Stockholm School of Economics
Department of Economics
Master's Thesis
October 2006

A Real-Time Data Set for Swedish GDP

GDP Revisions and the Output Gap in Real-Time

Abstract

In this thesis I construct a real-time data set for Swedish quarterly GDP. This data set is then used to analyze revisions of GDP and output gaps. The first part of this thesis follows Faust et al. (2005) who find some evidence for predictability of GDP revisions for the G-7 countries. I investigate revisions of Statistics Sweden's first announcements of quarterly GDP and find weak evidence suggesting that some of these revisions can be predicted. In the second part I follow Orphanides and van Norden (2002) who examine the accuracy and reliability of alternative output detrending methods in real-time. I apply this to Swedish output and investigate the performance of three different detrending methods for the period 1996–2003. The main reason for revisions of the output gap are not revisions in the national accounts but uncertainty about the trend at the end of the sample. Revisions due to changes in the national accounts are of less magnitude and persistence than revisions that occur because of new information about the position of the economy in the business cycle.

Author: Daniel Heller Sahlgren[1]

Tutor: Martin Flodén

Examinator: David Domeij

Discussant: David Salsbäck

Presentation: October 5, 2006, 10.15–12.00, in 336

---

[1] E-mail: 19037@student.hhs.se. Cell: +467 33 19 59 38

Contents

1 Introduction 1
2 A Review of Earlier Research 3
3 A General Description of the Data 4
4 Revisions of Quarterly GDP 7
4.1 Introduction to GDP revisions 7
4.2 A First Look on Bias in Revisions 8
4.3 News or Noise? 10
4.4 Sensitivity Analysis 12
5 A Real-Time Output Gap Series for Sweden 13
5.1 Introduction 13
5.2 Data 14
5.3 Measuring the Output Gap 14
6 Results 16
6.1 Revisions of Output Gaps 17
6.2 Decomposing the Revisions of Output Gaps 19
6.3 Sensitivity Analysis 20
8 Conclusions 22
References 40

Charts

Figure 1, q-o-q final growth vs. first observation, 1993:2–2005:4. 24
Figure 2, y-o-y final growth vs. first observation, 1984:1–2005:4. 24
Figure 3, development of q-o-q growth for 1994:2. 24
Figure 4, development of y-o-y growth. Final growth versus different points in time. 25
Figure 5, development of q-o-q growth. Final growth versus different points in time. 26
Figure 6, observation 1 vs. revisions, y-o-y. 27
Figure 7, observation 1 vs. revisions, q-o-q. 28
Figure 8, GDP gap estimated by the National Institute of Economic Research (KI) 29
Figure 9, real-time estimates of the business cycle, 1996–2003. 29
Figure 10, final estimates of the business cycle, 1996–2003. 29
Figure 11, total revisions in business cycle estimates, 1996–2003. 30
Figure 12, real-time Hodrick-Prescott output gaps and revisions, 1993–2003. 30
Figure 13, real-time linear trend model output gaps and revisions, 1993–2003. 31
Figure 14, real-time quadratic trend model output gaps and revisions, 1993–2003. 31
Figure 15, decomposition of Hodrick-Prescott revisions, 1996–2003. 31
Figure 16, decomposition of linear trend model revisions, 1996–2003. 32
Figure 17, decomposition of quadratic trend model revisions, 1996–2003. 32

Tables

Table 1, GDP announcements' publication scheme 33
Table 2, summary statistics for GDP revisions. 33
Table 3, Mincer-Zarnowitz regression results. 34
Table 4, Mincer-Zarnowitz regression results excluding outliers. 35
Table 5, Swedish results compared with results for the G-7 countries 36
Table 6, summary statistics for output gaps. 36
Table 7, reliability statistics for output gaps. 37
Table 8, summary statistics for output gap revisions. 37
Table 9, summary statistics for output gaps, quasi vintages. 38
Table 10, reliability statistics for output gaps, quasi vintages. 38
Table 11, summary statistics for output gap revisions, quasi vintages. 39

1 Introduction

Central Banks and other actors regularly face the problem of inaccurate measurements of macro variables such as inflation and GDP. National account figures often undergo revisions and the first estimate of a certain quarter's GDP can differ drastically from later estimates. Thus using real-time information about central macro economic variables such as GDP can be problematic, especially if real-time data tends to be biased and unreliable.

Macroeconomic actors need to make qualified judgments of the current state of the economy when deciding on monetary policy. While real-time data may suggest that the economy is in a through in the business cycle, it may very well be that revised figures tell the opposite. Of course actors such as the Riksbank base their decisions on many factors besides the current state of the economy, such as forecasts. But even these forecasts need data on the business cycle – if you don't know where you are how would you know where you are heading?

Other issues relevant to empirical macroeconomics are the need for accurate data when evaluating different models. Are models to help us explain changes in the economy we first need to know what the economy looks like. While researchers may be able to wait for more final data, economic policy issues can hardly wait for final data to arrive but need to use the information currently at hand. Since macro economic policy is executed in real-time decision makers need to base their decisions on data they more or less know will be revised. Because of this it is crucial to have an understanding of how these revisions behave and how trustworthy real-time data is.

Since it is real-time, and not final data that is at hand at the time for policy decision it is important for decision makers to make qualified judgments about the quality of the information available at the time for the decision-making. Even though the transmission mechanism of how the interest rate affects other variables such as GDP is not clearly understood it is important for decision makers to have as much knowledge as possible about the state of the economy.

A central macroeconomic variable is the output gap. In relation to these issues several rules have been suggested that aim to increase the transparency of the central bank's stabilization operations, the most widely known perhaps being the rule proposed by Taylor (1993). These rules often rely on real-time information about GDP. The evaluation of these rules appears to be difficult since using final data implies using revised GDP figures, i.e. with the bliss of hindsight models may perform better than they do in real-time. To be able to judge the performance of these rules many economists suggest that real-time data should be used. Orphanides (2000) shows that results when using this type of data can differ drastically from performance evaluation using final data.

The concept of the output gap is a practical one since it is used to refer to the current state of the economy. It is also a theoretical one since potential output is an unobservable theoretical concept. Previous research (Orphanides and van Norden 1999, 2002) shows that different methods of estimating potential output can yield different estimates of the output gap and also, that mismeasurements of output give wrong suggestions about the state of the economy.

The empirical focus on these issues has hitherto mainly concerned U.S. data. The problem however, is of universal character and was recently referred to by Stefan Ingves, governor of the Riksbank (Ingves 2006).

This thesis discusses issues concerning the case of Swedish real-time GDP. In an attempt to do so a wide range of quantitative investigations are carried out in hope of contributing to the understanding of Swedish GDP revisions and their implications for measuring the output gap. The thesis is divided into three parts. The first part is integrated in the two following parts. This part is the actual construction of the real-time data set used to carry out the studies. The construction of this data set is described in section 3. The second part examines whether Swedish GDP revisions are biased and if they follow systematic patterns and thus might be predicted. To test this I compute statistical properties for revisions and test if these revisions are rational forecast errors or not. Doing this involves testing the so-called *news* hypothesis first introduced by Mankiw et al. (1984). The third part investigates issues

concerning the measurement of the business cycle in real-time. Three methods for measuring the output gap are implemented and statistical properties are reported. To understand the driving forces of revisions of the output gaps revisions are also decomposed into two parts.

2 A Review of Earlier Research

The problems arising in relation with real-time measurements are not new and have been discussed for quiet a while, perhaps first mentioned by Friedman (1947). In recent years much attention has been directed towards different kinds of policy rules, perhaps mostly because of many central banks' introduction of explicit inflation targeting (Kuttner 2004). Somewhat artificially these rules are either instrument rules as the Taylor rule, relating the interest rate to macro variables such as inflation and GDP, or targeting rules derived from some loss function as proposed by e.g. Svensson (2002).[2]

This attention has resulted in a vast empirical literature evaluating different kinds of policy rules. Since GDP is often a central variable in these empirical studies the discussion about reliability of real-time estimates of e.g. the business cycle has in the last ten years become a recognized problem. The work by Croushore and Stark (1999) has set the guidelines for the literature concerning real-time data issues and most economists wishing to investigate economic policy issues in the U.S. in relation to real-time data have used their data sets. Orphanides (2000) triggered the debate further and strengthened the critique against using accurate historical data in the evaluation of policy rules such as the Taylor rule.

Orphanides shows that the use of real-time data generates quiet different policy recommendations than applying final data. Orphanides has since then extended the subject further investigating the reliability of real-time estimates of the output gap in the U.S. (Orphanides and van Norden 1999, 2002). According to their results real-time estimates are often unreliable. This further stresses that interest in the method of estimating output gap merits some concern due to its central role in economic policy decision-making.

Even though these real-time issues also have been discussed in Sweden the research in the field is limited. One study touching the topic is however the paper by Berg et al. (2004). Their focus is however not as much on real-time data issues as policy rules.

The latest major study of the quality of Swedish economic statistics was done in 2003 by Öller and Hansson (2003), covering the period 1980–1998. They find that Swedish GDP revisions are correlated with the final estimates. The authors of the study do not only concentrate on GDP but on other statistics as well produced by Statistics Sweden. Their methods for evaluating the revisions will in this thesis be extended following the work by Mankiw, et al. (1986) and Faust et al. (2005). Mankiw et al. suggested that GDP revisions could be tested using methods already suggested in the 60's by Mincer and Zarnowitz (1964). Their suggestion was that the nature of the revisions could be tested using ordinary regression techniques. This regression tests the so-called *news* hypothesis, which says that revisions of GDP are rational forecast errors. Faust et al. extend on this subject and test the *news* hypothesis for the G-7 countries. Their conclusion was that most in most of these countries, with an exception for the U.S., revisions reflect removal of noise rather than updating preliminary announcements with new information.

3 A General Description of the Data

If one wishes to use real-time data to evaluate the performance of the stabilization policy of the Riksbank one has to take into account the large structural changes that took place in Sweden during the middle of the 1990s. Including time periods before this date would yield dubious results because a number of factors perhaps, the most important one being Sweden changing its monetary regime to a flexible exchange rate regime in January 1993, with a focus on stabilizing the general price level. The operations of the Riksbank during the first couple of years after the introduction of the new regime cannot be characterized as being carried out under normal settings, especially having in mind the severe economic crisis Sweden faced in the beginning of the same decade. Having in mind that the Riksbank did not become fully independent until 1999 and that some time is required for the market expectations to adjust to the new targets of the bank, one must be cautious when choosing the time period of interest (Viotti 2004). Since it is commonly agreed upon that the

[2] For a more detailed description of different policy rules see Kuttner (2004)

operational framework of the Riksbank cannot be regarded as being normal until around 1996 I have chosen to investigate the output gap in real-time for the period 1996:1–2003:4.[3]

The raw data used in this study is unadjusted quarterly GDP for the period 1980:1–2005:4 and seasonally adjusted quarterly GDP for the period 1993:1–2005:4.[4] Since the purpose of this thesis is to construct a real time series it is not possible to use the last series for historical GDP at hand today. GDP is subject to revisions over time and the information available today is not the information that the policymakers had at the time when the growth rate for a certain quarter first was published (quarterly GDP is published with approximately a lag of three months).[5]

The research regarding Sweden's output gap is rather limited and to the author's knowledge the real-time data set of Swedish GDP presented here is the most extensive set available. By using GDP figures published by Statistics Sweden in the quarterly GDP reports (SCB 1980:1–1998:4) and with the help of Statistics Sweden I have been able to construct a real-time series of quarterly GDP year-on-year growth for the period 1984:1–2005:4. A corresponding set for the period 1993:2–2005:4 has been constructed for quarter-on-quarter growth using seasonally adjusted figures.[6]

The lack of long vintages for Swedish macro variables in general has posed some problems in the setting up of a real-time data set. The latest seasonally adjusted quarterly GDP vintages do not go further back than 1993:1. This could potentially be a problem if one wants to run regressions or detrend vintages, since they will tend to get shorter the earlier the vintages. To alleviate this problem I have also seasonally adjusted unadjusted figures for which vintages are longer (dating back to 1980) in order to conduct a sensitivity analysis.

