ORGANISATION DE COOPÉRATION ET DE DÉVELOPPEMENT ÉCONOMIQUES

OECD OCDE

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

OECD SHORT-TERM ECONOMIC STATISTICS EXPERT GROUP (STESEG)

# The OECD Project on Revisions Analysis: First Elements for Discussion

Paper prepared by Tommaso Di Fonzo
Statistics Directorate, OECD and University of Padova

Submitted to the Expert Group under item 5 of the draft agenda

Meeting:
27 – 28 June 2005

Room 3
OECD Headquarters, Paris
Starting at 9:30 a.m. on the first day

**OECD Short-term Economic Statistics Expert Group (STESEG), 27-28 June 2005**

# THE OECD PROJECT ON REVISIONS ANALYSIS
# FIRST ELEMENTS FOR DISCUSSION

Tommaso Di Fonzo
OECD and University of Padova

## 1. Introduction

One of the conclusions of the OECD-ONS workshop on "Assessing and improving statistical quality. Revision analysis for the national accounts"[1], held in Paris on October 7-8, 2004, states:

> "The OECD also agreed to explore the possibility to expand and maintain the revisions database that was set up for purposes discussed in the workshop. An international dataset and the resulting summary measures of revisions should help to put countries' individual revision analyses into an international context."

Furthermore, at the end of the 'Summary and conclusions' of the workshop, under the heading "Further development of best practices guidelines", it states:

> "Based on the present conclusions, the OECD will prepare a more comprehensive paper on 'best practice in revision analysis' that could ultimately be published as an OECD working paper for wider dissemination."

In this paper first elements of the ongoing project on revisions analysis, which encompasses both issues, that is (i) remarks on strategy and tools for revisions analysis, and (ii) building, maintaining and using revisions databases, are presented for discussion.

The considered issues are: a statement of the problem (section 2); the role of revisions analysis when dealing with the accuracy and reliability dimension of data quality (section 3); analysis of revisions as an approach to assessing statistical outputs in economic datasets (section 4); a first summary view about terminology, definitions and statistical tools, with a discussion on how to evaluate the significance of the mean revision (section 5 and technical appendix); the information contents, data limitations and possible problems in building a real-time database for Quarterly National Accounts aggregates (GDP and Expenditure components) published in the OECD monthly publication *Main Economic Indicators* (section 6).

Parts of this paper are still incomplete, or have a stylized form. These will be completed in the future, when refinements and changes to the present note are made. Other issues should also be considered, such as a more thorough discussion on terminology and definitions, a review of the international (best) practices, problems pertaining the presentation of revisions and any other aspect of revisions analysis which emerge from the discussion of this paper at the 2005 STESEG meeting.

---

[1] The agenda and documents for the meeting can be found on the OECD website http://www.oecd.org/findDocument/0,2350,en_2649_33715_1_119808_1_1_1,00.html.

## 2. Statement of the problem

The purpose of the short-term economic statistics is that they provide, together with annual statistics, a comprehensive and timely picture of economic processes such as production, income distribution, financing and expenditure, as well as their international and inter-temporal relation (Algera, 1999).

However, economic analysts and policy makers not only require accurate and timely information on the movements in and magnitude of the main economic variables, but they must also have confidence that these indicators are unlikely to change significantly as more complete data become available.

Questions like “Are the figures accurate?”, “Are these data reliable?” are important to the producers and users of any set of statistics. Getting answers is never easy, but is particularly difficult in the case of datasets such as national accounts, short term statistics and balance of payments. All these datasets are based on multiple, complex source data and typically undergo several routine revisions as more and better source data are incorporated into the final estimates. As a result, the estimated data cannot be subjected directly to the usual statistical measures of sampling biases, variances, and other measurement error properties. So, to shed light on the accuracy and reliability of these datasets, various approaches have been used and new ones must be explored.

The ideal situation is to present the relevant data as early as possible with a high accuracy. However, it seems not to be possible to fulfil all requirements at the same time. An argument is still observed between two alternatives. Which is better between 1) to present data as early as possible with less accuracy and 2) to make available accurate data later? Therefore, the art of a statistical organisation, either national or international, is to find the best balance among those three requirements – accuracy, timeliness and relevance of economic statistics.

National Statistical Institutes of most countries are continuously requested to improve timeliness and reliability in short-term statistics for economic policy purposes and, more generally, for monitoring the economic situation. Every statistical agency facing the problem of providing public information, frequently tackles the trade off between timeliness and precision: long-delayed information is likely to be useless despite being accurate, while timely but highly imprecise data inject unnecessary noise into market operations, increasing uncertainty of the economic subjects about the reliability of data.

The extent to which statistics are subject to revision is one of the more easily quantifiable aspects of quality. However, these measures do not provide an unambiguous guide to quality. A series may be subject to few revisions, but the series may be highly inaccurate due to poor data sources. Revisions can be reduced by delaying the release of statistics until all or most 'final' data sources are available, but this would mean that the statistics would be less relevant to users. On the other hand, it may be possible to compile timely statistics that are not subject to revision only by placing an unacceptable load on survey respondents or at great cost to the compiler.

An analysis of revisions can, however, identify the possibility of inaccurate initial data or inefficient compilation methods. If it can be established that revisions are significantly different from zero (i.e. consistently positive or negative) then initial estimates are unreliable. The information on revisions could then be used to improve compilation methods to remove systematic distortions arising from the estimation process. However, even if there are no systematic distortions in compilation processes, users may still consider certain statistics to be unreliable because the revisions are significantly dispersed (i.e. the mean absolute values of the revisions are large). Generally, it is only possible to deal with such problems by improving the quality of source data by, for example, increasing initial survey response rates.

From the data producers' perspective, a history of the differences between the later and the early estimates might lead either to a direct adjustment of the series in question so as to offset the presumed continuing bias (Rao *et al.*, 1989), or to more fundamental changes such as the collection of additional data or advances in the estimation technique.

From the users' perspective, a simple and clear communication about revisions (and other related quality issues, see Franchet and Grünenwald, 2002) should improve awareness and confidence on economic information, even if also in this field a balance between simplicity and detailed analysis should be found to avoid undesirable 'noise'.

As stated before, the analysis of revisions grounds mainly on quantitative procedures, whose usefulness, reasonability and applicability must be thoroughly evaluated. In this respect, the practice of those institutes which routinely perform the analysis of revisions for large systems of economic data such as Quarterly National Accounts (e.g., *inter alia*, BEA for the U.S.A., ONS for U.K., Statistics Canada for Canada and ABS for Australia) could help to identify some 'minimal requirements' for which this kind of analysis should meet.

Simple descriptive revisions' measures are certainly useful to get a first impression of the problem, and this is undoubtedly the starting point of every analysis. Yet at this stage it demonstrates the necessity to formalize the revision time scheduling, in order to assess the different information content of the various estimates, which must be dealt with accordingly in order to avoid poorly informative, spurious or even meaningless comparisons. The information that can be gained from this approach is rather rich, mainly for situations in which the production and/or collection of appropriate datasets for the analysis of revisions are not particularly extended in time.

It should be noted that, from another, but strictly related point of view, it would also be possible to exploit the temporal connections of the series involved in the revision process to get further insight about the capability of early estimates in 'forecasting' the target value. Simple test procedures, provided the available series are long enough, could thus be used to assess the rationality of early estimates, by this way obtaining other information about possible margins of gain in the estimation and revision processes.

## 3. The accuracy and reliability dimension of data quality

This paper takes as its frame of reference the dimension of data quality that deals with accuracy and reliability and, as a natural consequence, also timeliness has been kept into consideration. More precisely, it focuses on a subset of the elements and indicators in that dimension concerned with assessments of statistical outputs (as distinguished from assessments of source data or intermediate data), having in mind at least two needs for both the producers and the users of economic data:

1. to increase awareness of some of the approaches to assessing the quality of statistical output coming from official economic datasets;
2. to stimulate thinking about additional approaches that might be used to assess accuracy and reliability in such datasets.

In addition to the timeliness, accuracy and reliability are also regarded as an integral part of the quality of statistics. In fact, it is increasingly commonplace to say that statisticians used to equate the quality of statistics with accuracy, reliability, or both - sometimes these words were used as meaning the same thing and sometimes not. In the recent upswing in attention given to data quality taken to be multi-dimensional, accuracy and reliability take their place within the longer list of the dimensions of quality[2]. It is now rather common to put accuracy and reliability together in a single dimension, which is built around the following definitions (Carson and Laliberté, 2002).

- **Accuracy** refers to the closeness between the estimated value and the (unknown) true value that the statistics were intended to measure. Assessing the accuracy of an estimate involves evaluating the error associated with an estimate. In practical terms, there is no single aggregate or overall measure of accuracy; accuracy is evaluated in terms of the potential sources of error, and these potential sources of error differ across datasets.

- **Reliability** refers to the closeness of the initial estimated value(s) to the subsequent estimated values. Assessing reliability involves comparing estimates over time. *In other words, assessing reliability refers to revisions*. This feature is identified separately for two reasons. First, it is usually the initial estimate that captures attention, whence the importance of its accuracy. Second, the separation helps bring out the fact that data that are not revised are not necessarily the most accurate.

