

Session Number:8A
Session Title: *Topics in National and Historical Accounting and
Productivity Comparisons*
Paper Number: 4
Session Organizer: Edward Wolf
Discussant:

*Paper Prepared for the 28th General Conference of
The International Association for Research in Income and Wealth
Cork, Ireland, August 22 – 28, 2004*

REVISIONS OF NATIONAL ACCOUNTS 1980-1998.

Swedish Expenditure Accounts and GDP in Some Countries¹

**LARS-ERIK ÖLLER
KARL-GUSTAV HANSSON**

For additional information please contact:

Lars-Erik Öller
Statistics Sweden
lars-erik.oller@scb.se
Author FAX(es)
+46 8 5069 4333

This paper is posted on the following websites: <http://www.ariw.org>
<http://www.econ.nyu.edu/ariw>
<http://www.cso.ie/ariw/ariwhome.html>

¹ This is an abbreviated and elaborated version of the report Öller & Hansson (2002).

REVISIONS OF NATIONAL ACCOUNTS 1980-1998. Swedish Expenditure Accounts and GDP in Some Countries²

Lars-Erik Öller and Karl-Gustav Hansson

Statistics Sweden

Abstract. A revision generally augments a preliminary growth rate. The revision distributions are skew, often with fat tails of outliers. For many variables, revisions are correlated with the business cycle, i.e. growth rates are revised upwards in upturns and downwards in downturns. This results in a tendency toward bimodality in the frequency distributions. Different seasonal patterns further emphasize the incongruence between preliminary and final data. We identify where in the accounts the greatest benefits from increased reliability may be achieved. In the international comparison Statistics Canada has the smallest revisions.

March 2004

² This is an abbreviated and elaborated version of the report Öller & Hansson (2002).

1 Introduction

National Accounts are the most popular device in macroeconomic analysis. This is because they are designed to offer an internally consistent picture of the entire economy. Their weak point is accuracy. It is no trivial task to classify data collected from society into the precise framework of the accounts, and often the figures are approximations. A major cause of unreliability stems from the constant need to revise already published figures. National Accounts started in many countries more than half a century ago, and their inaccuracy has been the main topic of critical voices ever since.

Forecasters, analysts and planners, whether using econometric or judgmental methods, but whose activity is forward-looking, need early figures. Old figures, although they may be more accurate, may have lost their relevance when a decision has to be made. However, if the final figure gives an essentially different picture of the economy, the decision may be seriously sub-optimal. Unreliability of base data may be incorporated in the decision, but then, too, it would carry an extra cost, leading to a welfare loss, as compared to the ideal situation of exact early data. Quoting Cole (1969, p.3): "... (the revisions) may be considered a measure of the price, in terms of accuracy, of up-to-date GNP statistics". On the other hand, not revising figures that are known to be wrong is even more deceptive and requires even larger margins in decisions, thus carrying an even higher social cost.

The accuracy of macroeconomic forecasts is known to decrease rapidly as the horizon recedes. When the inaccuracy is criticized it is often forgotten that the forecaster has to project into the future from preliminary values, the reliability of which may be low, or even unknown. A forecaster can hardly be expected to be able to forecast revisions, which means that the revision figure forms a lower limit for how accurate a forecast can be made.

There are a number studies of European GDP forecast errors, e.g. Öller and Barot (2000). For Sweden they found that the root mean squared error of one-year-ahead GDP forecasts for the period 1971-1997 is 1.6 percentage points. The same measure for revisions is 0.7 pct. points. In a recent study, For many other countries the revisions were even larger. Since all errors are never detected it is hardly an exaggeration to say that half of the forecast error is due to inaccurate data. Dong (2002) studied consensus economic forecasts published in *Wall Street Journal*. These forecasts are made in December and are revised in June and cover that same year. Dong finds that financial forecasts improve when more data have accrued, while for GDP and inflation forecasts, no significant improvement can be found. A reason could be the inaccuracy of preliminary figures on GDP and inflation, while financial data are accurate and need no revision.

Another aim of the present study is to expose the shortcomings of the statistical production process and, if possible, to point out at least the most urgent needs for improvement.

1.1 Small revisions desirable but not an aim per se

In order to achieve timeliness and punctuality early statistical figures are published as preliminary information, which is eventually revised when more information becomes available. The statistical characteristics of these revisions are measures, albeit poor, of the reliability of preliminary figures. This is the aspect of quality that we are investigating in this study on revisions, measured as the difference between final and preliminary growth rates.

It is important not to mix the concept of general statistical quality and small revisions. One can never be sure that a revised figure really is more relevant and accurate than the preliminary one. In fact, there is plenty of evidence of revisions that introduce more error than present in the preliminary figure. Young (1995) warns for equating size of revisions and quality: "... an improvement in the current estimates results in a permanent decrease in revision size, while an improvement in the latest available estimates results in a permanent increase in revision size".

To some extent revisions are an ethical matter. Large revisions can be a sign of diligent statisticians who do their utmost to find errors in the data and who are brave enough to admit that an early figure was wrong, and vice versa for the case of small revisions. A study of revisions simply provides measures of the *nuisance* a consumer of National Accounts experiences when the figure, naively trusted to be correct, changes.

1.2 Revisions as a source of unreliability

Preliminary figures often have to be based on sample estimates that are revised when total annual account data become available. In this case quality measures like standard errors can easily be calculated. However, a much larger source of error is of another type. In Barklem (2000) the following sources are given³:

- Frame errors
- Measurement errors
- Processing errors
- Non-response errors
- Model assumption errors

For none of them can we easily produce some numerical measure of reliability.

For different reasons not all statistics are revised. The Consumption Price Index (CPI) is used as a base for commercial contracts and revisions would cause serious damage in the market where many contracts would have to be renegotiated. Business tendency surveys provide snapshot pictures of the economy using ordinal data that are never revised. Also, there are variables that do not need any revision because the first figure published is accurate. Stock market data is an example. Deals on the stock market are registered as they occur and both price and quantity are correct, and the data are delivered on-line. The high quality of some

³ See Ohlsson (1953), which contains an interesting and still up-to-date discussion of error sources in National Accounts.

financial data has made them popular in econometric research. Producers of National Accounts must accept this challenge and look for ways to increase the timeliness and reliability of statistics on the real economy.

A source of revision that will only be briefly touched upon here, is generated by seasonal adjustment. When new data accrue, old adjusted figures change and have to be revised. This is an important additional source of revisions. We give an example in Paragraph 4.2 where different seasonal profiles in preliminary and final figures emphasize the incongruence between these two data sets.

1.3 Approach and method

A revision is defined as the difference between a final and a preliminary growth rate. For Swedish annual data for year t , preliminary means March $t + 1$ ⁴, which is the time of publication of the first complete quarterly accounts for year t . Final is defined as of December $t + 2$. This is the time when the *first definite annual accounts* are published.

Revisions after $t + 2$ are ignored in this study, although both methodological and general revisions are made much later. By choosing $t + 2$ we try to avoid, as much as possible, revisions that are due to changes in definitions or methods. Leaving the final figure open-ended would further obstruct the meaning and comparability of the measures to be used.

It would be important to analyze the later revisions. According to Tengblad (1992) general revisions in the 1980s changed the Swedish final figures almost as much as the revisions between preliminary and final figures.

The international data, analyzed in Paragraph 5 were supposed to follow this publication scheme as closely as possible, but some deviations had to be accepted, because publishing schemes differ between countries. In some countries $t + 3$ comes closest to what we here mean by "final", and some studies, to which will be referred in the text, use the last available estimate.

In the Swedish study, we are mimicking the situation when a user of National Accounts receives a preliminary figure. Revisions are directly copied from published annual growth rates, which have neither been seasonally adjusted nor corrected for the number of trading days.

A fast and convenient way of conveying to the reader an impression of how revisions behave is to present them as histograms. If historical revisions have any relationship to future ones, a histogram can help a user of statistics to assess the reliability of a fresh preliminary figure. A histogram, in a concentrated form, says much about location, spread, non-normality and outliers. This visual overview is complemented by the following statistical characteristics:

Bias. Like in other studies we are interested in knowing if there is a systematic tendency for final figures to increase or decrease as compared to preliminary values. Here both the *mean* and the *median* (Med) of

⁴ Swedish preliminary quarterly data were published 70-85 days after the quarter has expired, during the period studied here. The publication delay varies between quarters and has become shorter in the 1990s.

revisions are necessary because histograms show that the distributions may be skew, in which case the mean is not a good measure of bias.

Dispersion. Another feature of revisions that could seriously hamper the usefulness of preliminary data is high dispersion, which translates into uncertainty. A practical measure that takes into account both bias and spread is the *mean of the absolute value of revisions* $m(\text{ABS})$ ⁵. The *standard deviation* (S)⁶ measures only dispersion (not bias), as does the distribution-free *range* (R), which is the largest (Max) minus the smallest (Min) value.

The coefficient of correlation between revision (F-P) and *(final) growth* (F), $R(\text{F}, \text{F-P})$.

Sign measures the number of times the signs of the preliminary and final figure differ.

Acceleration or deceleration (Ac/Dc)⁷ is a figure on how many times the two time series have disagreed on the second derivative.

Impact is the product of the mean absolute revision ($m(\text{ABS})$) and the share of the variable in total GDP.

The last four characteristics are important for those who monitor the business cycle. Positive correlation with the business cycle (growth rate) tells the user that a bias in a preliminary estimate may be larger than average in a period of strong growth, and smaller, or even negative in a sharp decline. “Sign” and “Ac/Dc” show if you can rely on the direction of the preliminary growth signal staying the same after a revision, or if the preliminary and final figures seem to indicate different stages of a business cycle. The *impact* on total GDP indicates to the producer of statistics where improved measurement would help the most in reducing revisions of total GDP.

1.4 Preliminary figures vs. forecasts

Following McNees (1989) one can say that “the process of estimating GNP starts with forecasts made many years before a quarter has begun and continues for years after it has ended as preliminary estimates are repeatedly revised”. A preliminary figure of a variable can be considered as an estimate, or more generally, since we will never know the true figure of GDP, all published figures of GDP are estimates, be they produced *ante* or *ex post*⁸.

Forecast studies assume that the eventual outcome is an exact figure that the forecast is trying to hit. But in National Accounts there is no true outcome and any measure of accuracy based on the closeness of a preliminary figure to an “outcome” only measures the distance between two *estimates*. In fact this applies not only to preliminary figures, but also to macroeconomic forecasts.

Over the years a standardized technique has been developed for assessing the accuracy of forecasts; see e.g. Holden and Peel (1982) and Öller and Barot (2000). These standards have been followed only to some

⁵ $m(\text{ABS}) = (1/N) \sum | \text{final} - \text{preliminary} |$, where N is the number of observations.

⁶ As all parametric characteristics, the standard deviation is not a good measure of dispersion for skew distributions.

⁷ Acceleration = change in change when positive, deceleration when negative.

⁸ Cole (1969) finds similarities between preliminary figures and extrapolative forecasts.

extent here. We use the mean absolute value of revisions, instead of root mean square forecast error, common in forecast studies. Few statistical tests are applied because the distributions presented in the following sections do not look like symmetric normal distributions. This complicates parametric testing, and for nonparametric tests more observations would be needed. Instead, the study aims at a description of the revisions during two decades.

1.5 Outline of this study

In Section 2 we start by reviewing some previous studies of revisions of National Accounts in Sweden and elsewhere. In Section 3 Swedish annual accounts in constant prices are studied. The GDP components to be analyzed are: *Private Consumption*, *Government Consumption (Central and Local)*, *Gross Fixed Capital Formation (henceforth: Investments)*, *Change in Inventories (henceforth: Inventories)*, *Exports (goods and services)*, *Imports (goods and services)*, and the total, *Gross Domestic Product (GDP)*. In Section 4 the revisions of these same variables, now in quarterly observations are commented on. In Section 5 the Swedish revisions are compared to GDP revisions of annual data from the following countries: Australia, Canada, Denmark, Finland, France, Germany, Netherlands, Norway, New Zealand, UK and USA. And finally, in Section 6 the results of this study are discussed and some suggestions are made for how to follow-up our study.

In all sections the data cover the period 1980 – 1998.

2 Previous studies

2.1 Revisions of National Accounts in Sweden

For Sweden we have found two studies that have relevance: Tengblad (1992) and Eklöf (1992). Tengblad (1992) starts by studying the statistical discrepancies that arise because data compiled from expenditure and production sources do not initially result in the same estimate. He finds that, in the balancing process, the discrepancies resulted in an upward revision for GDP of 0.15 % in the 1970s and 1980s. This is an important measure of unreliability that we have not had access to in this study, because the data have not been saved in a consistent and retrievable form in the 1990s. Comparing characteristics for the 70s and 80s *ibid.* finds that according to all measures, revisions became smaller in the 1980s. The biases were sizable in the 1970s. In Eklöf (1992) the main question asked is whether the preliminary figure is a better forecast than some naïve or model-based artifacts? The answer is in the affirmative⁹.