[3] I chose to end the series in 2003 in order to allow for revisions of the latest estimates. Otherwise the last real-time estimate would, by definition, have coincided with the corresponding estimate in the final series. This would in turn, among other things, have indicated too high a correlation between final and real-time estimates.

[4] The reason for this is purely the fact that this is the available data.

[5] See *table 1* for a detailed presentation of when GDP figures and revisions are published.

[6] For more information on real-time data series see Croushore & Stark (2000). I am grateful for the help to put some of this data together received from Göran Svensson at Statistics Sweden.

These tests are conducted to check whether the results are sensitive to the length of the vintages.

To do these tests with sufficiently long vintages, an old vintage going back to 1980 has been linked to each unadjusted vintage. To adjust for new base years these have been adjusted using an overlapping period, namely the first quarter of 1993. By using overlapping years as reference points I have coupled the vintages and thus made them longer, making it possible to go back to 1980:1. To avoid anachronistic seasonal adjustment I have adjusted these series using either X-11 ARIMA or TRAMO/SEATS, depending on the date of the vintage.[7] For this purpose I have used the program Eviews and Demetra for each respective technique.[8] A drawback is of course that these "quasi" adjusted series will not be perfect replicas of the real-time series. Statistics Sweden investigates series "manually" and employs the model that is judged to best fit the series. E.g. for vintages prior to 1999:2 this means that some are adjusted with a X-11 ARIMA additive and some with a multiplicative approach, depending on the nature of the series. In some cases an ordinary moving average is also used. The results are, nevertheless, quiet encouraging. In most cases the quasi-adjusted series appears to fit well.[9]

Major changes in the measurement of GDP have occurred during the period 1980:1–2005:4. For example, new definitions and seasonal adjustment techniques of GDP have been introduced. In May 1999 TRAMO/SEATS seasonal adjustment was introduced in a step towards European convergence. For the same reason the new definition SNA 93-ENS 95 was implemented.

Many researchers have requested that Statistics Sweden prioritize time series (SOU:2002:118, p.350). As it is today national accounts on a detailed level for seasonally

[7] Since 1999:2 Statistics Sweden uses TRAMO/SEATS for seasonal adjustment. Prior to this X-11 ARIMA was used. For more information about seasonal adjustment in this context see Öhlen (2003).

[8] Demetra is a free software published by the European Central Bank, ECB. For more information see Demetra (2002).

[9] I compare my seasonally adjusted figures with those from Statistics Sweden for the same time period and get an average correlation between corresponding vintages of 0.9955.

adjusted quarterly GDP only exist back to 1993. In this context this thesis can be seen as a forerunner to more accurate republished historical vintages.

4 Revisions of Quarterly GDP

4.1 Introduction to GDP revisions

In this section I investigate potential bias in the estimates of quarterly GDP growth on a quarterly basis. For the non-seasonal adjusted series the figures are all year-on-year growth (y-o-y), and for the seasonally adjusted they are quarter-on-quarter (q-o-q). The y-o-y growth gives us the growth rate of one quarter compared with the same quarters for the previous year. q-o-q compares quarter t's GDP with GDP in the previous quarter.[10] The publications of GDP estimates can be seen in *table 1*.

In this study *real-time* q-o-q growth, the growth corresponding to observation 1, for quarter t is defined as

$$GROWTH = \frac{(GDP_t - GDP_{t-1})}{GDP_{t-1}} * 100,$$

and for y-o-y as

$$GROWTH = \frac{(GDP_t - GDP_{t-4})}{GDP_{t-4}} * 100.$$

In *figure 1* and *2* the reader can see some of these estimates plotted against each other. It becomes rather evident just by mere visual inspection that real-time growth figures at times can differ substantially from the final ones.

It is also quiet interesting to see the development of a certain quarter's growth. In *figure 3* it can be seen that these changes are sometimes rather dramatic. The big jump in May 1999 coincides with the introduction of the new definition SNA 93-ENS 95 and the new seasonal adjustment techniques TRAMO/SEATS. The spontaneous question perhaps to pose is how

[10] One should note that y-o-y is in fact a way to seasonal adjust data.

real-time data corresponds to final estimates, or to some other point in time in the future, and if there are systematic biases in revised figures.

4.2 A First Look on Bias in Revisions

This chapter builds on Faust et al. (2005) who test whether differences between real-time and final estimates of GDP in the G-7 countries can be predicted or not.

Before moving on some definitions should be clarified. There are many ways to categorize a revision *r*. Using the following notation $r_t \equiv X_t^s - X_t^{obs1}$, where *s = t+1, definitive, t+8,* or *final*.

- *t+1*: the second estimate of the quarter's GDP.
- *definitive*: the revised estimate as of December two years after observation 1 was published. This definition is borrowed from Öller and Hansson (2003) in order to compute comparable statistics.
- *t+8*: the revised estimate after 8 quarters from observation 1.
- *final*: the data corresponding to the latest vintage available (2006:1).

This yields four different definitions of a revision:

- short-term: $r^{t+1} = X_t^{t+1} - X_t^{obs1}$
- medium-term: $r^{t+8} = X_t^{t+8} - X_t^{obs1}$
- definitive[11]: $r^{d} = X_t^{d} - X_t^{obs1}$
- long-term: $r^{f} = X_t^{f} - X_t^{obs1}$

Some general statistics are often suggested when evaluating revisions of macroeconomic variables such as GDP. In *table 2* some summary statistics of GDP revisions are presented.

- Mean: this is just the mean of the revisions

[11] The reason for including definitive, besides medium term revision, is the fact that this is the definition used by Öller and Hansson (2002). To get some comparable statistics I therefore computed these revisions as well.

- Absolute mean, A(MEAN): the mean of the absolute value of revisions. This measure also takes dispersion into account, which translates into uncertainty.
- Correlation coefficient (CORR): the correlation between revisions and "final" estimates.
- The mean square error of forecast (MSE): This is defined as the mean of $\left(X_t^s - X_t^{obs1}\right)^2$. MSE measures the variance around the 45-degree line in scatter plots in *figure 4* and *5*. This is one of many measures of accuracy of preliminary forecasts.
- SIGN: the number of times the preliminary figures and the revised figures show different signs.
- Acceleration/Deceleration (AC/DC)[12]: the number of times when the revised number and the preliminary differ with regards to the change of growth between one quarter and the next. Acceleration when the change is positive, deceleration when negative.

Some of these statistics should be of particular interest to macroeconomists. For example, a positive correlation between the business cycle and revisions tells us that in periods with high growth the bias in the preliminary estimate may be larger than in normal periods, i.e. with growth rates closer to the average. The number of signs gives an idea of whether or not the direction of the development of GDP is to be trusted or not, a high number thus would lead one to conclude that the preliminary estimates and the definitive ones tend to give opposite information about the growth for a certain quarter's GDP. AC/DC, just like SIGN tells us if the preliminary estimates give you a good indication of the development of GDP.

Table 2 gives us some information how q-o-q and y-o-y estimates differ. Basically all statistics are lower for q-o-q than for the corresponding y-o-y figures. It also compares these studies results with the ones from the study by Öller and Hansson (2003). Even though they cover partially different time periods no big differences can be detected. However Öller and Hansson only report y-o-y figures using definitive revisions as benchmarks. Worth noting is also that y-o-y and q-o-q short-term revisions and q-o-q medium term and definitive

[12] Acceleration is defined as change of change (acceleration when positive, deceleration when negative)

revisions have means close to zero. This suggests that these are the revisions most likely to follow the *news* hypothesis.[13]

In order to get an idea of how the estimates develop over time in relation to final estimates *figure 4* and *5* shows the development of growth over time. The black line is the regression line and the dashed line is the 45-degrees line where all observations would be if all observations were the right figures, measured without errors. Judging by the scatter plots presented in *figure 4* there seems to be some convergence of growth rates over time. This tendency is not as clear for q-o-q figures, judging by the scatter plots in *figure 5*. An F-test was conducted for both y-o-y and q-o-q figures and the null hypothesis of an intercept equal zero and a slope coefficient equal to one and was rejected for all four regression lines, this of course indicates systematic deviations from final figures.

4.3 News or Noise?

According to Mankiw et al. (1986) revisions can be categorized as either *news* or *noise*. In the noise case preliminary estimates are just the true values but contain measurement errors. The preliminary estimates under this assumption thus contain errors or *noise*, uncorrelated with the true value but correlated with the preliminary estimate.

In contrast with the noise view stands the *news* view. Under this assumption all relevant information is used efficiently when announcing a preliminary estimates. Therefore revisions reflect news that was not available at the first announcement. A revision will thus be uncorrelated with information available at the time the revision is done and revisions cannot be predicted. The preliminary estimate is here a rational forecast of the true value and the revision is a rational forecast error.

We are here interested to see whether the revisions of Swedish quarterly GDP contain a systematic pattern (the noise view) or not (the news view). Testing the news view involves

[13] If revisions reflect news their expected value is equal to zero since the revisions then are assumed to be white noise. See section 4.3 for a more detailed explanation. Note that none of the means is statistically different form zero.

running the so-called Mincer-Zarnowitz forecast efficiency test (Mincer 1964). This is done by running the regression in equation 4.1 under the assumption that u_t is white noise.

(4.1) $$r_t = \alpha + \beta X_t^P + u_t .$$

To test the news view I setup the following null hypothesis: $H_0 : \alpha = \beta = 0$.

Rejecting this hypothesis means that we accept the noise view and that preliminary estimates are correlated with the revision. If this is the case today's estimate could help us predict future revisions. Not rejecting the null hypothesis means that there is no relationship between preliminary estimates and future revisions. Before running the actual regression it is therefore useful to have a look at scatter plots and look for potential patterns that could indicate if we should expect to reject the null hypothesis of rational forecasts. Scatter plots of these relationships are shown in *figure 6* and *7*.

As can be seen in *figure 6* and *7* there are major differences in the relationship between revisions and preliminary estimates depending on whether the y-o-y or q-o-q measurements are used. For y-o-y figures the relationship tends to be slightly positive, save the short-term revisions, while for the q-o-q the opposite result seems to be true.

In *table 2* we saw that for the short-term revisions both the q-o-q and the y-o-y revisions had means close to zero. The same goes for q-o-q revisions for the medium-term and definitive revisions. Before running the regression, this suggests that these are the revisions for which the null hypothesis is most likely not to be rejected. Remember that forecast efficiency, corresponds to not rejecting the null hypothesis, i.e. $r_t = u_t$. Since u_t is assumed to be white noise the expected value of r_t is equal to zero, i.e. $E(r_t) = 0$. However, looking at *figure 7 b* and *c* gives us a clue that medium-term and definitive revisions might not pass the test of the news hypothesis.

The results for the Mincer-Zarnowitz regressions are reported in *table 3*. Inference from this regression is only possible if the residual is white noise. For q-o-q short-term, y-o-y and q-o-

q medium-term, q-o-q definitive and long-term revisions the news view can be rejected without testing the null hypothesis since these residuals appear to be auto-correlated.[14]. This leaves only the y-o-y short-term, definitive and long-term revisions to test.[15] The only revision to comply with the news view is the y-o-y short-term revisions, but with a caveat since the residuals according to the Jarque Bera test are not asymptotically normally distributed. For the other two revisions, y-o-y definitive and long-term revisions, the null hypothesis can be rejected at reassuring p-levels. This means that these revisions can somewhat be predicted, but here the extremely low R^2 is of course problematic. Disregarding this fact we have that for a one percentage point increases in preliminary growth the y-o-y definitive revision increases with approximately 0.1. For long-term revisions this figure is about 0.04.

