The Statistics Newsletter

for the extended OECD Statistical Network

Issue No. 48, April 2010
www.oecd.org/std/statisticsnewsletter

New ECB Statistics on Euro Area Investment Funds

Statistics Netherlands
Analysing the Sources of Change in Educational Expenditure: a Decomposition Analysis

Wikiprogress: The Global Project for Measuring the Progress of Societies

PLUS...

STATISTICS SWEDEN
Towards Efficient Statistical Data Editing

STATISTICS CANADA
Seasonal Adjustment and Identifying Economic Trends

OECD

Contents

The Statistics Newsletter is published by the Statistics Directorate of the OECD

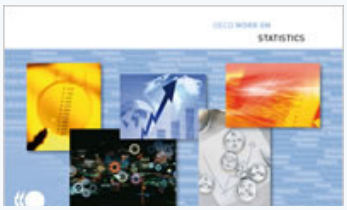
Editor-in-chief: Martine Durand
Editor: David Brackfield
Editorial assistant: Julie Marinho
Technical advice: Sonia Primot
Distribution: Martine Zaïda

For further information contact: The Editor, Statistics Newsletter, email : STD.STATNEWS@oecd.org

Readers are invited to send their articles or comments to the email address above.
Deadline for articles for the next issue:
02 May 2010

Contributors to the issue:
Hugo Elbers and Remco Kaashoek, Statistics Netherlands; Jani Matilainen and Barbara Zupancic, European Central Bank; Angela Hariche, OECD; Anders Norberg, Statistics Sweden; Diana Wyman, Statistics Canada.

OECD Work on Statistics



What kind of statistics can you find at the OECD?
What kind of statistics does OECD work on?
This new brochure will help you answer these questions.

Click **here** to download the brochure

or visit www.oecd.org/statistics/oecdwork for more information

STATISTICS NETHERLANDS

ANALYSING THE SOURCES OF CHANGE IN EDUCATIONAL EXPENDITURE: A DECOMPOSITION ANALYSIS

Hugo Elbers and Remco Kaashoek*

Statistics Netherlands has introduced a decomposition method to explain changes in spending on education. The method uses predefined factors which changes in expenditure are attributed. This article gives a brief explanation of the decomposition method and a number of results for Dutch expenditure on education for the period 1997-2007.

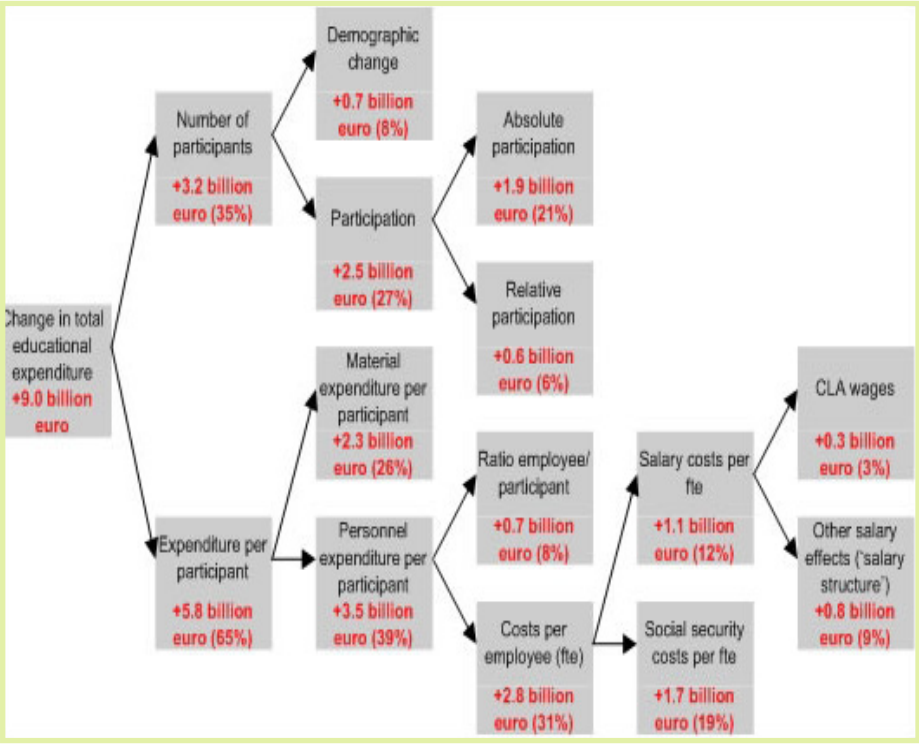
1. A realistic look at spending on education in the Netherlands

Expenditure on education as defined in this article is the expenditure by public and government dependent educational institutes in their responsibility of providing education and for R&D. The figures are deflated (using the GDP deflator) and expressed in the price level of 2007.

Between 1997 and 2007 real expenditure by public and government dependent educational institutions rose from 20 to 29 billion euro, an increase of 45%. In the same period Dutch GDP increased by 29%, so the total expenditure on education also increased significantly relative to GDP. Statistics Netherlands publishes annual statistics about spending on education. There is an increasing demand for explanations of the changes in expenditures, both nationally and internationally – based on the OECD publication Education at a Glance.

The question 'why' is not unique to this specific statistic, but in general it is not easy to come up with a good answer. There are different lines of approach in describing a 'cause' and mostly it is a matter of analysis rather than statistics. Inevitably the explanations for these changes – as provided by Statistics Netherlands at least – had to be fairly superficial. To provide some more in-depth insight, Statistics Netherlands has introduced a decomposition model with a standard set of indicators which influence the changes in spending on education. Statistics on educational expenditure are generally presented from a total spending and a total funding perspective. We focus here on what drives total spending.

Figure 1:
Decomposition of change in educational expenditure 1997-2007 (billion euro)



2. Decomposing spending on education

The model starts by a simple breakdown into two explanatory factors: the number of participants taking part in education and the expenditure per participant. These two factors are further decomposed into various underlying indicators. Figure 1 shows all factors and specifies what they have attributed to the total change of spending on education in the period 1997-2007.

The first factor underlying the effect of the number of participants is the demographic change. This accounts for changes in the population numbers per age category during the period investigated. It shows the effect on expenditure because the potential number of participants is increasing or decreasing. The participation rate shows how the potential number of participants translates into actual participants in education. The participation factor is divided further into two parts: absolute and relative participation. Absolute participation shows the effect on expenditure caused by an increase or decrease in percentages in the demand for education in general per age category. Relative participation shows the effect on expenditure caused by changing participation between the different types or levels of education.

In the expenditure per participant two kinds of expenditure are distinguished: the material and the personnel expenditure per participant. Material expenditure refers to investments and current expenditure such as rent and energy costs. The effect of personnel expenditure in the decomposition model depends on the number of employees in education relative to the number of participants and on the average labour costs per employee. The labour costs are further decomposed into the effects caused by changes in wages and changes in other employer's costs. In the wages we distinguish

the effects of collective labour agreements (CLA) and other effects on wages.

The results depicted in figure 1 show that all factors investigated have contributed to the real growth of the expenditure on education. The participation effect gave a substantial impulse due to the increased demand for education in the Netherlands. Almost all components of the material expenditure per participant increased, but two thirds of the total change occurred in secondary education, mainly because of the increased current expenditure in general secondary education. The rising costs per employee (FTE) also played a significant role. This is mainly due to the increased social security costs per employee, of which pension premiums are the most important component.

3. Results per level of education

The calculations presented under subheading 2 show the results for the Dutch educational system as a whole. The decomposition method also allows investigating the trends within the various levels of education. When primary education, secondary education and tertiary education are analysed individually, the results differ from the general view, especially in tertiary level education.

3.1 Primary education

Primary education has contributed 2.9 billion euro to the spending increase. This was caused almost entirely by the increased expenditure per participant. A reduction of the average class size at this level of education resulted in more jobs for teachers and classroom assistants. The effect is visible in the substantial growth of the employee / participant ratio.

3.2 Secondary education

Secondary education has contributed 4.3 billion euro to the increased expenditure on education. Three quarters of this is due to increased expenditure per participant. The material expenditure per participant has increased more than in primary education, mainly because the spending on buildings, hardware, educational tools, rent and energy doubled. Demographic changes played a significant role in general secondary education, while vocational secondary education showed the effects of increased participation. The employee / participant ratio increased mainly because of the successful efforts to reduce the number of dropouts, resulting in more expenditure.