Tengblad notes that the National Accounts are built up from pieces that are produced by different departments. Some of the pieces, he continues, are notoriously unreliable as preliminary statistics and have to be corrected judgmentally by the National Accounts department. The main

⁹ Cole (1969) arrives at the same conclusion for U.S. data.

task of that unit, he argues, should be to do the compilation of the data into the current SNA framework. This being a huge work, correcting mistakes by others should be discontinued, and the responsibility transferred to those who compile the bits and pieces from which the accounts are formed. Tengblad also calls for more open explanations to why revisions are made.

2.2 Other studies

One of the first to criticize National Accounts for inconsistency and poor empirical applicability was Oskar Morgenstern (1950). Ever since researchers have encountered these problems. They have looked at inconsistencies of GDP data based on expenditures, production and income, on revisions due to changes in definitions, base years of price indices and other revisions. But all efforts boil down to the same issue.

When reading the next paragraphs the reader should be aware of the differences in measuring revisions in different countries.

2.2.1 Australia

Three studies by the Australian Bureau of Statistics (ABS) should be mentioned. ABS(1997) covers the three measures of GDP: expenditures, production and income revisions for the period 1984-1993. The data are quarterly and measure quarterly change, not annual as in the present study. Initial estimates of Private Consumption were found to be close to final figures. Preliminary figures of Government Consumption were overestimated and indicated the wrong direction in almost half of the observations, c.f. Paragraph 4.1. Revisions seemed to be independent of the growth rates, in accordance with our findings, see Table 3.

The study ABS (1998) is a bold effort to assess the quality of Balance of Payments data. It contains a classification according to reliability of the components, based on a survey among ABS statisticians producing the data. A similar classification of National Accounts data followed in ABS (2000). One of the variables regarded as most reliable (by the statisticians themselves) is Private Consumption, while the group of least reliable variables are some service variables as well as Central Government Consumption; a problematic variable also in the Swedish accounts as will be shown in Section 3.

2.2.2 Netherlands

In Kazemier and van Rooijen (2001), some aspects of reliability of annual National Accounts data are assessed. Dutch revisions have larger bias than the Swedish ones. Government consumption again stands out as problematic. Dispersion is also smaller in Swedish data, but both cases can at least in part be explained by the fact that the Dutch final figures come from year $t + 3$ and hence could be expected to be more revised than their Swedish counterparts of time $t + 2$.

2.2.3 UK

Two recent studies have been made of U.K. revisions. Barklem (2000) covers quarterly National Accounts figures of income, expenditure and production components, both in current and in constant prices. Additionally revisions of some other economic indicators are studied. The report also analyses the first sum of four quarters (our definition of a preliminary

annual figure), which is compared to a *three* year later outcome. The revisions are calculated in a different way than used in our study so we refrain from numerical comparisons and just report the main findings. No bias is found in annual current price GDP, but constant price GDP is underestimated in preliminary figures. Preliminary values of many variables were found to be underestimating growth in expansion periods. The mean absolute value (m(ABS)) of some revisions had declined considerably during the past decade. For GDP the results are corroborated in Section 5.

Symons (2001) investigates only annual figures of GDP in constant prices, and its major components, both expenditures and incomes. The preliminary estimate is the first figure based on annual data issued in August $t + 1$ and the benchmark is the revised figure from August $t + 2$. Note that this is not the same data our study is analyzing, so here, too, we refrain from numerical comparisons. The main results were mostly the same as in Barklem (2000) above:

1. Preliminary figures were on the average too low, i.e. positively biased revisions.
2. The means of the component revisions were larger than the GDP revision, which suggests that some revisions tended to cancel across components.
3. The bias has been falling during the last three decades.
4. Revisions correlate positively with the business cycle.

These studies make references to earlier studies of U.K. revisions, indicating that the Office of National Statistics has repeatedly analyzed its revisions.

2.2.4 USA

No other country's revisions of National Accounts have been analyzed so much as those of the U.S.A. In fact such studies belong to the publication scheme in the BEA¹⁰ series *Survey of Current Business*. The latest is Fixler and Grimm (2002)¹¹, which is said to be the 14th revision study of U.S. National Accounts data, and updates Grimm and Parker (1998). Jaszi (1965) covered the period 1942-1962. Young (1995) discusses five of the earlier studies.

There are at least three differences between the US and the Swedish data:

- (1) Most US studies are made using only quarterly data, Fixler and Grimm (2002) being a notable exception, where annual revisions are also analyzed. The quarterly figures were decomposed into current and constant price figures.
- (2) The quarterly US data are *seasonally adjusted* quarterly growth rates that are normalized to annual growth rates. This automatically introduces a substantial source of revision: the updating of the seasonal estimates.
- (3) A revision is calculated as the difference between the preliminary figure and the best *available* estimate.

Point (2) merits some discussion. The Swedish figures are just pure annual growth rates. The reason to the US practice is that it is thought that

¹⁰ Bureau of Economic Analysis.

¹¹ Moulton et al. (2001) contains a closer study of the most recent revisions.

the quarterly seasonally adjusted differences lead annual growth rates (by half a year) and thus provide earlier information. This may hold for reasonably well-behaved time series, not too much contaminated by random noise, if the seasonal and trend components are estimated by maximum likelihood methods¹², but not generally. Since the 'most accurate' latest estimate is seasonally adjusted in the U.S. case and the raw figure in the Swedish case, the 'true' value being unknown, there is no easy way of comparing which technique would provide e.g. better turning point signals.

The third point was discussed in Paragraph 1.3. Note that the differences in calculation make comparisons difficult by inflating the U.S. revisions in comparison with the corresponding Swedish data.

2.2.5 Academic studies

Some academic writers have taken an interest in revisions of U.S. National Accounts, the best-known and first being Oskar Morgenstern. In the 2nd edition of his book (1963) he devotes two paragraphs, pp. 261-275 to studying revisions in the U.S.A and the U.K. He claims that revisions are so large in the U.S. quarterly GNI that the business cycle fits entirely into the uncertainty interval. Investments, although presumably a good indicator of future production, are found to be of little use as preliminary figures. Many, but not all, of Morgenstern's basic critical statements on economic statistics are still relevant today; see Kenessay (1997).

Zellner (1958) remarks that directional errors often occur in turning points. Stekler (1967 and 1987) argues that preliminary data contain meaningful information; contrary to Morgenstern's statement, and that early data present an approximation to the true pattern of economic movements. But he also emphasizes that there are substantial errors in the GDP components. The quality of the underlying data seem to have improved very little since Morgenstern's studies, according to Christianson and Tortora (1995).

An advanced book, entirely on errors in provisional estimates, is Cole (1969). The author presents evidence of preliminary figures having similar features as extrapolations.

Much later there were some studies on how errors in preliminary figures affect econometric models, two of them being Holden and Peel (1982) and Stekler (1987). There have also been attempts at modeling revisions, see Harvey et al. (1981) and Howrey (1984)¹³.

A researcher who during more recent years has written extensively on revisions and their rationality is K.D. Patterson. He has studied if Different vintages of statistical data are co-integrated, with only one common stochastic trend. The common trend should then be very close to the final figures, in which case we could say that the final figure is the best. For U.K. production data Patterson (2002) finds that there is one common trend for all vintages, but that the final vintage is a linear function of the vintages, and in this sense no better than they are. In U.S. data one trend seems to prevail, but here the last revision is exogenous to the rest, thus representing the common trend, and is in this respect to be regarded as the "best" estimate, see Patterson and Heravi (2004). It is shown that the accuracy of a forecast depends essentially on which vintage of data is

¹² Ad hoc type methods such as X-11 have been used in all countries until recently.

¹³ This is not possible with final data from $t + 2$, see Paragraph 4.1.

used to construct the model and which vintage is used to evaluate the forecasts.

3 Results: Swedish Annual Accounts Revisions

3.1 Frequency distributions of revisions

Diagrams 1-12 show histograms of the revisions of GDP components and 13-15 of total GDP. According to the law of large numbers one expects larger variation in the smaller aggregates than in larger ones, and indeed, this can be seen from the diagrams.

A typical distribution of a large aggregate is that of Private Consumption in Diagram 1. It has small dispersion, positive bias, a tendency to be skew, and an outlier. Smaller aggregates like Exports and Imports of Services are scattered across a large interval on the horizontal axis. They had to be provided with different broken scales to accommodate all the data points. Government Consumption, both Central and Local are also problematic, the former with two, the latter with one observation outside the horizontal scale.

[Diagrams 1-15]

Revisions of Investments can be criticised for being more or less uniformly distributed in the interval -1.1 and 2.7 percentage units. It means that a preliminary figure of 1% growth in fixed investments could with approximately the same probability be either -0.1% or 3.7% . The former might signal a recession, while a figure as high as 3.7% or more has actually been recorded in less than half of the years studied. Morgenstern (1963, p. 272), writing about U.S. data from the 1940s and the 1950s, expressed the same finding when stating that the preliminary figures showing the general direction of the economy were what he called "firm" statistics, but other important preliminary information too often was "weak": "Corporate profits and gross private domestic investment on the other hand would be very interesting for estimating, say, future activity on the stock market, but they are definitely weak series and therefore of little use when needed".

There is a distinct bias in the revisions of Exports (Diagram 7) and the distribution has thick tails, but the dispersion is moderate. The relatively good accuracy is due to Exports of Goods. The revisions of Imports are distributed not much differently from Exports, but the tails are even longer and thicker. A tendency toward bimodality can be seen in all foreign trade revisions. Imports of Goods have a tendency toward bimodality that reflects on total Imports. Preliminary figures of both Exports and Imports of Services are to be considered extremely unreliable, historically, and their revisions could potentially have a considerable effect on GDP. However, the revisions of these two variables are correlated and hence they tend to cancel in Net Exports and in GDP, and may therefore go unnoticed. In Paragraph 3.3 we present evidence that the revisions of these variables have declined in the 1990s.

3.2 Revision characteristics

All variables have at least a moderate positive bias. In two cases the revision bias is extreme: Central Government Consumption and Investments. A possible explanation is that there may be an element of substituting planned for realised consumption. Defence-spending is a major contributor to inaccuracy of Central Government preliminary figures. There may also be widespread imperfections in the accounting systems. Local Government seems to have a better control of data, but note the large bias and skew distribution. Statistics Sweden has been aware of these shortcomings for a long time and has tried to cope with the problems.

[Table 1a]

Confirming what was observed from the histograms, Private Consumption, being the largest sub aggregate, has the smallest revisions, while the largest discrepancies between final and preliminary figures can be found in the variables Central Government Consumption and foreign trade in services. Skewness is reflected in the difference between mean and median.

Are booms associated with underestimating the growth in preliminary figures and do we see overestimation in busts? Four revision variables seem to be correlated in such a way with the business cycle: Central and Local Government Consumption, Imports of Services and GDP. According to a Student's *t* test the correlations are in these cases statistically significant, but because of indications of non-normality, these test results should not be taken literally.

What is the reason to high correlation between revision and growth? What first comes to mind is that preliminary figures can only partially be based on real data. In order to be able to publish aggregate figures the statistician has to use approximations such as interpolation and extrapolation. The former means that a class average is imputed, while in the latter case a trend or the latest observed value is chosen. A class average misses accelerations and decelerations if the missing values represent variables that are most sensitive to cyclical fluctuations, while substitutes like a trend or the previous observation are predetermined to miss a turn, up or down. To the extent that preliminary figures are judgemental, correlation would be an indication of cautious behaviour or judgemental compromises¹⁴. We return to this question in Paragraph 4.1.

The two last columns of Table 2 provide a palpable illustration of the correlation between revision and cycle. In the shaded entries preliminary GDP figures have been overestimating the outcome exactly where the final outcome indicates strong decline in growth. The bold figures in the revision column are associated with strong acceleration in production. As was the case with the other characteristics, the relationship between

¹⁴Nordhaus (1987) discusses possible reasons to "forecast smoothing". Where judgement enters the imputation process the same mechanisms could be at work. Another explanation is given by Mork (1987): "Suppose for example that the available observations and other information indicate 9 % growth. By releasing an estimate of, say 6 %, the BEA can both signal strong growth and hedge against a potential embarrassment should the current signals turn out to be misleading".

revision and growth can also be seen in the histograms as bimodality in GDP revisions in Diagram 13. The left hand top is associated with retarding and the right hand one with accelerating production. The tendency toward bimodality is even more distinct in current price data, see Diagram 14, and also discernable in Diagram 15 showing more numerous quarterly data on revisions. Analyzing distributions of GDP *forecast errors* in Chile, Chumacero (2001) finds bimodality that can be explained by underestimations in upturns and overestimations in downturns, corroborating the similarity between forecasts and preliminary figures, already observed by Cole (1969). Continuing the parallel with forecasts, Harvey and Newbold (2003) report non-normality of macroeconomic forecast errors.