In summary this presents evidence for the noise view of Swedish GDP revisions, and some tentative suggestions of predictability in the case of y-o-y definitive and long-term revisions. These conclusions are line with Öller and Hansson (2003) who also conclude that revisions are biased in the case of definitive revisions.

In *table 5* the results for the q-o-q medium term revision are compared with the G-7 countries studied by Faust et al (2005). We see that the results are similar with those for Germany.[16]

4.4 Sensitivity Analysis

To test the sensitivity of these results regarding outliers the same test was carried out excluding outliers according to the definition used by Faust et al. (2005). A preliminary y-o-y observation is thus an outlier if it deviates more than three percentage points from the average. For q-o-q the definition is somewhat modified to one percentage point deviation. This is due to the systematic difference in the magnitude of growth figures between these

[14] I follow Roodenburg (2004) in the evaluation of the white noise process. Auto correlated residuals suggest that they are not independent of past information, which in effect means that the news hypothesis cannot hold.

[15] Despite the low p-values for the ARCH(4) test for the y-o-y long-term I don't reject the hypothesis of white-noise error terms.

[16] This comparison should not be taken too seriously since the assumption of white noise error terms for q-o-q medium-term revisions. Thus a comparison is only legitimate when relaxing the ARCH-test for these revisions.

two growth definitions. The results are presented in *table 4*. The results do not differ substantially from the ones in *table 3*, and the overall results remain. Worth noting is that the residuals from the short-term y-o-y regression now seem to be normally distributed, which strengthens the news view hypothesis for these revisions. All in all one can conclude that the results are not sensitive to outliers in the dataset.

5 A Real-Time Output Gap Series for Sweden[17]

5.1 Introduction

One central variable in macroeconomic policy is the economy's position in the business cycle, often referred to as the difference between actual output and its trend value. Difficulties in estimation of the business cycle and evaluating the reliability of output gap estimates pose problems to an efficient macroeconomic stabilization policy. In this chapter three different methods of estimating the output gap are presented, as well as corresponding reliability measures.

Also, in empirical studies much research is done on different models' capability to explain changes in the business cycle. But to make reasonable comparisons between these models one needs reliable measures of the underlying economy's development. Since many models use the output gap as input it is vital to have reliable data on this input.

In Sweden, the only published estimates of the output gap come from the National Institute of Economic Research (Konjunkturinstitutet, KI).[18] The measurement of the output gap is however subject to major uncertainties and the estimates need not be reliable, whatever the method employed. This fact is perhaps mirrored in the debate concerning the alternative business cycle estimate provided by the National Institute of Economic Research – the labor gap.[19]

---

[17] The second part of this thesis builds on Orphanides & van Norden (2002).

[18] The Riksbank also estimates the output gap but does not publish the results.

[19] See Jansson and Palmqvist (2005), Dillén, Nilsson and Zettergren (2005) and Flodén (2005).

5.2 Data

The real-time output gap series for Sweden presented in this section covers the period 1996:1–2003:4. Final data is from 2006:1. The figures and data discussed are produced using the real-time data set with vintages going back to 1993. For reasons discussed earlier in section 3 I also provide the reader with alternative figures using the quasi real-time data set I have constructed by seasonally adjusting the vintages going back to 1980. In the end of this chapter robustness is discussed and a sensitivity analysis is done using these quasi vintages.

5.3 Measuring the Output Gap

The output gap is generally defined as the difference between actual (real) and potential GDP. Since potential output is not observable is has to be estimated. This can be done using several different methods. The methods employed here are restricted to detrending using deterministic trend models and the Hodrick-Prescott filter. Other techniques may of course also be employed; implementations of state-space models and the Kalman filter are becoming popular in economic research but are beyond the scope of this thesis.[20] There are of course pros and cons with each and every estimation technique but the discussion here will be limited to those relating to the Hodrick Prescott filter and two deterministic trend models.

To construct the real-time output gap series I first detrend each vintage and create output gap series for each vintage. Then the last estimate of the output gap in every vintage is used to compile a new series, which constitutes the real-time series of the output gap.

[20] For a discussion of different methods of estimating the output gap see e.g. Orphanides and van Norden (1999).

The Hodrick-Prescott filter[21]

The Hodrick-Prescott filter has become a popular way to detrend different kinds of time series data and is today incorporated in most software designed to deal with time series. The technique was first suggested by Hodrick and Prescott in a working paper from 1981 but not officially published until 1997 (Hodrick and Prescott 1997).

The Hodrick-Prescott filter takes a time series y_t and decomposes it into a trend component, τ_t and a cyclical component, ς_t, this can be written as:

$$y_t = \tau_t + \varsigma_t$$

More formally the Hodrick Prescott filter calculates the trend in order to minimize the variance of the cyclical component but with a penalty regarding the variation of the second difference of the trend component, i.e. that potential GDP cannot vary too much:

$$\tau_t = \arg\min \sum_{t=1}^{T} \left(y_t - \tau\right)^2 + \lambda\left(\Delta^2 \tau_{t+1}\right),$$

where λ is a parameter saying how much weight is to be given to the penalty restriction.[22] Some drawbacks exist with the use of this filter. First, it tends to exaggerate corrections at the end of the sample period. Second, it does not yield confidence intervals. If the sample ends in a recession the filter thus produces downward biased estimates of potential output. Despite this the Hodrick-Prescott filter is an accepted method widely employed in econometrical and economic research.

Deterministic trend models

The simplest deterministic trend model is the linear trend model, which models the trend as a linear function of time:

(5.1) $y_t = \alpha + \beta t + \varepsilon_t$, where $E(y_t) = \tau_t$

Extending this model gives us the quadratic trend model:

(5.2) $y_t = \alpha + \beta t + \mu t^2 + \varepsilon_t$,

[21] This section builds on Orphanides and van Norden (1999) and Apel, Jansson and Linberg (1996).

[22] The most common parameter value used for quarterly data is λ = 1600, which is also the chosen parameter value here.

These models present a simple alternative to the Hodrick-Prescott filter. Moreover, equation 5.2 lets the user model changes in the trend over time, embedded in the quadratic function.

6 Results

In *figure 8* the final estimates using the three models presented above are compared with the most current GDP gap series published by the National Institute of Economic Research. Just by looking at this graph it becomes quiet clear that the estimated business cycle varies a lot just by using different techniques and models. Note that the quasi vintages seem to produce an output gap series more in line with the business cycle as it is estimated by KI. This suggests that the length of the time period is important since the quasi vintages are much longer. This is in line with logic reasoning - intuitively it is easier to understand the position in the business cycle when more information of historical development is available. In the next section a sensitivity analysis related to these questions is performed.

In *figure 9* the real-time estimates of the output gap from *figure 8 are* presented, now excluding the KI gap for a better comparison between final and real-time estimates. The three methods appear to be strongly correlated, mostly moving in the same direction at roughly the same time. For the real-time estimates the quadratic trend method seems to produce bigger positive gaps in the beginning of the period than the other two, while the linear model appears to produce bigger negative estimates than the other two at the end of the period. All three methods are clustered in the time of the peak in the year 2000.

In *figure 10* the final estimates from *figure 8* are shown, excluding the KI gap. The final estimates share some of the characteristics of the real-time series. Co-movements tend to be strong, with all methods giving estimates moving in the same direction at roughly the same time. Here the linear trend model produces greater output gap than the two others over almost the whole period. The difference, in spite of strong co-movement, gives rise to large differences in the estimated business cycle. The largest difference is found at the peak in the year 2000. Here the difference between the linear trend and the Hodrick-Prescott model estimates is about one percentage point, which is approximately half of the gap itself, as

measured by the Hodrick-Prescott filter.

In *figure 11* the total revisions of the output gaps are shown (the real-time estimate minus the final estimate). The dispersion of the revisions is largest in the first half of the period and shows a tentative sign of reducing the difference in the second half of the period. The greater dispersion and volatility in revisions with regards to the three models in the 90's suggests difficulties in estimating the accuracy of the estimated business cycle at that time. A possible reason for this is that the Swedish economy at that time was beginning to recover from a deep recession in the beginning of the decade at the same time as a new monetary regime was introduced.

6.1 Revisions of Output Gaps

In order to better understand the differences between real-time and final figures *table 6* provides some summary statistics for comparison between different methods. We see that the total revisions are of roughly the same order of magnitude as the gaps themselves, with the linear trend showing the greatest difference between real-time the final series' standard deviation and the standard deviation of the total revisions for the corresponding method. The correlation between final and real-time estimates is, as a comparison between *figure 9* and *10* suggests, quite low for all three methods. In the last column the first order autocorrelation function for total revisions are reported. The persistence in revisions is high for all three methods, ranging from 0.74 to almost 0.9.

Table 7 shows some reliability statistics for easier comparison between the different methods. In the first column the correlation (CORR) between final and real-time estimates of the business cycle is presented. This figure should of course be equal to unity in the ideal case. The second column, named NS, shows the ratio between the standard deviation of the total revisions to that of the standard deviation of the final estimate. The third column named NSR shows the ratio between the root mean square of total revisions to the standard deviation of the final estimate, and thus captures the effect of persistent revisions downwards or upwards. Hence, these two columns provide proxies for the noise-to-signal-ratio in real-time estimates (Orphanides and van Norden 2002). Both these ratios are fairly

high for all three methods, the highest corresponding to the quadratic trend model, which actually has a ratio above one.[23] In the last two columns the reader is provided with some frequency statistics showing whether the real-time estimates give "good" or "bad" measures of the business cycle. OPSIGN is the frequency at which the final and the real-time series show opposite signs. Again the quadratic trend model yields the worst results, but with frequencies around one third also the two other methods are far from giving reliable estimates of the output gap's sign in real-time. XSIZE is a similar statistic showing the frequency at which the absolute value of the total revision is greater than the absolute value of the final estimate. Just as for the U.S. figures presented by Orphanides and van Norden (1999 and 2002) these figures are closer to each other than OPSIGN, hovering around 50 percent for all three methods.

Having in mind the caveat pointed out by Orphanides and van Norden (1999) the revisions here are probably underestimated, resulting in overestimated precision and accuracy. The fact that the vintages available are rather short of course adds to this uncertainty since valuable information of the historical development of GDP is omitted. To test how this affects these statistics I perform a sensitivity analysis in the next section, using the longer quasi vintages that incorporate more historical information than the official vintages. Generally the results seem to depend on the length of the series, i.e. historical information tends to be helpful in positioning the economy in the business cycle.

Despite this, one can draw some conclusions from the approach presented here. In general the reliability for all three methods is quiet low. The quadratic trend model is clearly the worst of the three combining the lowest correlation (0.22) between the final and the real-time estimates of the output gap, the highest persistence of revisions (0.89) and the highest noise-to-signal ratio (1.19 or 1.17). The pattern is similar for all other reliability statistics provided in *table 2*.

Of the two remaining models neither stands out as superior to the other. The correlation between final and real-time estimates is a bit higher for the linear trend model (0.71 vs. 0.54). The same is true for OPSIGN (0.28 vs. 0.31), but for XSIZE the Hodrick-Prescott

[23] A high noise-to-signal ratio implies that the information embedded in the signals will be hard to extract.

filter is somewhat lower (0,41 vs. 0,47). The same is true for the persistence in revisions being somewhat lower for the Hodrick-Prescott method (0.74 vs. 0.77). The differences are however relatively small in comparison with the quadratic trend model.