The Accuracy and Reliability dimension of a data quality assessment framework is applicable to practically all economic datasets. Table 1 shows the generic framework developed by IMF (Carson and Laliberté, 2002).

To summarize, accuracy and reliability refer to the discrepancy between the compiled data and the unknown "true" figure (the target value). This discrepancy (the total error) can be separated into two components (Bier and Ahnert, 2001):

1. The discrepancy between the final figure and the target value. This quantity is not precisely measurable (although for estimates produced from probability sample surveys the sampling error of the estimated can be evaluated). Some indication of this discrepancy can be obtained by measuring the consistency between related statistics. Consistency may be compared over time, between related variables, across countries or between data at different frequencies.[3]
2. The difference between the first published data and the final figure, which can be quantified by comparing these two: it is a measure of inaccuracy which is due to the release of provisional results.

Large economic datasets as short term economic statistics, national accounts and balance of payments are alike in that they are based on multiple, complex source data and typically undergo several routine revisions as more and better source data are incorporated into the statistical outputs. These features have important implications for the assessment of data quality.

[2] The IMF's Data Quality Assessment Framework (DQAF) can be found on the IMF website at http://www.imf.org/external/np/sta/dsbb/2001/supp.htm. The DQAF comprises a generic framework, which is applicable to all datasets, and a set of specific frameworks that provide more detail (Carson, 2000).

[3] A well-known consistency requirement is the accounting identity of GDP compiled from the output, expenditure and income side.

**Table 1: The Accuracy and Reliability Dimension of the Data Quality***

| Quality dimensions | Elements | Indicators |
|---|---|---|
| **3. Accuracy and reliability**<br><br>*Source data and statistical techniques are sound and statistical outputs sufficiently portray reality.* | **3.1 Source data** - *Source data available provide an adequate basis to compile statistics.* | 3.1.1 Source data are collected from comprehensive data collection programs that take into account country-specific conditions.<br>3.1.2 Source data reasonably approximate the definitions, scope, classifications, valuation, and time of recording required.<br>3.1.3 Source data are timely. |
| | 3.2 **Statistical techniques** - *Statistical techniques employed conform to sound statistical procedures.* | 3.2.1 Data compilation employs sound statistical techniques.<br>3.2.2 Other statistical procedures (e.g., data adjustments and transformations, and statistical analysis) employ sound statistical techniques. |
| | **3.3 Assessment and validation of source data -** *Source data are regularly assessed and validated.* | 3.3.1 Source data—including censuses, sample surveys and administrative records—are routinely assessed, e.g., for coverage, sample error, response error, and non-sampling error; the results of the assessments are monitored and made available to guide planning. |
| | **3.4 Assessment and validation of intermediate data and statistical outputs -** *Intermediate results and statistical outputs are regularly assessed and validated.* | 3.4.1 Main intermediate data are validated against other information where applicable.<br>3.4.2 Statistical discrepancies in intermediate data are assessed and investigated.<br>3.4.3 Statistical discrepancies and other potential indicators of problems in statistical outputs are investigated. |
| | **3.5 Revision studies -** *Revisions, as a gauge of reliability, are tracked and mined for the information they may provide.* | 3.5.1 Studies and analyses of revisions are carried out routinely and used to inform statistical processes. |

*Other quality dimensions are 1. Integrity, 2. Methodological Soundness, 4. Serviceability, and 5. Accessibility. See http/www.imf.org/external/np/sta/dsbb/2001/supp.htm.
*Source: Carson and Laliberté (2002)*

Carson and Laliberté (2002) discuss four approaches that can be used in assessing the accuracy and reliability of statistical output for national accounts and balance of payments, whose validity is however more general and can be extended to the majority of systems of economic data:

1. examination of statistical discrepancies;
2. comparison with other data;
3. judgemental evaluation;
4. analysis of revisions.

On this last point we focus in the next sections.

## 4. The analysis of revisions as an approach to assessing statistical outputs in economic datasets

The analysis of revisions is a means of assessing the quality of the first (or other relatively early) estimate in relation to later, sometimes, final estimates. A sensible (and possibly simple) way for doing it, is to prepare and examine measures of revision. These studies are typically done by statistical agencies and draw on databases that record the various vintages of estimates. As well, there are several other methods, including use of statistical and econometric models.

### 4.1. Reasons for revisions and revision patterns

There are several reasons for revisions: improved data coverage, correction of errors, seasonal adjustment and so on. In particular, the need for revisions to follow timely publication is justified by the fact that initial data are often based on incomplete or provisional information, with some series interpolated in the absence of data which will become available only in a subsequent time. Typically, each subsequent series of the variable of interest should incorporate more information than the previous ones until the statistical agency regards the process as complete. It is obvious that the characteristics of these vintages of data are of great interest. Some important questions concern, for example, the relationship between initial and final vintages and the sensitivity of time series model to successive revisions of the initial data.

The main aim of revisions is to improve the quality of the preliminary figures published. It is generally known that revisions can be caused by methodological changes, the correction of mistakes, or by the supply of new and more complete data. Due to non-response, corrections and other reasons revisions are in many cases unavoidable.

The specific reason for revisions can differ across the calculated indicators and between countries. The relevance of revisions of a main aggregate indicator on the one hand and revisions of the underlying components on the other should be taken into account. However, the implications of revisions for policymaking are much greater in the first case.

In general, revisions are made due to:

- receipt of additional or more comprehensive data;
- replacement of model-based estimates with source data;
- methodological changes;
- changes to base weights or classifications.

In addition, we may distinguish between common revisions which are common practice (e.g. reweighting/methodological changes) and revisions that are due to the requirement of improved timeliness (caused mainly by incomplete data).

A more detailed list of reasons for revision has been provided by a US study (Office of Management and Budget, 1983), which identifies as major reasons of revisions (i) tight deadlines for the release of data resulting mainly from pressure from users, (ii) late responses or corrections from respondents, (iii) errors detected by the statistical agencies, (iv) the update of seasonal adjustment, (v) change in definition of the variable being measured, (vi) benchmarking, and (vii) change in the sample frame.

Taking quarterly national accounts as an example, revisions are a natural consequence of the processes used to compile the national accounts. At country level, initial estimates are based on survey responses received and processed before a particular cut-off time. Following the cut-off, imputations are made for the non-respondents based on the responses of similar businesses and the responses of the non-respondents in the previous quarter. Subsequently, when the non-respondents finally respond the imputations are replaced and revisions to the estimates result. For many QNA aggregates, the estimates are compiled by applying indicators to annual (or less frequent) benchmarks based on superior data sources.

This benchmarking process typically leads to revisions over an extended period of time. Often the first benchmark data to become available are preliminary estimates and are therefore themselves subject to revisions. For the most part, benchmarks are considered 'final' three years after the period to which they relate has passed.

Another source of revisions is the availability of a major new data source or the development of an improved estimation methodology. Sometimes, the resultant revisions may even affect estimates for periods prior to those for which the benchmark estimates would have otherwise been considered 'final'.

Finally, seasonally adjusted and trend estimates will usually experience some degree of revision over several years, due to the prolonged period required to finalise the estimation of seasonal adjustment factors.

As regards the revision patterns, following ECB (2001), the following four categories of revision patterns can be identified:

1. non-regular and minor revisions of the time series;
2. continuous revisions of the latest value(s);
3. revision of the time series at intervals (e.g., every year) over longer time spans, and
4. continuous revisions over longer time spans.

Where possible, revisions should be broken down by type of source of revision. If quantification is not possible, sources of revisions should at least be explained so as to avoid the exclusive association of 'revision' with 'error'.

At the OECD-ONS workshop of October 2004 the ONS of UK presented a possible classification of sources of revision:

A. As *indicator data evolve* (= increasing quantity of data)
- data replacing forecasts
- increased response rates
- revised source data
- seasonal adjustment

B. As *benchmark data available* (= increasing quantity of data)
- annual surveys, tax data, government expenditure, etc.
- re-basing - 5-yearly, annual chain-linking
- population census

C. As *data sources improved* (= one-off improvements)
- new survey or administrative data sources appear/developed
- improved survey forms

D. As *compilation/balancing methods improved* (= one-off)
- moving to use of supply-use tables on quarterly basis

E. As *accounts brought in line with conceptual targets* (= one-off)
- SNA93, ESA95
- allocation of FISIM
- capitalisation of software

Such a list is useful for several purposes, such as explaining revisions to users, and deciding how to break down total revisions for analytical purposes. With the aim of producing an exhaustive and generally agreed list, the OECD agreed to circulate the draft list so that countries can add examples from their own experience and comment/improve on the classification.

For example, seasonal adjustment, re-basing, and the move to annual chain-linking might warrant separate categories.

### 4.2. Presentation of revisions

Keeping in mind the principle that the main aim of revisions should be the improvement of the preliminary results, it makes sense to inform users about

- the status of the published data, and
- the revision history of major economic indicators.

This information could be very helpful for the expectation building process of the economic actors. In addition, revision history could be an important tool for economic forecasts.

The way the revisions are communicated to the public is another important issue related to the impact of revisions on the user's perception of data quality. The Statistical Policy Directive Number 3 of the US Office of Management and Budget also includes the following six guidelines when issuing and evaluating preliminary data and revisions.

