increase in inflation, to which the central bank responds by raising the interest rate. The central bank therefore interprets the shock as a cost-push shock. Since inflation is persistent, the effect of the shock is amplified in subsequent periods, so that inflation, and thus the interest rate, continues to rise (see figures 8(a) and 8(b)). The initial rise in the interest rate is, however, insufficient compared to what the central bank would have done if it had observed the gap perfectly, as illustrated by the dotted line in Figure 8. Since the process for the measurement error is stationary, the central bank gradually "learns" about the true gap and the estimate of the gap is adjusted upwards, as illustrated by the dashed line in Figure 8(b).

Within the New Keynesian literature, the output gap measure that enters the welfare loss function is the deviation of output from the level that would have occurred if prices and wages were fully flexible. Estimating potential output by various trending techniques, as discussed in Section 3, implies in practice a smoother series for potential output than is likely to be the case with the theoretical flex price concept, which will tend to jump around as, e.g., technology shocks occur. The case discussed above, where the true output gap changes while the central bank's estimate of it remains (almost) constant, is likely to be a case that could happen quite frequently. The above impulse-responses depend of course crucially on the model and parameter values, and one should therefore be cautious when generalizing the results. However, the above figures illustrate a more general point, namely that the main welfare costs associated with failing to capture movements in potential output arises because monetary policy does not respond quickly enough to changes in the true output gap, thereby allowing inflation to move too far away from the target and closing the output gap too slowly.

In the discussion above, the comparison between monetary policy with and without output gap mismeasurements was made under the same monetary policy rule. However, the optimal coefficients of the variables in the interest rate rule depend on the degree of uncertainty. In order to analyse optimal simple rules, we apply the following standard (period) loss function:

$$L = \pi_t^2 + \lambda y_t^2 + \omega(\Delta i_t)^2 \tag{12}$$

We then find the coefficients in the simple rule (8) that minimizes the unconditional expectations of the loss for various weights in the loss function. As discussed in section 3, the revisions of the gap estimates based on the production function method display less persistence than the estimates based on HP. To illustrate the role of persistence in output gap mismeasurements, we therefore consider optimal coefficients when the true mismeasurements are assumed to follow the same process as the revisions in the PF gaps, i.e.

$$\varepsilon_t^0 = 0.38\varepsilon_{t-1}^0 + \eta_t^0, \quad \hat{\sigma}_{\eta_t^0} = 1.41 \tag{13}$$

The optimal simple rules are reported in Tables 7 and 8 (HP and PF), while Table 9 reports the optimal coefficients in the hypothetical case with perfectly observable output gaps.

Table 7: Optimal simple sules. Hodrick Prescott with λ =1600 (HP)

	Loss Function			Optimal Weights				Measures of Macro Variability		
	λ	ω	E[L]	α_i	α_π	α_y	$\alpha_{\Delta y}$	σ_π	σ_y	$\sigma_{\Delta i}$
0.	0.00	0.50	1.25	0.40	3.10	0.20	0.00	0.87	1.50	0.99
1.	0.50	0.50	2.13	0.20	2.40	0.30	0.00	0.93	1.25	0.99
2.	1.00	0.50	2.86	0.10	2.10	0.30	0.00	0.99	1.17	0.99
3.	1.50	0.50	3.51	0.10	2.00	0.40	0.00	1.03	1.13	1.04
4.	2.00	0.50	4.13	0.00	1.90	0.40	0.00	1.06	1.10	1.10
6.	0.00	0.00	0.50	0.00	4.60	0.40	0.00	0.71	1.53	1.64
7.	0.50	0.00	1.44	0.00	3.50	0.80	0.90	0.81	1.25	3.40
8.	1.00	0.00	2.15	0.00	2.70	0.70	0.90	0.91	1.15	3.20
9.	1.50	0.00	2.78	0.00	2.30	0.60	0.90	0.98	1.10	3.03
10.	2.00	0.00	3.37	0.00	2.10	0.60	1.00	1.03	1.07	3.25

Table 8: Optimal simple rules. Production Function (PF)