6.2 Decomposing the Revisions of Output Gaps

In order to evaluate revisions of the output gap two more series are constructed. Total revisions, as stated before, are simply the difference between the real-time and final estimates. Total revisions generally occur due to either revisions of the national accounts or new observations, becoming available as time goes by. To investigate the driving forces behind the revisions a quasi-real time gap estimate is constructed as well.[24] Differences between real-time and the quasi-real time output gap must exist due to data revisions. The other kind of revision is defined as the difference between the quasi-final estimate minus the final estimate. These revisions occur because new information is being added when new observations are included in the estimation procedure. Just as Bernhardsen et al. (2005), I call these “other revisions”.

Figures 12–14 show the real-time output gap together with total revisions and data revisions. Studying these revisions it becomes quiet clear that data revisions (i.e. revised national accounts) do not stand for the bulk of total revisions. In any point in time where the real-time and total revisions meet the final estimate of the business cycle for that point in time has to be zero. For example in the trough following the peak in year 2000, the real-time estimate and total revisions as measured with the Hodrick-Prescott method coincide. This means that despite the real-time estimate suggesting that the economy was operating under its potential in that time, the final estimate suggests that the economy was actually just in line with its potential.[25]

Figures 15–17 provide similar graphs, showing total revisions with both its two components for all three methods. The summary statistics in *table 8* confirm what could be seen in the graphs. For all three methods adding new information in the form of new observations is

[24] The procedure is the same as for real-time estimates, the only difference is that the final series is now being used with truncation at time *t*.

[25] Remember that total revisions are defined as the real-time estimate minus the final estimate. This implies that when real-time and total revisions are of the same magnitude the final estimate must be equal to zero.

much more important than revised data. The largest revisions are given by the linear trend (-0.78), whereas the mean of revisions in the quadratic trend model is closest to zero (0,05). For this model almost all revisions are due to other revisions[26].

Further, *table 8* provides persistence measures for all three models. As noted before, revisions tend to be persistent. For all three models other revisions are considerably more persistent than data revisions. The difference noted here is much more pronounced than for the U.S. figures presented by Orphanides and van Norden (2002), but less apparent than in the Norwegian data as estimated by Bernhardsen et al. (2005). The fact that data revisions display so little persistence is somewhat surprising regarding the results in section 4.3, which suggest that future revisions of quarterly GDP could, to some extent, be predicted. The noise-to-signal ratios are also much higher for other revisions than for data revisions.

For all three models other revisions are small in the middle of the period and increase in the beginning and the end of the period. As Bernhardsen et al. (2005) remark, this suggests problems with end-point instability

The conclusion is that other revisions stand for the bulk of revisions for all three models. This means that the adding of new observations is responsible for most of the revisions of the position in the business cycle.

6.3 Sensitivity Analysis

The fact that the official vintages for Swedish GDP only go back to 1993 is of course a concern when conducting investigations such as these, since the results might very well vary depending on how much historic information is taken into account. To test whether the results above are sensitive to the fact that the vintages used are quiet short (dating back to 1993) I provide alternative vintages constructed using the much longer non-seasonally adjusted vintages provided by Statistics Sweden. The construction of this series is explained in detail in section 3.

[26] The fact that other revisions and data revisions don’t sum up to total revisions is due to rounding of.

All figures for the output gaps are reproduced using these quasi vintages instead. Comparing *table 6* with *table 9* one can immediately see that the figures are sensitive to these changes. Not even the relationship between different models always remains. First of all the standard deviations in *table 9* are clearly all much larger than those in *table 6*, the orders of magnitude of revisions for these series are however roughly the same size as the output gaps themselves, with an exception for the linear trend model, where the final series has a standard deviation almost twice as big as the one for total revisions.

The revisions are still persistent, the Hodrick-Prescott model increasing its persistency from 0.74 to 0.82, and the linear model decreasing somewhat from 0.77 to 0.63. Correlations are still fairly low, except the linear trend, where the correlation between the real-time series and the final series now jumps from 0.71 to 0.93. Looking at *table 10* provides additional insight in the performance of the three models using the alternative vintages. Using the quasi vintages clearly produces a winner – the linear model is now the "best" performing model: all statistics have improved, some drastically such as correlation and the noise-to-signal proxies.

The ranking of the three models, in terms of revisions, thus seems to be a more clear cut case using the alternative vintages, with the linear trend model being the best and the quadratic trend model being the worst. There are of course many caveats to this conclusion, one being the fact that not officially seasonally adjusted figures are used. These statistics should therefore not be taken too seriously since they probably overestimate the precision and underestimates the total error of the estimates. One reason is of course the same as mentioned above. Another reason is the additional inaccuracy incorporated in the "unofficial" seasonal adjustment of the series. One way to deal with this problem is of course to construct similar output gap series using yearly data, i.e. comparing each quarter with the same quarter one year ago. That would reduce measurement errors due to "wrong" seasonal adjustment.

It is of course important to note that the ranking here does not necessarily give the best

model in terms of which model best measures the real business cycle. The reason that the linear model performs so much better with the longer quasi vintages is probably due to the way the way the cycle is measured by the linear model. Just as noted before, the linear model in contrast with the two other models, does not allow for changes in the trend. Events such as the deep recession in the beginning of the 90's could thus affect the model's performance when switching from official to quasi vintages, which cover the year of the recession. Thus, one should not interpret "best performance" as the model that best describes reality.

Despite the caveats posed by these quasi vintages it seems very likely that the results are sensitive to the length of the vintages, i.e. historical information of the development of the business cycle. This should of course not come as a big surprise and further stresses the need for longer seasonally adjusted vintages from Statistics Sweden.

8 Conclusions

In this thesis I conclude that Swedish GDP revisions in general are somewhat predictable and that the news hypothesis can be rejected for virtually all revisions, save the short-term y-o-y revisions. This suggests that the GDP announcements by Statistics Sweden are not efficient in the sense that revisions are not due to new information being available as times goes by but more due to the "cleaning up" of noise include in the first preliminary estimates. These results tend to be robust.

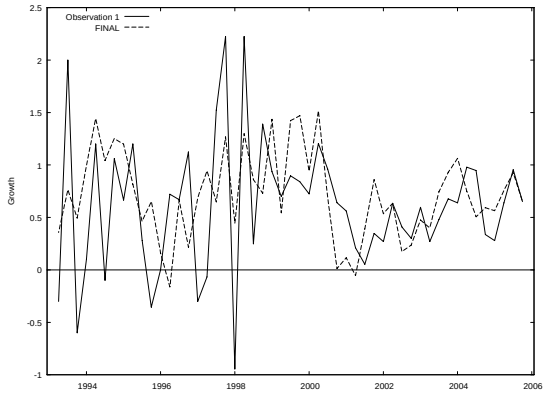
In the second part of this thesis I conclude that real-time output gap measurements of GDP are also unreliable and depend much on which model is used. The results are not very robust but fairly sensitive to length and seasonal adjustment of the series. This suggests the need of longer vintages to be able to draw robust conclusions.

The next step in research concerning real-time GDP would thus be for Statistics Sweden to publish longer historical vintages in the future. For empirical historical research the re-publishing of longer historical vintages is of interest for further research in the field. That

would eliminate the need to use quasi figures.

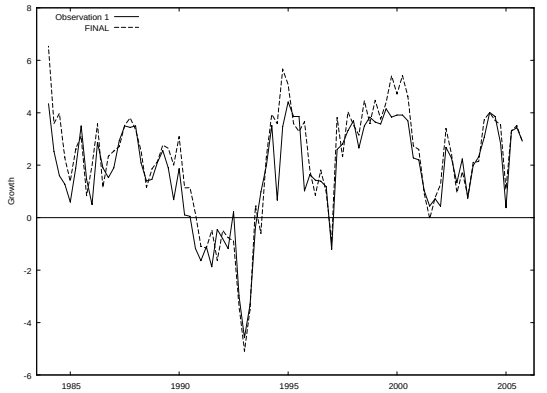
In this thesis I use three relatively uncomplicated models. The study could be extended to the incorporation of UC-models, using the Kalman-filter. A more detailed analysis of KIs figures in comparison with alternative measures of the output gap, such as those presented in this thesis is of course also an issue that merits further concern. Evaluating KIs figures formally this way could give an indication about the reliability of these publications.

Figure 1, q-o-q final growth vs. first observation, 1993:2–2005:4.



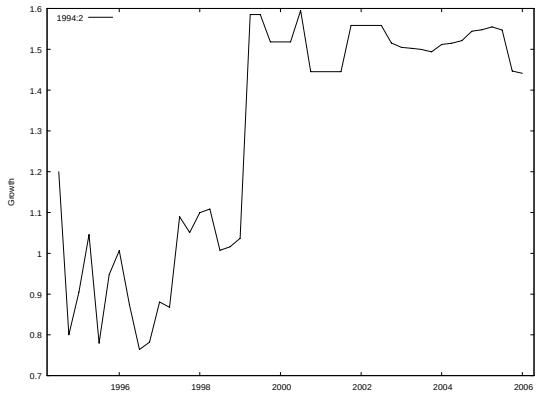
Note: The growth rate is quarter-on-quarter using seasonally adjusted GDP. Observation 1 is the first announcement and final is the latest estimate available.

Figure 2, y-o-y final growth vs. first observation, 1984:1–2005:4.



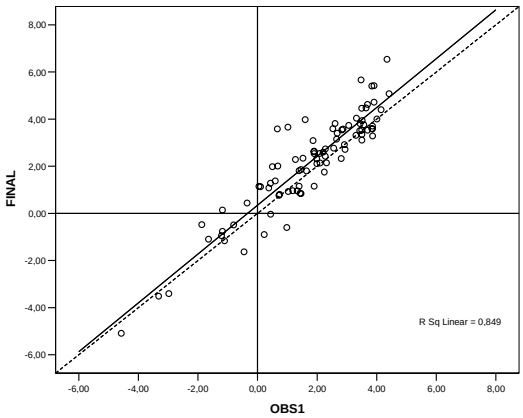
Note: The growth rate is year-on-year using unadjusted GDP. Observation 1 is the first announcement and final is the latest estimate available.

Figure 3, development of q-o-q growth for 1994:2.

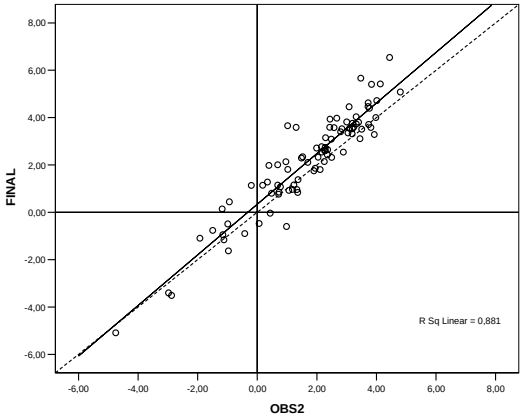


Note: The big jump in growth (about a 0,5 percentage points) coincides with Statistics Sweden's introduction of the new seasonal adjustment technique (TRAMO/SEATS) and introduced the new GDP definition ENS 95.

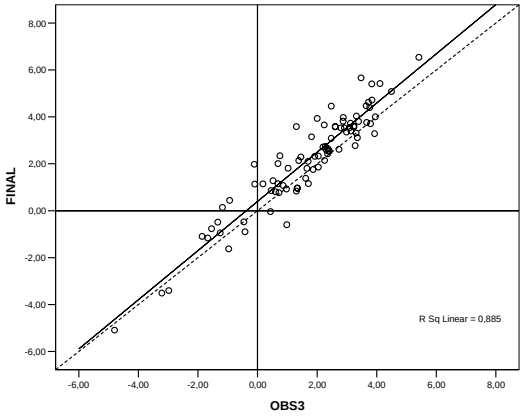
Figure 4, development of y-o-y growth. Final growth versus different points in time.