1. Agencies shall clearly identify figures as preliminary or revised.
2. Agencies shall only release routine revisions of a principal economic indicator as part of the regular reporting schedule.
3. If the difference between preliminary and final aggregate figures is large relative to average period-to-period differences, the agency must either take steps to improve the accuracy of preliminary estimates or delay the release of estimates until a reliable estimate can be made.
4. If preliminary estimates show signs of a consistent bias (for example, if revisions are consistently in the same direction), the agency shall take steps to correct this bias.
5. Revisions occurring for routine reasons, such as benchmarking and updating of seasonality factors, shall be consolidated and released simultaneously.
6. Revisions occurring for other than routine reasons shall be fully explained and shall be released as soon as correction can be completed.

Recently ONS (2004) has established 9 guidelines in form of "A guide to putting the principles into practice” on the following issues:

1. Published statement on revisions;
2. Key outputs – published revisions policy;
3. Key First Releases – explanatory statement;
4. Same principles as all new information;
5. Timeliness balanced with frequency;
6. Methodological changes;
7. Unexpected revisions;
8. Explanations;
9. Monitoring long-term effects.

Obviously the observance of these guidelines should be tested in daily practice by the National Statistical Institutes to appreciate eventual negative impact on the data production process due to the involved burden and, at the same time, its value and usefulness with respect to users' needs.

## 5. A first summary view on terminology, definitions and statistical tools

### Reliability *vs.* accuracy

In line with IMF definitions (Carson and Laliberté, 2002), revisions analysis provides a measure of the quality of a statistic or indicator, and that the dimension of quality being measured is reliability, not accuracy. It should be carefully explained to users that revisions do not indicate that one statistic is more or less accurate than another; on the contrary, a statistic that is not revised could easily be less accurate than one that is.

### Revision

Revision should be defined as the difference between a later (more recent) estimate and an earlier estimate. This is often expressed as Latest minus Preliminary and written as *L-P*. On the other hand, in the past Eurostat and the OECD have used *P-L*. When calculating the mean revision, this results in an estimate with the opposite sign to that of a revision. For this reason, in the OECD-ONS workshop of October 2004 it was agreed that if an organization chooses to use *P-L* then they should not use the phrase ‘revision’ but ‘deviance’ or ‘error’ or some other term.

### Latest and final

There is some inconsistency between practitioners in the use of terms such as 'latest' and 'final'. For example, for quarterly GDP estimates, some countries use the term final to distinguish between 'firm' quarterly estimates (probably published around 90 days after the end of the quarter in question) and preliminary or flash estimates (published much more quickly). Other countries take the view that whilst there is a possibility that an estimate could be revised due to methodological improvements or re-basing, the estimates can never be considered as final.

### Bias

Following from the definition of revision above, it seems that the term 'bias' is often used incorrectly; the estimate of mean revision is called bias, whereas (i) the two concepts have opposite signs[4], and (ii) when the mean revision is not significantly different from zero, no bias is present at all in the preliminary estimates. At the OECD-ONS workshop of October 2004 there was general agreement that the term bias should be avoided. If a country has a need to use the term, then it should be used in a limited and specific sense only, and its usage should be explained to users.

**Summary statistics**

Summary indicators that are found in most revision studies are:

- *Mean revision*

$$\bar{R} = \frac{1}{n}\sum_{t=1}^{n}\left(L_t - P_t\right) = \frac{1}{n}\sum_{t=1}^{n}R_t$$

  where $L_t$ is the latest estimate, $P_t$ is the preliminary (or earlier) estimate, $R_t = L_t - P_t$ is the revision and *n* is the number of observations

- *Mean absolute revision*

$$MAR = \frac{1}{n}\sum_{t=1}^{n}\left|L_t - P_t\right| = \frac{1}{n}\sum_{t=1}^{n}\left|R_t\right|$$

- *Standard deviation of revision*

$$SDR = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(R_t - \bar{R}\right)^2}$$

- *Mean squared revision*

$$MSR = \frac{1}{n}\sum_{t=1}^{n}\left(L_t - P_t\right)^2 = \frac{1}{n}\sum_{t=1}^{n}R_t^2$$

The mean squared revision is a summary measure based on a symmetric and quadratic loss function. It has some properties which merit to be quoted. These properties and a numerical example are presented in Appendices A and B, respectively.

Less widely used, but potentially very useful are statistics on range of revisions, median revision and information about the amount of positive and negative revisions.

Another indicator which has been kept into consideration (see, for example, Ahmad *et al.*, 2004) is the *Relative mean absolute revision*, defined as

[4] For example, if a country's GDP estimates have a positive mean revision, then the preliminary estimates could be negatively biased, as they have a tendency to be estimating growth to be lower than it actually is.

$$RMAR = \frac{\sum_{t=1}^{n} |L_t - P_t|}{\sum_{t=1}^{n} |L_t|} = \frac{\sum_{t=1}^{n} |R_t|}{\sum_{t=1}^{n} |L_t|}.$$

### 5.1. Statistical significance of the mean revision

As far as the evaluation of the significance of the mean revision is concerned, a simple and robust approach based on the Heteroskedasticity Autocorrelation Consistent estimate`s variance proposed by Newey and West can be considered. In this case the standard *t* test is applied using a robust estimate of the variance of the mean revision, which is calculated in a different way with respect to what is currently done by ONS (Jenkinson and Stuttard, 2004). Technical details on this point are presented in the appendix.

A point that merits to be stressed is that that mean revisions, even if found to be statistically significant different from zero, should not, in general, be used to adjust preliminary estimates at the aggregate level. Efforts should be made to reduce the systematic revision by setting up a research plan to improve the source data and estimation methods. Revision analysis is, therefore, a useful diagnostic tool for improving the quality of estimates, and should be built into the production processes for national accounts and other indicators.

## 6. Building an OECD-MEI real-time dataset for GDP and Expenditure components. Starting considerations and first evidences

Due to unsynchronised national release policies, the OECD publishes both timely data on key macroeconomic variables (e.g. Quarterly National Accounts series) and revised versions of the same time series. In other words, almost all published variables are first released on a preliminary basis to satisfy the users' need for timely information, and are then revised at a later period to incorporate information that was not available at the time of the preliminary release.

Coverage problems and different revision practices across countries and across statistical indicators give rise to a revision process which is generally characterized by a preliminary (or provisional) estimate and by a subsequent series of routinely produced[5] revisions, that follow closely upon one another.

A comprehensive analysis of revisions requires considerable time and resources. It would encompass the full set of published data, ideally including both raw and seasonally adjusted data, and, where it is the case, in both current and constant prices. Growth rates as well as levels would be analyzed[6] and a complete documentation about the revisions themselves and about the reasons why they were necessary would be collected and investigated.

The ongoing OECD project on revisions analysis assumes these indications in studying the feasibility of

[5] It should be noted that, from time to time, additional, more unusual revisions could also occur, which take place at infrequent and irregular intervals. These ‘extraordinary’ revisions can include changes in concepts and definitions of the aggregates and/or information coming from decennial censuses and improved estimation procedures.
[6] Even if it is generally agreed that the main concern of users is much more on growth rates than on levels.

building a real-time dataset for revisions analysis coming from the OECD monthly publication *Main Economic Indicators* (MEI).

**Main characteristics of MEI**

- Published in paper format since the early 1960's.
- Long time series for key short-term economic statistics for most OECD countries, e.g.:
  - Quarterly GDP and expenditure components
  - Index of industrial production & Composite Leading Indicators
  - Retail Trade, Consumer and Producer Prices
  - Wages, employment and unemployment
  - Interest rates, exchange rates, monetary aggregates
  - Business tendency and consumer opinion surveys
  - International trade
- Electronic monthly 'snapshot' databases available back at least until 1995.

The MEI database has time series going back until the late 1950's and early 1960's for several of the larger OECD countries for many key economic variables.

The MEI has been published as a CD Rom for many years. We have established a catalogue of these back until 1995 and are in the process of trying to retrieve earlier information saved to disks back to the 1980's.

MEI monthly snapshots in principle provide historical data on subsequent releases on a wide range of key variables. They therefore provide a coherent set of variables representing the 'information set' available at successive monthly intervals. It should be noted that a monthly snapshot is most likely frequent enough to pick up all releases of economic statistics as they are seldom revised twice within one month – although there are likely to be some exceptions.

**OECD interest in a real-time database**

- Analysis – are revisions in general becoming smaller? Are they random (i.e. centred around 0)?
- Support internationally coordinated research work to improve the quality of early estimation methods;
- Encourage transparency in the statistical process;
- Assess real-time performance of the OECD Composite Leading Indicators.

The OECD-ONS workshop of October 2004 has been a major step forward for promoting the construction of real time databases and revisions analysis across OECD countries. Three key outcomes from the conference were:

- Systematic archiving of all vintages of data is key for revision analysis (NSOs should start doing this).
- Revision analysis should be seen as an integral part of the statistical production process.
- OECD would explore the possibility of maintaining and expanding a revisions database for the

expenditure components of quarterly GDP.