3.3 Tertiary education

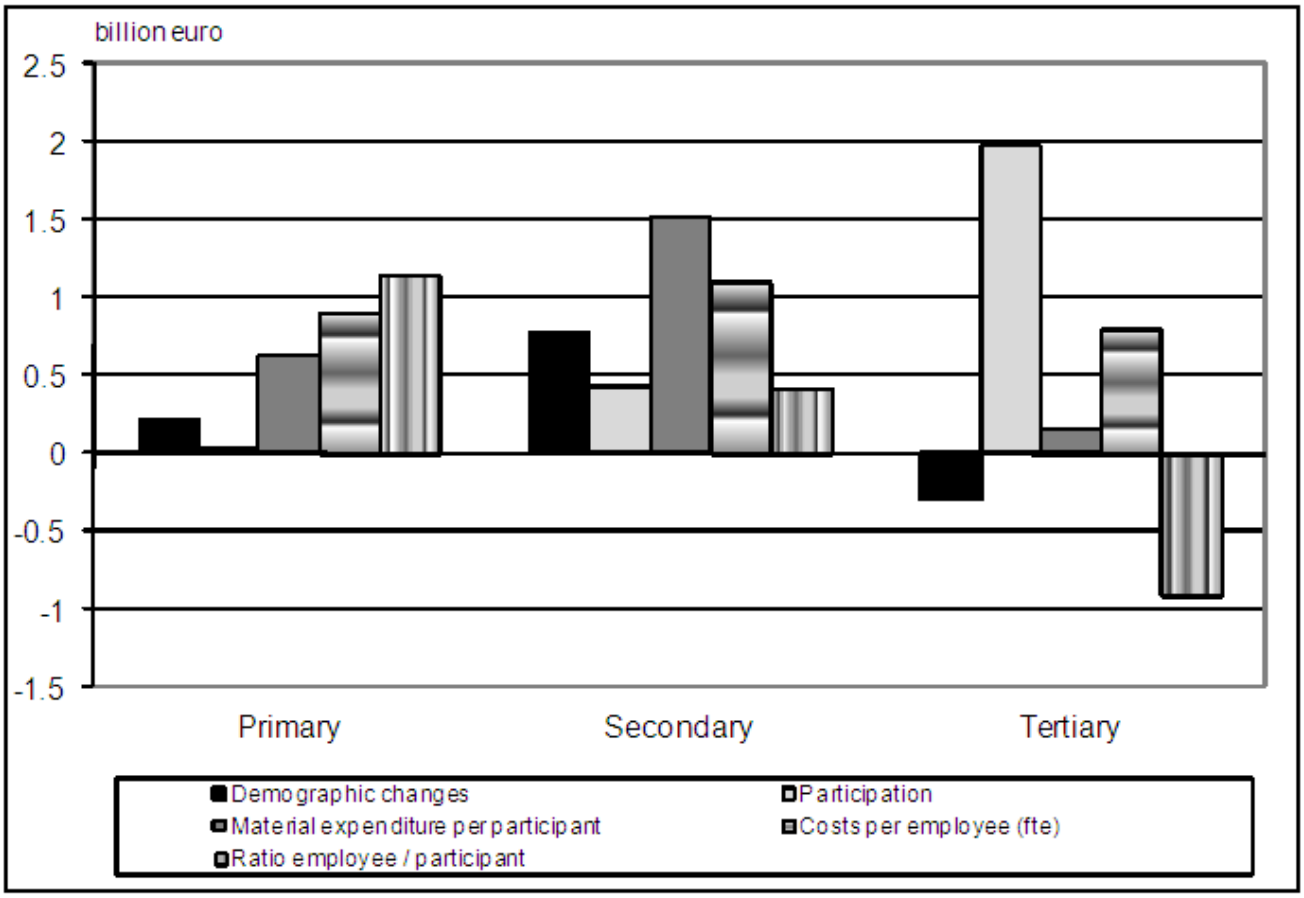
Tertiary education has contributed 1.7 billion euro to the increase in spending on education. Here, the main driving force was the significant rise in student number because the participation rate at this level of education has grown considerably. In the lower education levels the employment expressed in full-time equivalents has grown relatively more than the number of participants. In tertiary education, however, the effect was negative, because the increase in student numbers was not matched by an increase in teaching staff.

For further information:

http://www.cbs.nl/en-GB/menu/themas/overheid-politiek/publicaties/artikelen/archief/2010/2010-educational-expenditures-pub.htm

*Hugo Elbers and Remco Kaashoek both work for Statistics Netherlands.

Chart 1:
Decomposition of change in educational expenditure 1997-2007 by level of education



EUROPEAN CENTRAL BANK

NEW ECB STATISTICS ON EURO AREA INVESTMENT FUNDS

Jani Matilainen and Barbara Zupancic*

Background

Non-bank financial intermediaries have substantially grown in the euro area over the past two decades in terms of financial assets [See Gonnard, E, Kim, E.J. and Ynesta, I. (2008), Recent Trends in Institutional Investors Statistics. Financial Market Trends, OECD 2008.]. Their role in the funding of non-financial corporations is also increasing. As a consequence, non-bank financial intermediaries' balance sheet statistics are becoming more relevant for monetary analysis and the other tasks of the European System of Central Banks (ESCB). This is also true in view of the steadily growing inter-linkages between the various financial intermediaries.

The European Central Bank (ECB) and the European System of Central Banks (ESCB) have developed three sets of statistics for some of the most important sub-sectors of non-bank financial intermediaries (investment funds, insurance corporations and pension funds, and financial vehicle corporations engaged in securitisation transactions) in order to improve their statistical coverage.

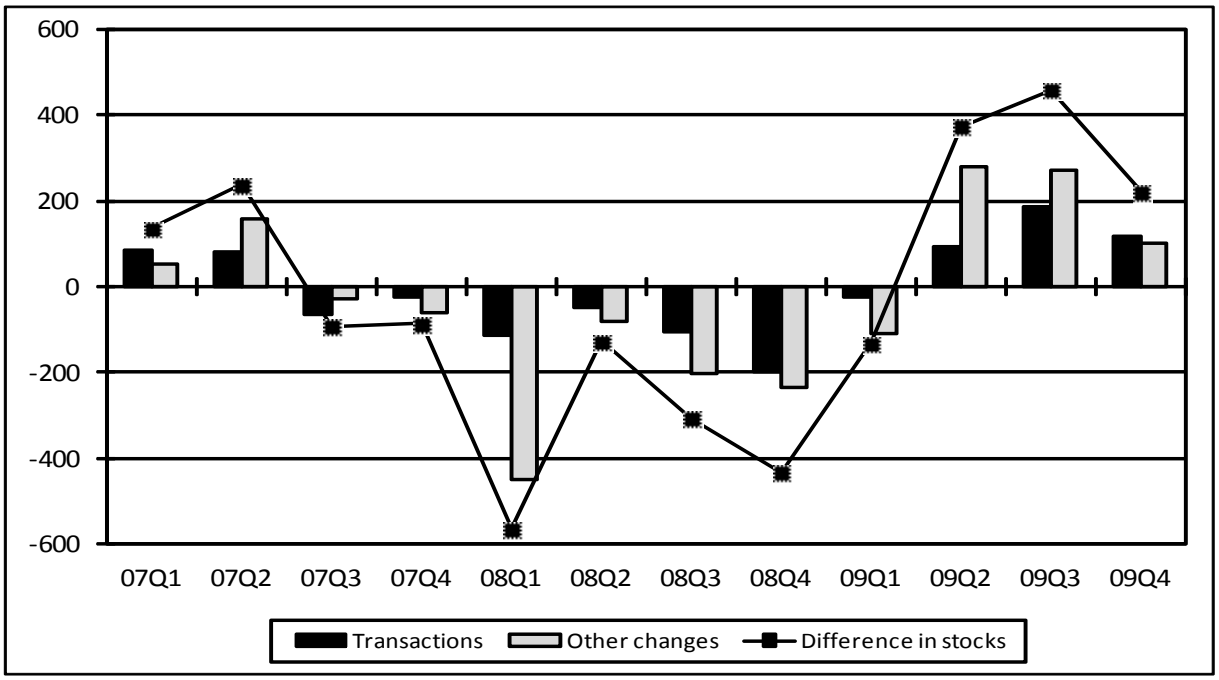
This article presents the new harmonised statistics on the assets and liabilities of investment funds (IFs) resident in the euro area, which the ECB released for the first time in December 2009. These statistics do not per se cover money market funds (MMFs), which form part of the monetary financial institutions (MFI) sector. Data on euro area MMF balance sheets are compiled every quarter. More details on MMFs can be found

Table 1:
Summary of published euro area aggregates for IF statistics

		Further breakdowns			
	Total	**Geographical[1]**	**Sector[2]**	**Maturity[3]**	**Currency[4]**
Assets					
Deposit and loan claims	Q and M	Q	Q		
Securities other than shares	Q and M	Q and M	Q and M	Q and M	Q and M
Shares and other equity	Q and M	Q	Q		
o/w quoted shares	Q	Q	Q		
o/w IF and MMF shares/units	Q and M	Q and M	Q and M		
Non-financial assets	Q and M	Q			
Remaining assets	Q and M				
Liabilities					
Investment fund shares/units	M*	Q and M			
Loans and deposits received	Q and M				
Remaining liabilities	Q and M				

Q: quarterly stocks and transactions; M: monthly stocks; * including transactions.
1) Domestic; Other euro area Member States (other than domestic); Rest of the World (in the case of securities further broken down into issuers resident in non-euro area EU Member States, the United States and Japan).
2) For domestic and euro area counterparts: MFIs, general government, other financial intermediaries, insurance corporations and pension funds, non-financial corporations, and households.
3) up to one year; one to two years; over two years.
4) euro; total foreign currencies.