	Loss Function			Optimal Weights				Measures of Macro Variability		
	λ	ω	E[L]	α_i	α_π	α_y	$\alpha_{\Delta y}$	σ_π	σ_y	$\sigma_{\Delta i}$
0.	0.00	0.50	1.24	0.40	3.00	0.20	0.00	0.88	1.48	0.97
1.	0.50	0.50	2.09	0.30	2.40	0.40	0.00	0.93	1.21	1.00
2.	1.00	0.50	2.76	0.30	2.10	0.50	0.00	1.00	1.12	1.02
3.	1.50	0.50	3.35	0.30	2.00	0.60	0.00	1.03	1.07	1.08
4.	2.00	0.50	3.90	0.30	1.90	0.60	0.00	1.06	1.05	1.07
6.	0.00	0.00	0.49	0.00	4.60	0.60	0.00	0.70	1.44	1.86
7.	0.50	0.00	1.28	0.00	3.70	1.20	0.60	0.79	1.15	4.07
8.	1.00	0.00	1.87	0.00	2.80	1.20	0.60	0.90	1.03	4.02
9.	1.50	0.00	2.37	0.00	2.50	1.20	0.60	0.96	0.98	4.00
10.	2.00	0.00	2.84	0.00	2.30	1.20	0.60	1.01	0.96	3.99

Table 9: Optimal simple rules. No output gap uncertainty

	Loss Function			Optimal Weights				Measures of Macro Variability		
	λ	ω	E[L]	α_i	α_π	α_y	$\alpha_{\Delta y}$	σ_π	σ_y	$\sigma_{\Delta i}$
0.	0.00	0.50	1.12	0.70	4.00	1.40	1.90	0.84	1.24	0.91
1.	0.50	0.50	1.69	0.60	3.30	1.90	1.70	0.89	0.95	0.94
2.	1.00	0.50	2.08	0.60	3.10	2.40	2.00	0.96	0.84	0.96
3.	1.50	0.50	2.40	0.60	3.00	2.80	2.20	1.01	0.77	0.99
4.	2.00	0.50	2.67	0.60	2.90	3.00	2.40	1.05	0.73	1.01
6.	0.00	0.00	0.37	0.00	9.90	3.00	2.00	0.61	1.28	3.22
7.	0.50	0.00	0.96	0.00	9.90	6.80	3.20	0.75	0.90	5.37
8.	1.00	0.00	1.30	0.00	7.90	7.60	2.60	0.86	0.75	4.96
9.	1.50	0.00	1.56	0.00	7.10	8.20	2.20	0.92	0.69	4.83
10.	2.00	0.00	1.77	0.00	6.50	8.60	2.00	0.98	0.64	4.80

A striking result is that output gap mismeasurements reduce the optimal coefficients considerably compared with the full information case. Although the quantitative magnitude of the reduction depends on the particular parameter values in the model, among other things, the qualitative result confirms the findings in Smets (2002) and thus seems quite robust. Another result is that in the loss function with zero weight on interest rate smoothing, the relative importance of the change in the (observed) output gap becomes greater when the degree of persistence in output gap mismeasurement increases, as seen by comparing Tables 7 and 8. This result thus confirms the results by Orphanides et al. (2000). Note, however, that the coefficient of the change in the output gap is strictly positive in the case with perfectly observable gaps. This reflects the role of history-dependence in forward-looking models. Since the optimal coefficient of the lagged interest rate is positive only when interest rate smoothing enters the loss function, it follows that in this particular model history-dependence is more efficiently introduced through responding to the change in the output gap rather than through smoothing the interest rate directly. However, this result is likely to be less robust to alternative model specifications. The expected losses, E[L], in Tables 7 – 9 suggest that there is a substantial excess loss stemming from output gap mismeasurement. Improving existing measures of the output gap may therefore bring about significant welfare effects through improved monetary policy.

5 Conclusions

Interestingly, the tests of the *news* hypothesis indicate that future revisions to output growth for the Norwegian mainland economy are unpredictable. When we augmented the tests with macroeconomic variables which were observable at the same time as the preliminary growth estimates, none of these turned out to be significant (at the 10 per cent level of significance).

The different output gap estimates obtained from final data were compared with those from real-time data and we have decomposed total revisions into the sum of data revisions and other revisions. In general, other revisions are relatively more important than data revisions.

The reliability of the various univariate output gap models is poor. Total revisions are large and persistent, and the correlations between real-time and final estimates are generally low. By comparison with the US data analysed by Orphanides and van Norden, the real-time estimates of the Norwegian data based on univariate models seem to be even less reliable than the real-time estimates for the US. The PF model, however, stands out as more reliable than the other models.

Analysing the consequences for monetary policy within a small New Keynesian macroeconomic model, the main welfare costs associated with failing to capture movements in potential output arise because monetary policy does not respond quickly enough to changes in the true output gap, thereby allowing inflation to move too far away from the target and the output gap to close too slowly. Furthermore, output gap mismeasurement reduces the optimal coefficients in generalized Taylor rules considerably compared with the full information case.