**Current OECD project: real time database & revisions analysis of GDP for OECD countries**

The initial OECD work (Ahmad *et al.*, 2004) was for G7 countries, period 1996 – 2000. The current project being undertaken can be summarized as follows:

- Expand coverage to all OECD countries;
- Expand time period covered;
- Devise a scheme which can support ongoing production of real time database and summary revision analysis indicators;
- Consider quarterly GDP and its components.

Possible future extensions should cover other key variables, and most likely the next variable would be the Index of Industrial Production. However, we need first to establish operational process for GDP and its components, which is a priority.

Some obstacles to extending the work can be:

- Possible gaps in our electronic sources pre-1995, only resort may be paper copies. A related problem is how far back should the real time database go.
- Operational problems (e.g. series code changes).
- Methodological issues:
  - constancy of national series over time (e.g. various methodological changes causing breaks and revisions).
  - OECD procedures (linking, seasonal adjustment) .
- Data management (produces big databases)

Methodological issues can affect the validity of assessing revisions. As revisions caused by statistical processes e.g. changing response rate, data editing, changes to estimation methods, benchmarking etc. may need to be considered very differently to revisions caused by other changes such as scope, definitions (e.g. SNA 93) etc. Hence it's a very resource intensive process to establish a historical real time database for each variable.

Table 2 presents the current data availability for GDP in an initial OECD real-time database, as this work is still in progress. Entries of the tables will be subjected to revisions until the 'final' version of the database will be obtained.

**Table 2: Data availability for GDP in the OECD real-time database (preliminary, work in progress)**

| | Country | | NOT CURRENT PRICES | | | | CURRENT PRICES | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | **Seasonally adjusted** | | **Raw data** | | **Seasonally adjusted** | | **Raw data** | |
| | | | MEI edition | Vintage | MEI edition | Vintage | MEI edition | Vintage | MEI edition | Vintage |
| 1 | AUS | First | Jan-97 | 96q3 | - | - | Jan-97 | 96q3 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 2 | AUT | First | Jan-97 | 95q4 | Jan-97 | 95q4 | Jan-97 | 95q4 | Jan-97 | 95q4 |
| | | Last | May-05 | 04q4 | **Oct-03** | **03q2** | May-05 | 04q4 | **Oct-03** | **03q2** |
| 3 | BEL | First | **Mar-97** | 96q3 | - | - | **Apr-99** | **98q4** | - | - |
| | | Last | May-05 | 05q1 | - | - | May-05 | 04q4 | - | - |
| 4 | CAN | First | Jan-97 | 96q3 | - | - | Jan-97 | 96q3 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 5 | CHE | First | Jan-97 | 96q3 | - | - | Jan-97 | 96q3 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 6 | CZE | First | **Jul-99** | **99q1** | Jan-97 | 96q3 | **Jul-99** | **99q1** | Jan-97 | 96q3 |
| | | Last | May-05 | 04q4 | **Sep-03** | **03q1** | May-05 | 04q4 | **Sep-03** | **03q1** |
| 7 | DEU | First | Jan-97 | 96q3 | - | - | Jan-97 | 96q3 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 8 | DNK | First | Jan-97 | 96q2 | - | - | Jan-97 | 96q2 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 9 | ESP | First | Jan-97 | 96q3 | - | - | Jan-97 | 96q3 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 10 | FIN | First | Jan-97 | 96q3 | Jan-97 | 96q3 | Jan-97 | 96q3 | Jan-97 | 96q3 |
| | | Last | May-05 | 04q4 | **Apr-99** | **98q3** | May-05 | 04q4 | **Apr-99** | **98q3** |
| 11 | FRA | First | Jan-97 | 96q3 | - | - | Jan-97 | 96q3 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 12 | GBR | First | Jan-97 | 96q3 | - | - | Jan-97 | 96q3 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 13 | GRC | First | **Apr-04** | **03q4** | - | - | **Apr-04** | **03q4** | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 14 | HUN | First | - | - | **Apr-03** | **02q3** | - | - | **Apr-03** | **02q3** |
| | | Last | - | - | May-05 | 04q4 | | | May-05 | 04q4 |
| 15 | IRL | First | **Nov-04** | **04q2** | - | - | **Nov-04** | **04q2** | **Jan-03** | **02q2** |
| | | Last | May-05 | 04q4 | | | May-05 | 04q4 | **Oct-04** | **04q2** |
| 16 | ISL | First | - | - | - | - | - | - | **Apr-03** | **02q4** |
| | | Last | - | - | - | - | - | - | May-05 | 04q4 |
| 17 | ITA | First | Jan-97 | 96q2 | - | - | Jan-97 | 96q2 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 18 | JPN | First | Jan-97 | 96q3 | - | - | Jan-97 | 96q3 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 19 | KOR | First | **May-97** | 96q4 | - | - | **May-97** | 96q4 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 20 | LUX | | - | - | - | - | - | - | - | - |
| 21 | MEX | First | Jan-97 | 96q3 | Jan-97 | 96q3 | Jan-97 | 96q3 | Jan-97 | 96q3 |
| | | Last | **Sep-00** | **00q1** | May-05 | 04q4 | **Sep-00** | **00q1** | May-05 | |
| 22 | NLD | First | Jan-97 | 96q3 | - | - | Jan-97 | 96q2 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 23 | NOR | First | Jan-97 | 96q3 | Jan-97 | 96q3 | Jan-97 | 96q3 | Jan-97 | 96q3 |
| | | Last | May-05 | 04q4 | **Sep-00** | **00q1** | May-05 | 04q4 | **Sep-00** | **00q1** |
| 24 | NZL | First | Jan-97 | 96q2 | - | - | Jan-97 | 96q2 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 25 | POL | First | - | - | **Sep-02** | **02q1** | - | - | **Sep-02** | **02q1** |
| | | Last | - | - | May-05 | 04q4 | - | - | May-05 | 04q4 |
| 26 | PRT | First | Jan-97 | 95q4 | - | - | Jan-97 | 95q4 | - | - |
| | | Last | May-05 | 04q4 | - | - | May-05 | 04q4 | - | - |
| 27 | SVK | First | - | - | **Feb-01** | **00q3** | - | - | **Feb-01** | **00q3** |
| | | Last | - | - | May-05 | 04q4 | - | - | May-05 | 04q4 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 28 | SWE | First | Jan-97 | 96q3 | Jan-97 | 96q3 | Jan-97 | 96q3 | Jan-97 | 96q3 |
| | | Last | **Sep-00** | **00q2** | May-05 | 04q4 | **Sep-00** | **00q1** | May-05 | 04q4 |
| 29 | TUR | First | Jan-97 | 96q1 | Jan-97 | 96q2 | Jan-97 | 96q1 | Jan-97 | 96q2 |
| | | Last | **Sep-00** | **00q2** | May-05 | 04q4 | **Sep-00** | **00q2** | May-05 | 04q4 |
| 30 | USA | First | Jan-97 | 96q3 | - | - | Jan-97 | 96q3 | - | - |
| | | Last | May-05 | **05q1** | - | - | May-05 | 04q4 | - | - |

The remaining tables present an example of a simple revisions analysis which basically assumes the ONS framework (Jenkinson and George, 2005). The real-time database of GDP in levels (table 3) is used to obtain the real-time database of QoQ rates of change (table 4), based on which the revisions triangle is calculated. Table 5 and figure 1 present the summary statistics and a chart to visualize the revisions, respectively.

**Table 3: An excerpt from the OECD real time data set: Canada GDP (levels, seasonally adjusted, constant prices)**

| | 1996Q4 | 1997Q1 | 1997Q2 | 1997Q3 | 1997Q4 | 1998Q1 | 1998Q2 | 1998Q3 | 1998Q4 |
|---|---|---|---|---|---|---|---|---|---|
| **First estimate** | 624.2 | 629.5 | 637.6 | 803.8 | 809.9 | 816.0 | 819.0 | 830.7 | 840.4 |
| **1 year later** | 777.4 | 786.5 | 794.6 | 812.2 | 817.9 | 824.4 | 833.7 | 839.0 | 848.9 |
| **2 years later** | 783.5 | 793.4 | 807.8 | 819.3 | 825.8 | 834.1 | 836.4 | 842.4 | 855.1 |
| **3 years later** | 790.8 | 800.0 | 810.0 | 821.5 | 825.8 | 834.1 | 911.9 | 922.2 | 936.7 |
| Jun-97 | **624.2** | **629.5** | | | | | | | |
| Jul-97 | 624.2 | 629.5 | | | | | | | |
| Aug-97 | 624.2 | 629.5 | | | | | | | |
| Sep-97 | 624.2 | **629.9** | **637.6** | | | | | | |
| Oct-97 | 624.2 | 629.9 | 637.6 | | | | | | |
| Nov-97 | 624.2 | 629.9 | 637.6 | | | | | | |
| Dec-97 | 624.2 | 629.9 | 637.6 | | | | | | |
| Jan-98 | **777.4** | **785.3** | **795.7** | **803.8** | | | | | |
| Feb-98 | 777.4 | 785.3 | 795.7 | 803.8 | | | | | |
| Mar-98 | 777.4 | **786.4** | **796.3** | **804.0** | **809.9** | | | | |
| Apr-98 | 777.4 | 786.4 | 796.3 | 804.0 | 809.9 | | | | |
| May-98 | 777.4 | 786.4 | 796.3 | 804.0 | 809.9 | | | | |
| Jun-98 | 777.4 | **786.5** | **794.6** | **803.0** | **808.6** | **816.0** | | | |
| Jul-98 | 777.4 | 786.5 | 794.6 | 803.0 | 808.6 | 816.0 | | | |
| Aug-98 | 777.4 | 786.5 | 794.6 | 803.0 | 808.6 | 816.0 | | | |
| Sep-98 | 777.4 | 786.5 | 794.6 | 803.0 | 808.6 | **815.4** | **819.0** | | |
| Oct-98 | 777.4 | 786.5 | 794.6 | 803.0 | 808.6 | 815.4 | 819.0 | | |
| Nov-98 | 777.4 | 786.5 | 794.6 | 803.0 | 808.6 | 815.4 | 819.0 | | |
| Dec-98 | **783.5** | **793.4** | **803.5** | **812.2** | **817.9** | **824.0** | **827.0** | **830.7** | |
| Jan-99 | 783.5 | 793.4 | 803.5 | 812.2 | 817.9 | 824.0 | 827.0 | 830.7 | |
| Feb-99 | 783.5 | 793.4 | 803.5 | 812.2 | 817.9 | 824.0 | 827.0 | 830.7 | |
| Mar-99 | 783.5 | 793.4 | 803.5 | 812.2 | 817.9 | **824.4** | **827.5** | **831.1** | **840.4** |