Chart 1. Euro area investment funds shares/units issued



in the statistics section of the ECB's website at www.ecb.europa.eu. These statistics replace the non-harmonised euro area IF statistics that were previously published by the ECB on a quarterly basis. The new statistics provide fully harmonised, complete and timely information at a monthly frequency.

Overview of the new euro area investment fund statistics

The reporting population covers IFs resident in the euro area, including those managed from outside the euro area. Conversely, the statistics do not cover IFs established outside the euro area, even if they are operated by management companies resident in the euro area.

IFs resident in a particular euro area country report data to their respective national central bank (NCB). Regulation (EC) No 958/2007 of the European Central Bank of 27 July 2007 concerning statistics on the assets and liabilities of investment funds (ECB/2007/8). IFs may, subject to the prior approval of the relevant NCB, report their securities portfolio on a security-by-security basis, instead of on an aggregated basis. The NCBs then

Chart 2:
Euro Area Investment Fund Shares/Units Issued by Country (end-2009)

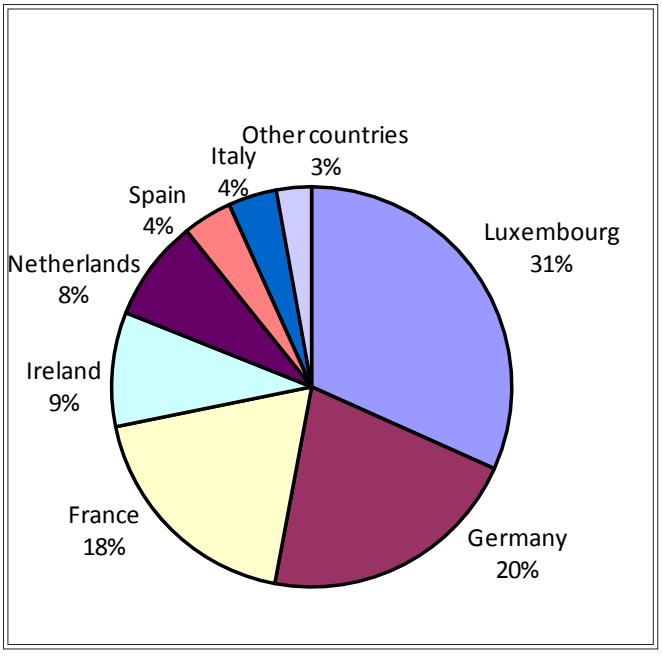
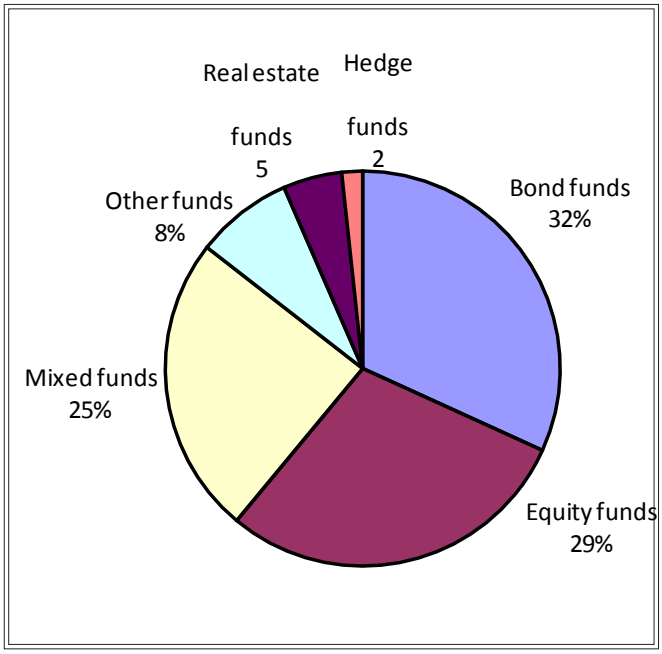


Chart 3:
Euro Area Investment Fund Shares/Units Issued by Type of Funds (end-2009)



compile the necessary aggregates using a common ESCB securities reference database (the Centralised Securities Database), or a compatible local securities database. This reduces the costs for reporting agents and gives more flexibility in aggregating the data. The NCBs compile nationally aggregated statistics and send them to the ECB. Guideline of the ECB of 1 August 2007 on monetary, financial institutions and markets statistics (recast) (ECB/2007/9). The ECB then produces euro area aggregates.

The data published by the ECB consist of statistics on various categories of assets held and liabilities issued by IFs resident in the euro area, broken down by residency and economic sector of the counterpart, by maturity and by currency.

The data cover key variables at a monthly frequency, while more detailed data are provided at a quarterly frequency. In particular, the quarterly dataset contains financial transactions for each asset and liability item, whereas the monthly dataset is mainly based on stock data and transactions are only available for the shares/units issued by IFs. Work is ongoing in the NCBs and further information, such as more detailed monthly transaction data, may become available in the future.

Table 1 summarises the euro area aggregates that are published on a monthly and quarterly basis.

Euro area aggregates for the assets and liabilities of IFs are available for total funds and for the following six sub-categories: equity funds, bond funds, mixed funds, real estate funds, hedge funds and other funds (i.e. breakdown by investment policy). Each sub-sector is further broken down into open and closed-end funds (i.e. by type of fund). The classification criteria for different types of IF are set out in Guideline ECB/2007/9, with further guidance in the Manual on investment fund statistics, May 2009, available on the ECB's website.

The statistics are published with a time lag of around one and a half months after the reference period. The harmonised time series date back to the reference period December 2008. Further non-harmonised quarterly euro area IF statistics are available for the reference periods 1998Q4 to 2009Q2. However, the coverage and definition of the new statistics differ from the previously published IF statistics. Therefore, a direct comparison of the new and old statistics is not always possible. Estimated longer historical time series consistent with the new statistics will be released in due course.

Some characteristics of the euro area investment fund industry

The value of euro area IFs' total outstanding shares/units stood at EUR 4,959 billion at the end of 2009, up 23% from a year earlier when it stood at EUR 4,038 billion. This development reflects both increases in the prices of assets held by euro area IFs, accounting for around 14 percentage points of the increase, and net issuances of IF shares/units, representing around 9 percentage points. Significant net issuances were observed in the last three quarters of 2009 (see Chart 1). These include sizeable, specific purchases of investment fund shares by two Dutch pension funds in June and July 2009. However, the net issuances remain positive even without taking these operations into account. This suggests a recovery in the euro area IF industry, following a prolonged period of net redemptions in IF shares/units that started in the third quarter of 2007 and lasted until the first quarter of 2009.

The countries with the largest IF sector, in terms of IF shares/units outstanding, are Luxembourg, Germany and France, which together accounted for 69% of the total euro area at the end of 2009 (see Chart 2). They are followed by Ireland and the Netherlands, adding a further 17% and bringing these five countries to 86% of the euro area total.

Breaking down the issuance of euro area IF shares/units by investment policy shows that bond, equity and mixed funds accounted for a total of 86% of IF shares/units at the end of 2009 (see Chart 3). Hedge funds accounted for a mere 2% of total IF shares/units. As mentioned above, a hedge fund resident outside the euro area does not form part of the reporting population, even if it is managed from within the euro area, which is often the case in this industry.