The 'Sign' row in Table 1a indicates that the preliminary figure for GDP has had another sign than the final one just once, in 1982 (the revision is 0.9%), but note that there are only four negative annual growth rates.

According to what has been said above it could easily happen that preliminary figures indicate acceleration but final figures show that there was a deceleration, and vice versa. The row (Ac/Dc) in Table 1a shows that for GDP this has *never* happened during the two decades investigated.

The last row of Table 1a presents a measure of the "Impact" on the total, i.e. GDP, of revisions of the components. The impact is calculated as the product of the share in GDP (see top of table) and the mean of absolute revisions, $m(\text{ABS})$. If e.g. the absolute value of revisions of the shakiest preliminary estimates of GDP components (*Investments, Exports, Exports of Services, Imports of Services*) could be halved, the $m(\text{ABS})$ of GDP revisions could be reduced by more than a quarter (from 0.4 to less than 0.3). In other words, this row tells us where in the accounts the largest benefit in terms of smaller GDP revisions can be achieved. Here we do not calculate the cost of implementing such improvements so that this is not a cost-benefit analysis, only a "benefit-analysis". We return to this matter in Paragraph 3.4.

Relative bias ($r\text{Mean}$), i.e. the ratio between the bias in the revisions of a variable and the mean of the variable being revised, is shown in Table 1b. For GDP the figure 0.25 means that the growth rate indicated by a preliminary figure is on average 25 % too low, so that e.g. a preliminary figure showing that GDP increased by 2 % will on average change to 2.5 % after the final revision¹⁵. The correlation with the cycle could change this expectation in times of strong acceleration or deceleration of the business cycle, see Table 2.

The relative biases (rM) in the foreign trade revisions are negligible in Table 1b. This is because of the strong growth trends in the variables themselves. The problem with these variables can be located to the dispersion of revisions of Exports and Imports of Services.

For Central Government Consumption the bias is more than one and a half times larger than the average growth rate in that variable ($r\text{Mean} = 1.7$)! The standard deviation and $m(\text{ABS})$ of this variable is more or less equal to that of the revisions. This means that the standard deviation of revisions is as large as if one would use average growth as a constant preliminary figure, ignoring the published preliminary figures altogether.

Relating revision dispersion characteristics to those of the variable being revised can be ambiguous, because one doesn't know if high

¹⁵ Note that Bjerke (1974) reports $r\text{Mean} = 27\%$ for Denmark's GDP revisions 1959-1968.

volatility in final figures reflects high volatility in the phenomenon recorded in the statistics or if it is simply a more general sign of poor statistics.

3.3 Have revisions decreased?

In Table 1b we have calculated average relative absolute revisions ($rm(ABS)$), separately for the 1980s and the 1990s. Note that the only variables for which the revisions have decreased considerably, both in numerical terms and in relation to the variable being revised, are the two problem variables: Exports and Imports of Services. The opposite has occurred in Government Consumption. Both $m(ABS)$ and $rm(ABS)$ of Central Government Consumption was four times larger in the 1990s than in the 1980s and these measures doubled for Local Government Consumption. In the 1980s Central Government Consumption preliminary data were so shaky that Statistics Sweden often used models to correct obviously erroneous figures. In the beginning of the 1990s an effort was made to improve the data. At the same time the model-based adjustment procedure was abandoned. These changes do not seem to have been a success. The unreliable preliminary figures of this variable are known to have made current year forecasts *less* inaccurate than one-year-ahead forecasts (!), see Borg (1996). The possible effect on fiscal policy is less well known.

Note that Exports and Imports of Goods had much larger $m(ABS)$ values in the 1990s than in the 1980s, while the relative measures remained unchanged. Whether this is a result of the major change in data collection from total statistics to sampling that was introduced when Sweden joined the EU in 1995, is hard to say.

The tendencies and changes between the 1980s and 1990s can be studied in detail in Table 2 showing revisions over time. Extreme values have been marked in bold figures. For Central Government Consumption the bold figures dominate the 1990s while for Exports and Imports of Services, they are mostly to be found in the 1980s.

3.4 Where to start improving preliminary data

Diagram 16 shows $m(ABS)$ of revisions vs. share of GDP of the variables studied. The horizontal distance to the origin shows how large the revisions are in each variable. Some people would be more interested in a certain variable's reliability as a preliminary figure and less so in its overall effect on GDP. Then it is easy to pick out Exports and Imports of Services as those in most urgent need of improved recording procedures. On the other hand, variables with the greatest distance to the origin should most urgently be looked over for also achieving better preliminary GDP figures. This is a graphical description of the impact figures in Table 1a, but here the GDP share and $m(ABS)$ can be studied together. Private Consumption is used as a standard for the indifference curve

$$\text{GDP Share} \times m(ABS) = 0.2 = 0.5 \times 0.4$$

shown in Diagram 16, where 0.5 is the share of Private Consumption in GDP, and 0.4 its $m(ABS)$ from Table 1a. This determines the position of

the indifference curve. The points to the right of and above this line are variables whose revisions have greater impact than the largest component Private Consumption on GDP revisions. We see that Imports and Exports of Services are now joined by Exports and Imports (total) and by Investments, but Government Consumption and its both components lie underneath the indifference curve, together with Exports and Imports of Goods. The same message is of course delivered by comparing the impact factors in Table 1a to that of Private Consumption.

Diagram 17 is the same comparison, but with data only from the 1990s. There is a remarkable change: Six out of 11 variables were below the Private Consumption indifference curve in Diagram 16; in Diagram 17 only four! Variables that were before below the indifference curve, but now are above it are: Exports and Imports of Goods, Government Consumption, and both its components. Only Investments have moved from above to below. This diagram modifies what was said above if the trends during the 1990s have continued since then. The main target of improvement should now be *Government Consumption* and *Exports and Imports of Goods*.

[Diagrams 16 & 17]

4 Some Remarks on Revisions of Quarterly Data

The analysis was repeated for constant price quarterly figures¹⁶, measuring annual growth rates. The conclusions from the study of annual data mostly hold when the frequency is changed to quarterly. Hence, there is no point in reporting the results as extensively as in Section 3. We will just pick a few features, which may explain some of the results of the previous section and highlight aspects for which quarterly figures are particularly suitable, such as monitoring the business cycle and studying seasonality.

The previous study by Tengblad (1992) covering the 70 and 80s was based on quarterly figures. This allows for a comparison of changes during three decades. There has been a steady decrease in the characteristics of Swedish GDP revisions. In particular, relative bias and relative mean absolute revision have halved since the 70s. But from Section 3 we know that the decrease is far from uniform in the components.

4.1 The business cycle

Opposite signs of changes or accelerations/decelerations signals can easily result in wrong inference concerning the phase of the business cycle. The revisions correlating with the business cycle may also have this effect.

As expected, volatile quarterly figures would more often have different signs for preliminary and final data. From pure numbers one would expect four times more different signs as compared to annual figures. It is slightly surprising that different signs occur only twice with quarterly GDP data, i.e.

¹⁶ In Öller and Hansson (2002) revisions of current value annual figures are also studied, as well as implicit price revisions.

in the first two quarters of 1981, when the economy was turning into a contraction. It must have been confusing for economic forecasters to specify the turning point from the statistics at the time. Indeed, both the National Institute of Economic Research (Konjunkturinstitutet) and OECD grossly misjudged the Swedish business cycle 1981-1982, see Öller and Barot (2000). During the next 17 years different signs did not turn up again.

But if one were to look at the components of GDP, the direction of activity was not so well reflected in preliminary figures. The worst is Central Government Consumption. There were 26 wrong signs, despite the fact that 12 % of the observations were missing. As an example, the direction of the preliminary figure was not the same as that of the final figure during the entire period 1987, 3rd quarter to 1988 4th quarter. Private Consumption, the rather reliable variable in Section 3, has different signs seven times, four of which occur in the business cycle 1990-1994, which was very severe in Sweden.

Acceleration/deceleration of GDP was never different in preliminary and final *annual* figures (Table 1a). The situation is quite different in quarterly data: 14 misses¹⁷. The most critical cases occurred in the first two quarters of 1981 (see above), in 1991, 4th quarter, and in 1992, 1st quarter, when the length and depth of the recession were underestimated by many analysts, and finally at the start of the mini-recession in 1996.

Diagram 18 shows preliminary and final quarterly figures of GDP for the recession in the 90s. Growth 1994-1995 was systematically underestimated in preliminary figures, which in the uncertain situation in the beginning of the expansion might have postponed the growth signals. Subsequently, the boom was understated. Fortunately, this time messy data did not cause bad forecasts, see Öller and Barot (2000).

[Diagram 18]

High correlations between revisions and the business cycle were found in quarterly data, too, and for the same variables, including GDP. In addition to the reason given in the previous section we here investigate a technical explanation. Preliminary growth rates have the previous year's level in the denominator. This figure is yet to be revised once more before becoming final. Then a revision has this uncertainty in common with a multi-step forecast, which is known to be autocorrelated. Now, the business cycle in the final figures may be expected to be autocorrelated, a common feature among macroeconomic growth rates. It is well-known that two autocorrelated time series are also *contemporaneously correlated*. We tested if there would be autocorrelation in the revisions and found it in some variables, but there was no significant autocorrelation in GDP revisions. Hence, this is hardly a major reason for the correlation between the revisions and the business cycle of this variable.

Quarterly data on investments revealed clearly how a bad imputation may cause correlation with the business cycle and a biased preliminary figure. There is a large *negative* bias in the second quarter (-2.2), for which statistics have to be released so early that investment statistics for that quarter are not yet available. Instead, first quarter data on *expectations* for the next quarter are used as an estimate. Investment expectations appear

¹⁷ That is 1/5 of the observations. Fixler and Grimm (2002) report ¼ misses in US data, but recall that their final figures are the latest available data.

to be overoptimistic, as plans often are, especially in the beginning of the year¹⁸.

We conclude that, by and large, the Swedish preliminary figures would seldom have led business cycle analysts astray, but at some occasions and for some variables it would happen. Going 45 years back in time we quote Zellner (1958) who writes about revisions of US GDP data 1947-1955: "This good showing, however, must be tempered by the fact that a sizable proportion of the errors in direction of change occurred in the neighborhood of cyclical turning points".

Later writers, studying more recent data, come to more or less the same conclusion, namely that preliminary US data give a satisfactory description of the business cycle, but will occasionally transmit a false signal close to a turning point, see Stekler (1967), Young (1995) and Grimm and Parker (1998). They find that preliminary GDP data often overstate a decline before a trough and understate growth in recoveries.

4.2 Looking at seasonal patterns in levels and testing for cointegration

Here we have performed a consistency study of Swedish data, in the spirit of Siklos (1996) and Patterson and Heravi (2004) for US data and Patterson (2002) for UK data. We have also looked at seasonal similarities, or rather dissimilarities as they appear in the lack of seasonal cointegration.

In Diagram 19 we have plotted preliminary and final GDP in levels. No such uniform series existed for preliminary figures so we had to reconstruct it using level figures for 1979 and projecting the series forward according to the annual growth rates.

The general opinion seems to be that a preliminary figure is an unbiased estimate that has a larger error than the final figure only because it is based on less data. The thinking here follows the same lines as with sample estimates: a smaller sample generates less accurate estimates than a larger sample, but if the estimates are correctly constructed both will be consistent estimates of the unknown population value. As emphasized by Patterson (2002), one should expect that a preliminary and a final level estimate both have the same underlying stochastic trend. As can be seen from Diagram 19 this is not the case. The positive bias means that preliminary annual level figures have a lower trend than the final figures. In addition to a unit root on the trend frequency, both series seem to have seasonal patterns that wander over time, indicating unit roots also on the seasonal frequencies. They could still reflect a common seasonal (non-stationary) pattern if the two series were seasonally cointegrated. Just a look at Diagram 19 raises strong doubts that this would be the case. While the final figures most of the time have a seasonal profile like an inverted U, the profile of the preliminary data look like an M.

We performed several HEGY tests, with and without augmentations, and none of them supported the null hypothesis of common unit roots. The conclusion is that the final and the preliminary figures do not have a

¹⁸ Note that there was large positive bias in annual investment revisions, so that the high imputation was eventually overcompensated for.

common trend, nor do they have a common underlying seasonal development.