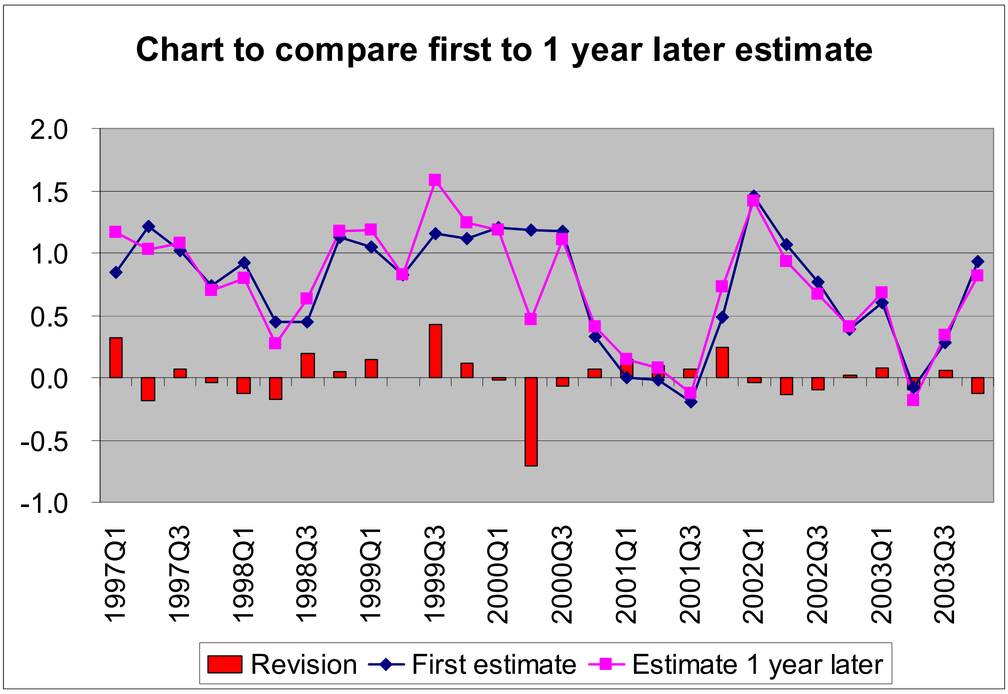
**Table 4: An excerpt from the OECD real time data set: Canada GDP (Quarter-on-Quarter rates of change, seasonally adjusted, constant prices)**

| | 1997Q1 | 1997Q2 | 1997Q3 | 1997Q4 | 1998Q1 | 1998Q2 | 1998Q3 | 1998Q4 |
|---|---|---|---|---|---|---|---|---|
| **First estimate** | 0.8 | 1.2 | 1.0 | 0.7 | 0.9 | 0.4 | 0.4 | 1.1 |
| **1 year later** | 1.2 | 1.0 | 1.1 | 0.7 | 0.8 | 0.3 | 0.6 | 1.2 |
| **2 years later** | 1.3 | 1.1 | 1.4 | 0.8 | 0.7 | 0.3 | 0.7 | 1.5 |
| **3 years later** | 1.2 | 1.2 | 1.4 | 0.9 | 0.7 | 0.4 | 1.1 | 1.6 |
| Jun-97 | **0.8** | | | | | | | |
| Jul-97 | 0.8 | | | | | | | |
| Aug-97 | 0.8 | | | | | | | |
| Sep-97 | **0.9** | **1.2** | | | | | | |
| Oct-97 | 0.9 | 1.2 | | | | | | |
| Nov-97 | 0.9 | 1.2 | | | | | | |
| Dec-97 | 0.9 | 1.2 | | | | | | |
| Jan-98 | **1.0** | **1.3** | **1.0** | | | | | |
| Feb-98 | 1.0 | 1.3 | 1.0 | | | | | |
| Mar-98 | **1.1** | **1.3** | **1.0** | **0.7** | | | | |
| Apr-98 | 1.1 | 1.3 | 1.0 | 0.7 | | | | |
| May-98 | 1.1 | 1.3 | 1.0 | 0.7 | | | | |
| Jun-98 | **1.2** | **1.0** | **1.1** | **0.7** | **0.9** | | | |
| Jul-98 | 1.2 | 1.0 | 1.1 | 0.7 | 0.9 | | | |
| Aug-98 | 1.2 | 1.0 | 1.1 | 0.7 | 0.9 | | | |
| Sep-98 | 1.2 | 1.0 | 1.1 | 0.7 | **0.8** | **0.4** | | |
| Oct-98 | 1.2 | 1.0 | 1.1 | 0.7 | 0.8 | 0.4 | | |
| Nov-98 | 1.2 | 1.0 | 1.1 | 0.7 | 0.8 | 0.4 | | |
| Dec-98 | **1.3** | **1.3** | **1.1** | **0.7** | **0.8** | **0.4** | **0.4** | |
| Jan-99 | 1.3 | 1.3 | 1.1 | 0.7 | 0.8 | 0.4 | 0.4 | |
| Feb-99 | 1.3 | 1.3 | 1.1 | 0.7 | 0.8 | 0.4 | 0.4 | |
| Mar-99 | 1.3 | 1.3 | 1.1 | 0.7 | **0.8** | **0.4** | **0.4** | **1.1** |

**Table 5: An excerpt from the OECD real time data set: Revisions analysis for Canada GDP (seasonally adjusted, constant prices) – First estimates *vs.* 1 year later estimates**

| Basic indicators | Whole sample | Last 20 quarters |
|---|---|---|
| number of observations (n) | 28 | 20 |
| mean absolute revision | 0.1396 | 0.1382 |
| mean revision | 0.0099 | 0.0086 |
| std. dev. of mean revision | 0.0396 | 0.0522 |
| t-statistic | 0.2502 | 0.1647 |
| t-crit (n-1 degrees of freedom) | 2.0518 | 2.0930 |
| significant mean revision or not? | NO | NO |
| Other indicators | | |
| Correlation | 0.901 | 0.908 |
| Min revision | -0.7 | -0.7 |
| Max revision | 0.4 | 0.4 |
| Range | 1.1 | 1.1 |
| % Later > Early | 57.1 | 60.0 |
| % Sign(Later) = Sign(Early) | 92.9 | 90.0 |

**Figure 1: An excerpt from the OECD real time data set: Revisions analysis for Canada GDP (seasonally adjusted, constant prices) – Graphical analysis**



