

Performing Revisions and Real-time Data Analysis Introducing the Main Economic Indicators Original Release Data and Revisions Database

By Richard McKenzie

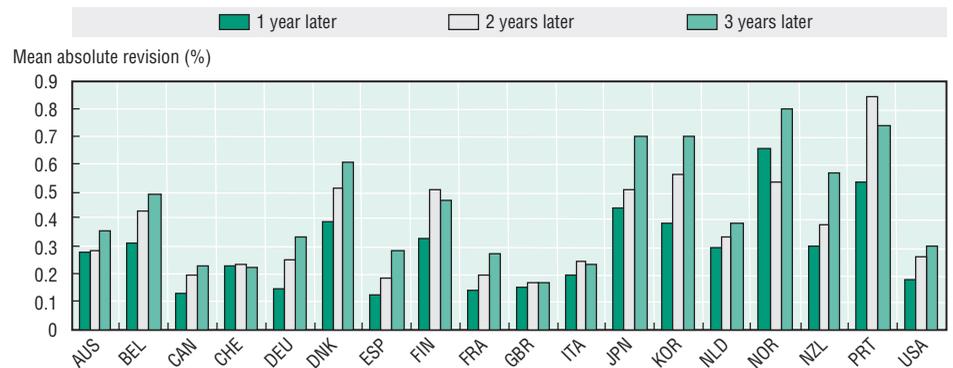
First releases of official statistics are often revised in subsequent releases, sometimes substantially. Such revisions can impact on policy decisions, as revisions to first published data may alter the previous assessment of the state of the economy. This may occur through a changed interpretation based on the revised data themselves or the impact the revisions may have on econometric models which may incorporate several statistics, each subject to revision. Whilst this is a recognised issue of key importance, most producers of official statistics do not quantify expected revisions to their data and economists do not have the required data to test the sensitivity of their econometric models to revisions in input data. This important gap in knowledge required to effectively use official statistics motivated the OECD to develop a unique new product: the *Main Economic Indicators Original Release Data and Revisions Database*, freely available at <http://stats.oecd.org/mei/default.asp?rev=1>.

Accessing this source of originally published data will enable analysts to test the likely effectiveness of econometric models in simulated real-time. It will also enable producers of official statistics to study the magnitude and direction of subsequent revisions to published data, which can lead to a better understanding of the statistical compilation process, enabling problems to be identified and improvements to be made. The OECD has already exploited this database through undertaking comprehensive revisions analyses for a range of OECD member countries and selected non-member economies for gross domestic product (GDP), index of industrial production (IIP) and retail trade volume (RTV). Figure 1 shows the mean absolute revision between first published estimates of seasonally adjusted

In this issue

- 2 **The Main Economic Indicators Original Release Data and Revisions Database**
- 2 **Revisions analysis of official statistics**
- 3 **Main findings from revisions analysis studies undertaken by the OECD**
- 6 **Demand by central banks for a “real-time” database**
- 8 **Further information**

Figure 1. Mean absolute revision to first published estimates of quarter-on-previous-quarter growth rates for GDP at different intervals, 1995-2004



GDP quarter-on-previous-quarter (QoQ) growth rates and estimates published one, two and three years later. The data used for the eighteen countries included in this study relate to the reference periods from 1995Q1¹ to 2004Q4, as published in successive monthly editions of the MEI from June 1995 to June 2006. These studies have stimulated national statistical institutes (NSIs) to pay more attention to revisions analysis studies and to improve existing revisions policies in the publication of official statistics.

The Main Economic Indicators Original Release Data and Revisions Database

The *Main Economic Indicators (MEI) Original Release Data and Revisions Database* is a user friendly web-interface² providing access to a wide range of data, information and automated programs. Its main features can be described by the following points.