There is a very practical consequence. Since the two series have significantly different seasonal patterns, revisions due to seasonal adjustment become large. Indeed, the two series follow two different ARIMA models. The model for the model-based adjustment procedure (TRAMO/SEATS) uses almost entirely final figures so that the forecasts needed for adjusting the last figures will be biased and inconsistent estimates of the preliminary figures.

Here we have discussed the seasonal patterns of just the GDP figures. Its components look very much the same; seasonal cointegration can be strongly doubted in each case. The graphs can be obtained from the authors upon request.

[Diagram 19]

5 International Comparison of Revisions of Annual GDP Figures

A brief international comparison will complete this study. The countries whose GDP revisions are studied are: Australia (AUS), Canada (CAN), Denmark (DEN), Finland (FIN), France (FRA), Germany (GER), Netherlands (NL), Norway (NOR), New Zealand (NZ), Sweden (SWE), the United Kingdom (UK), and the United States of America (USA). We will use the devices of the previous sections to show the characteristic features of revisions in the countries participating in this study.

As compared to the previous sections there is an additional complication here. Preliminary, and especially final figures, are not published on the same dates in all countries, and for some countries the dates have varied over the period studied. For others, we do not know the dates. Still, by and large the figures should be comparable.

5.1 Frequency distributions of revisions

The histograms in Diagrams 20-32 have all the features one has become familiar with in the previous sections. Note that the horizontal scales are different here, because the spread in the international GDP data is much smaller than for the Swedish disaggregated variables.

The revisions generally have underestimation bias, the distributions tend to be skew, and to have kurtosis and bimodality. The law of large numbers seems to be working here, too. The distributions of small countries (Denmark, Finland, New Zealand, Norway) have slightly larger spreads than that of larger economies. Possibly, there are two exceptions: The rather large spread of GDP revisions in Germany that has a resemblance to smaller countries, and Sweden's distribution that may be better thought of as belonging to the group of larger countries. Outliers are common. Even some countries with generally small revisions, notably Canada and the UK, have outliers.

In Diagram 32 all revisions have been pooled as if the statistics had been made by the same agency. We see that with 238 data points some skewness still remains.

5.2 Revision characteristics

The characteristics shown in Table 3a resemble those of Swedish data, see Table 1a. Only Canada and Germany have means (bias) of zero, but if bias is measured by the more appropriate median, only the Canadian preliminary GDP revisions are unbiased. The average bias is between 0.2 and 0.3 percentage points, depending on which measure is used. For Australian, Finnish, Dutch and Swedish preliminary figures, one would be better off adding 0.4 pct. points to the preliminary figures.

The standard deviation does not vary much around half a pct. point, whereas measured by the outlier-sensitive range (R), the spread varies between 1.3 pct. points for Sweden and 3 pct. points for Finland.

The coefficient of correlation $R(F,F-P)$ between revisions and final figures is considerable for seven European countries: Finland, France, Germany, Netherlands, Norway, Sweden and the UK. But for Australia, Canada, Denmark, New Zealand and USA there is only low correlation or none at all. The divide goes between Europe and the rest of the countries in this sample, the only exception being Denmark. This may indicate a difference in compilation methods. Recalling the discussion in the previous section on the figure of comparison still being subject to further revision, one could speculate that the previous year's figure is closer to the final in the countries with no autocorrelation.

In Table 3a the *Mean Absolute Value of Revisions* ($m(ABS)$) is in all countries quite close to the average of half a pct. point. The smallest values are achieved by Canada and France (0.3). For most countries the signs of preliminary and final data have been the same, but for some small countries different signs have occurred once (Norway and Sweden) or twice (Denmark and New Zealand). In most countries acceleration and deceleration was indicated differently in preliminary and final data just once, but this never happened in Canada, Finland or Sweden. In Australia two such cases occurred.

[Table 3]

5.3 Have revisions decreased?

According to Table 3b, the Netherlands and Norway have been able to lower the average absolute amount of revision, still not achieving the level of Canada and France, whereas UK has moved from a group of countries with the largest $m(ABS)$ in the 1980s to the lowest value in the 1990s. Note that in some countries the spread of revisions has increased in magnitude.

The smallest value of the relative standard deviation, rS , 0.2 pct. points, is achieved by Canada, Denmark, Finland, Netherlands, Sweden, UK and USA. Note that France has low revision spread in Table 3a, but in relation to growth variations, the uncertainty of preliminary figures is rather high, while the opposite is true for Finland.

Canada has the lowest $rS = 0.1$ (0.3 for $m(\text{ABS})$). For some countries like Finland and Norway the uncertainty of preliminary figures is much lower when related to the uncertainty associated with the (volatile) growth rate in the respective country.

The decrease in $m(\text{ABS})$ in the 1990s as compared to the 1980s for the Netherlands and Norway is also reflected in the relative uncertainty measure $rm(\text{ABS})$. With just one exception, $m(\text{ABS})$ and $rm(\text{ABS})$ decreased between the two decades. For USA no decrease was registered, but comparing our estimate $rm(\text{ABS}) = 0.2$ to the one reported by Morgenstern (1963, p. 270) for the period 1947-1958, $rm(\text{ABS}) = 0.4$, the decrease is substantial.

For completeness, in Table 4 we give the time series of revisions in this study. These will show when the outliers (marked in bold) occurred that can be seen in the diagrams.

[Table 4 and Diagrams 20-32]

6 Conclusions and Suggestions

This study was mostly on Swedish revisions of National Accounts. The citations of international studies and the comparison in Section 5 are intended to give some general relevance. Swedish statistics is hardly among the least accurate internationally. We have seen in Section 2 that variables which were found to be problematic in Swedish statistics were so also in other countries.

Revisions were found to be positively biased (preliminary figures underestimate) and their distributions are skew, often with thick tails. In Swedish data, some variables were revised so much and so erratically that the use of their preliminary figures can seriously be questioned. Exports and Imports of Services in the 1980s and Central Government Consumptions in the 1990s are cases in point. By and large the preliminary GDP figures convey the same message of the business cycle as final data, but in some cases, notably in critical turning points the preliminary and the final statistics have not agreed, obstructing the task of the analyst and the forecaster. All these features can also be found in reports from other countries.

A related phenomenon is the positive correlation between revisions and the business cycle. This could be due to old structures being projected forward in time, either by imputation, where preliminary data are unavailable, or through a database that does not stay up to date in major turning points. An example is the birth of new firms in upturns and the deaths that occur in downturns, both being registered only when annual accounts are available. Caution in judgmental assessments of large changes has been cited as another reason.

There is some evidence of a decrease in bias and $m(\text{ABS})$ of GDP from the 70s to the 90s in the Swedish data. For the components the results are mixed. In the international studies there is also some evidence of decreasing revisions from the early times of National Accounts, but only

some evidence exists for a continuation of that general tendency in later decades.

Many features of preliminary figures were found to be the same as for forecasts. Since macroeconomic forecasts are mainly judgmental, this affinity may hint at a judgmental feature in preliminary statistics.

Not all countries publish any revision studies at all. The Bureau of Economic Analysis in USA has hitherto published no less than 14 such reports. The Office for National Statistics in UK has also published some reports, as has OECD, see York and Atkinson (1997). In the U.S. and the U.K. academic researchers have found an interest in this area and have made some studies. Various international studies were highlighted in Paragraph 2.2.

Some basic information on how much a preliminary figure is expected to change, and if biased, in what direction and how much, should be attached to the very time series when published. A minimum would be a histogram, and/or the mean or median and $m(\text{ABS})$ of revisions, all these being continuously updated. Also, it would be informative to the user of National Accounts if the statistical office would publish the discrepancy that emerges as the difference between the Expenditure and Production Accounts estimates of GDP. In fact, this would be real quality information on reliability, not just comparing two estimates, as we have done in this study. The magnitude of this item measures something nobody knows where it belongs, if anywhere. Furthermore, the user would like to know what has happened to this discrepancy. Has it been added to (or subtracted from) the change in Inventories? What procedures have been followed in preliminary, quarterly accounts, vs. annual National Accounts?

Ideally, revision studies would result in better statistics. However, it is important to realize that the target is *not* to reduce revisions as such. They can readily be set to zero just by discontinuing all revisions, leaving errors and ambiguities in the data. One might also mechanically “improve” the preliminary figures by utilizing the systematic features in the revisions. In fact, many academic studies have shown that preliminary data could be brought closer to final figures through modeling; see e.g. Harvey et al. (1981) and Howrey (1984).

Modeling is a second best solution, the first best being trying to find the source of large and especially systematic discrepancies between preliminary and final figures. More attention should be paid to making the best use of modern data transferring techniques so as to get more timely and accurate figures. Quoting Carson (1995): “Electronic data collection and transfer methods are beginning to improve data collection and editing of data. One of the most important improvements BEA can make to address the problems in this paper is the reengineering of the information technology system”. 10 years of progress in IT should now open even better opportunities. The study also suggests where the greatest benefits are to be attained in the form of smaller revisions. Still, we know little about the cost to implement them. A complete cost-benefit analysis of reducing uncertainty of preliminary estimates remains still to be done.

We stated in the introduction that those who forecast could hardly be required to guess on the revisions that the producer of statistics will make, and hence the spread of revisions form a lower limit to the accuracy of macroeconomic forecasts. The chain can be continued two more steps. The National Accounts cannot be more accurate than the data that are fed into them and that have been produced outside the National Accounts

Unit. We like to repeat what was said in Tengblad (1992) that those who produce the input data should be better aware of the use of that data in the National Accounts, and should carry the full responsibility for the accuracy of the figures delivered. Revision studies should be extended to these data sources.

The second step is to co-ordinate accounting practices in companies and government agencies with the needs of Official Statistics. Now, primary data in many cases are more or less black boxes to both producers and users of statistics. Wrong primary data can be corrected at Statistics Sweden only in cases where the error is obvious and fairly easily detected, and even then at a considerable cost. Where necessary, accounting rules in companies, organizations and government agencies should be made to comply with the need of Statistical Offices so that all necessary data are easily accessible. When this has been achieved the next step is to have the data automatically transferred online; Hansson (1996 and 2002) are examples of such attempts.

Thus two benefits can be achieved from revision studies: better information to the users and improved statistics. Undocumented judgment that is so common in forecasting should be avoided in official statistics. Where imputations have to be done optimal model forecasts should be made instead of ad hoc estimates (hot deck, cold deck etc.).

An attempt should be made to improve the quarterly accounts so that preliminary and final figures would have more or less the same seasonal pattern. Swedish investments during the second quarter is a case in point.

This is only a first attempt at an international investigation of revisions. The authors hope this will inspire other researchers to make more thorough analyses that hopefully would help the statisticians to improve the quality of preliminary macroeconomic statistics. The first task would be to build an international database on revisions. The present study is a very modest start. Other recent contributions of data of all vintages are, for the U.S.A., Croushore and Stark (2001) for the U.K., Egginton et al. (2002). It would be important also to study later revisions than the ones under scrutiny here.

Acknowledgements

We want to thank Svante Öberg for initiating this project and for presenting constructive comments during the work. Per Ericson, Göran, Svensson, Hans Svensson and Olof Bergström from the Division of National Accounts of Statistics Sweden made valuable contributions by gathering our data, commenting extensively on the text. We are also indebted to our foreign colleagues who helped us with Section 5 and to all those who have provided constructive criticism in seminars; in particular to Svein Longva, Steve Landefeld and Tim Holt. Special thanks to Åke Tengblad for valuable comments on our text and to Anders Norberg who helped us with programming. Jonas Nilsson and Olof Rosander made the time series analysis of the revisions. None of these persons carry any responsibility for mistakes and shortcomings in this paper, for which only the authors are to blame.

References

Australian Bureau of Statistics , ABS (1997): Revisions to Quarterly Economic Growth Rates 1984 to 1993, website: www.abs.gov.au.

- “ - (1998): Balance of Payments, Australia, Concepts, Sources and Methods. Paragraph 15. Data quality.

- “ - (2000): Australian National Accounts: Concepts, Sources and Methods. Chapter 29: *Quality of the National Accounts*.

Barklem, A.J.E. (2000): Revision analysis of initial estimates of key economic indicators and GDP components, *Economic Trends*, No. 556 , March, 31-52.

Bjerke, K. (1974): The preliminary and final figures of Danish National Accounts, *The Review of Income and Wealth* (20), 337-347.

Borg, H. (1996): Hur pålitliga är Konjunkturinstitutets prognoser? Analys av prognoser för försörjningsbalansen (How reliable are the forecasts of the National Institute of Economic Research? Analysis of forecasts of the expenditure accounts), *Konjunkturläget*, June, 33-39.

Carson, C.S. (1995): Mid-Decade Strategic Review of BEA's Economic Accounts: Maintaining and Improving Their Performance, *Survey of Current Business*, February 1995.