As regards the most relevant types of asset, euro area IFs held EUR 2,379 billion worth of shares and other equity (including shares/units issued by IFs and money market funds) and EUR 2,076 billion worth of debt securities at the end of 2009, which accounted for 44% and 39% of total assets respectively.

The geographical breakdown of the issuers shows that 57% of the shares and other equity held by euro area IFs were issued by euro area residents, mainly other IFs and non-financial corporations, with both accounting for around 40% (see Charts 4 and 5). Of the holdings of shares and other equity issued by residents outside the euro area, 44% were issued by the United States and non-euro area EU Member States.

In the case of debt securities, 68% of those held by euro area IFs were issued by euro area residents. Approximately half were government bonds and more than a quarter bank bonds. Bonds issued by other financial institutions and non-financial corporations accounted for another quarter (see Charts 6 and 7). Two-thirds of the debt securities held by euro area IFs and issued by residents outside the euro area were issued by residents in the United States and non-euro area EU Member States.

Dissemination of investment fund statistics

The euro area aggregates are published in regular monthly and quarterly ECB statistical press releases. The quarterly press releases contain detailed information on the assets held and liabilities issued by IFs. The monthly press releases show data on total IF shares/units issued for the different types of fund. The press releases also include data on euro area money market fund balance sheets, which form part of the monetary financial institutions (MFI) population. These enable users to perform analyses on the entire fund industry.

A more detailed set of euro area data, together with explanatory notes, is released to the public in the "Statistics" section of the ECB's website: www.ecb.europa.eu. The Statistical Data Warehouse – the ECB's online data delivery service for statistics – can be consulted on the website: http://sdw.ecb.europa.eu/. The updated statistics are also published in the ECB's Monthly Bulletin, tables 2.9 and 2.10 in the "Euro area statistics" section.

In addition to the euro area aggregates, national data on IF shares/units issued broken down by investment policy and by type of fund are available on the ECB's website. Further national results are published by the NCBs as appropriate.

In parallel to the new statistics for the euro area, several non-euro area EU Member States are implementing the same new statistical standards for collecting and compiling IF statistics as part of their preparation for future membership of the euro area.

*Barbara Zupancic is Senior Economist-Statistician and Jani Matilainen is Economist-Statistician, European Central Bank.

Chart 4:
Geographical breakdown of Euro area IF holdings of shares and other equity (end-2009)

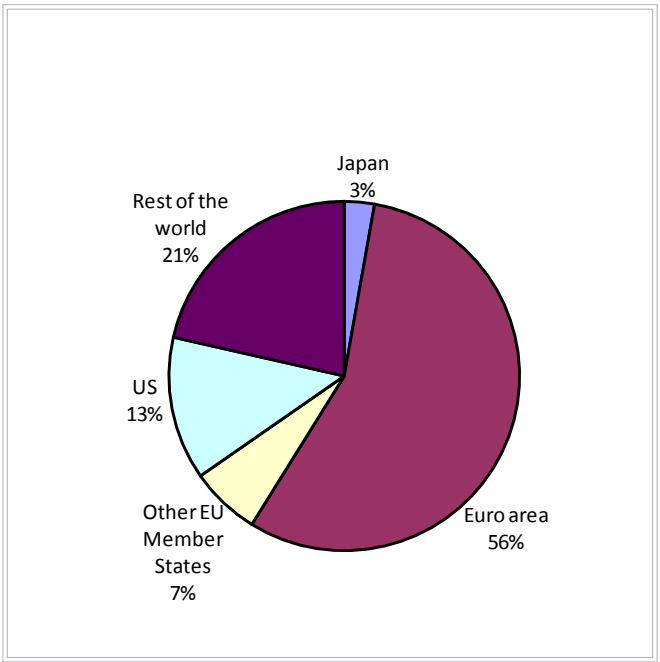


Chart 5:
Sectoral breakdown of Euro area IF holdings of shares and other equity issued in Euro area (end-2009)

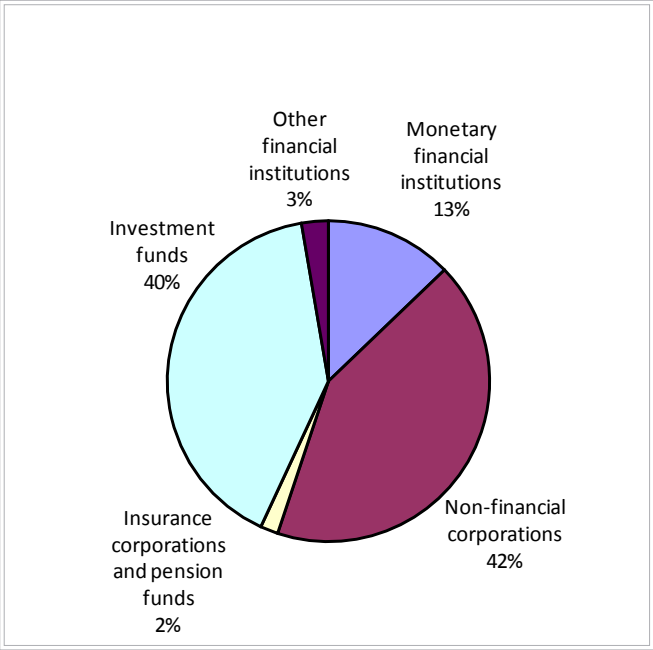


Chart 6:
Geographical breakdown of Euro area IF holdings of securities other than shares (end-2009)

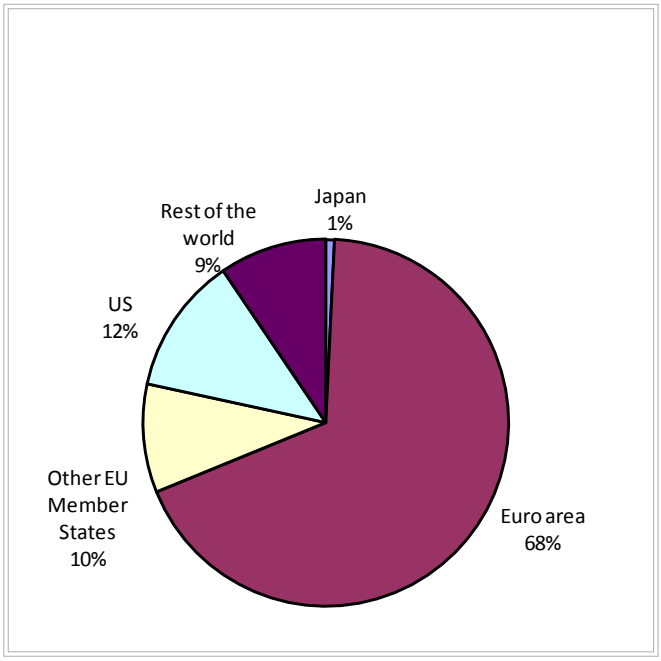
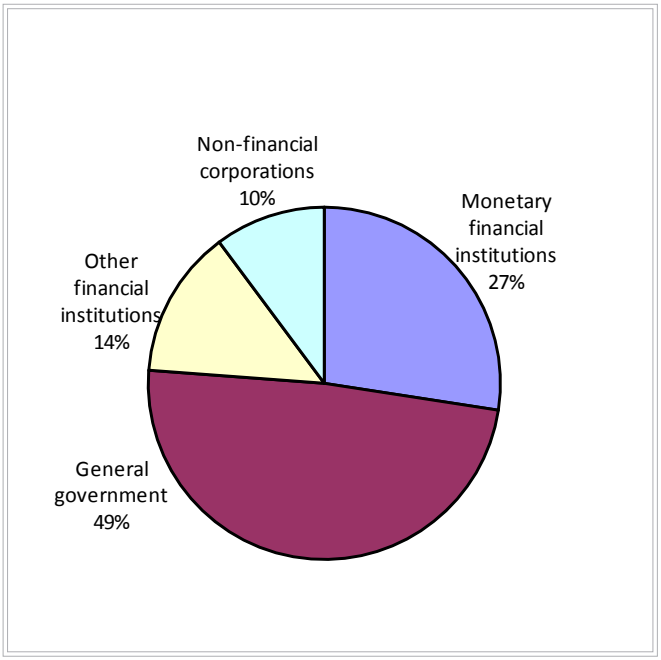


Chart 7:
Sectoral breakdown of Euro area IF holdings of securities other than shares issued in Euro area (end-2009)



THE GLOBAL PROJECT FOR MEASURING THE PROGRESS OF SOCIETIES

WIKIPROGRESS

Angela Hariche*, OECD

The Global Project for Measuring the Progress of Societies fosters the development of economic, social and environmental indicators to provide a comprehensive picture of how the well-being of a society is evolving. It also seeks to encourage the use of indicators to inform and promote evidence-based decision-making, within and across the public, private and citizen sectors. The project is open to all sectors of society and particularly invites members of civil society to get involved, building both on good practice and innovative research work.