## APPENDIX A
## Properties of the Mean Squared Revision

For unbiased preliminary estimates (that is, $E(R_t)=0$, t=1,…,*n*, where the symbol $E(\cdot)$ denotes the expected value), *MSR* is the variance of the revision. Taking the square root gives the Root Mean Squared Revision (*RMSR*), which is a suitable measure of accuracy when using a quadratic loss function. The disadvantages of the *RMSR* are that (i) while it has a minimum value of zero, it is unbounded and that (ii) it is unrelated to the variation in the latest estimate. Theil (1966) proposed an inequality coefficient (*U*) defined as the positive square root of

$$U^2 = \frac{MSR}{\sum_{t=1}^{n} L_t^2 \Big/ n},$$

which takes account of the variation in the most recent series. It has the value one when all the latest figures are zero, which is an interesting property if the index is used on growth rates. For, a value of *U* between zero and one indicates that the early estimates of growth rates are an improvement over a no-change estimate, while larger values imply poor preliminary estimates. Theil (1961) proposed two alternative decompositions of *MSR* and, following Granger and Newbold (1973), we consider the following:

$$MSR = \overline{R}^2 + \left(S_P - \rho S_L\right)^2 + \left(1-\rho^2\right)S_L^2,$$

where $\overline{R}$ is the mean revision, $S_L$ and $S_P$ are the standard deviations of the latest and preliminary estimates, respectively, and $\rho$ is their correlation. Dividing throughout by *MSR* gives

$$1 = UM + UR + UD$$

where

$$UM = \frac{\overline{R}^2}{MSR}$$

$$UR = \frac{\left(S_P - \rho S_L\right)^2}{MSR}$$

$$UD = \frac{\left(1-\rho^2\right)S_L^2}{MSR}$$

The interpretation of these is helped by consideration of the regression model in which the latest estimate is linked to the preliminary estimate as

$$L_t = \alpha + \beta P_t + u_t, \qquad (1)$$

for which the least squares estimators are $\hat{\beta} = \frac{S_{LP}}{S_P^2}$ and $\hat{\alpha} = \overline{L} - \hat{\beta}\overline{P}$, where $S_{LP}$ is the covariance between $L_t$ and $P_t$. If the preliminary estimates are unbiased, $\alpha = 0$ and $\beta = 1$, so that $\overline{R}$, and hence *UM*, is zero. That is, *UM* gives the proportion of *MSR* due to systematic differences between the preliminary and the latest estimates. For *UR* we note that

$$S_P - \rho S_L = S_L \left(1 - \hat{\beta}\right),$$

so that *UR* is the proportion of *MSR* due to the slope coefficient in (1) differing from one. Finally, if (1) gave a perfect fit, then *UD* would be zero, so *UD* can be interpreted as the disturbance proportion of *MSR* or that part of the observed revision which is not explained by the mean or slope error. 'Good' preliminary estimates will have low values of *UM* and *UR* and a high value of *UD* (a numerical example is shown in table B2).

## APPENDIX B
## How to judge the significance of the mean revision: A note

### Introduction

The evaluation of the significance of the mean revision is becoming a common way in revisions analysis to assess (a dimension of) the quality of economic data for which different releases are made available as time goes by. In order to evaluate the significance of the mean revision to early estimates of economic aggregates, simple inferential procedures are generally used, like a $t$-test to verify the maintained hypothesis that the mean revision is zero against the alternative of (either positive or negative) non-zero value. At this regard, the practice at ONS is described, *inter alia*, by Jenkinson and Stuttard (2004, p. 71, from now onward, JS). They write:

**What is the modified t-statistic?**
Revisions are considered to be biased if the mean revision is statistically significantly different from zero. A t-test compares the calculated bias in the series (that is, the mean revision) with the variability of the revisions to test whether the bias is significant. However, the standard t-test is based on the assumption that the revisions are independent of each other. This is not true for a time series as revisions made for one period may be associated with revisions made to previous periods. The modified t-test corrects for this lack of independence by adjusting the estimate of the variability of the revisions to take into account the serial correlation, that is the extent of the association between successive revisions.

In other words, for "standard $t$ test" it is meant the test-statistic calculated assuming a (gaussian and) independently distributed sample of data, while the "modified $t$ test" takes into account the serial correlation (which is assumed coming from an AR(1) process) of the data.

In the Excel files containing the "Revision Triangles" - released by ONS starting from 2005 (Jenkinson and George, 2005) - for a number of quarterly and monthly variables, there are "Revisions spreadsheets" in which all the calculations needed to get the $t$-statistics (standard and adjusted) are performed (see table A1).

In this technical appendix we discuss the choices made by ONS in calculating both the standard and the modified $t$-statistics. We then propose a simple approach to evaluate the significance of the mean revision which stands on the *Heteroskedasticity Autocorrelation Consistent* (HAC) estimate's variance by Newey and West (1987).

All the involved formulae are made explicit and an example of implementation in an Excel spreadsheet is presented to make it clear the feasibility of the proposed approach for routine revisions analysis.

### Standard *t*-statistic

The standard $t$-statistic is defined as (JS, p. 72)

$$t = \frac{\bar{R} - \mu}{\sqrt{\sigma^2 / n}}$$

where $\bar{R}$ is the sample mean revision, $\mu$ is the population mean revision (which is assumed zero by the null hypothesis), $\sigma^2$ is the variance and *n* is the number of observations. It should be noted that (i) the gaussianity of the revisions $R_t$, *t*=1,…,*n*, is implicitly assumed, and (ii) if $\sigma^2$ was known, under the null hypothesis the previous test-statistic would be gaussian and not a Student's *t*. However, the variance $\sigma^2$ being unknown, it has to be estimated. In the standard case JS (p. 72, step 3 of calculation) estimate $\sigma^2$ as

$$s^2 = \frac{\sum_{t=1}^{n}\left(R_t - \overline{R}\right)^2}{n}$$

and the standard *t*-statistic in the "Revisions spreadsheet" is calculated as

$$\text{standard } t = \frac{\overline{R}}{\sqrt{s^2 / n}} = \frac{\overline{R}}{s}\sqrt{n}\ .$$

However, in the case of an independent random sample, $s^2$ is a biased estimate of $\sigma^2$, and an unbiased estimate is obtained by dividing $\sum_{t=1}^{n}\left(R_t - \overline{R}\right)^2$ by *n*-1 instead of *n*. As a consequence, the standard *t*-statistic should be calculated as

$$\text{updated standard } t = \frac{\overline{R}}{\sqrt{\tilde{s}^2 / n}} = \frac{\overline{R}}{\tilde{s}}\sqrt{n}\ ,$$

where

$$\tilde{s}^2 = \frac{\sum_{t=1}^{n}\left(R_t - \overline{R}\right)^2}{n-1}$$

is the unbiased estimate of $\sigma^2$ for independent sample data.

The reported critical value (t-crit in the "Revisions spreadsheet", see table B1) comes from the *t* distribution with *n* degrees of freedom. In line with the previous proposal, a critical value for the standard *t*-statistic coming from Student's *t* distribution with *n*-1 degrees of freedom should be considered.

**Adjusted *t*-statistic**

JS (p. 72) assume that the revisions fit a regression model of the form

$$R_t = \mu + \varepsilon_t\ , \quad t=1,\ldots,n.$$

If the errors are thought to be serially correlated according to an AR(1) model, that is

$$\varepsilon_t = \alpha\varepsilon_{t-1} + u_t \qquad\qquad |\alpha| < 1$$

**Table B1: An example of Revisions Spreadsheet by ONS**

## Revisions spreadsheet for quarterly data *M1 estimates, after 2 months*

(Please note that each worksheet in this file contains calculations for one indicator)

| Relating to Period | Month 1 estimate | Value 1 month later | Revision | r(t-1) | Abs (rev) |
|---|---|---|---|---|---|
| 2000Q1 | 0.4 | 0.5 | 0.1 | | 0.1 |
| 2000Q2 | 0.9 | 0.9 | 0.0 | 0.1 | 0.0 |
| 2000Q3 | 0.7 | 0.7 | 0.0 | 0.0 | 0.0 |
| 2000Q4 | 0.3 | 0.3 | 0.0 | 0.0 | 0.0 |
| 2001Q1 | 0.3 | 0.4 | 0.1 | 0.0 | 0.1 |
| 2001Q2 | 0.3 | 0.3 | 0.0 | 0.1 | 0.0 |
| 2001Q3 | 0.6 | 0.5 | -0.1 | 0.0 | 0.1 |
| 2001Q4 | 0.2 | 0.0 | -0.2 | -0.1 | 0.2 |
| 2002Q1 | 0.1 | 0.0 | -0.1 | -0.2 | 0.1 |
| 2002Q2 | 0.9 | 0.6 | -0.3 | -0.1 | 0.3 |
| 2002Q3 | 0.7 | 0.8 | 0.1 | -0.3 | 0.1 |
| 2002Q4 | 0.4 | 0.4 | 0.0 | 0.1 | 0.0 |
| 2003Q1 | 0.2 | 0.2 | 0.0 | 0.0 | 0.0 |
| 2003Q2 | 0.3 | 0.3 | 0.0 | 0.0 | 0.0 |
| 2003Q3 | 0.6 | 0.7 | 0.1 | 0.0 | 0.1 |
| 2003Q4 | 0.9 | 0.9 | 0.0 | 0.1 | 0.0 |
| 2004Q1 | 0.6 | 0.6 | 0.0 | 0.0 | 0.0 |
| 2004Q2 | 0.9 | 0.9 | 0.0 | 0.0 | 0.0 |
| 2004Q3 | 0.4 | 0.4 | 0.0 | 0.0 | 0.0 |
| 2004Q4 | 0.7 | 0.7 | 0.0 | 0.0 | 0.0 |

| Standard t-test | | Adjusted t-stat | |
|---|---|---|---|
| n | 20 | **alpha** | 0.1889 |
| rbar | -0.0150 | **s*^2** | 0.0136 |
| **s^2** | 0.0093 | **n*** | 19 |
| **t** | **-0.6965** | **t*** | **-0.5753** |
| **t-crit** | 2.0860 | **t*-crit** | 2.0930 |
| **significant or not?** | **No** | **significant or not?** | **No** |

| | | | |
|---|---|---|---|
| **Test for significance of Mean Revision** | | **Mean Revision =** | **- 0.02** |
| **Which test to use** | **Adjusted** | **Abs av rev =** | **0.06** |
| **Is test significant?** | **No** | **Is test significant?** | **No** |

where $u_t$ are independent, JS, following Priestley (1981), claim that "where auto-correlation is present, the equivalent number of independent observations for estimating the mean should be reduced to $n\frac{(1-\alpha)}{(1+\alpha)}$, and thus the variance of the mean should be adjusted by increasing it to $\frac{\sigma^2(1+\alpha)}{n(1-\alpha)}$".[7]