Chumacero, R. (2001): Empirical analysis of systematic errors in Chilean GDP forecasts, *Journal of Forecasting* (20), 37-45.

Cole, R. (1969): *Errors in Provisional Estimates of Gross National Product*, NBER and Columbia University Press.

Christianson A. and Tortora, R.D. (1995): Issues in surveying business: an international survey, in *Business Survey Methods*, Editors: B.G. Cox, D.A. Binder, B.N. Chinnappa, A. Christianson, M.J. Colledge, P.S. Kott, John Wiley & Sons.

Croushore, D. and Stark, T. (2001): A real-time data set for macroeconomists, *Journal of Econometrics* 105, pp. 11-130.

Dong, W.C. (2002): Do revisions improve forecasts?, Dpt. Of Economics, Wichita State University, Kansas, database EBSCO.

Egginton, D.M., Pick, A. and Vahey, S.P. (2002): 'Keep it real!': a real-time UK macro data set, *Economics Letters* (77), 15-20.

Eklöf, Jan A. (1992): *Varying Data Quality and Effects in Economic Analysis and Planning*, Chapter 5. Stockholm School of Economics.

Fixler, D.J. and Grimm, B.T. (2002): Reliability of GDP and related NIPA estimates, *Survey of Current Business*, January, pp. 9-27.

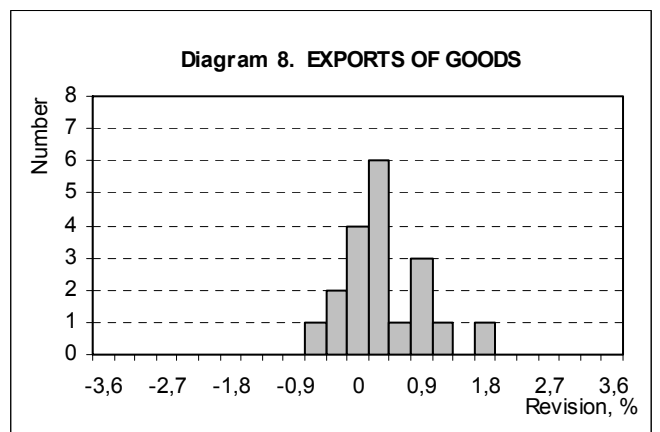
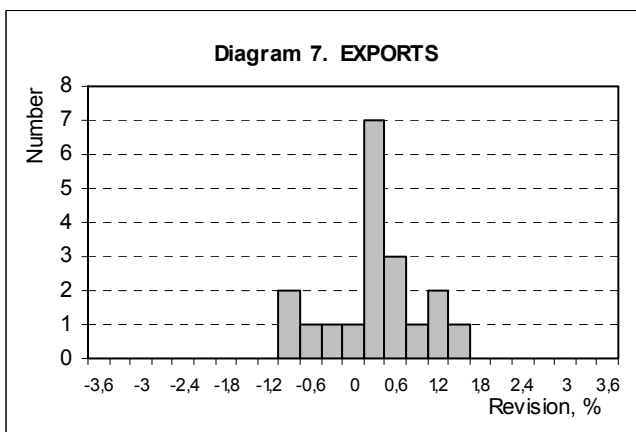
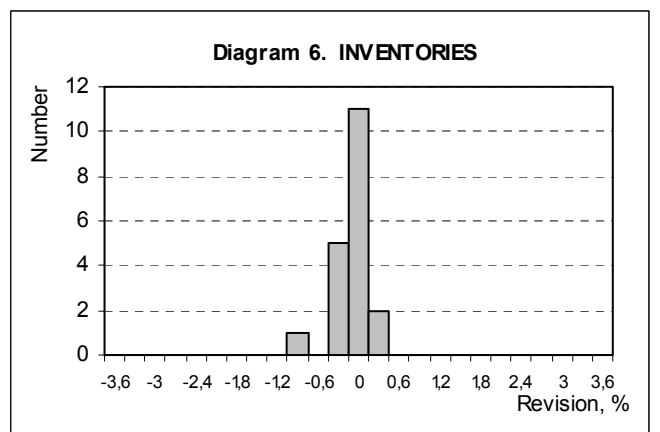
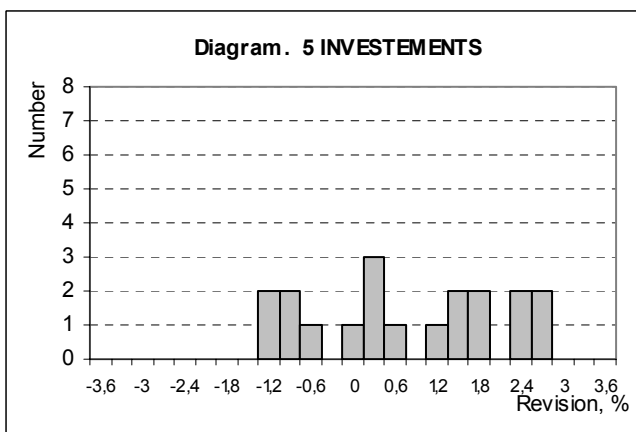
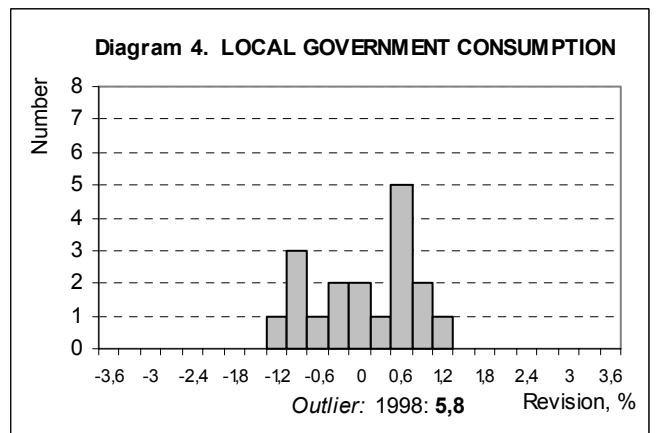
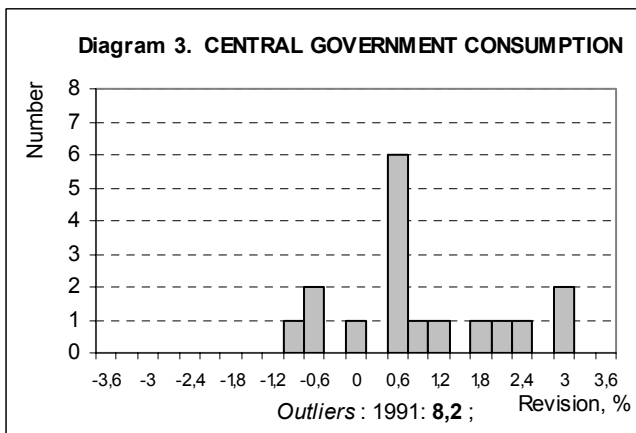
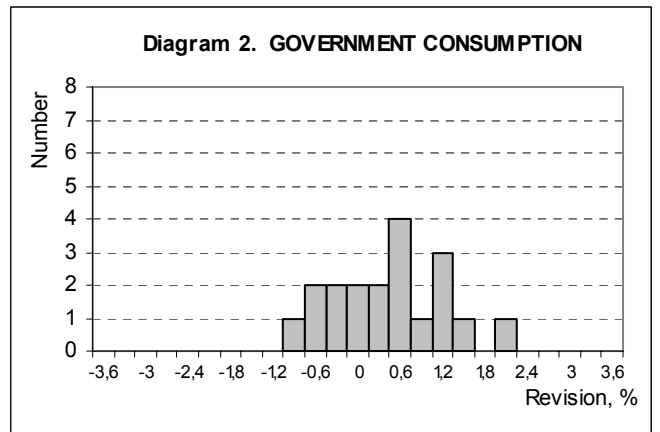
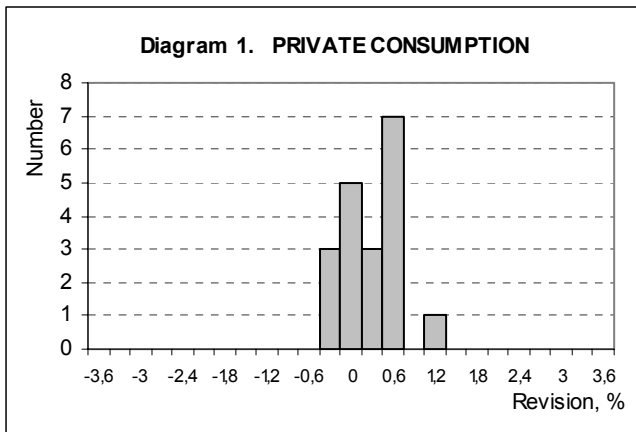
- Grimm, B.T. and Parker, R.P.(1998): Reliability of quarterly and annual estimates of GDP and Gross Domestic Income, *Survey of Current Business*, December, pp. 13-21.
- Hansson, K-G. (1996): Index problems when introducing total screening and on-line reporting in National Accounts (in Swedish with English summary), Arbetspapper Nr. 5, National Institute of Economic Research , Sweden.
- Hansson, K.-G. (2002): Reconciliation of quarterly and annual national QNA, Paper presented at the The 27th IARIV Conference, www.iariw.org.
- Harvey, A.C., McKenzie, C.R., Blake, D.P.C., Desai, M.J. (1981): Irregular data revisions, in *Applied Time Series Analysis of Economic Data*, ed. A. Zellner, Bureau of Census, pp. 329-343.
- Harvey, D.I. and Newbold, P. (2003): The non-normality of some macroeconomic forecasts, *International Journal of Forecasting* (19), 635-653.
- Holden, K. and Peel D.A. (1982): The Relationship between Preliminary and Revised Estimates of GDP and Consumers' Expenditure in the UK, *The Statistician*, Vol. 31, No. 3, 259-266.
- Howrey, P.E.(1984): Data revision, reconstruction, and prediction: An application to inventory investment, *The Review of Economics and Statistics*, pp. 386-393.
- Jaszi, G. (1965): The Quarterly National Income and Product 1942-1962, *The Review of Income and Wealth*, Series XI, Ed. Simon Goldberg and Phyllis Deane, pp. 100-187.
- Kazemier, B. and van Rooijen, R. (2001): Assessment of the reliability of the Dutch provisional National Accounts, Statistics Netherlands Research Report.
- Kenessay, Z. (1997): A perspective on the accuracy of economic observations, *International Statistical Review* (65.2), 247-259.
- De Leeuw, F. (1990): The reliability of U.S. Gross National Product, *Journal of Business & Economic Statistics*, Vol. 8, No. 2, pp. 191-203.
- McNees, S.K. (1989): Forecasts and actuals: The trade-off between timeliness and accuracy, *International Journal of Forecasting* 5, 409-416.
- Morgenstern, O. (1963): *On the Accuracy of Economic Observations*, Princeton University Press, Princeton, Second Edition, First Edition 1950, New Jersey.
- Mork, K.A. (1987): Ain't behavin': Forecast errors and measurement errors in early GNP estimates, *Journal of Business & Economic Statistics*, Vol. 5, No. 2, pp. 165-175.