- Contains full time series, as far back as 1960 in some cases, for 21 key economic variables as originally published in each monthly edition of the OECD Main Economic Indicators (MEI) CD-Rom from February 1999 onwards for OECD countries, the Euro area, China, India, Brazil, South Africa and the Russian Federation. This database is updated on a monthly basis and provides the raw data needed by economists to test the performance of their econometric models in simulated real-time.
- Provides access to comprehensive revisions analysis studies performed by the OECD for GDP, IIP and RTV.
- Contains automated programs and a detailed user guide allowing either producers or users of official statistics to perform their own revisions analysis based on the OECD methodology for any country and variable combination available in the database. Alternatively, producers of official statistics can use the automated programs to perform revisions analysis on their own data for any variable, provided they have access to their own source of vintage datasets.
- Provides information on reasons for revisions, together with recommended practices to aid producers of official statistics in establishing a transparent revisions policy for economic statistics.

The following variables are included in the database: GDP and its expenditure components; industrial production and production in construction; OECD composite leading

indicators; retail trade, consumer price index; standardised unemployment rate; civilian employment; hourly earnings in manufacturing; monetary aggregates; international trade in goods and current account balance. Market based financial variables also published in the MEI³ (e.g. interest rates, exchange rates), which are often part of econometric models, are not included in the database as they are not revised.

Revisions analysis of official statistics

Analysis of revisions for key economic variables such as those included in the database enables NSIs to evaluate their performance against a key dimension of statistical quality – reliability. Ultimately, this can lead to a better understanding of the statistical compilation process and enable problems to be identified and improvements to be made. This is one of the reasons why the International Monetary Fund's Special Data Dissemination Standards (SDDS) encourages countries to undertake revisions analyses and gives considerable prominence to the need for NSIs to develop a revisions policy that is both transparent (as to the underlying cause(s) of revisions) and consistent across the range of economic statistics (both structural and short-term) compiled.

It should be noted that the existence of revisions per se should not be viewed in a negative light. They are a necessary by-product of the statistical production process, which aims to produce the highest quality outputs given the available information at a particular point in time. Therefore, if a particular statistic is never revised one should not assume that the data are of high quality without reviewing the associated compilation methods and data sources used and the NSI's revisions policy. Revisions can occur due to a variety of reasons as summarised in Table 1 below. The extent to which countries undertake the processes outlined in Table 1 can also affect the size of revisions to their statistics, and this can affect the comparability of results of revisions analyses across countries.

Despite the recognised importance of revisions analysis, few national NSIs undertake them on an ongoing basis for their key economic statistics. Presumably, the reason is that obtaining the necessary data and developing the programs required is a resource intensive task. However, this new OECD facility provides the opportunity for NSIs to easily and quickly undertake such analyses.

1. Reference periods covered start later for Germany (95Q2), Portugal (96Q1), Belgium (96Q4) and Korea (99Q4).
2. <http://stats.oecd.org/mei/default.asp?rev=1/>.

3. A wide range of these market based financial variables can be downloaded at <http://stats.oecd.org/mei/default.asp>.

Main findings from revisions analysis studies undertaken by the OECD

Three comprehensive revisions analysis studies using the data available in the *MEI Original Release Data and Revisions Database* were undertaken by the OECD in 2005 and 2006. These were for GDP (Di Fonzo 2005, Tosetta and Lequiller 2006), IIP (McKenzie and Park 2006) and RTV (McKenzie 2006). The detailed results from these studies and data used can be accessed through the database interface. For each of these studies detailed revisions analysis spreadsheets and a wide range of summary statistics were produced for each country⁴ based on the methodology established by Di Fonzo (2005).

The periods for which the measures of revisions are derived are relatively short and it is possible that the data for some countries have been more subject to major revision events, such as the introduction of chain-linked volume estimates or major benchmark revisions in the period than others by chance. In addition, the data used for the analysis are those published by the OECD, and they can differ from those nationally published in some cases.⁵ Consequently, countries were invited to provide comments on the reasons for revisions and this information is provided together with the detailed results from each study.

Figure 2 shows the mean absolute revision between first published estimates of seasonally adjusted IIP year-on-year (YoY) growth rates and estimates published 3 months later, 1 year later and two years later for all OECD countries (except Iceland), Brazil, India, the Russian Federation and South Africa. The analysis is derived from data published in each successive monthly edition of the MEI from February 1999 to February 2006.