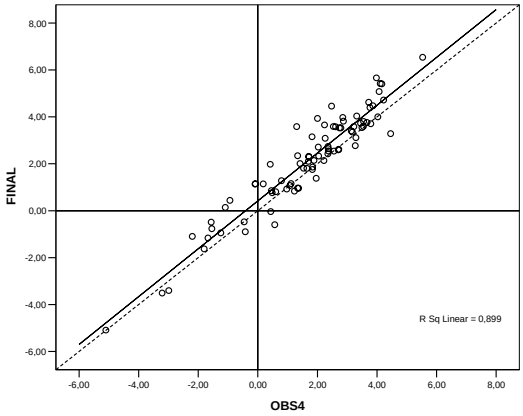
a, final growth versus observation



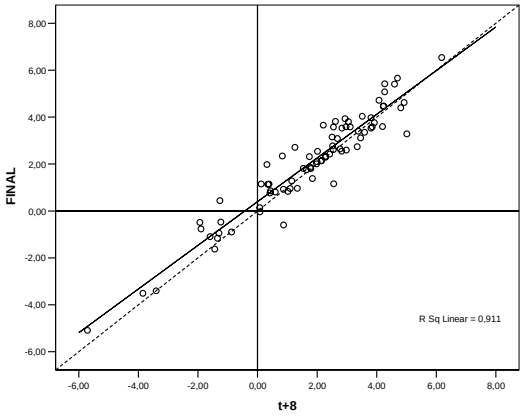
b, final growth versus observation 2



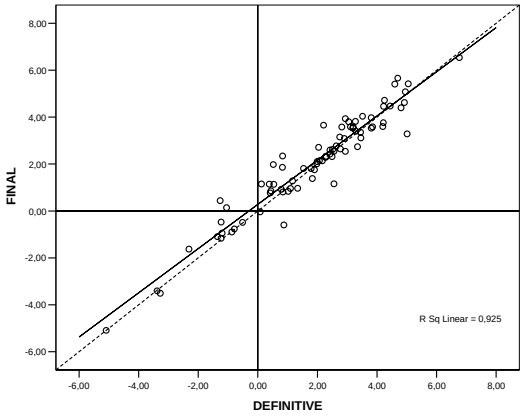
c, final growth versus observation 3



d, final growth versus observation 4



e, final growth versus observation t+8

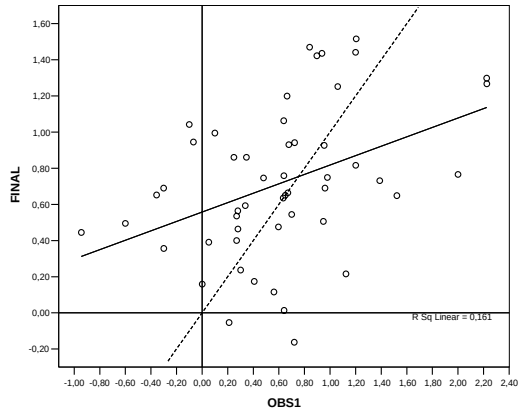


f, final growth versus definitive observation

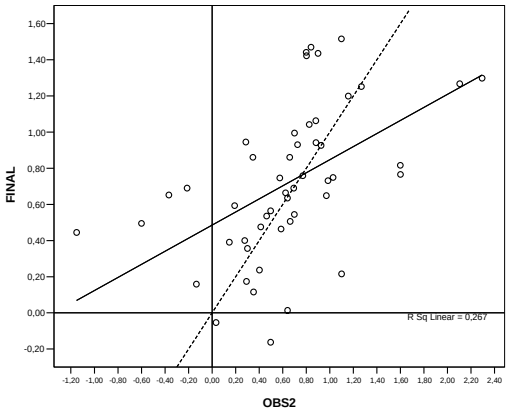
Note: Scatter plots *a–f* compare different year-on-year growth rates with final estimates. As time precedes the dispersion around the dashed 45-degree line should diminish if revisions adjust preliminary announcements.

Perfect preliminary announcements would imply that all points were on the dashed line.

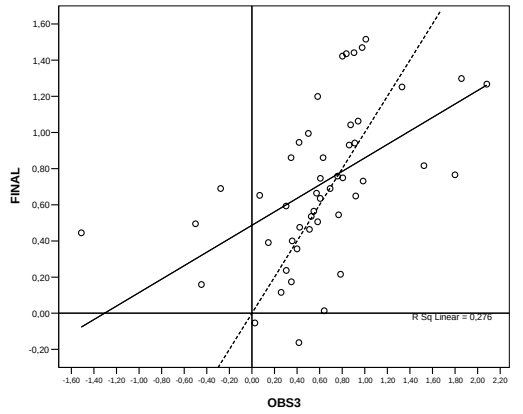
Figure 5, development of q-o-q growth. Final growth versus different points in time.



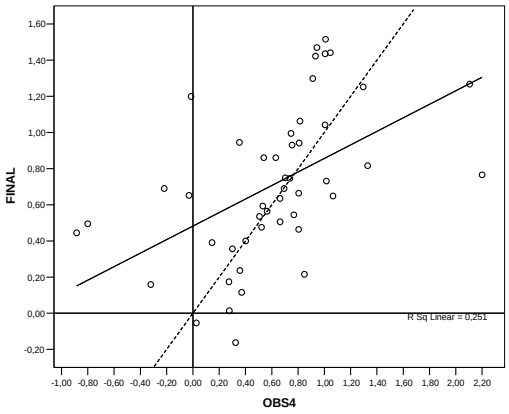
a, final growth versus observation



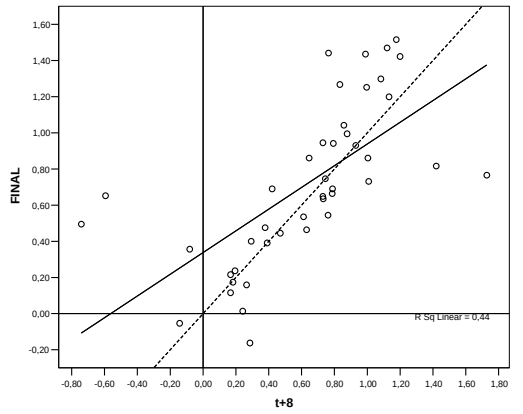
b, final growth versus observation 2



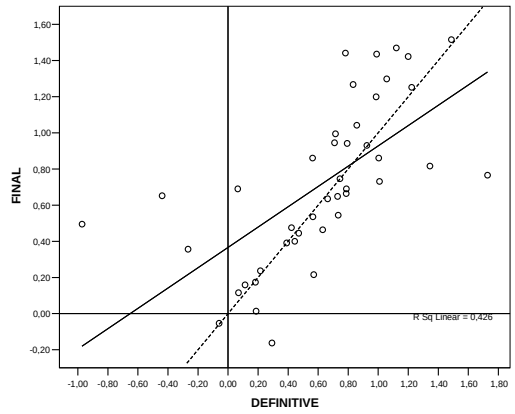
c, final growth versus observation 3



d, final growth versus observation 4

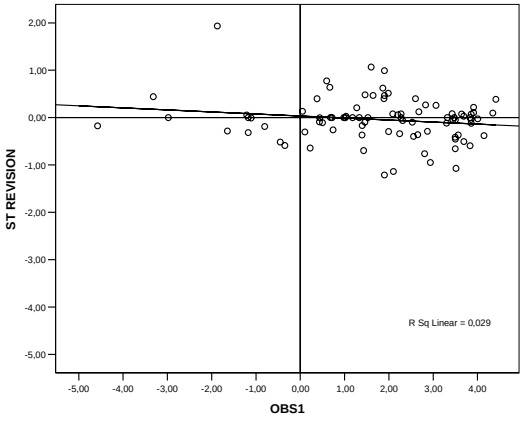


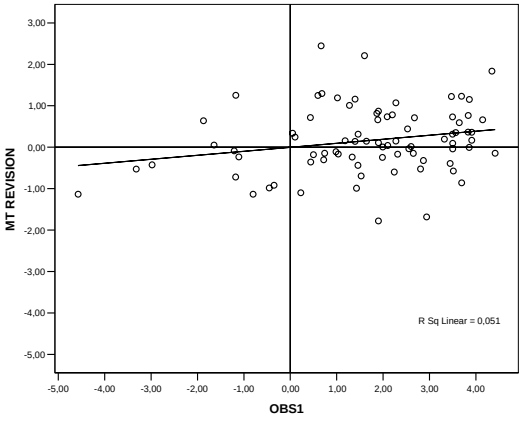
e, final growth versus t+8



f, final growth versus definitive

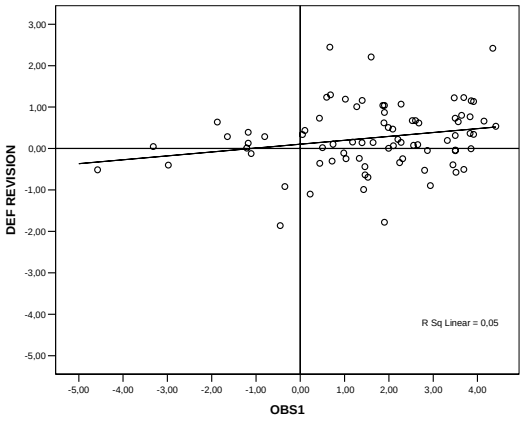
Figure 6, observation 1 vs. revisions, y-o-y.

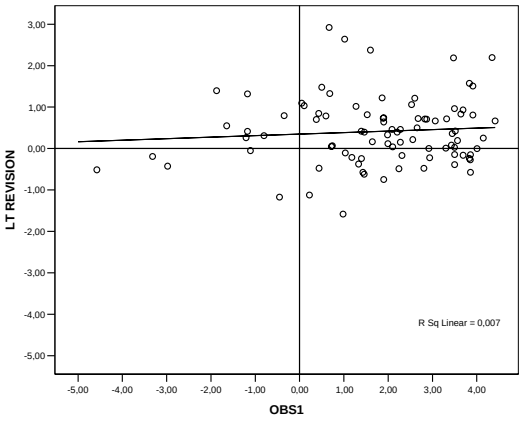




a, short-term revision versus observation 1

b, medium term revision versus observation 1



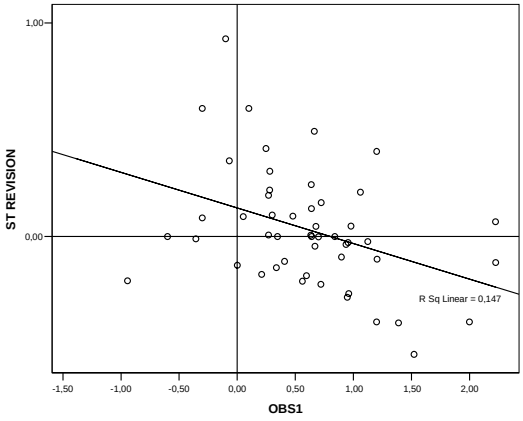


c, definitive revision versus observation 1

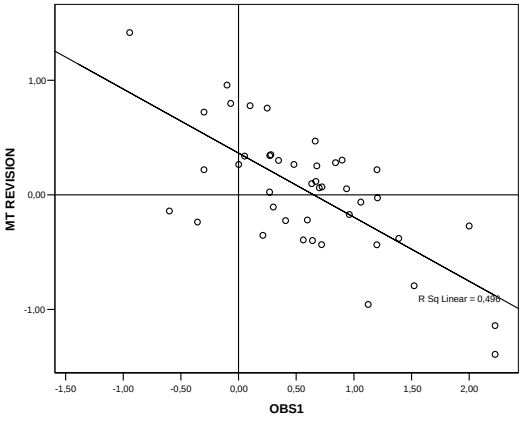
d, long-term revision versus observation 1

Note: If the news hypothesis holds there is no systematic relationship between revisions and first announcements. This corresponds to a zero intercept and a slope equal to zero.