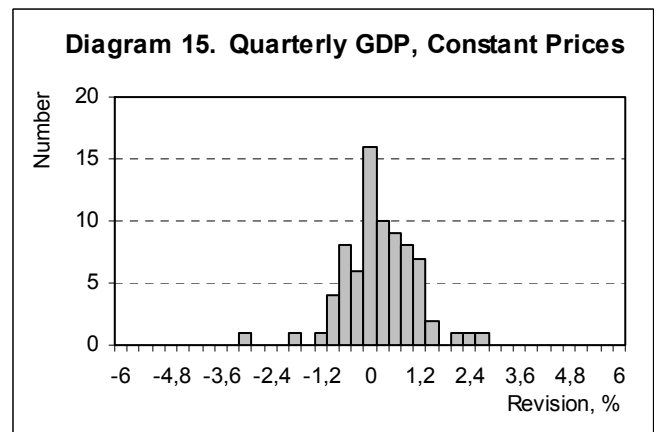
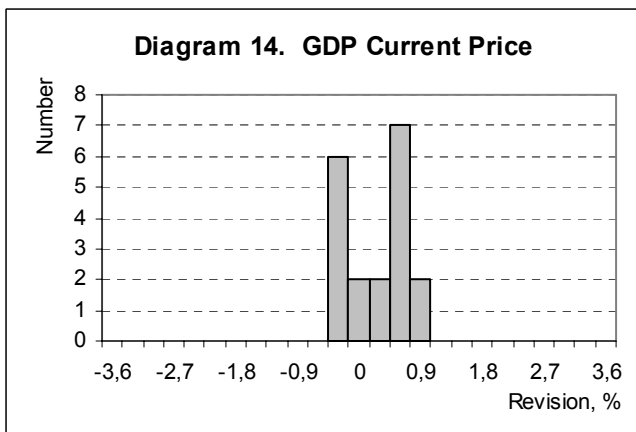
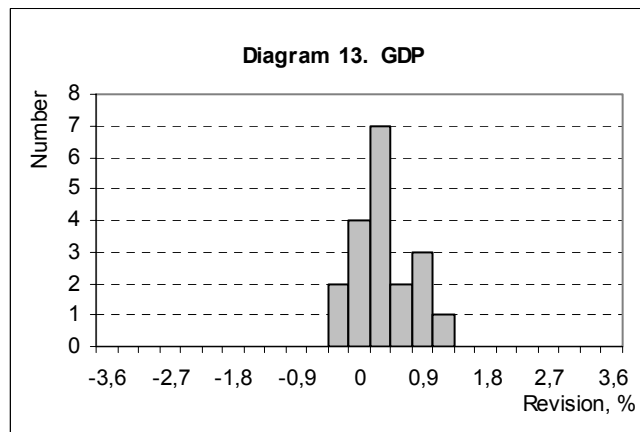
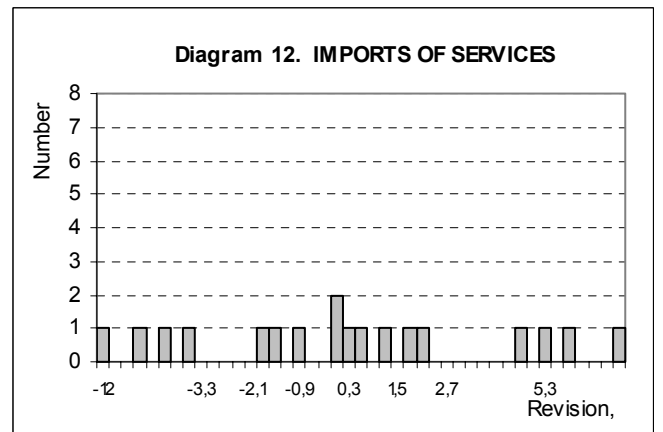
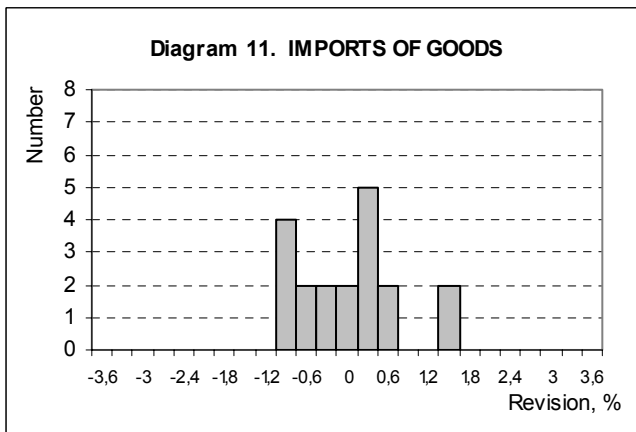
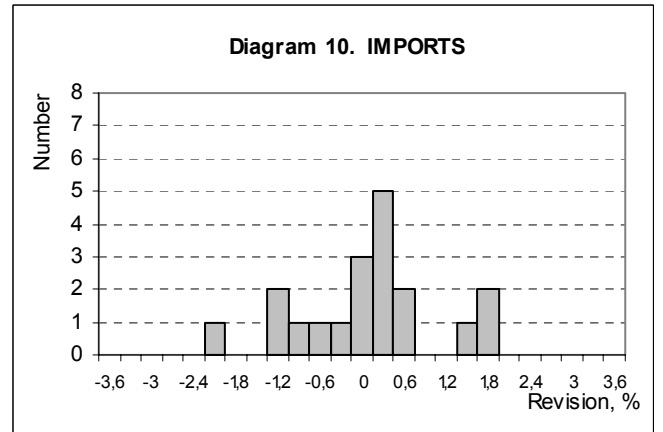
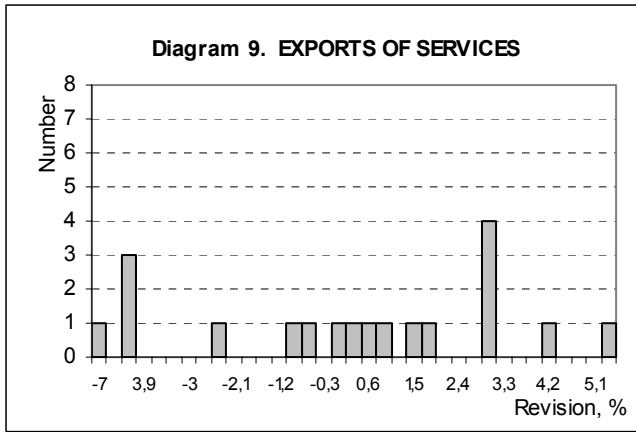
- Moulton, B.R., Seskin, E.P., and Sullivan, D.F.(2001): Annual revision of the National Income and Product Accounts, *Survey of Current Business*, August, pp. 7-32.
- Nordhaus, W.D. (1987): Forecasting efficiency: concepts and applications, *The Review of Economics and Statistics*, 69, 667-674.
- Ohlsson, I. (1953): *On National Accounting*, thesis, reprinted by the National Institute of Economic Research, Sweden.
- Öller, L.-E. and Barot, B. (2000): The accuracy of European growth and inflation forecasts, *International Journal of Forecasting* (16), pp. 293-315.
- Öller, L.-E. and Hansson, K.-G. (2003): Revisions of Swedish National Accounts 1980-1998, and an international comparison, in *Development and Improvement of Economic Statistics*, Whitebook of the Swedish Government, SOU 2002:118.
- Patterson, K.D. (2002): Modelling the data measurement process for the index of production, *Journal of the Royal Statistical Society, Series A*, 279-296.
- Patterson, K.D. and Heravi, S.M. (2004): Revisions of official data on US GNP: A multivariate assessment of different vintages, forthcoming in *Journal of Official Statistics* .
- Siklos, P.L. (1996): An empirical exploration of revisions in US national income aggregates, *Applied Financial Economics* (6), 59-70
- Stekler, H.O.(1967): Data revisions and economic forecasting, *Journal of the American Statistical Association*, June, pp. 470-483.
- Stekler, H.O. (1987): The effect of data revisions and additional observations on time series estimates, *Applied Economics*, 19, 347-353.
- Symons, P. (2001): Revisions analysis of initial estimates of annual constant price GDP and its components, *Economic Trends* No. 568, 40-55.
- Tengblad, Å. (1992): The reliability of the quarterly accounts (in Swedish), Arbetspapper nr 2, National Institute of Economic Research, Sweden.
- York, R. and Atkinson, P. (1997) The reliability of quarterly National Accounts in seven major countries: A user's perspective, OACD, Economics Department, Working Paper no. 171.
- Young, A.H. (1974): Reliability of the quarterly national income and product accounts of the United States 1947-71, *The Review of Income and Wealth* (20), pp. 1-39.

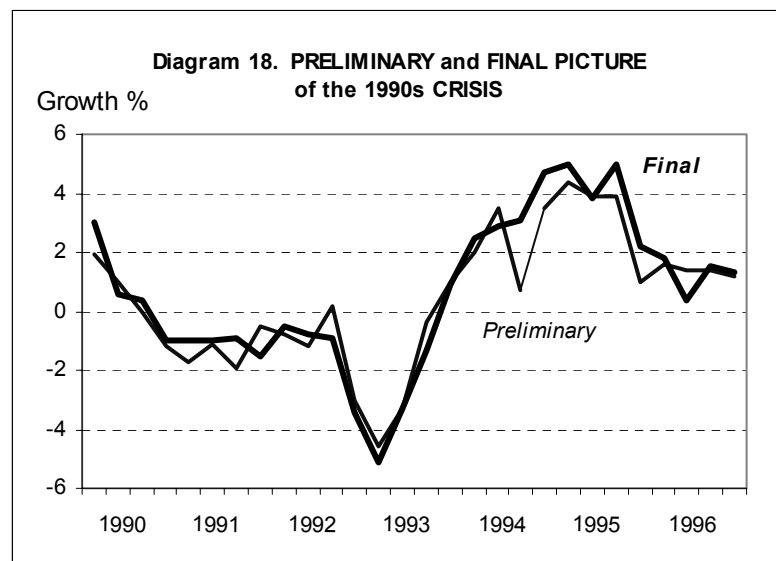
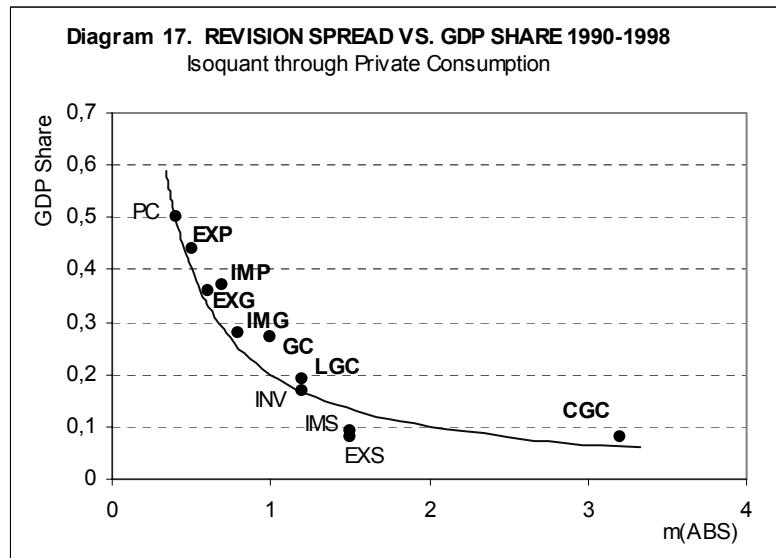
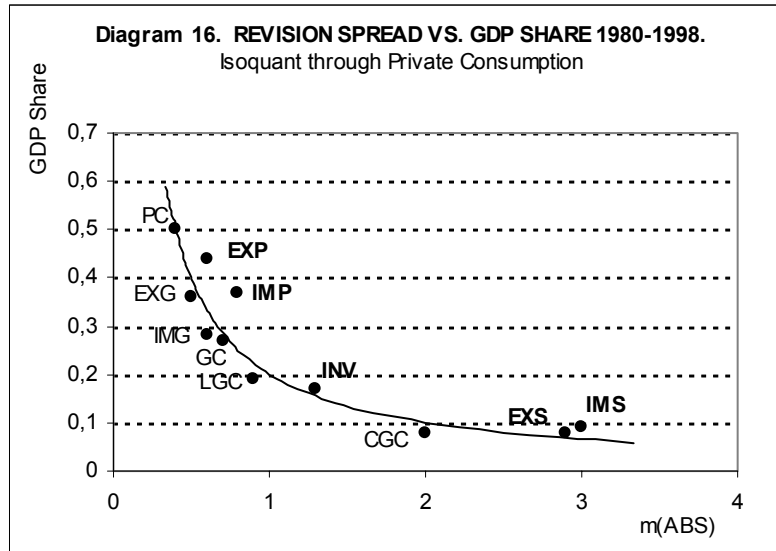
Young, A.H. (1995): Reliability and accuracy of quarterly GDP estimates: A review, in *The New System of National Accounts*, ed. J.W. Kendrick, Kluwer Academic Publishers, pp. 423-455.