Figure 3 shows the mean absolute revision between first published estimates of seasonally adjusted Retail trade volume month-on-previous-month (MoM) growth rates and estimates published 3 months later, 1 year later and two years later for twenty five OECD countries, the Russian Federation and South Africa. The analysis is derived from

4. The key for country acronyms included in the graphs is: AUS = Australia; AUT = Austria; BEL = Belgium; BRA = Brazil; CAN = Canada; CHE = Switzerland; CZE = Czech Republic; DEU = Germany; DNK = Denmark; EMU = Euro area; ESP = Spain; FIN = Finland; FRA = France; GBR = United Kingdom; GRC = Greece; HUN = Hungary; IND = India; IRL = Ireland; ITA = Italy; JPN = Japan; KOR = Korea; LUX = Luxembourg; MEX = Mexico; NLD = Netherlands; NOR = Norway; POL = Poland; PRT = Portugal; RUS = Russian Federation; SVK = Slovak Republic; TUR = Turkey; USA = United States; ZAF = South Africa.

5. For example, data have been seasonally adjusted by the OECD rather than the country. In addition, delays and/or errors in the data transmission process can have an impact in rare circumstances.

Table 1. Reasons for revision of official statistics

1. Incorporation of source data with more complete or otherwise better reporting (e.g. including late respondents) in subsequent estimates.
2. Correction of errors in source data (e.g. from editing) and computations (e.g. revised imputation).
3. Replacement of first estimates derived from incomplete samples (e.g. sub-samples) judgmental or statistical techniques when firmer data become available.
4. Incorporation of source data that more closely match the concepts and/or benchmarking to conceptually more accurate but less frequent statistics.
5. Incorporation of updated seasonal factors.
6. Updating of the base period of constant price estimates.
7. Changes in statistical methodology (such as the introduction of chain-linked volume estimates), concepts, definitions, and classifications.
8. Revisions to national accounts statistics arising from the confrontation of data in supply and use tables.

data published in each successive monthly edition of the MEI from February 1999 to May 2006.

In almost all countries included in the revisions analyses studied, the size of mean absolute revisions to first published estimates of QoQ (for GDP), MoM and YoY (for IIP and RTV) growth rates are non-ignorable and increase the longer the interval from the first estimate. It is relatively difficult to distinctively group countries into those with say high, medium or low mean absolute revisions as there appears to be a degree of similarity across a large number of countries and the ordering of the size of mean absolute revisions across countries can differ depending on the revisions interval considered (e.g. revisions after 3 months, 1 year, 2 years, etc.). Nonetheless, it is apparent in each of the studies that some countries have noticeably higher⁶ mean absolute revision (MAR) than the majority of

6. Whilst the mean absolute revision (MAR) has been used here as a basis for comparison across countries, it is also useful to observe the mean absolute revisions to growth rates relative to the mean size of the growth rates themselves – referred to as the relative mean absolute revision (RMAR) – which may often give a different ordering for the size of revisions across countries. OECD studies show results for both MAR and RMAR across countries.

others. In particular, Japan, Korea, Norway and Portugal stand out as having high MAR for GDP, Belgium and the Russian Federation for IIP and Belgium, Greece, Czech Republic and the Netherlands for RTV.

Assessing the statistical significance of mean revisions

Ideally, revisions should have a tendency to be random; that is, equally likely to be positive or negative and centred around zero. Therefore, an important output of a revisions analysis study are the tests to determine whether mean revisions (calculated at a range of different time lengths from the first estimate) are statistically significantly different from zero. Statistically significant mean revisions imply that the data have a significant tendency to be revised in a particular direction (*i.e.* up or down) and that the compilation methodology for early estimates should be reviewed. This is just one good example of how revisions analysis studies can serve as a catalyst for identifying

problems with compilation practices that require investigation by the producers of official statistics.

Mean revisions to GDP (for QoQ growth rates), IIP and RTV (for MoM and YoY growth rates) between first estimates and those published one⁷ year later were assessed for statistical significance for all countries included in the respective studies. Those countries found to be statistically significantly different from zero were: in the case of GDP, Denmark; in the case of IIP, Greece, Belgium and India for MoM growth rates and Belgium, India, Russia, Turkey, Germany, Euro area, France, United Kingdom and Korea

7. Assessing the statistical significance of mean revisions between first published estimates and those published one year later was considered the most appropriate interval to detect any possible systematic tendencies for first estimates to be revised in a particular direction that should be addressed in compilation procedures. This is because revisions may have a greater tendency to be in a particular direction the longer the revision interval (*e.g.* after 3 years) if, for example, there is a systematic impact of major methodological changes applied to the entire time series.