In October 2009 at the The 3rd OECD World Forum on "Statistics, Knowledge and Policy" Busan Korea, the Global Project announced the launch of www.wikiprogress.org, which provides a unique forum for sharing data and analysis on the measurement of environmental, social and economic well-being and communicating these tools to policy makers and the wider public. Developing better indicators and awareness is not enough. We also need to ensure that the measures become widely used, and widely understood – not just by statisticians, but by all those interested in societal progress.

We invite the OECD statistics community to go to wikiprogress and become a part of the community for societal progress by sharing in initiatives, best practices and data. Wikiprogress.stat has also just been launched which is bringing together progress-related data into one place. Examples of such data currently include The Happy Planet Index and the Global Peace Index.

Friends of the wiki

The Friends of the wiki are an informal advisory board who are closely related to the wiki and advise on data, editorial, quality assurance, multimedia and IT issues.

Recent contributions by the friends include:

- Ruut Veenhoven, Director of the World Database of Happiness, contributed a full country study on happiness in 159 countries. See Happiness in Argentina.

> "And the OECD has firmly embraced the web 2.0 revolution. I am pleased to announce the launch of Wikiprogress, a website especially designed to share and undertake analyses of progress, and disseminate this knowledge worldwide".
>
> Remarks by Angel Gurría, OECD Secretary-General, during the OECD World Forum, Busan, Korea, October 2009

- Gene Shackman, Ph.D., Applied Sociologist for The Global Social Change Research Project contributed a dataset to wikiprogress.stat called is "Is life getting better" which includes a series of indicators from different sources for the measurement of progress. The project is also preparing a series of pamphlets which discusses measuring progress, and the issues involved in measuring progress.
- Anat Itay, Project Leader of Israel's Progress Index, contributed an article on Israel's Progress Index, which is the leading project of the Israeli Society for Sustainable Economics, and marks its development to national economic-environmental-social decision making. The goal of the project is to establish an index for progress and quality of life (QOL), and promote it as a central tool for decision making in Israel.

Wiki statistics (as of 29 March, 2010)

# of registered editors – 210
# of articles – 489
# of edits in articles – 6766
# of visits in March – 2515 with 15,205 page views
# of data sets - 6

Call for country studies

We are looking to improve the articles on Italy, France and India. If you know of indices or other initiatives that are measuring progress in these countries or would simply like to improve the page, please log in (or register) and start editing. The administrators and the friends of the wiki will do a light quality assurance process after you have edited.

We are actively looking for new community members in the movement for progress in terms of editors, data sharing, multimedia, quality assurance, etc.

If you would like more information, please feel free to contact angela.hariche@oecd.org

*Angela Hariche, Project Manager, OECD.

STATISTICS SWEDEN

TOWARDS EFFICIENT STATISTICAL DATA EDITING: THE SWEDISH EXPERIENCE

Anders Norberg*

Statistical data editing is a resource-demanding process in business surveys. A 2004 study at Statistics Sweden demonstrated that around one third of resources were spent on editing (somewhat more for annual and periodic surveys than for monthly and quarterly surveys). Most resources were spent on the traditional editing of micro data.

The use of web-questionnaires makes it possible to include some form of editing for respondents at the point of data capture. In fact, many respondents today expect to meet "intelligent" communication via the web. So far, most such systems lack techniques to store process data (paradata) from the response process. Output (macro) editing is another sub-process that has the potential to be improved and to be more important. Output editing can detect errors introduced in the production and compilation processes. When resources can be released from the large micro editing process, some of these same resources should be invested into these two types of editing.

The role of editing

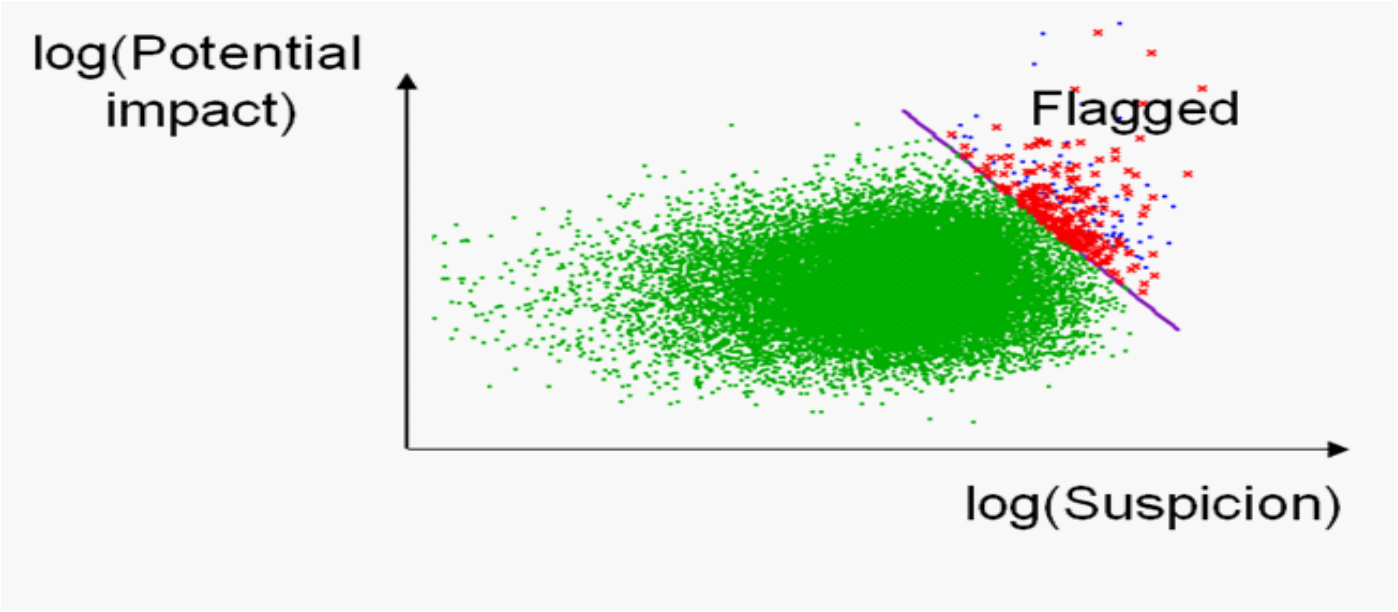
A new role of editing is slowly being implemented at statistical institutes. Its focus is on collecting process data on problem areas and causes of errors in the measurement process. These data will provide a basis for a continuous improvement of the measurement process and the whole survey vehicle in general. The old paradigm, "…the more and tighter the edit checks and re-contacts, the better the quality", should be replaced [Granquist, L. (1997) "The New View on Editing" International Statistical Review]. However, when editing primarily is quality control of the measurement process, it is still needed to adjust (change/correct) significant errors in the current survey round and to contribute to quality declarations.

The role of the query edit checks should be designed to focus on errors influencing the estimates. The effects of the edit checks should be continuously evaluated by analysis of performance measures, which the editing process should be designed to produce. The software SNOWDON-X developed jointly by the UK Office for National Statistics and Southampton University of the UK, is a good tool for this.