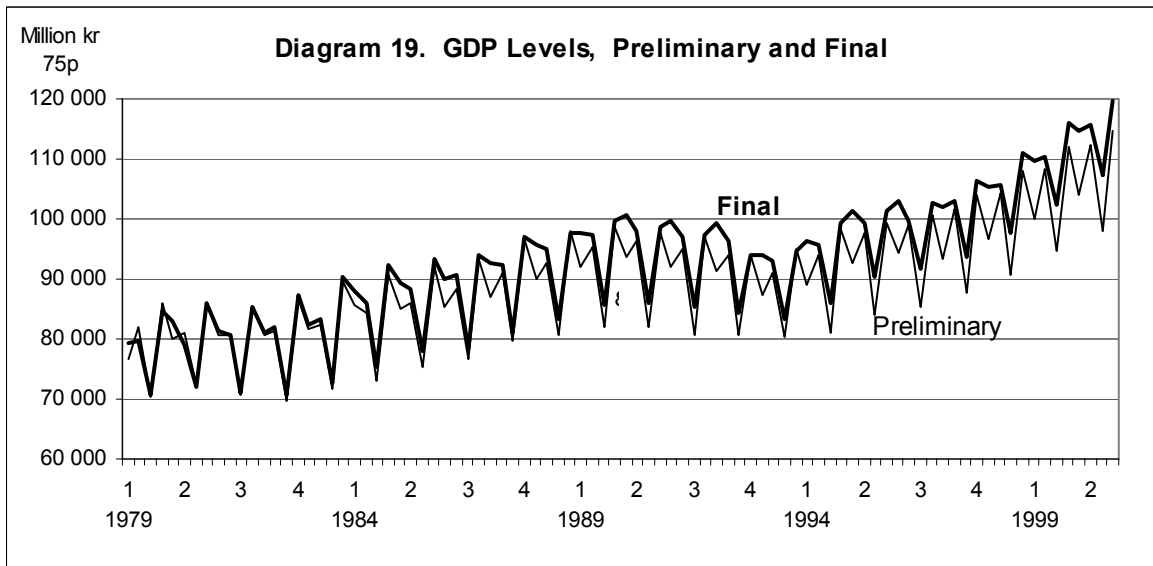
Zellner, A. (1958): A statistical analysis of provisional estimates of Gross National Product and its components, of selected National Income components, and of Personal Saving, *Journal of the American Statistical Association*, March, pp. 54-65

Diagrams 1-13. Distributions of Revisions 1980-1998.
Constant Prices, Annual Growth Rates

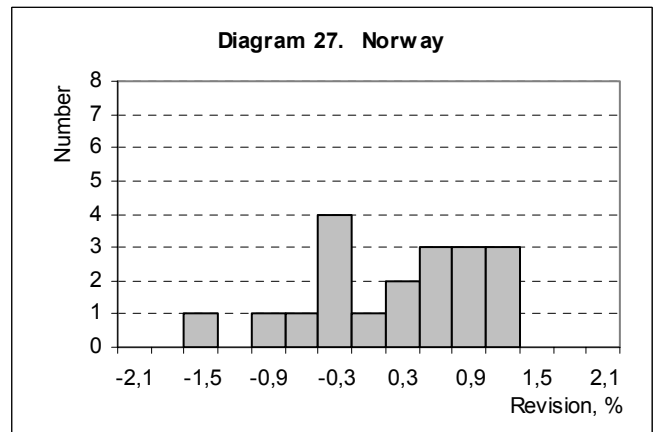
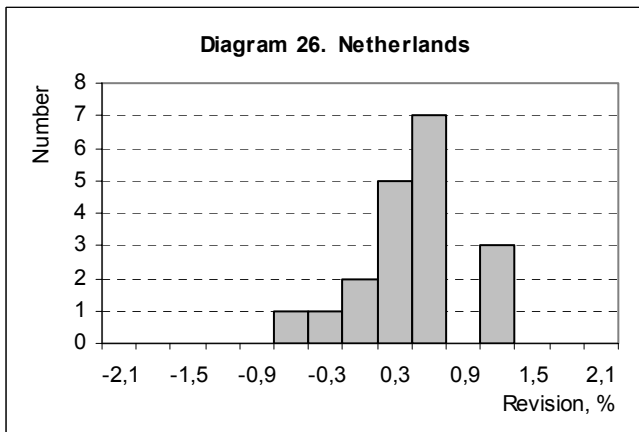
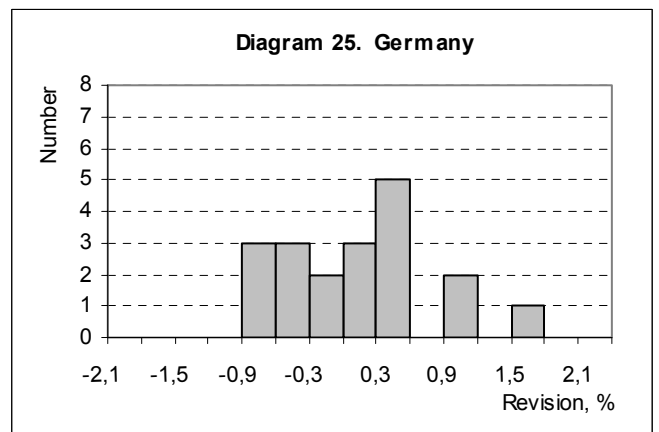
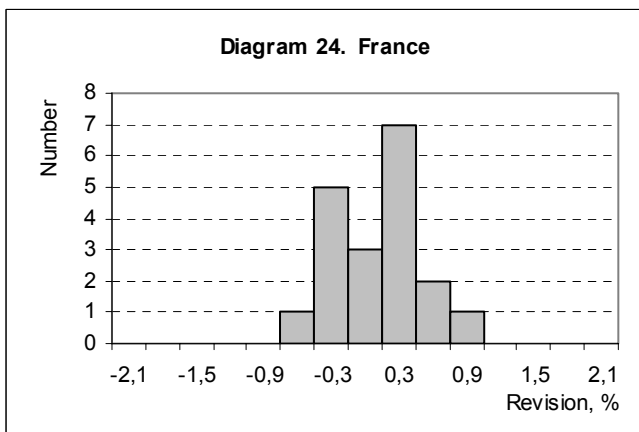
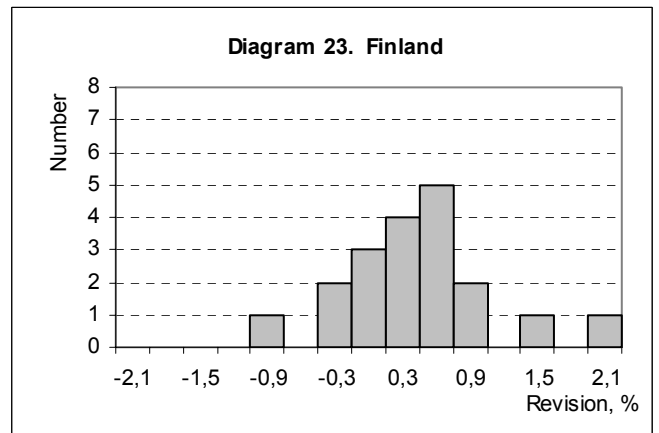
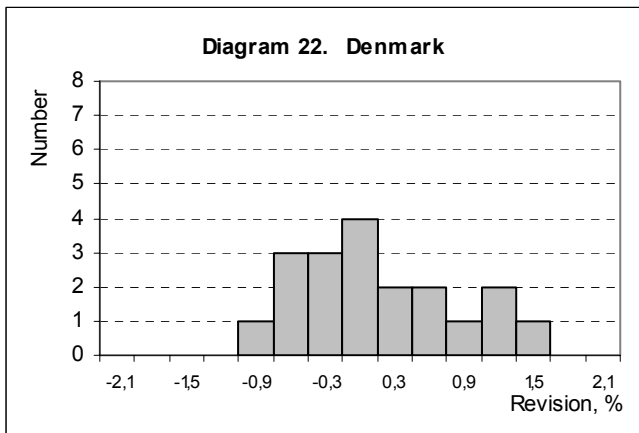
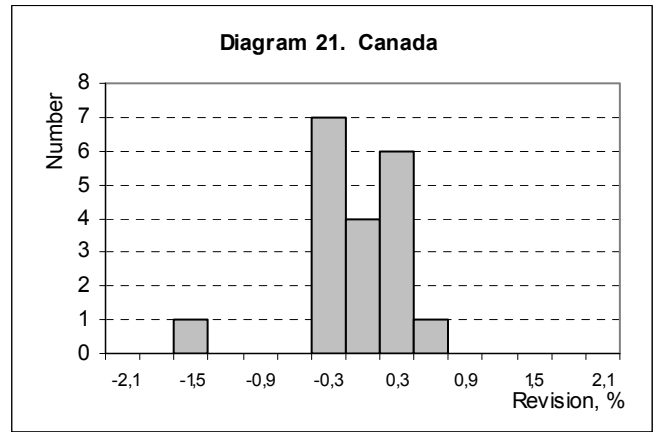
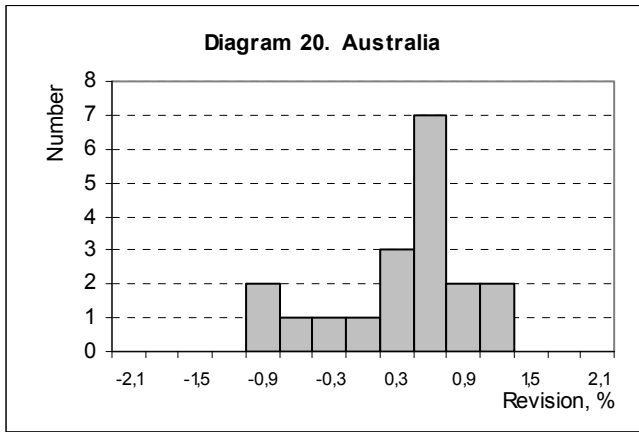








Diagrams 20-32. International Distributions of GDP Revisions 1980-1998.
Constant Prices, Annual Growth Rates



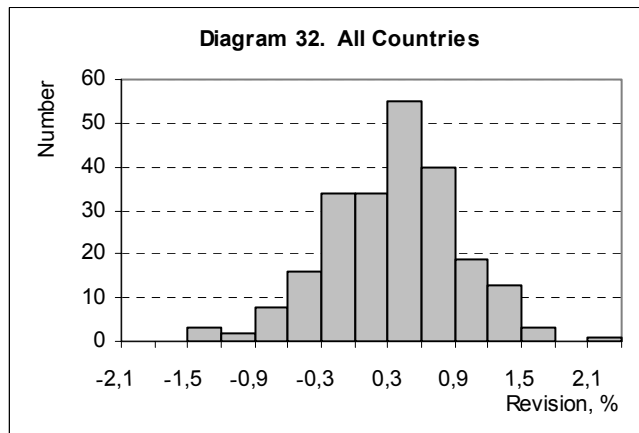
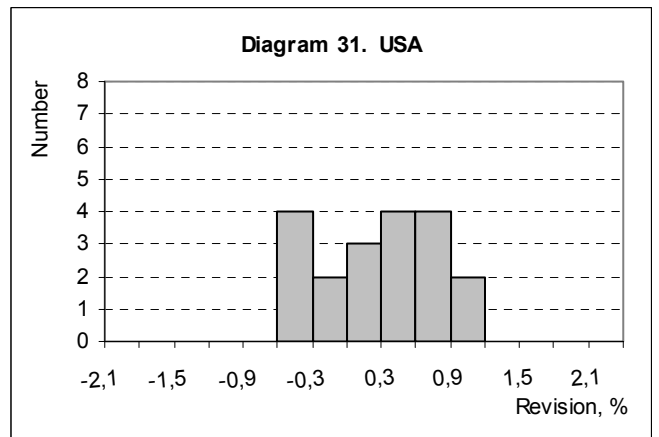
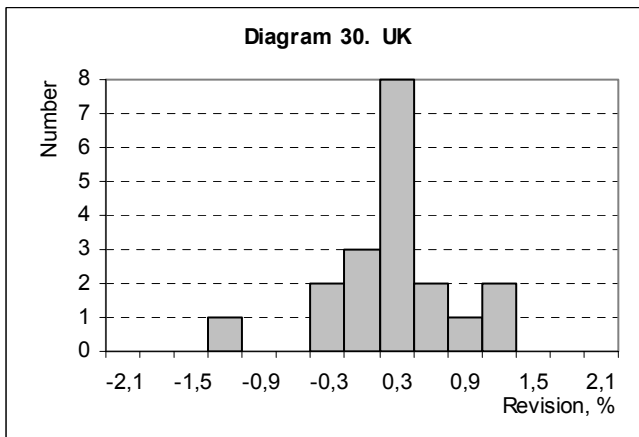
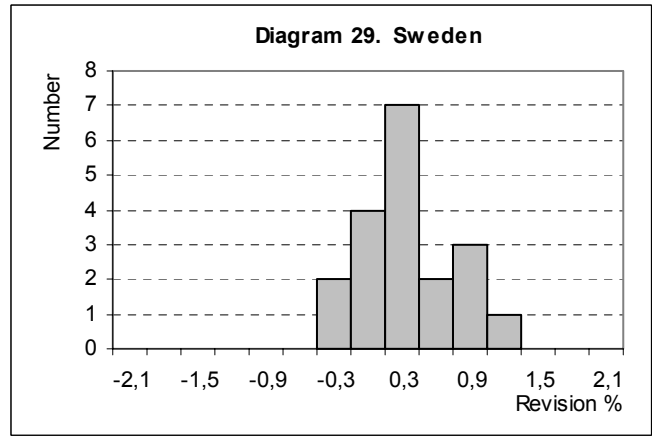
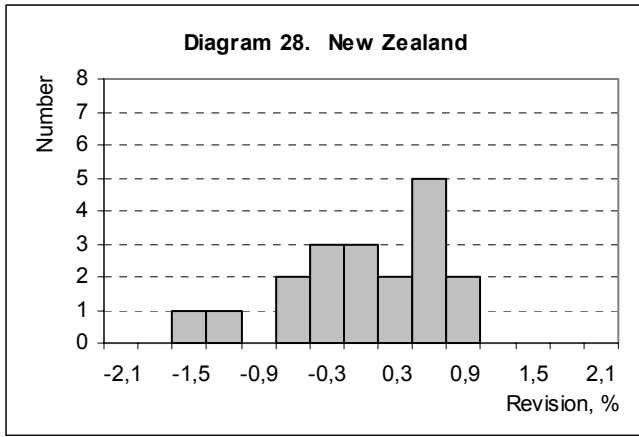