Figure 2. Mean absolute revision to first published estimates of year-on-year growth rates for the index of industrial production (February 1999-February 2006)



for YoY growth rates; and in the case of RTV, Canada for MoM growth rates, and Germany, Czech Republic, Korea and the Netherlands for YoY growth rates.

Assessing the robustness of first published estimates of MoM growth rates

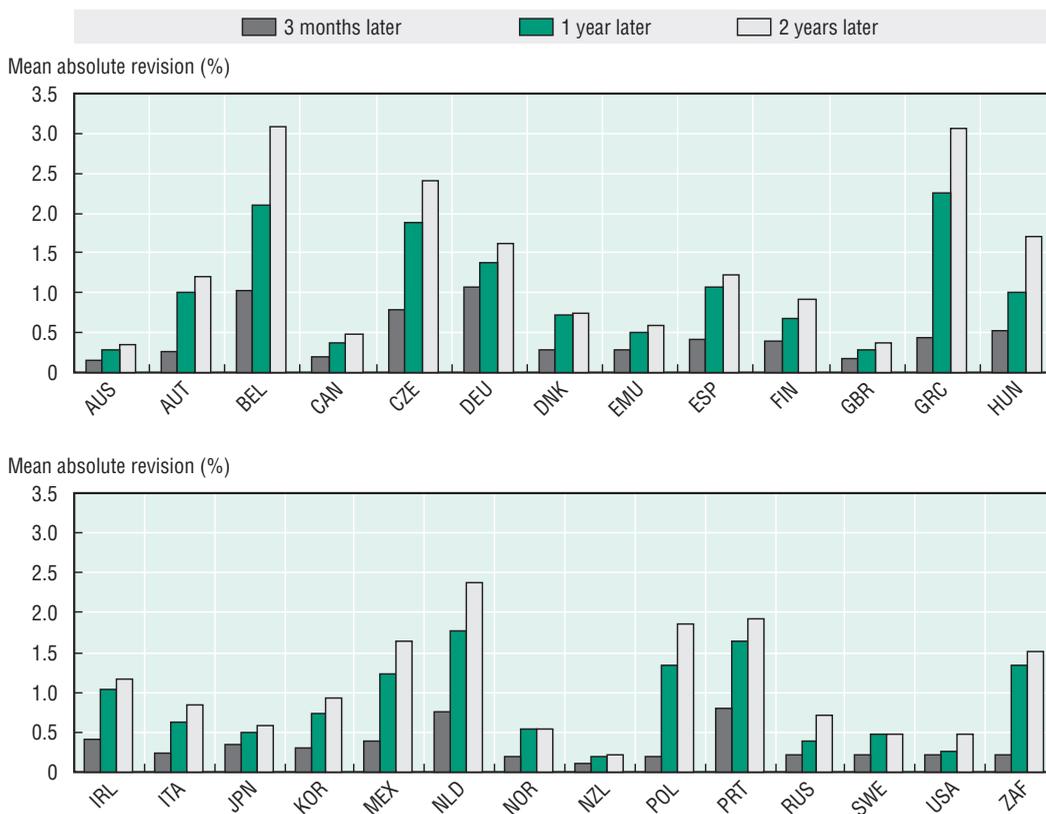
Analysts often refer to the latest available (*i.e.* first published) estimates of the MoM growth rate of seasonally adjusted IIP as an indicator to assess the current state of the economy – in regards to expansions and contractions in production activity. Likewise, RTV estimates are used as an early indicator of consumer demand. The degree of confidence that users should place on these statistics depends largely on the extent to which the initial estimates are likely to be revised in future months, such that the subsequent revisions may paint a different picture of the current state of industrial activity (IIP) or consumer demand (RTV). If first estimates of MoM growth rates for a country have a history of being largely revised, then users would

be advised to focus on a more robust statistic such as the YoY growth rate or some other measure⁸ for short-term analysis.