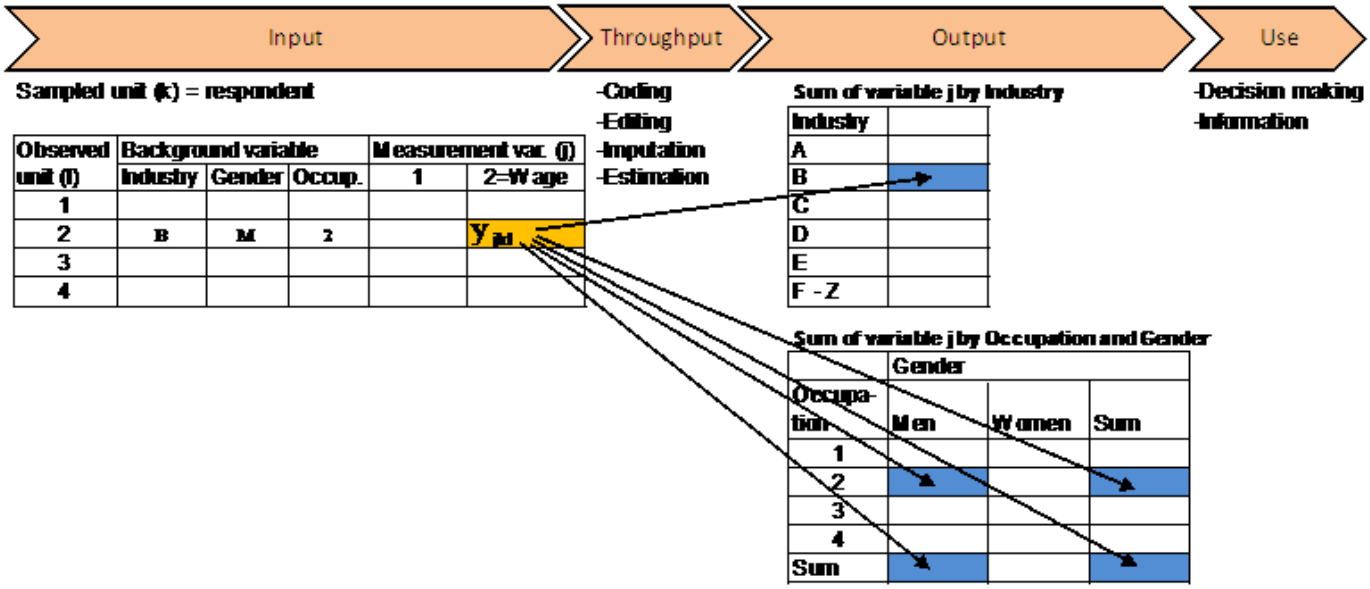
Editing staff debriefing is a method for collecting and analysing information on problems the respondents report from notes and contacts. At Statistics Sweden editing staff working on a particular survey meet and discuss their experiences in presence of a moderator from the Unit for Cognitive Methods. The purpose is to find out how the respondents understand the questions, which questions are problematic and what kind of error indicators are turning up in the editing process. As such, although the editing staff debriefings are a qualitative technique in nature, they can provide ideas about how common certain problems arise.

Good methods, generic IT-tools and a structured collection and analysis of process data have the potential to give better meaning to the editing job and provide for a better working environment for the editing staff. This manpower-demanding work will be easier to plan for in the business surveys area

Simultaneous significance editing of foreign trade statistics by suspicion and potential impact



Input data, production of statistics and use of statistics



as staff can work with and on several surveys collectively once they are well acquainted with the new IT editing tools. The editing staff will not have to contact or re-contact as many respondents (some annoyed) that have considered their delivered data (questionnaire) to be correct in the first place. This will happen when high hit rates of edit checks become part of the quality process.

Selective / significance editing

Traditional edit checks often only focus on "suspicion" towards a unit's value for a single variable. Flagged data are suspected whereas un-flagged data are accepted. There is a dichotomisation (in other words, the data have been divided into two opposing groups) of suspicion. Selective editing is a procedure which targets only some of the flagged variables or records that failed at least one edit check for manual review. This selection is based on the potential impact on estimates from the suspected error.

We see the dichotomisation as a waste of information when it is possible to measure suspicion on a continuous scale. The suspicion grows with the distance from the expected a priori distribution of the un-edited variable value. Suspicion and potential impacts can be treated simultaneously to form a score in significance editing. In foreign trade statistics the statistics produced are values of imports and exports and consequently the potential impact of a suspected error is expressed as an error in transaction value. Suspicion for the record is based on price per quantity, and as these two have a poor correlation both dimensions are important.

Respondents, producers and customers

One erroneous input data value can have an impact on several output statistical values. This is so when output is spread by more than one variable, for example when wages are presented by industrial sector, gender and occupation. Here, as in design in general, it is necessary that the national statistics institute can assess the quality demands of each output table from the users point of view.

Suspicion on a data value yjkl can be estimated by a variety of robust methods and from the saved edited "cold" data. The potential impact on statistical output, if input data is erroneous, is the difference between the received data value and an expected/ predicted value, weighted according to the estimation formula.

We have adopted the concept of relative pseudobias (RPB) to evaluate the quality of editing. This bias of an estimate is due to the follow-up of only a selected subset of input data (assuming that there are some errors left in the output data): the bias is analysed relative to the standard error of the estimate. A 20 percent RPB has little contribution to the total error in most statistics.

Generic tools

Tests of selective/significance editing were performed for nine of the most editing intensive business surveys in 2007 at Statistics Sweden. We saw likely efficiency gains and likely cost reductions. We also realised that the introduction of new methods demand intensive testing in every specific survey because of the variation between the surveys regarding data structure, users demands of the statistics, etc. Generic tools for editing must therefore be very flexible to be able to deal with these different situations.

The method and the IT tools for flagging of incorrect or suspected data values through traditional, selective and significance editing at Statistics Sweden is called SELEKT. Necessary parameters, several of these can be set to the default values, are stored in a table with the module PRE-SELEKT and need to be maintained on a regular basis. PRE-SELEKT also computes expected/ predicted values and measures of variation on cold deck data to be used in the edits. AUTO-SELEKT calculates scores according to the parameter table, indicating the expected impact on all important output. In a laboratory environment and supported by modules in PRE- and AUTO-SELEKT, tests will be undertaken before implementation and up-dates to surveys to evaluate the earlier production rounds.

Expected/predicted values and variation are computed for homogenous groups. These may, but need not, correspond to strata or domains of study. In SELEKT, the groups can be formed by a set of auxiliary variables, the detail of classification (number of digits) and a fixed minimum number of observations

required for the computation. Estimation of totals, functions of totals and their estimated standard errors is done by a generic tool, outside the 'selekt' family. Several different types of software can undertake this calculation, Statistics Sweden uses its own software: CLAN.

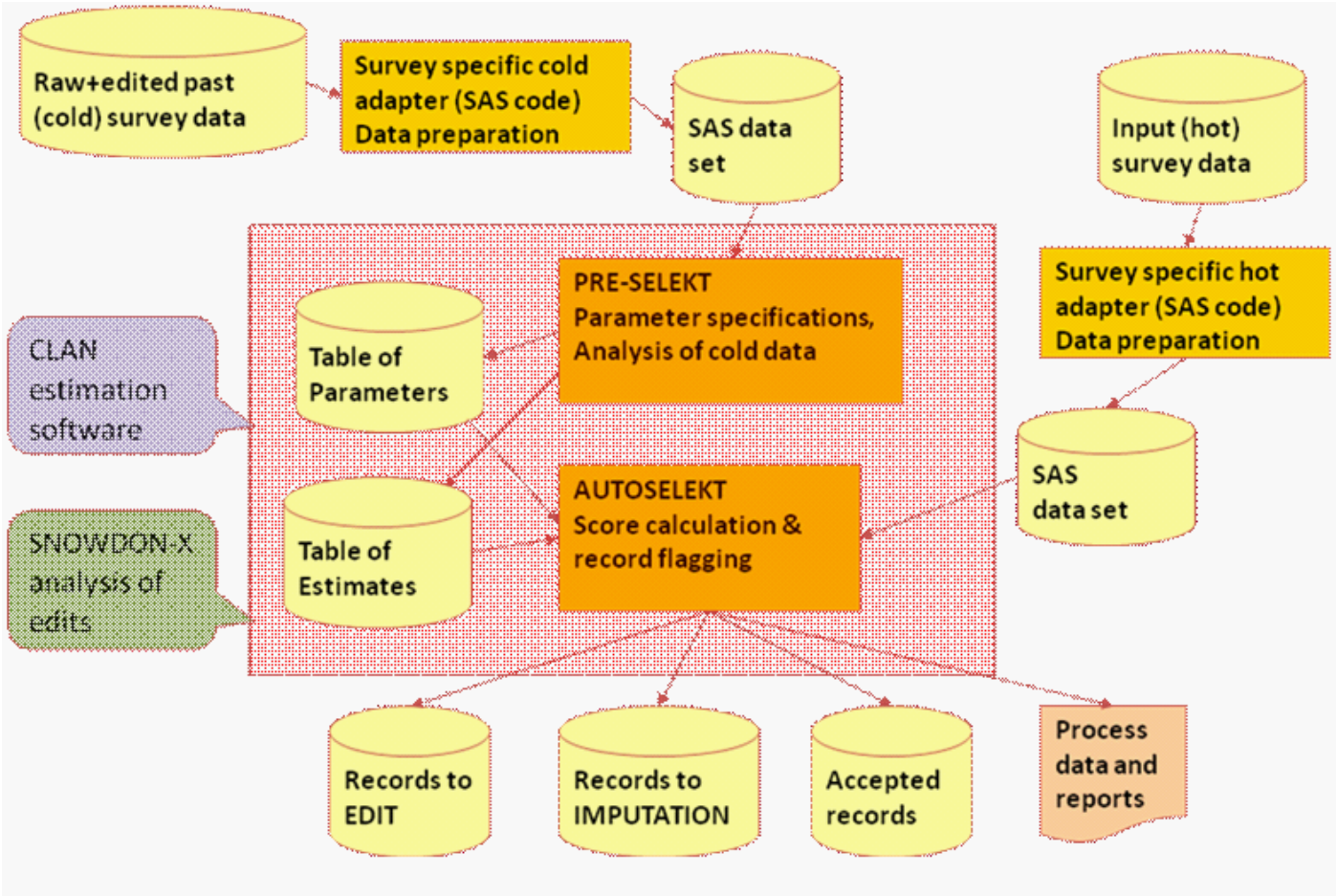
EDIT is the tool editing staff use to follow-up flagged items. To be generic it must be flexible for different types of survey data. In this sense EDIT will have a standard interface, a windows look with a lot of tabs, functionality that presents all of the information needed such as previous data and analysis thereof, register data look-up ability, etc. It must also be possible to ask SELEKT to check a specific batch of data and this needs to be undertaken quickly.