Table 1. Characteristic Features of Revisions 1980-1998.**1a. The Whole Period**

	Private Cons.	Government Consumption			Gross Invest.	Inven- tories	Exports			Imports			GDP
		Total	Central	Local			Total	Goods	Services	Total	Goods	Services	
% of GDP	50	27	8	19	17	1	44	36	8	37	28	9	100
Mean	0,3	0,5	0,9	0,4	0,8	-0,1	0,3	0,4	0,1	0,1	0	0,3	0,4
Median	0,3	0,6	0,7	0,2	0,7	-0,1	0,3	0,2	0,7	0,2	0,1	0,1	0,3
m(ABS)	0,4	0,7	2	0,9	1,3	0,2	0,6	0,5	2,9	0,8	0,6	3	0,4
S	0,4	0,8	3	1,5	1,3	0,3	0,6	0,6	3,7	1	0,8	5,6	0,7
R	1,6	3	16,7	6,9	3,8	1,2	2,4	2,3	14,8	4	2,6	26	1,3
Min	-0,3	-0,8	-8,5	-1,1	-1,1	-0,8	-0,9	-0,6	-7	-2,1	-1	-12	-0,2
Max	1,3	2,2	8,2	5,8	2,7	0,4	1,5	1,7	7,8	1,9	1,6	14	1,1
R(F,F-P)	0,17	0,1	0,64*	0,57*	0,32	0,11	0,09	0,2	0,32	0,27	0	0,47*	0,49*
Sign	1	1	5	2	2	2	0	1	4	0	2	4	1
Ac/Dc	0	5	2	9	0	2	2	3	6	1	0	8	0
Impact	0,2	0,19	0,16	0,17	0,22	0	0,26	0,18	0,23	0,3	0,17	0,27	0,4

1b. Mean Absolute Revisions and Relative Measures. Comparing the 1980s and 1990s.

m(ABS)80	0,4	0,4	0,8	0,6	1,1	0,2	0,6	0,4	4,2	0,8	0,4	4,2	0,4
m(ABS)90	0,4	1	3,2	1,2	1,2	0,2	0,5	0,6	1,5	0,7	0,8	1,5	0,4
rMean	0,3	0,4	1,7	0,2	0,6	-9	0,1	0,1	0	0	0	0	0,25
rS	0,2	0,6	1	0,7	0,2	0,3	0,1	0,1	0,8	0,2	0,1	0,7	0,2
rm(ABS)	0,2	0,5	0,9	0,4	0,2	0,3	0,1	0,1	0,5	0,1	0,1	0,4	0,2
rm(ABS)80	0,2	0,3	0,5	0,3	0,2	0,3	0,2	0,1	1	0,2	0,1	1	0,2
rm(ABS)90	0,3	0,8	1	0,7	0,2	0,2	0,1	0,1	0,2	0,1	0,1	0,1	0,2

m(ABS) = Mean Absolute Value of Revisions,

R() = Coefficient of Correlation, * = Statistically significant, F = Final (t+2), P = Preliminary (t+1), S = Standard Deviation, R = Range

Impact = GDP Share times m(ABS)

Sign=Number of times preliminary and final figures and different signs

Ac/Dc = Acceleration/Deceleration, i.e. the number of times that preliminary and final figures show different developments.

m(ABS)80 = as above 1980-1989

m(ABS)90 = as above 1990-1998

r = relative; rMean= relative mean, i.e. mean of revision of variable x divided by mean of x ; rS and rm(ABS) analogously.

Table 2. Time Series of Revisions 1980-1998.

	Private Cons.	Government Consumption			Gross Invest.	Invent. Change	Exports			Imports			GDP	GDP Final*
		Total	Central	Local			Total	Goods	Services	Total	Goods	Services		
1980	0,3	-0,6	-0,6	-0,4	1,7	0	0,9	0,4	2,9	0,4	0,3	0,7	0,5	1,9
1981	-0,1	0,3	0,9	0,1	0,3	-0,3	0,4	0,9	-2,4	-1,3	-1	-3,5	0,4	-0,5
1982	0	-0,3	0,6	-0,7	1,4	0,4	0,6	0,1	3	-0,6	-1	2	0,9	0,8
1983	-0,2	-0,1	-0,1	0	2,7	-0,3	-0,8	0,1	-5,2	-1,2	0,4	-12	0,1	2,4
1984	0,5	0,6	0,5	0,5	1,5	0	0,6	0,2	3	-0,4	0,4	-5,6	1	4
1985	0,3	0,6	0,6	0,6	-0,5	-0,1	-0,6	0,2	-5,1	0,5	-0,1	5,3	-0,2	2,1
1986	0,7	-0,5	0,6	-0,9	0,2	-0,3	1,2	0	7,8	1,9	0,3	14	-0,2	1,1
1987	0,5	0,6	-0,5	1,3	0,1	0,1	-0,9	1,2	-7	0,3	-0,6	6	0,1	2,9
1988	0,3	-0,8	-0,9	-0,8	0,7	-0,1	0	-0,2	0,4	-0,8	0,2	-8	0,2	2,3
1989	0,7	0	3,1	-1,1	2,3	-0,4	-0,4	0,4	-4,9	0,3	0,7	-1,4	0,3	2,4
1990	0	1	1,9	0,7	1,3	0	0,3	0,5	-0,7	0,6	0,7	-0,1	0,4	0,7
1991	-0,1	2,2	8,2	-0,2	-1,1	0,4	0,3	-0,3	2,9	1,9	1,6	2,8	0	-1,1
1992	0,5	-0,3	0,7	-0,9	0,2	-0,1	0,3	0	1,6	0	-0,7	1,3	0,3	-1,4
1993	0,7	0,7	1,3	0,7	-1	-0,8	0,4	0,9	-1	-2,1	-1	-1,9	-0,1	-2,2
1994	1,3	0,3	0,7	0,2	2,4	0	0,2	0,2	0,1	0	0,1	-0,8	1,1	3,3
1995	0,5	1,4	2,4	0,8	1,8	0	1,5	1,7	0,7	1,5	1,4	1,9	0,9	3,9
1996	-0,2	1,5	3,1	0,9	-1	-0,1	0,5	0,3	0,9	0,2	0,2	0,4	0,2	1,3
1997	-0,3	1,1	2,3	0,5	2,6	-0,2	0,2	-0,6	4,1	0,1	-0,3	0,1	0,2	2
1998	0,1	1,3	-8,5	5,8	-1,1	0,1	1,1	0,9	1,7	0,2	-0,9	4,6	0,7	3,6

Table 3. Characteristic Features of Revisions in Some Countries 1980-1998.**3a. The Whole Period**

	AUS	CAN	DEN	FIN	FRA	GER	NL	NOR	NZ	SWE	UK	USA	ALL
Mean	0,4	0	0,2	0,4	0,1	0	0,4	0,2	0,1	0,4	0,3	0,1	0,2
Median	0,5	0	0,1	0,4	0,2	0,1	0,4	0,3	0,1	0,3	0,3	0,2	0,3
m(ABS)	0,6	0,3	0,6	0,6	0,3	0,5	0,5	0,7	0,6	0,4	0,4	0,4	0,5
S	0,6	0,4	0,4	0,7	0,4	0,7	0,3	0,8	0,7	0,4	0,5	0,5	0,5
R	2,2	2	2,3	3	1,5	2,6	1,9	2,7	2,6	1,3	2,3	1,7	3,8
Min	-1	-0,8	-0,9	-0,9	-0,5	-1	-0,7	-1,4	-1,6	-0,2	-1,2	-0,7	-1,7
Max	1,2	0,6	1,4	2,1	1	1,6	1,2	1,3	1	1,1	1,1	1	2,1
R(F,F-P)	0,06	0,16	0,06	0,42	0,41	0,52*	0,38	0,6*	0	0,49*	0,54*	0,28	..
Sign	0	0	2	0	0	0	0	1	2	1	0	0	..
Ac/Dc	2	0	1	0	1	1	1	1	1	0	1	1	..

3b. Mean Absolute Revisions and Relative Measures. Comparing the 1980s and 1990s.

m(ABS)80	0,7	0,3	0,6	0,6	0,4	0,4	0,6	0,8	0,5	0,4	0,6	0,5	0,5
m(ABS)90	0,6	0,3	0,5	0,5	0,3	0,6	0,4	0,6	0,7	0,4	0,2	0,4	0,5
rMean	0,1	0	0,1	0,2	0,1	0	0,2	0,1	0	0,25	0,15	0	0,1
rS	0,3	0,2	0,2	0,2	0,3	0,5	0,2	0,4	0,3	0,2	0,2	0,2	0,3
rm(ABS)	0,2	0,1	0,3	0,2	0,2	0,2	0,2	0,3	0,3	0,2	0,2	0,1	0,2
rm(ABS)80	0,2	0,1	0,3	0,2	0,2	0,2	0,3	0,4	0,25	0,2	0,2	0,2	0,2
rm(ABS)90	0,2	0,1	0,2	0,1	0,2	0,2	0,1	0,2	0,3	0,2	0,1	0,2	0,2

For acronyms, see Table 1.

Table 4. Time Series of Revisions in Some Countries 1980-1998.*Final (t+2) - Preliminary (t+1)*

	AUS	CAN	DEN	FIN	FRA	GER	NL	NOR	NZ	SWE	UK	USA
1980	-1,0	0,3	-0,6	0,7	0,3	-1,0	0,3	1,0	0,1	0,5	-1,2	-0,2
1981	0,7	0,3	-0,6	0,5	0,3	0,1	0,5	-0,1	0,4	0,4	0,9	0,7
1982	-0,2	0,2	-0,1	2,1	0,4	0,1	0,1	-0,4	0,5	0,9	0,0	-0,4
1983	1,0	0,2	0,1	-0,4	-0,1	0,8	0,7	-0,3	0,7	0,1	0,3	0,2
1984	-0,5	0,4	0,5	-0,2	0,2	0,2	1,1	0,9	-0,4	1,0	0,4	-0,5
1985	0,3	-0,2	0,5	0,5	0,5	-0,6	1,1	1,1	-1,1	-0,2	0,3	0,8
1986	0,6	0,0	0,3	0,1	0,5	-0,2	0,3	-0,6	-0,1	-0,2	1,1	0,3
1987	0,6	0,6	1,3	0,8	0,2	-0,1	-0,7	1,3	-0,3	0,1	1,1	0,7
1988	1,2	-0,1	1,4	0,8	1,0	0,3	-0,2	-1,4	0,6	0,2	0,6	0,4
1989	0,9	-0,4	-0,6	0,4	0,2	0,2	1,2	-0,8	0,5	0,3	-0,2	-0,5
1990	0,0	-1,4	-0,2	0,0	-0,2	1,0	0,6	-0,2	-0,6	0,4	-0,2	-0,1
1991	0,2	-0,2	0,3	-0,9	-0,4	1,6	0,2	1,2	0,7	0,0	0,2	0,0
1992	0,5	-0,3	-0,9	-0,1	-0,2	0,3	0,5	0,9	0,0	0,3	0,1	0,2
1993	0,7	-0,2	-0,4	1,4	0,1	0,2	0,6	0,6	-1,6	-0,1	0,1	-0,7
1994	0,7	-0,4	1,3	0,6	-0,5	-0,5	0,7	0,2	0,9	1,1	0,3	-0,6
1995	-1,0	0,1	1,0	0,7	-0,4	-0,2	-0,1	-0,4	-0,6	0,9	0,2	0,0
1996	0,4	-0,2	0,1	0,4	-0,1	-0,6	0,3	0,6	1,0	0,2	0,3	1,0
1997	1,1	0,1	-0,4	0,4	-0,4	-0,8	0,4	0,3	0,4	0,2	0,2	0,6
1998	0,5	0,3	-0,1	0,4	0,2	-0,8	0,5	0,7	-0,3	0,7	0,5	0,5

Revisions >1 pct. unit are marked in bold.