One way to assess the relative robustness of first estimates of MoM and YoY growth rates is to calculate the relative mean absolute revision. This statistic compares the mean absolute revision to the mean absolute size of later estimates of growth. Results from such an analysis (McKenzie 2006) show that in the majority of countries' first estimates of MoM growth rates for the IIP and RTV

8. For example, the United Kingdom's Office for National Statistics encourages users to focus on the growth rate of the three month moving average of the IIP which is the headline indicator for their press release, although their press releases also quotes figures of the month-on-previous-month growth rate. Another example is the publication of trend estimates by the Australian Bureau of Statistics, which have proved particularly popular for monthly data. Averaging or trending monthly data not only reduces the effect of revisions but also reduces the "noise" to which monthly data are particularly susceptible.

Figure 3. Mean absolute revision to first published estimates of month-on-previous-month growth rates for retail trade volume (February 1999-May 2006)



should not be considered as reliable early indicators of the magnitude of short-term changes in the volume of industrial output or consumer demand, respectively. This is due to the fact that on average, across all countries, first estimates of MoM growth rates for both the IIP and RTV are revised by approximately two thirds of their initial value within one year. On the other hand, first estimates of YoY growth rates are shown to provide a more robust measure in terms of magnitude, being revised on average across all countries by only 24% (IIP) and 22% (RTV) of their initial value within one year.

Demand by central banks for a “real-time” database

Central bankers and economic forecasters have been paying increasing attention to the fact that their econometric models may be sensitive to revisions to input variables, and thereby affect the usefulness of their models. This reflects the need to test the robustness of their models on historical originally released (and then subsequently revised) data – thus simulating the likely performance of their models in “real-time”. This issue has been a subject of debate over the last few years at the Euro Area Central Bankers Network (EABCN) meetings, and is the subject of a special conference being organised by the Federal Reserve Bank of Philadelphia in April 2007 titled *Real-Time Data Analysis and Methods in Economics*. Consequently this has generated a considerable demand for databases which provide access to first (and subsequent) releases of data for key economic variables. The OECD *Main Economic Indicators Original Release Data and Revisions Database* is the first continuously updated database freely available to the general public that provides the required data to perform this type of analysis across such a wide range of countries.

Real-time analysis for econometric models

In general when undertaking research to construct an econometric model (e.g. to forecast GDP growth) economists are restricted to using the latest available time series of variables under consideration for inclusion in the model. A general approach is to undertake empirical analysis, based on economic theory, to construct the model of best fit, and then test the performance of this model using a set of “out-of-sample” data, that is, observations of the variables which were not included in the empirical analysis which determined the parameters of the model.

This assessment of model performance using out-of-sample data is intended to show how the model can be

expected to perform when applied in practice. However, the out-of-sample data used to evaluate the model consists of variables where each data point will most likely have been revised from its first published value – whereas it is these first published values that will be used by the model when applied in practice. Having access to first-published estimates such as those available in the *MEI Original Release Data and Revisions Database* to use for the out-of-sample analysis overcomes this problem and allows a more realistic assessment of the expected performance of the model in practice. Such analysis is referred to in the economic literature as “real-time” analysis. Furthermore, this data source provides the necessary information for economists to test the sensitivity of model parameters to revisions in the independent variables when constructing their models.

Real-time analysis of a model for private final consumption expenditure in the United States

To demonstrate the use of the database for real-time analysis, McKenzie (2006) developed a simple model to predict QoQ growth rates for private final consumption expenditure using RTV as the independent variable for the United States. This showed that the out-of-sample assessment of the model using first published estimates indicated a noticeably worse model performance (RMSE = 0.36) than if the latest available time series was used for this assessment (RMSE = 0.30).