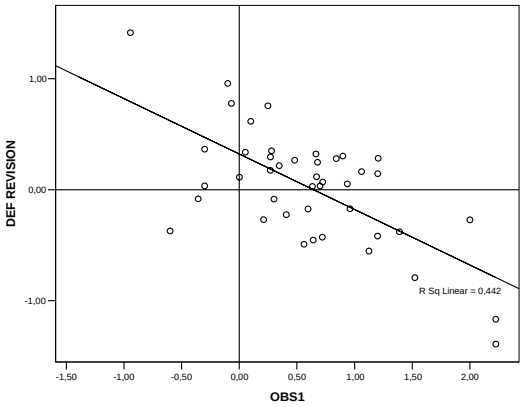
Figure 7, observation 1 vs. revisions, q-o-q.



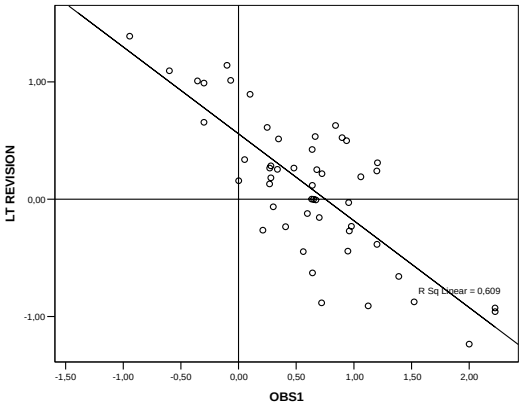
a, short-term revision versus observation 1



b, medium term revision versus observation 1

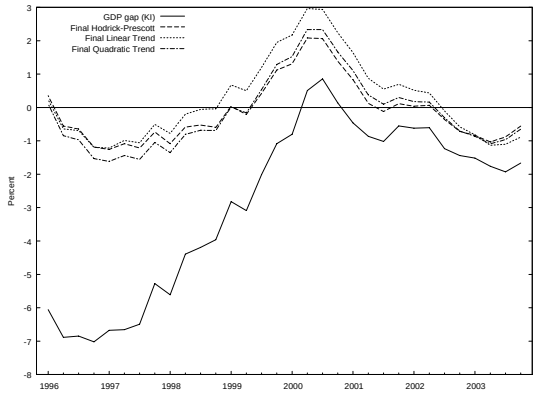


c, definitive revision versus observation 1

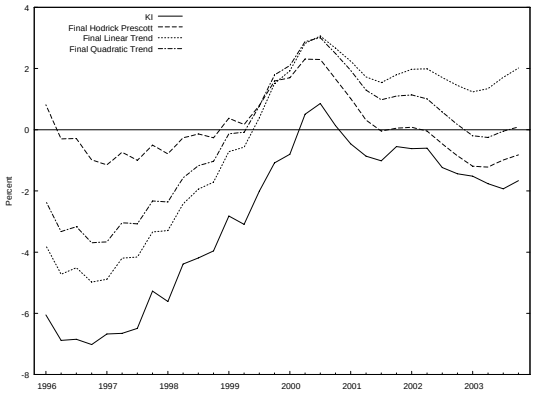


d, long-term revision versus observation 1

Figure 8, GDP gap estimated by the National Institute of Economic Research (KI) and final output gap estimates using the Hodrick-Prescott Filter and deterministic trend models, 1996–2003.

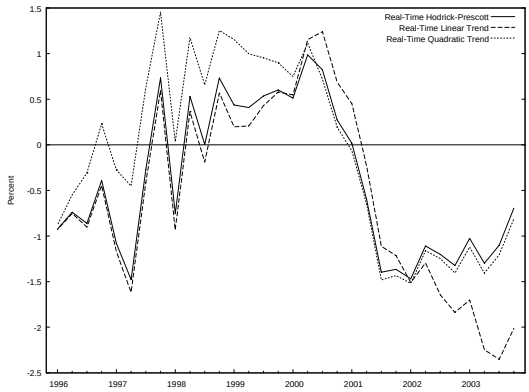


a, official vintages



b, quasi vintages

Figure 9, real-time estimates of the business cycle, 1996–2003.

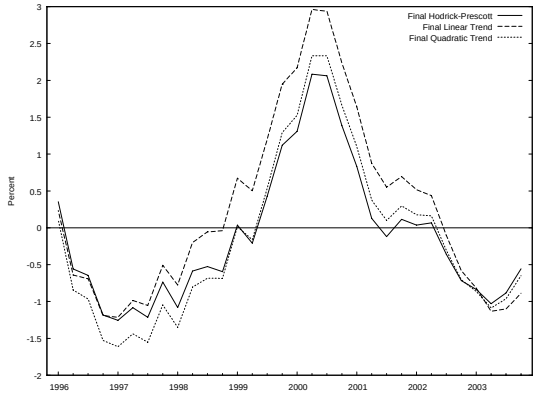


a, official vintages

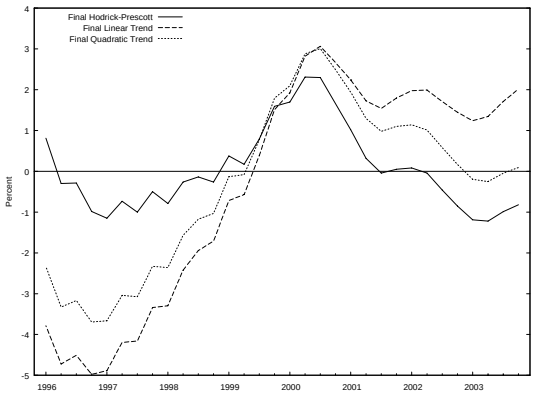


b, quasi vintages

Figure 10, final estimates of the business cycle, 1996–2003.

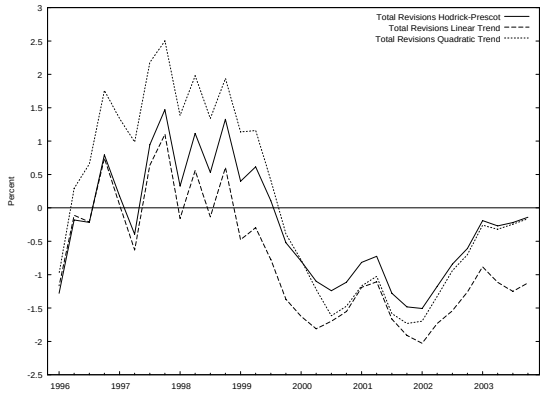


a, official vintages

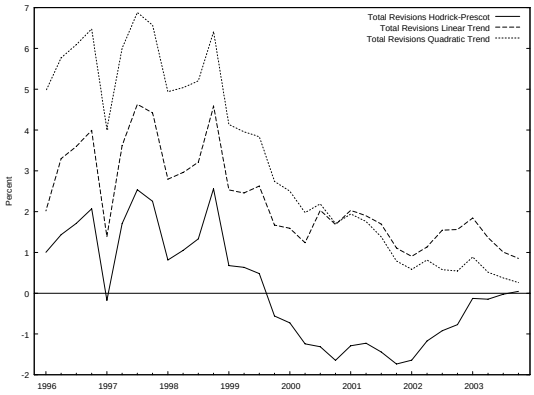


b, quasi vintages

Figure 11, total revisions in business cycle estimates, 1996–2003.

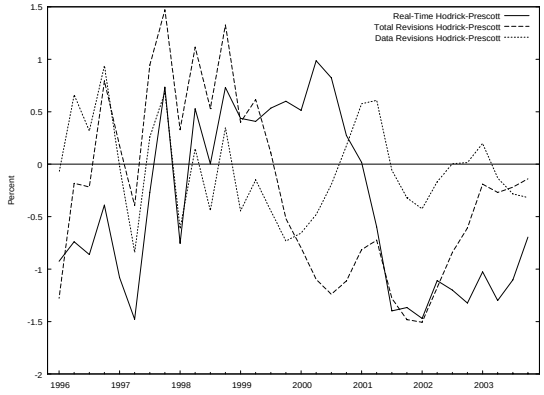


a, official vintages

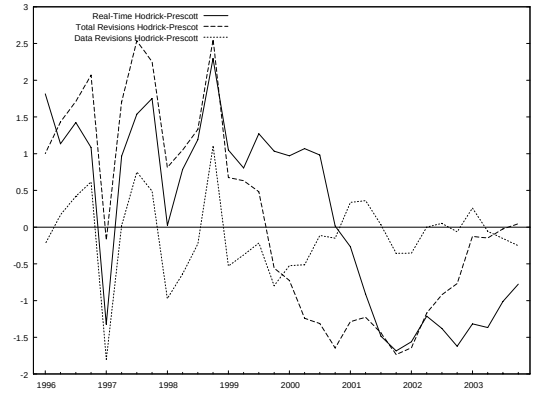


b, quasi vintages

Figure 12, real-time Hodrick-Prescott output gaps and revisions, 1993–2003.



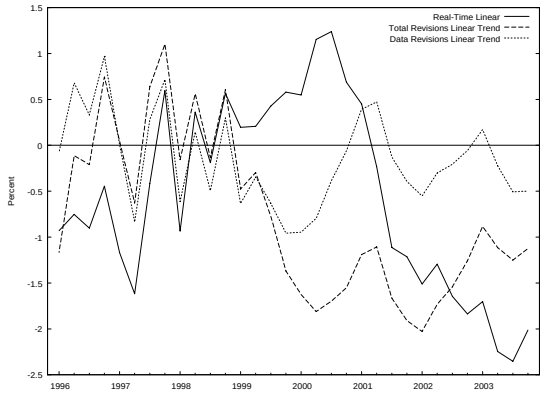
a, official vintages



b, quasi vintages

Note: Even though the real-time output gap at times appears to be large, final estimates of the output gap may actually tell us that the gap is close to zero. Where real-time estimates of the output gap and total revisions are of the same magnitude, i.e. the distance between the lines is zero, the final gap is zero (this is due to how total revisions are defined). For example, in *figure 12 a*, in the middle of 2001 the real-time estimate suggests that the economy is operating below it's potential but since the distance to total revisions is virtually zero, final estimates tell us that the economy was in fact just in line with the trend, i.e. the final output gap is zero.

Figure 13, real-time linear trend model output gaps and revisions, 1993–2003.

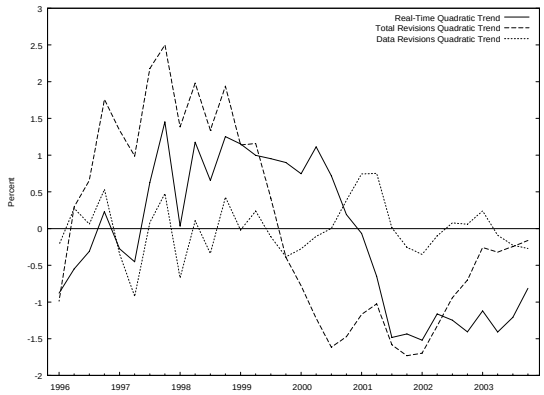


a, official vintages



b, quasi vintages

Figure 14, real-time quadratic trend model output gaps and revisions, 1993–2003.