This serves as starting point of a "modified t-adjusted statistic" (referred to as "Adjusted t-stat" in the Revisions spreadsheet, see table A1), which is defined by JS as:

$$t-adj=\frac{\bar{R}}{\sqrt{\text{adjusted variance}}}\left[\text{using } n^* \text{ degrees of freedom}\right]$$

where $n^*$ is the "equivalent number of independent observations for estimating the variance", given by

$$n^*=n\frac{(1-\alpha^2)}{(1+\alpha^2)}.$$

As for the estimation of $\alpha$, it is estimated by JS (p. 72, step 2 of calculation) as

$$\hat{\alpha}=\frac{Cov(R_{t-1},R_t)}{\sqrt{Var(R_{t-1})Var(R_t)}}=\frac{\sum_{i=2}^{n}(R_{i-1}-\bar{R}_{t-1})(R_i-\bar{R}_t)}{\sqrt{\sum_{i=2}^{n}(R_{i-1}-\bar{R}_{t-1})^2\sum_{i=2}^{n}(R_i-\bar{R}_t)^2}},$$

where $\bar{R}_{t-1}=\frac{1}{n-1}\sum_{i=1}^{n-1}R_i$ and $\bar{R}_t=\frac{1}{n-1}\sum_{i=2}^{n}R_i$ . It should be noted that Priestley (1981, p. 331) suggests that $\alpha$ should be estimated as

$$\tilde{\alpha}=\frac{\sum_{t=1}^{n-1}(R_t-\bar{R})(R_{t+1}-\bar{R}_t)}{\sum_{t=1}^{n}(R_t-\bar{R})^2},$$

which is a common standard practice in time series analysis, rather than as $\hat{\alpha}$: "Its disadvantages are that (…) it uses two different estimates of the mean (…) and two different estimates of the variance (…) (derived separately from the first (*N*-*r*) and last (*N*-*r*) observations). It does not, therefore, make full use of the stationary character of $\{X_t\}$)" (Priestley, 1981, p. 331), where for the problem in hand and in our notation, it is $N\equiv n$, $r\equiv 1$ and $X_t\equiv R_t$.

In practical terms, the modified *t*-statistic, *t-adj*, is calculated as (JS, p. 72, step 6 of calculation)

[7] It should be noted that an increase in the variance of the mean occurs for $\alpha>0$, while the variance reduces when $\alpha<0$. In this last case ONS uses the standard *t*-statistic.

$$t - adj = \frac{\overline{R}}{s^*}$$

where $s^* = \sqrt{\frac{s^2}{n}\frac{(1+\hat{\alpha})}{(1-\hat{\alpha})}}$, and the degrees of freedom of the reference Student`s *t* are $n^*$.

**Some first comments**

This way of taking into account the possible auto-correlation of the data series seems to be rather 'weak' from both a theoretical and a practical point of view.

First of all, it should be noticed that only auto-correlation of order 1 has been considered, but this is only one of the various forms of possible violations to the standard assumption (basically, omoskedasticity and lack of serial correlation) which can affect the regression model assumed as the basis for our analysis.

Secondly, in line with the chosen framework, the use of the modified *t*-statistic would be justified only when $\alpha$ is significantly different from zero, otherwise the standard *t*-statistic should be used. So, a significance test on $\alpha$ should be performed, for example using the classical test-statistic

$$\frac{\hat{\alpha}}{\sqrt{1-\hat{\alpha}^2}}\sqrt{n-2}$$

which in the Gaussian case[8], under the null $\alpha = 0$, distributes as a Student's *t* with *n*-2 degrees of freedom. An alternative could be using the portmanteau statistic

$$\hat{\alpha}^2 \cdot n$$

which under the null of zero autocorrelation asymptotically follows a $\chi^2(1)$ distribution.

Thirdly, the choice of using the standard *t*-statistic instead of the adjusted one when the estimated $\alpha$ is negative does not seem to have a sensible statistical meaning.

However, there exists a more comprehensive non-parametric statistical framework able to take into account (heteroskedasticity and) autocorrelation of unknown forms when estimating the variances and covariances of the estimated parameters, namely the *Heteroskedasticity and Autocorrelation Consistent* (HAC) solution given by Newey and West (1987, see also Andrews, 1991, Andrews and Monahan, 1992, and Newey and West, 1994).

In the next section it is discussed the derivation of a robust estimate according to Newey and West (1987) of the variance of the mean revision, which can be used to verify the hypothesis that the parameter of interest is zero against the alternative of non-zero mean revision.

[8] The normality hypothesis can be tested through the Jarque-Bera (1980, 1987, see George, 2005, p. 34).

**Heteroskedasticity and Autocorrelation Consistent (HAC) variance of the mean revision**

Our starting point is again the simple regression model $R_t = \mu + \varepsilon_t$, $t$=1,…,$n$, used by JS. In this case, however, we don't assume any specific form for the disturbance term.

When in a regression model the forms of heteroskedasticity and/or autocorrelation are not known, it may not be possible to obtain efficient estimates of the parameters using weighted least squares. Ordinary Least Squares (OLS) still provides consistent parameter estimates in the presence of eteroskedasticity and/or autocorrelation, but the usual OLS standard errors will be incorrect and should not be used for inference.

Newey and West (1987) have proposed a general covariance estimator that is consistent in the presence of both eteroskedasticity and autocorrelation of unknown form. Consider the multiple linear regression model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where $\mathbf{y}$ is a ($n$x1) vector of observations on the dependent variable, $\mathbf{X}$ is a ($n$x$k$) matrix of regressors, $\boldsymbol{\beta}$ is a ($k$x1) vector of unknown coefficient and $\boldsymbol{\varepsilon}$ is a ($n$x1) random vector with mean zero and unknown covariance matrix $\boldsymbol{\Omega}$. Let $\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y}$ be the OLS estimated coefficient vector, and $\hat{\varepsilon}_t = \left(y_t - \mathbf{x}_t'\hat{\boldsymbol{\beta}}\right)$, $t$=1,…,$n$, the OLS residuals. The Newey-West estimator of the covariance matrix of the estimated coefficients is given by

$$\widehat{Var\left(\hat{\boldsymbol{\beta}}\right)} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\hat{\boldsymbol{\Omega}}\left(\mathbf{X}'\mathbf{X}\right)^{-1}$$

where

$$\hat{\boldsymbol{\Omega}} = \frac{n}{n-k}\left\{\sum_{t=1}^{n}\hat{\varepsilon}_t\mathbf{x}_t\mathbf{x}_t' + \sum_{v=1}^{q}\left(\left(1-\frac{v}{q+1}\right)\sum_{t=v+1}^{n}\left(\mathbf{x}_t\hat{\varepsilon}_t\hat{\varepsilon}_{t-v}\mathbf{x}_{t-v}' + \mathbf{x}_{t-v}\hat{\varepsilon}_{t-v}\hat{\varepsilon}_t\mathbf{x}_t'\right)\right)\right\}$$

and $q$, the truncation lag, is a parameter representing the number of autococrrelations used in evaluating the dynamics of the OLS residuals $\hat{\varepsilon}_t$. Following the suggestion by Stock and Watson (2003, p. 505), $q$ can be calculated as[9]

$$q = \left[0.75 \cdot n^{1/3}\right],$$

where $[x]$ stands for integer part of $x$. In practice, it turns out $q$=1 for $n$≤18, $q$=2 for 19≤$n$≤63, and 64≤$n$≤151.

When the parameter of interest is the mean, that is the constant in the regression model $R_t = \mu + \varepsilon_t$, the

[9] The popular econometric package EViews (Quantitative Micro Software, 2004, p. 457) sets $q = \left[4 \cdot \left(n/100\right)^{2/9}\right]$. This expression was used by Newey and West (1994, p. 634) in the prewhitening phase of their automatic lag selection procedure.

calculations are notably simplified. For, in this case $k$=1, the OLS estimate of $\mu$ is $\overline{R} = \frac{1}{n}\sum_{t=1}^{n} R_t$ , matrix **X** is a ($n$x1) vector of one, and $\mathbf{X}'\mathbf{X} = n$ . Then, the Newey-West HAC estimator of the variance of the estimated mean revision is given by

$$\widehat{Var\left(\overline{R}\right)} = \frac{1}{n\left(n-1\right)}\left\{\sum_{t=1}^{n}\hat{\varepsilon}_t^2 + 2\sum_{v=1}^{q}\left(\left(1-\frac{v}{q+1}\right)\sum_{t=v+1}^{n}\hat{\varepsilon}_t\hat{\varepsilon}_{t-v}\right)\right\},$$

where $\hat{\varepsilon}_t = R_t - \overline{R}$ . For $q$=2, which is the most usual case for revisions analysis purposes, this formula becomes

$$\widehat{Var\left(\overline{R}\right)} = \frac{1}{n\left(n-1\right)}\left\{\sum_{t=1}^{n}\hat{\varepsilon}_t^2 + \frac{4}{3}\sum_{t=2}^{n}\hat{\varepsilon}_t\hat{\varepsilon}_{t-1} + \frac{2}{3}\sum_{t=3}^{n}\hat{\varepsilon}_t\hat{\varepsilon}_{t-2}\right\}.$$

An example of spreadsheet containing a prospect of all the needed calculations for the revisions series of table B1, is presented in table B2.