Process data are generated in an ongoing process. They can be used both for continuous monitoring and for analysis and evaluation in order to improve the production cycle and reach an optimal resource allocation.

The first version of the SELEKT software is in place to begin the implemented phase. Prototype versions have been implemented and tested in a few surveys to date. Experience will bring us forward to efficient editing. A new project, TRITON, aiming to make all tools for several current processes in data capture and data processing communicate with others has just started at Statistics Sweden.

*Anders Norberg is a Senior Statistician at Statistics Sweden.

Data flow and software



STATISTICS CANADA

SEASONAL ADJUSTMENT AND IDENTIFYING ECONOMIC TRENDS

Diana Wyman*

Overview

There has been heightened interest in understanding the movements in monthly and quarterly economic data as a result of the recent recession. This makes a review of the basics and the subtler aspects of seasonal adjustment both timely and relevant for analysts who are grappling with the repetitive ups and downs and added volatility of the unadjusted data. This article aims to explain seasonal adjustment and highlight how seasonally adjusted data help analysts home in on the underlying economic trend.

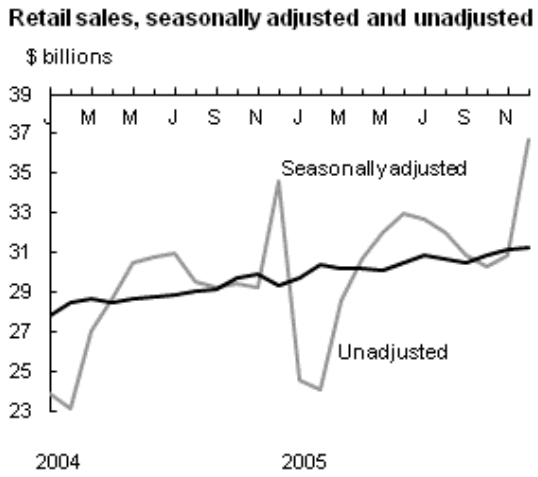
Most examples will use Canadian data. Given the four distinct seasons and highly variable weather in Canada, accounting for "typical weather patterns" means that seasonal swings are more pronounced in Canada than other major countries. Because of the importance of seasonality in Canadian data, Statistics Canada has been a leading force in the statistical community in the development of seasonal adjustment.

The Basics

The unadjusted data can be broken down into four elements: the long-term trend (T), the cyclical component (C), the irregular component (I) and the seasonal component (S). Of these, the cyclical component, or business cycle, is the focus of current economic analysts, as it is movements in this component that show whether the economy is contracting or expanding and at what rate. The long-term trend is a relatively stable component, changing gradually to reflect phenomenon such as demographic change. Because it has little effect on short-term movements, it is grouped with the cyclical component to form what is called the trend-cycle. The two remaining components are those which obscure the current status of the trend-cycle – the seasonal and irregular components.

The seasonal component refers to the recurring fluctuations in economic data that occur as a result of "average" weather, holidays and calendar norms within a year. Retail sales are one example of a series that exhibits a clear seasonal pattern. In the unadjusted retail sales data shown in Figure 1, January sales fall nearly one-quarter on average after December and then drop a bit further to their yearly low in February. Retail sales then rise throughout the spring and summer as milder weather encourages shoppers and drivers to venture out. Sales dip as cool weather returns, before shooting upward in November and December when holiday shopping is in full swing.

Figure 1



The irregular component encompasses the random element introduced by unexpected events. One recent example of an irregular component occurred as a result of the Toronto municipal strike in July 2009. The stoppage in the approval of building permits led to a large decline in the number of permits issued in July and a subsequent recovery in the autumn as the city caught up with the backlog of applications. More frequently, unscheduled maintenance, labour strikes, or atypical weather are reflected in this component.

Adjusting for seasonality

The concept of "seasonality" is not limited to the effect of changing seasons. It encompasses any regularly recurring movements in the unadjusted data. There are three parts to seasonal patterns: the climatic (weather-related) effect; the institutional component; and the calendar effect. Only the first part, the climatic effect, reflects the change in economic activity resulting from recurring weather changes of the four seasons; for example, the drop in agriculture following the autumn harvest. The second part, the institutional component, refers to industry-specific norms, such as regular annual vacations and scheduled shutdowns, and statutory holidays. The impact of Christmas, which fuels retail sales through November and December, is one example.

The third part of seasonality, the calendar effect, relates to the composition of the calendar. The main calendar effect is the trading day effect, which refers to the

Figure 2a

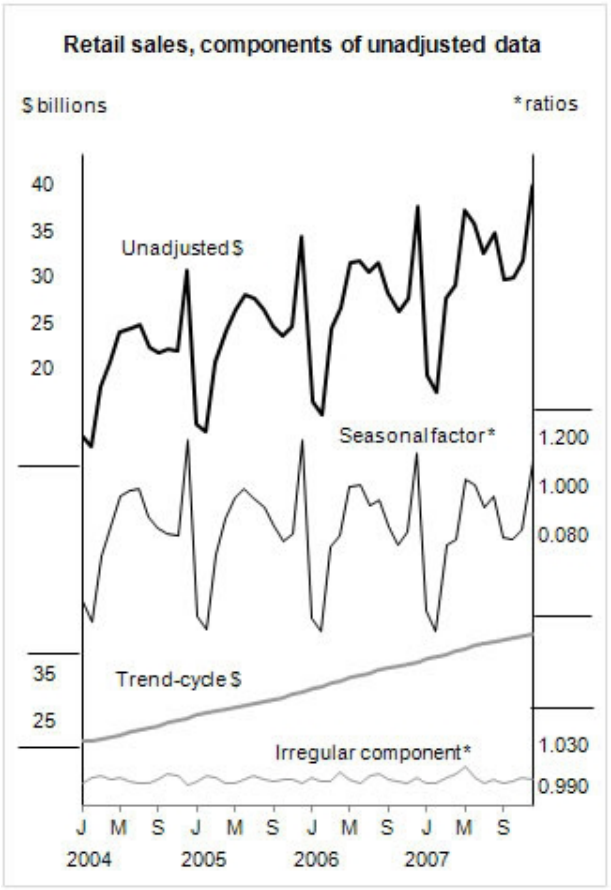
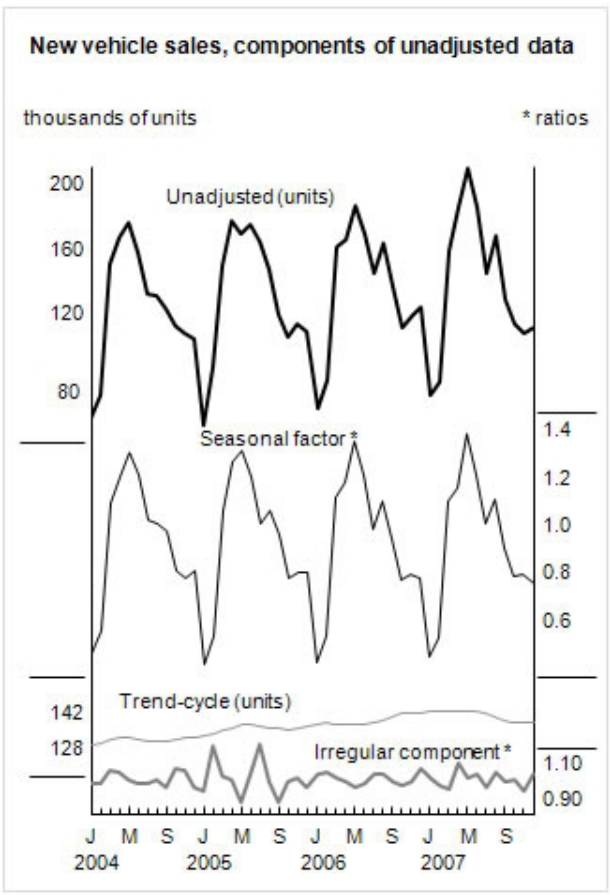