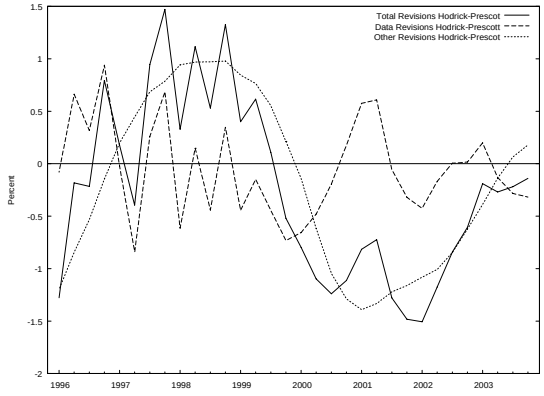


a, official vintages

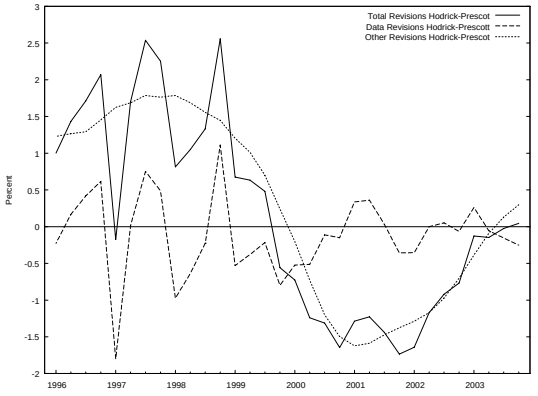


b, quasi vintages

Figure 15, decomposition of Hodrick-Prescott revisions, 1996–2003.



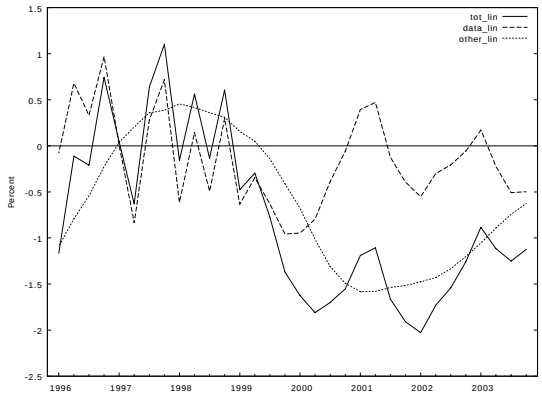
a, official vintages



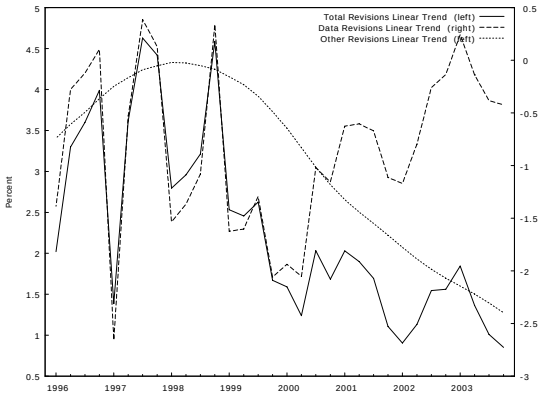
b, quasi vintages

Note: Total revisions are decomposed into two parts: data revisions and other revisions. Data revisions are the part of the revision of the output gap due to revised national accounts. Other revisions are revisions of the output gap due to new information about the development of the economy.

Figure 16, decomposition of linear trend model revisions, 1996–2003.

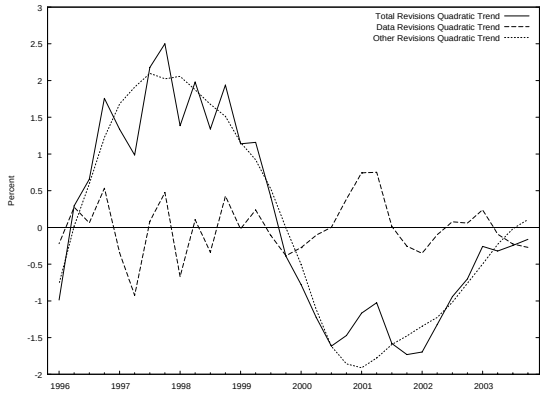


a, official vintages

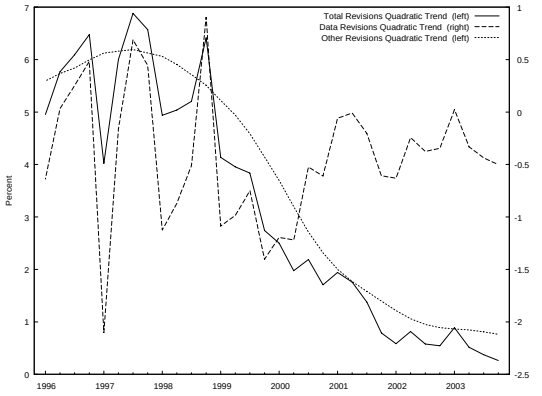


b, quasi vintages

Figure 17, decomposition of quadratic trend model revisions, 1996–2003.



a, official vintages



b, quasi vintages

Table 1, GDP announcements' publication scheme.

Quarter in T:	**1 Quarter**	**2 Quarter**	**3 Quarter**	**4 Quarter**
Observation 1	June **T**	Sept. **T**	Dec. **T**	March **T+1**
Observation 2	Sept. **T**	Dec. **T**	March **T+1**	June T+1
Observation 3	Dec. **T**	March **T+1**	June T+1	Sept. T+1
Observation 4	March **T+1**	June T+1	Sept. T+1	Dec. T+1
t+8	June T+2	Sept. T+2	Dec. T+2	March T+2
Definitive.	Dec. T+2	Dec. T+2	Dec. T+2	Dec. T+2
Final	2006Q1	2006Q1	2006Q1	2006Q1

Note T: indicating years, t: indicates quarters.

Table 2, summary statistics for GDP revisions.

	Short-Term		**Medium-Term**		**Long-Term**		**Definitive**		**Öller & Hansson**	
	y-o-y	**q-o-q**	**y-o-y**	**q-o-q**	**y-o-y**	**q-o-q**	**y-o-y**	**q-o-q**	**y-o-y**	-
MEAN	-0.04	0.03	0.16	0.03	0.42	0.11	0.26	0.02	0.2	-
MEDIAN	0.00	0.00	0.10	0.06	0.38	0.18	0.18	0.07	0.2	-
A(MEAN)	0.32	0.20	0.62	0.40	0.69	0.49	0.62	0.37	0.7	-
SIGN	3	4	2	5	5	7	1	5	-	-
CORR[27]	0.08	0.05	0.56	0.12	0.46	0.26	0.55	0.12	0.4	-
MSE	0.22	0.08	0.66	0.28	0.82	0.37	0.68	0.39	-	-
MIN	-1.21	-0.55	-1.78	-1.39	-1.58	-1.23	-1.86	-1.39	-1.2	-
MAX	1.93	0.93	2.45	1.42	2.92	1.39	2.45	1.42	1.8	-
SD	0.47	0.28	0.80	0.54	0.81	0.60	0.79	0.52	0.9	-
AD/DC	16	13	22	13	27	18	18	13	14	-

Note: short term y-o-y: 1984:4-2005:3, q-o-q: 1993:2-2005:3. Medium term y-o-y: 1984:4-2003:4, q-o-q: 1993:2-2003:4. Long term y-o-y: 1984:4-2005:4, q-o-q: 1993:2-2005:4: Definitive y-o-y: 1984:4-2003:4, q-o-q: 1993:2-2003:4. Öller and Hansson: 1980:1-1998:4.

[27] This is correlation between each respective revision and the final estimate. For example, in the third column CORR measures correlation between medium-term y-o-y revisions and estimates in t+8.

Table 3, Mincer-Zarnowitz regression results.

	Short-Term		Medium-Term		Definitive		Long-Term	
	y-o-y	**q-o-q**	**y-o-y**	**q-o-q**	**y-o-y**	**q-o-q**	**y-o-y**	**q-o-q**
α	0.0336	0.1339	-0.0028	0.3638	0.1038	0.3213	0.3468	0.5577
SE	0.0680	0.0663	0.1582	0.1171	0.1039	0.1261	0.1490	0.0800
p-value	0.6225	0.0491**	0.9860	0.0034***	0.3211	0.0147**	0.0223**	0.0000***
β	-0.0428	-0.1666	0.0966	-0.5587	0.0941	-0.4999	0.0365	-0.7399
Std.e	0.0251	0.0689*	0.0517	0.1588	0.0354	0.1747	0.0524	0.0636*
p-value	0.0924*	0.0195**	0.0654*	0.0011***	0.0095***	0.0066***	0.4877	0.0000***
JB	0.0004***	0.1455	0.1090	0.1548	0.0449**	0.4550	0.0452**	0.8811
F	2.0137	2.9193	3.0747	6.2096	5.1419	4.0971	9.9361	70.3712
P-value	0.1398	0.0636*	0.0518*	0.0044***	0.0080***	0.0239**	0.0001***	0.0000***
R^2 adjusted	0.0172	0.1296	0.0390	0.4838	0.0382	0.4288	-0.0046	0.6007
BG(1)	0.3820	0.6830	0.0017***	0.7660	0.1090	0.7400	0.1130	0.0039***
BG(4)	0.7620	0.0989*	0.0375**	0.1170	0.7510	0.0361**	0.6460	0.0022***
Ramsey's Reset Test	0.8300	0.0145**	0.4530	0.3960	0.5060	0.1280	0.2720	0.6430
Heteroscedasticity	0.3318	0.0772*	0.4816	0.0114**	0.4237	0.0007***	0.4244	0.1573
ARCH(1)	0.9690	0.2250	0.9369	0.0000***	0.7824	0.0000***	0.5169	0.0031***
ARCH(4)	0.9960	0.0856*	0.9818	0.0002***	0.9413	0.0001***	0.0639*	0.0013***

Note: All standard errors are Newey-West HAC robust with a lag truncation parameter of 4.28 For normality test the Jarque-Bera (JB) p-value is reported. The F-value reported is testing if both coefficients are equal to zero. The Breusch Godfrey test reports for autocorrelation of order 1 and 4. For model specification the Ramsey's reset test was used with two terms. White's test was used for the heteroscedastic test and Engel's test for ARCH(h).
* significant at the 10% level., ** significant at the 5% level, *** significant at the 1% level.

[28] The choice of truncation parameter does not affect the calculated standard errors; hence they are insensitive to the choice of lags.

Table 4, Mincer-Zarnowitz regression results excluding outliers.

	Short-Term		Medium-Term		Definitive		Long-Term	
	y-o-y	**q-o-q**	**y-o-y**	**q-o-q**	**y-o-y**	**q-o-q**	**y-o-y**	**q-o-q**
α	-0.0222	0.2340	0.0255	0.3533	0.0565	0.2926	0.4073	0.5364
SE	0.0892	0.0986	0.2184	0.1241	0.1564	0.1085	0.1750	0.1023
p-value	0.8043	0.0222**	0.9074	0.0073***	0.7190	0.0106**	0.0224**	0.0000***
β	-0.0222	-0.3229	0.0821**	-0.5100	0.1122	-0.4020	0.0094	-0.7104
Std.e	0.0320	0.1249	0.0744	0.1778	0.0564	0.1511	0.0627	0.1471
p-value	0.4889	0.0132**	0.2731	0.0069***	0.0505**	0.0116**	0.8808	0.0000***
JB	0.1325	0.1103	0.1563	0.3044	0.0832	0.4029	0.0553**	0.8431
F	1.3655	3.3539	2.3232	4.4208	4.9623	3.9239	9.1688	15.3774
P-value	0.2611	0.0443**	0.1052	0.0192**	0.0095***	0.0287**	0.0003	0.0000***
R^2 adjusted	-0.0065	0.2541	0.0095	0.3223	0.0299	0.2331	-0.0120	0.3862
BG(1)	0.8230	0.0353**	0.0007***	0.2900	0.0397**	0.2570	0.0823	0.0077***
BG(4)	0.8230	0.0221**	0.0188**	0.3950	0.4530	0.0853*	0.6140	0.0170**
Ramsey's Reset Test	0.2440	0.7100	0.3350	0.6680	0.2840	0.6320		0.1830
Heteroscedasticity	0.1035	0.0610*	0.5495	0.1787	0.6269	0.2319	0.4744	0.4260
ARCH(1)	0.3781	0.0591*	0.9825	0.0522*	0.8617	0.0302**	0.7252	0.5195
ARCH(4)	0.2279	0.1706	0.9451	0.0740*	0.7885	0.1299	0.3094	0.0011***

Note: outliers for y-o-y corresponds to the following dates were excluded 1991:1 1991:3 1992:4 1993:1 1993:2. The definition of an outlier follows Faust et al. (2005), i.e. an outlier is a preliminary growth figure deviating from the mean (0,0177) with more than 3 percentage points. For q-o-q figures the following preliminary observations were excluded: 1993:2, 1993:4, 1997:4, 1998:1, and 1998:2. Preliminary q-o-q growth figures for these deviated more than one percentage points from the mean (0,61). Since the average magnitude of growth figures for y-o-y and q-o-q differ, the definition of an outlier also differs. The deviation has been set approximately to the absolute value of the mean multiplied with two.
* significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level.