## Table B2: Prospect of calculations of summary statistics for the revisions analysis

| Period | Month 1 estimate (P) | Value 1 month later (L) | R(t)=L(t) -P(t) | abs(R(t)) | R(t)^2 | R(t-1) | R(t-2) | R(t)-Rbar | R(t-1)-Rbar | R(t-2)-Rbar | (R(t)-Rbar)^2 | (R(t)-Rbar)* (R(t-1)-Rbar) | ((R(t)-Rbar)* (R(t-2)-Rbar) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2000Q1 | 0.4 | 0.5 | 0.1 | 0.1 | 0.010 | | | 0,115 | | | 0,013 | | |
| 2000Q2 | 0.9 | 0.9 | 0.0 | 0.0 | 0.000 | 0.1 | | 0,015 | 0,12 | | 0,000 | 0,002 | |
| 2000Q3 | 0.7 | 0.7 | 0.0 | 0.0 | 0.000 | 0.0 | 0.1 | 0,015 | 0,02 | 0,12 | 0,000 | 0,000 | 0,002 |
| 2000Q4 | 0.3 | 0.3 | 0.0 | 0.0 | 0.000 | 0.0 | 0.0 | 0,015 | 0,02 | 0,02 | 0,000 | 0,000 | 0,000 |
| 2001Q1 | 0.3 | 0.4 | 0.1 | 0.1 | 0.010 | 0.0 | 0.0 | 0,115 | 0,02 | 0,02 | 0,013 | 0,002 | 0,002 |
| 2001Q2 | 0.3 | 0.3 | 0.0 | 0.0 | 0.000 | 0.1 | 0.0 | 0,015 | 0,12 | 0,02 | 0,000 | 0,002 | 0,000 |
| 2001Q3 | 0.6 | 0.5 | -0.1 | 0.1 | 0.010 | 0.0 | 0.1 | -0,085 | 0,02 | 0,12 | 0,007 | -0,001 | -0,010 |
| 2001Q4 | 0.2 | 0.0 | -0.2 | 0.2 | 0.040 | -0.1 | 0.0 | -0,185 | -0,09 | 0,02 | 0,034 | 0,016 | -0,003 |
| 2002Q1 | 0.1 | 0.0 | -0.1 | 0.1 | 0.010 | -0.2 | 0.1 | -0,085 | -0,19 | -0,09 | 0,007 | 0,016 | 0,007 |
| 2002Q2 | 0.9 | 0.6 | -0.3 | 0.3 | 0.090 | -0.1 | 0.2 | -0,285 | -0,09 | -0,19 | 0,081 | 0,024 | 0,053 |
| 2002Q3 | 0.7 | 0.8 | 0.1 | 0.1 | 0.010 | -0.3 | 0.1 | 0,115 | -0,29 | -0,09 | 0,013 | -0,033 | -0,010 |
| 2002Q4 | 0.4 | 0.4 | 0.0 | 0.0 | 0.000 | 0.1 | 0.3 | 0,015 | 0,12 | -0,29 | 0,000 | 0,002 | -0,004 |
| 2003Q1 | 0.2 | 0.2 | 0.0 | 0.0 | 0.000 | 0.0 | 0.1 | 0,015 | 0,02 | 0,12 | 0,000 | 0,000 | 0,002 |
| 2003Q2 | 0.3 | 0.3 | 0.0 | 0.0 | 0.000 | 0.0 | 0.0 | 0,015 | 0,02 | 0,02 | 0,000 | 0,000 | 0,000 |
| 2003Q3 | 0.6 | 0.7 | 0.1 | 0.1 | 0.010 | 0.0 | 0.0 | 0,115 | 0,02 | 0,02 | 0,013 | 0,002 | 0,002 |
| 2003Q4 | 0.9 | 0.9 | 0.0 | 0.0 | 0.000 | 0.1 | 0.0 | 0,015 | 0,12 | 0,02 | 0,000 | 0,002 | 0,000 |
| 2004Q1 | 0.6 | 0.6 | 0.0 | 0.0 | 0.000 | 0.0 | 0.1 | 0,015 | 0,02 | 0,12 | 0,000 | 0,000 | 0,002 |
| 2004Q2 | 0.9 | 0.9 | 0.0 | 0.0 | 0.000 | 0.0 | 0.0 | 0,015 | 0,02 | 0,02 | 0,000 | 0,000 | 0,000 |
| 2004Q3 | 0.4 | 0.4 | 0.0 | 0.0 | 0.000 | 0.0 | 0.0 | 0,015 | 0,02 | 0,02 | 0,000 | 0,000 | 0,000 |
| 2004Q4 | 0.7 | 0.7 | 0.0 | 0.0 | 0.000 | 0.0 | 0.0 | 0,015 | 0,02 | 0,02 | 0,000 | 0,000 | 0,000 |
| | | | | | | | | | | **Total** | 0.186 | 0.034 | 0.044 |

### Summary statistics

| Basic statistics | | Other statistics | |
|---|---|---|---|
| n | 20 | Correlation between L and P | 0.9340 |
| mean absolute revision | 0.0550 | Min revision | -0.3 |
| mean revision (Rbar) | -0.0150 | Max revision | 0.1 |
| st. dev(Rbar) - HAC formula | 0.0261 | Range | 0.4 |
| mean squared revision | 0.0095 | % L > P | 20 |
| t-stat | -0.5739 | % Sign(L) = Sign(P) | 90 |
| t-crit | 2.0930 | UM % | 2.37 |
| degrees of freedom | 19 | UR % | 0.23 |
| Is mean revision significant?* | NO | UD % | 97.40 |

**References**


Ahmad N., S. Bournot and F. Koechlin (2004), *Revisions to quarterly GDP estimates: a comparative analysis for seven large OECD countries*, paper presented at the OECD-ONS workshop Paris, 7-8 October 2004, http://www.oecd.org/dataoecd/20/26/34350524.pdf.

Algera S. (1999), *Structural business statistics: a key role in the system of economic statistics, a Dutch case*, paper presented at the seminar on Structural Business Statistics, Warsaw, 7-9 June 1999.

Andrews D.W.K. (1991), Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica*, 59: 817-858.

Andrews D.W.K. and J.C. Monahan (1992), An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator, *Econometrica*, 61: 821-856.

Bier W. and H. Ahnert (2001), Trade-off between timeliness and accuracy. ECB requirements for general economic statistics, *Economisch Statistische Berichten (ESB)*, 15 March 2001, (English traduction from Dutch).

Carson C.S. (2000), *What is data quality? A distillation of experience*, Statistics Department, International Monetary Fund (*mimeo*).

Carson C.S. and L. Laliberté (2002), Assessing accuracy and reliability: A note based on approaches used in National Accounts and Balance of Payments statistics, *IMF working paper 02/24*.

ECB (2001), Revisions to quarterly national accounts data for the euro data, *Monthly Bulletin*, August 2001, pp. 25-27.

Franchet Y. and W. Grünewald (2002), Eurostat's approach to quality, *OECD Statistics Newsletter*, 8.

Jarque C.M. and A.K. Bera (1980), Efficient tests for normality, homoscedasticity and serial independence of regression residuals, *Economics Letters*, 6: 255-259

Jarque C.M. and A.K. Bera (1987), A test for normality of observations and regression residuals, *International Statistical Review*, 55: 163-172.

Jenkinson G. and E. George (2005), Publication of revisions triangles on the National Statistics website, *Economic Trends*, 614: 43-44.

Jenkinson G. and N. Stuttard (2004), Revisions information in ONS first releases, *Economic Trends*, 604: 70-72.

Newey W.K. and K.D. West (1987), A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55: 703-708.

Newey W.K. and K.D. West (1994), Automatic lag selection in covariance matrix estimation, *Review of Economic Studies*, 61: 631-653.

Office of Management and Budget (1983), *An interagency review of time-series revision policy*, Federal Committee on Statistical Methodology, http://www.fcsm.gov/working-papers/sw7.html.

ONS (2004), *Protocol on revisions*, London, ONS.

Priestley M.B. (1981), *Spectral analysis and time series*, London, Academic Press.

Quantitative Micro Software (2004), *EViews 5.1 User's Guide*, Irvine CA.

Rao J.N.K., K.P. Srinath and B. Quenneville (1989), Estimation of level and change using current preliminary data, in D. Kasprzyk, G. Duncan, G. Kalton and M.P. Singh (eds.), *Panel surveys*, New York, Wiley, pp. 457-479.

Stock J.H. and M.W. Watson (2003), *Introduction to econometrics*, Boston, Addison Wesley.

Theil H. (1961), *Economic forecasts and policy*, Amsterdam, North-Holland.