References

Bernhardsen, T., Ø. Eitrheim, A.S. Jore and Ø. Røisland (2004), Real-time data for Norway: Challenges for monetary policy, Discussion Paper 26/2004, Deutche Bundesbank, (http://www.bundesbank.de/download/volkswirtschaft/dkp/2004/200426dkp.pdf)

Cayen, J.-P. and S. van Norden (2004). The Reliability of Canadian Output Gap Estimates. Unpublished paper. Ecole des Hautes Etudes Commerciales, Montreal, Bank of Canada, CIRANO and CIREQ.

Clark, P. (1987). The cyclical component of U.S. economic activity. *Quarterly Journal of Economics, 102*(4), 797 -814.

Croushore, D. and T. Stark (1999). A real-time data set for macroeconomists. Working paper 1999/4, Federal Reserve Bank of Philadelphia.

Croushore, D. and T. Stark (2000). A funny thing happened on the way to the data bank: a real-time data set for macroeconomists. Business Review september/october 2000, Federal Reserve Bank of Philadelphia.

Croushore, D. and T. Stark (2001). A real-time data set for macroeconomists. *Journal of Econometrics, 105*, 111-130.

Croushore, D. and T. Stark (2002). Forecasting with a real-time data set for macroeconomists. *Journal of Macroeconomics, 24*, 507-531.

Faust, J., J.H. Rogers and J.H. Wright (2000). News and noise in G-7 GDP announcements. International Finance Discussion Paper 2000/690, Federal Reserve Board of Governors.

Giorno, C., P. Richardson, D. Roseveara and P van den Nord (1995). Estimating potential output, output gaps and structural budget balances, Working papers 152, OECD.

Granger, C. W. J. and Y. Jeon (2004). Thick modeling. *Economic Modelling, 21*, 323-343.

Harvey, A. C. (1985). Trends and cycles in macroeconomic time series. *Journal of Business and Economics Statistics, 3*, 216-227.

Husebø, T.A., S. McCaw, K. Olsen and Ø. Røisland (2004), A small calibrated macromodel to support inflation targeting at Norges Bank, Staff Memo 2004/3, Norges Bank, (http://www.norges-bank.no/publikasjoner/staff_memo/memo-2004-03.pdf)

Kozicki, S. (2004). How do data revisions affect the evaluation and conduct of monetary policy? Economic Review first quarter 2004, Federal Reserve Bank of Kansas City.

Mankiw, N.G., D.E. Runkle and M.D. Shapiro (1984). Are preliminary announcements of the money stock rational forecasts? *Journal of Monetary Economics, 14*, 15-27.

Mankiw, N.G. and M.D. Shapiro (1986). News or noise: An analysis of GNP revisions. *Survey of Current Business*, (May), 20-25.

Nymoen, R. and E. Frøyland (2000). Output gap in the Norwegian economy – different methodologies, same result? *Economic Bulletin 2000/2, 71*, 46-52.

Olsen, K. and F. Wulfsberg (2001). The role of assessments and judgment in the macroeconomic model RIMINI. *Economic Bulletin 2001/2, 72*, 55-64.

Orphanides, A. (2001). Monetary policy rules based on real-time data. *American Economic Review, 91*(4), 964-968.

Orphanides, A. (2003). Monetary policy evaluation with noisy information. *Journal of Monetary Economics*, 50, 605-631.

Orphanides, A. and S. van Norden (1999). The reliability of output gap estimates in real time. Unpublished Paper, Board of Governors of the Federal Reserve System and Ecole des Hautes Etudes Commerciales, Montreal.

Orphanides, A., R.D. Porter, D. Reifschneider, R.J. Tetlow and F. Finan (2000). Errors in the measurement of the output gap and the design of monetary policy. *Journal of Economics and Business, 52*, 117-141.

Orphanides, A. and S. van Norden (2001). The reliability of inflation forecasts based on output gap estimates in real time. Unpublished Paper, Board of Governors of the Federal Reserve System and Ecole des Hautes Etudes Commerciales, Montreal and CIRANO.

Orphanides, A. and S. van Norden (2002). The unreliability of output-gap estimates in real time. *Review of Economics and Statistics, 84*(4), 569-583.

Smets, F. (2002). Output gap uncertainty: Does it matter for the Taylor rule?, *Empirical Economics*, Vol. 27 (1) pp. 113-129

Svensson, L. and M. Woodford (2003). Indicator Variables for Optimal Policy. *Journal of Monetary Economics, 50*, 691-720.

Walsh, C. (2003a) Minding the speed limit. FRBSF Economic Letter. 2003-14, 3 p.

Walsh, C. (2003b) Implications of a changing economic structure for the strategy of monetary policy, in Monetary Policy and Uncertainty: Adapting to a Changing Economy, Federal Reserve Bank of Kansas City, 2003.

Woodford, M. (1999). Optimal monetary policy inertia. NBER Working Paper no. 7261.