For more information on this product, contact:
Richard.mckenzie@oecd.org. ■

Previous issues of Statistics Brief

www.oecd.org/statistics

No. 1, October 2001 – Trade in Goods and Services: Statistical Trends and Measurement Challenges

No. 2, February 2002 – Creation of the Euro Area: Implications for Economic Statistics

No. 3, March 2002 – Purchasing Power Parities – Measurement and Uses

No. 4, June 2002 – Measuring Student Knowledge and Skills: The PISA 2000 Experience

No. 5, November 2002 – Measuring the Non-observed Economy

No. 6, October 2003 – Measuring Regional Economies

No. 7, December 2003 – Comparing Growth in GDP and Labour Productivity: Measurement Issues

No. 8, June 2004 – Comparison of Household Saving Ratios: Euro area/United States/Japan

No. 9, July 2005 – The Comparability of International Migration Statistics: Problems and Prospects

No. 10, September 2005 – Measuring Sustainable Development

No. 11, May 2006 – Alternative Measures of Well-being

Further information

Di Fonzo T. (2005), *The OECD project on revisions analysis: First elements for discussion*. Paper presented at the OECD STESEG Meeting, Paris, 27-28 June. Available at www.oecd.org/dataoecd/55/17/35010765.pdf.

Di Fonzo T. (2005), *Revisions in Quarterly GDP of OECD countries*. Paper presented at the OECD Working Party on National Accounts meeting, 11-14 October. Available at www.oecd.org/dataoecd/13/49/35440080.pdf.

McKenzie, R. (2006), *Undertaking Revisions and Real-Time Data Analysis using the OECD Main Economic Indicators Original Release Data and Revisions Database*. OECD Statistics Working Paper. OECD, Paris, August 2006.

McKenzie, R. and Park, S. (2006), *Revisions Analysis of the Index of Industrial Production for OECD Countries and Major Non-Member Economies*. Paper presented at the OECD Short-Term Economic Statistics Working Party meeting, June. Available at www.oecd.org/dataoecd/45/29/36561675.pdf.

OECD (2004), Proceedings from the OECD Revisions Analysis Workshop, October, available at www.oecd.org/document/23/0,2340,en_2825_497146_33729303_1_1_1_1,00.html.

Tosetto, E. and Lequiller, F. (2006), *Revisions in Quarterly GDP of OECD Countries: An Update*. For presentation at the OECD Working Party.

The **Statistics Brief** is published by the Statistics Directorate of the OECD

Editor in chief: Enrico Giovannini

Editor: Charles Aspden

Editorial team: Benoît Arnaud, Eileen Capponi, Julien Dupont

For further information contact:

The Editor

STATISTICS BRIEF

Statistics Directorate

OECD

2, rue André-Pascal

75775 Paris France

e-mail: STD.STATBRIEF@oecd.org

Do you want to subscribe to *Statistics Brief*?

Register now at www.oecd.org/OECDdirect and choose *Statistics Brief* or e-mail/write to the Editor.

Observer oecd

OECD Worldwide

OECD NORTH AMERICA

OECD Washington Center
2001 L Street N.W., Suite 650
Washington, DC 20036-4922, USA
Toll free: +1 (800) 456-6323
Fax: +1 (202) 785-0350
General information:
+1 (202) 785-6323
E-mail: washington.contact@oecd.org
Internet: www.oecdwash.org

OECD JAPAN

OECD Tokyo Centre
Nippon Press Center Bldg., 3rd floor
2-2-1 Uchisaiwaicho, Chiyoda-ku
Tokyo 100-0011, Japan
Tel.: +81 (3) 5532 0021
Fax: +81 (3) 5532 0035
E-mail: center@oecdtokyo.org
Internet: www.oecdtokyo.org

OECD GERMANY

OECD Berlin Centre
Schumannstrasse 10
10117 Berlin
Tel: +49 30 288 8353
Fax: +49 30 288 83545
E-mail: berlin.contact@oecd.org
Internet: www.oecd.org/deutschland

OECD MEXICO

OECD Mexico Centre
Av. Presidente Mazaryk 526
Colonia: Polanco
C.P. 11560
Mexico D.F., Mexico
Tel: +52 55 9138 6233
Fax: +52 55 5280 0480
E-mail: mexico.contact@oecd.org
Internet: www.oedemexico.org.mx

OECD PARIS

OECD Paris Centre
2, rue André-Pascal
75775 Paris Cedex 16, France
Tel: +33 1 45 24 81 67
Fax: +33 1 45 24 19 50
E-mail: sales@oecd.org
Internet: www.oecd.org