Table 5, Swedish results compared with results for the G-7 countries as reported by Faust et al. (2005).

	Canada	France	Germany	Italy	Japan	UK	US	Sweden
α	0.29	0.02	0.31	0.19	0.26	0.27	0.07	0.36
β	-0.3	0.01	-0.76	-0.27	-0.25	-0.32	-0.01	-0.56
F	9.5	0.10	18.50	7.90	25.8	43.00	3.9	6.21
P-value	0.01	0.94	0.00	0.02	0.00	0.00	0.14	0.00
R^2 adjusted	0.23	-0.03	0.44	0.20	0.27	0.26	-0.01	0.48

Note: These are q-o-q medium-term revisions. The figures Faust et al. (2005) present do not cover exactly the same time periods. Their samples begin in 1965:1, 1987:4, 1979:4, 1979:4, 1970:1, 1965:1 and 1964:1. They all end in 1997:4. This comparison should not be taken too seriously since the assumption of white noise error terms for q-o-q medium-term revisions. Thus a comparison is only legitimate when relaxing the ARCH-test for these revisions.

Table 6, summary statistics for output gaps.

	MEAN	MIN	MAX	SD	CORR	ACF
Hodrick-Prescott						
Final	-0.13	-1.26	2.08	0.93	1.00	
Quasi Real-Time	-0.33	-1.40	1.47	0.93	0.60	
Real-Time	-0.39	-1.48	0.99	0.82	0.54	
Total Revisions	-0.26	1.47	-1.51	0.85		0.74
Linear Trend						
Final	0.24	-1.21	2.96	1.23	1.00	
Quasi Real-Time	-0.39	-2.02	1.95	1.15	0.82	
Real-Time	-0.55	-2.35	1.24	1.04	0.71	
Total Revisions	-0.78	1.10	-2.03	0.88		0.77
Quadratic Trend						
Final	-0.16	-1.61	2.33	1.11	1.00	
Quasi Real-Time	-0.11	-1.49	1.28	1.00	0.18	
Real-Time	-0.12	-1.52	1.45	0.99	0.22	
Total Revisions	0.05	2.50	-1.73	1.32		0.89

Table 7, reliability statistics for output gaps.

	CORR	NS	NSR	OPSIGN	XSIZE
Hodrick-Prescott	0.54	0.91	0.94	0.31	0.41
Linear Trend	0.71	0.71	0.95	0.28	0.47
Quadratic Trend	0.22	1.19	1.17	0.47	0.53

Table 8, summary statistics for output gap revisions.

	MEAN	MIN	MAX	SD	ACF	RMS	NSR	NS
Hodrick-Prescott								
Total Revision	-0.26	-1.51	1.47	0.85	0.74	0.87	0.94	0.91
Data Revision	-0.06	-0.84	0.94	0.45	0.24	0.44	0.48	0.48
Other Revision	-0.2	-1.39	0.98	0.82	0.93	0.84	0.90	0.89
Linear Trend								
Total Revision	-0.78	-2.03	1.1	0.88	0.77	1.17	0.95	0.71
Data Revision	-0.16	-0.96	0.97	0.50	0.35	0.52	0.42	0.41
Other Revision	-0.62	-1.58	0.46	0.72	0.96	0.95	0.77	0.59
Quadratic Trend								
Total Revision	0.05	-1.73	2.5	1.32	0.89	1.30	1.17	1.19
Data Revision	-0.01	-0.92	0.75	0.37	0.16	0.37	0.33	0.34
Other Revision	0.05	-1.91	2.1	1.36	0.96	1.34	1.20	1.22

Table 9, summary statistics for output gaps, quasi vintages.

	MEAN	MIN	MAX	SD	CORR	ACF
Hodrick-Prescott						
Final	0.04	-1.22	2.31	1.01	1.00	
Quasi Real-Time	0.28	-1.58	2.04	1.24	0.40	
Real-Time	0.17	-1.68	2.30	1.26	0.32	
Total Revisions	0.13	-1.73	2.56	1.34		0.82
Linear Trend						
Final	-0.38	-4.98	3.06	2.77	1.00	
Quasi Real-Time	2.75	-1.15	6.11	2.19	0.94	
Real-Time	1.91	-3.50	5.10	2.08	0.93	
Total Revisions	2.29	0.85	4.63	1.12		0.63
Quadratic Trend						
Final	-0.32	-3.69	3.01	2.03	1.00	
Quasi Real-Time	3.31	0.60	6.07	1.66	0.32	
Real-Time	2.87	0.26	5.38	1.57	0.23	
Total Revisions	3.18	0.26	6.88	2.26		0.89

Table 10, reliability statistics for output gaps, quasi vintages.

	CORR	NS	NSR	OPSIGN	XSIZE
Hodrick-Prescott	0.32	1.32	0.89	0.44	0.63
Linear Trend	0.93	0.40	0.29	0.22	0.44
Quadratic Trend	0.23	1.12	0.47	0.53	0.78

Table 11, summary statistics for output gap revisions, quasi vintages.

	MEAN	MIN	MAX	SD	ACF	RMS	NSR	NS
Hodrick-Prescott								
Total Revision	0.13	-1.73	2.56	1.34	0.82	0.90	0.89	1.32
Data Revision	-0.12	-1.80	1.11	0.55	0.00	0.05	0.05	0.54
Other Revision	0.25	-1.62	1.79	1.25	0.97	0.99	0.97	1.23
Linear Trend								
Total Revision	2.29	0.85	4.63	1.12	0.63	0.79	0.29	0.40
Data Revision	-0.84	-2.66	0.39	0.78	0.30	0.55	0.20	0.28
Other Revision	3.13	1.28	4.33	1.05	0.94	0.97	0.35	0.38
Quadratic Trend								
Total Revision	3.18	0.26	6.88	2.26	0.89	0.94	0.47	1.12
Data Revision	-0.44	-2.10	0.90	0.64	0.11	0.34	0.17	0.32
Other Revision	3.62	0.76	6.19	2.16	0.95	0.98	0.48	1.07

References

Apel, M., Hansen J. and Lindberg, H. (1996), "Potentiell produktion och produktionsgap." *Penning och valutapolitik,* No .3, 24–36.

Berg, C., Jansson, P. and Vredin, A. (2004), "How Useful Are Simple Rules for Monetary Policy? The Swedish Experience." Working Paper No. 169. Sveriges Riksbank Working Papers Series.

Bernhardsen, T., Eitrheim, O., Jore, A.S, Roisland, O. (2005), "Real-time data for Norway: Challenges for Monetary Policy." *North American Journal of Economics and Finance*, Vol. 16, No. 3, 333-49.

Croushore, D., Stark, T. (1999), "A Real-Time Data set for Macro Economists." Working Paper Federal Reserve Bank of Philadelphia, No. 99:4.

Croushore, D., Stark, T. (2000) "A Funny Thing Happened On the Way to the Data Bank: A Real-Time Data Set for Macroeconomists." *Federal Reserve Bank of Philadelphia Business Review*, Vol. 0, No. 0, 15–27.

Demtra User Manual (2002), Release Version 2.0.

Dillén, M., Nilsson, K., Zettergren, G. (2005) "Konjunkturinstitutets syn på resursutnyttjandet." *Ekonomisk Debatt,* No. 7.

Faust, J., Rogers, J.H., Wright J.H. (2005), "News and Noise in G-7 GDP Announcements." *The Journal of Money, Credit and Banking*, Vol. 37, No. 3, 403–419.

Flodén, M. (2005), "Angående KIs syn på resursutnyttjandet och arbetsmarknadsgapet." *Ekonomisk Debatt,* 8.

Friedman M. (1947), "Lerner on the Economics of Control." *The Journal of Political Economy*, Vol. 55, 405–416.

Hodrick, R., Prescott, E. (1997), "Post-war Business Cycles : An Empirical Investigation." *Journal of Money, Credit and Banking* , Vol. 29, No. 1, 1–16.

Ingves, S. (2006), Speech at the Swedish Economics Association, "The inflation target and monetary policy.", available [online]:
http://www.riksbank.com/templates/Page.aspx?id=21100 [2006-08-17].

Jansson, P.J., Palmqvist, S. (2005), ”Mer om Konjunkturinstitutets arbetsmarknadsgap.” *Ekonomisk Debatt* No. 6, 53–57.

Kuttner, K. (2004), “The Role of Policy Rules in Inflation Targeting.” *Federal Reserve Bank of St. Louis Review*, Vol. 86, No. 4, 89–111.

Mankiw, G., Runkle, D., Shapiro, M (1984), “Are preliminary announcements of the money Stock rational forecasts?” *Journal of Monetary Economics*, Vol. 14, 15–27.

Mankiw, G., Runkle, D., Shapiro, M. (1986), “News and Noise: an analysis of GNP revisions.” *Survey of Current Business*, Vol. 66, 20–25.

Mincer, J., Zarnowitz, V. (1964), “The Evaluation of Economic Forecasts.” In Mincer, J. (ed.), *Economic Forecasts and Expectations*: New York : Columbia University Press.

Orphanides, A. and van Norden S. (1999), “The Reliability of Output Gap Estimates in Real Time.” The Federal Reserve Board Working Paper, *Finance and Economics Discussion Series*, No. 38

Orphanides, A. (2000), The Quest for Prosperity without Inflation, European Central Bank Working Series No 15.

Orphanides, A. and van Norden S. (2002), “The Unreliability of Output-Gap Estimates in Real Time.” *The Review of Economics and Statistics*, Vol. 84, No. 4,

Roodenburg, Olivier (2004), “On the predictability of GDP data revisions in the Netherlands.” Working Paper , No. 4, July. Netherlands Central Bank (DNB).

SCB, *BNP Kvartal* 1980:1–1998:4.

Svensson, Lars E.O. (2002) “The Inflation Forecast and the Loss Function.” CEPR Discussion Papers No 3365.

Taylor, J. (1993), “Discretion versus Policy Rules in Practice.” *Carnegie-Rochester Conference Series on Public Policy*, Vol. 39, No. 0, 195–214.

Viotti, S. (2004), “Swedish monetary policy.” Sveriges Riksbank Economic Review No. 2, also available [online]: http://www.riksbank.com/upload/Dokument_riksbank/Kat_publicerat/Artiklar_PV/er04_2_artikel3.pdf [2006-08-17].

Öhlén, S. (2003), Säsongrensning av Nationalräkenskaperna – Översikt, published by Statistics Sweden, available [online]: http://www.scb.se/statistik/NR/AA9999/2003M00/NRFT0301.pdf [2006-08-07].

Öller, L., Hansson K-G (2003), ”Revisions of Swedish National Accounts 1980–1998 and an International Comparison.” In SOU 2002:118, *Utveckling och förbättring av den ekonomiska statistiken*, Bilaga 4. Stockholm : Fritzes offentliga publikationer, 2002.