Figure 2b



variation in the number of days in a month (for example, unadjusted retail sales data are lower for February relative to the other months, which have 30 or 31 days) and the distribution of particular week days in a month. For example, retail sales are highest on Saturdays. A month of unadjusted retail sales data with five Saturdays as opposed to four may seem to indicate that there has been an improvement in sales even though demand was unchanged. The switching of Easter between March and April is another calendar effect, shifting sales and output between the two months.

Seasonally adjusted data are the unadjusted (or 'raw' in the lingo of statisticians) data with an estimate of these regularly recurring effects removed. This estimate is based on the patterns observed in the past. Seasonal adjustment entails estimating the usual monthly seasonal pattern, which is reflected in the seasonal factor. The seasonal factor is expressed as a ratio of the monthly seasonally adjusted data and the monthly unadjusted data for that month. This is the case only for series for which a multiplicative model of seasonal adjustment, the most common, applies. An additive model is used for series such as the change in inventories where negative numbers occur. Dividing the original series by this factor yields the seasonally adjusted series. Seasonal factors for a given month can be greater or less than 1.0. For example, retail sales in December are typically 15% higher than those in an average month, as indicated by its seasonal factor of 1.15. January's sales are below average, as indicated by its seasonal factor of 0.82, which shows that sales are usually about 18% less than an average month.

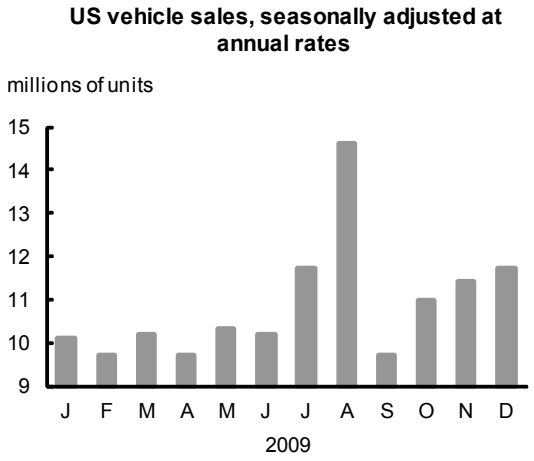
Seasonally adjustment and volatility

While the seasonal component is the most consistent source of volatility in the unadjusted data, there is a common misunderstanding that data that are seasonally adjusted reflect only the underlying trend. Seasonal adjustment does not remove all seasonal effects but rather the 'average' seasonal pattern determined by observing the past.

Shifting seasonality, which refers to the movement of a seasonal pattern into a new time period or a change in the magnitude of a seasonal pattern, is not addressed immediately by seasonal adjustment. One reason for this is that a new seasonal pattern may not obviously be a pattern until it recurs several times. Until then, shifting seasonality is reflected in the irregular component. The introduction of retail gift cards provides one example. As a result of their adoption as Christmas presents, gift cards dampened the rise of retail sales in December as well as the fall in January as the purchase of the gift card is included in retail sales when the gift card is redeemed.

Also, though there is a weather component to seasonality, not all weather-related changes are seasonal. For example, a prolonged heat wave may increase seasonally adjusted energy use as the seasonal factor reflects 'average' weather as determined by that observed in past years. Any divergence from this norm will remain in the seasonally adjusted data,

Figure 3



as part of the irregular component. In most months, the seasonally adjusted series is a good indicator of the movements in the trend-cycle. However, when there are large changes in the irregular component, the seasonally adjusted series diverges from the trend-cycle.

The irregular component is always present in seasonally adjusted data and can produce a large amount of volatility in any given month. The magnitude of the irregular component varies among different economic series. This is demonstrated in Figures 2a and 2b, which show the unadjusted data for monthly retail sales and new vehicle sales and their breakdown into their seasonal, irregular and trend-cycle components. Monthly (or quarterly) movements in the irregular component can even exceed changes in the trend-cycle as in the case of new motor vehicles, partly due to the recurring use of rebates and other incentives to purchase. In comparison, the irregular component of retail sales is relatively small, and its trend-cycle is much smoother.

In addition to the irregular component creating challenges for analysts, the trend-cycle itself can be a source of volatility. Since the economy contracted sharply in late 2008, data users have keenly followed short-term changes in the economy and found themselves wading through many volatile datasets. For example, unadjusted retail sales in December 2008 were below those of May 2008, a rare shift from the normal pattern of December sales being above those of all other months of the year by a sizeable margin. On average, between 2004 and 2007, retail sales were 13% higher in December than in May. One of the reasons for this volatility is that the seasonal pattern was overridden by the cyclical movement of a severe recession.

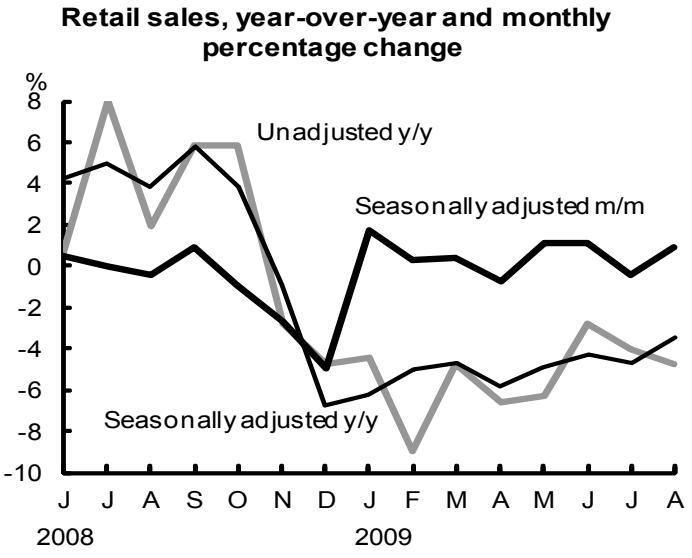
Year-over-year changes

Analysts have always searched for ways to cope with the recurring ups and downs in the unadjusted data as a first step in isolating the trend-cycle. One commonly used method is the same-month year-ago percentage change, which compares the same two months one year apart.

Using unadjusted data is useful in several instances such as when seasonally adjusted data are not yet released (one of the drawbacks of seasonally adjusted data is the time required to process and publish these estimates) but the unadjusted indicator is available early (such as daily or weekly changes in financial markets or company sales). Often, seasonally adjusted data are not available at all for the level of detail required, such as employment by occupation. Unadjusted data are also useful in trying to understand seasonal patterns. Some data are not adjusted for seasonality, either because the irregular pattern is so large that it overwhelms the seasonal pattern, which consequently cannot be identified, or because there is no seasonal pattern to that particular economic activity.

There are several weaknesses in using the year-over-year percentage change for unadjusted data. The main limitation is that the story it tells is an outdated one, a weakness that is especially important now given the current focus on identifying the point at which the economy may have moved from contraction to recovery. A forthcoming paper by Benoit Quenneville, "Gain and phase shift of the Annual Difference Operator" Statistics Canada, Working Paper of the Methodology Branch, demonstrates that the trend of the

Figure 4



year-over-year calculation is dominated by what happened, on average, six months earlier rather than what is taking place in the current month. By comparing the latest month to the same month one year ago, the year-over-year calculation is the sum of the twelve monthly changes during that year. For example, the economy in late 2008 and early 2009 contracted sharply. Since the spring of 2009, the Canadian economy has stabilized and begun improving, with moderate gains observed in most series. As a result, the same-month year-ago change throughout 2009 was a mixture of two very different trends.

Given that the central objective of most analysts is to identify the underlying trend-cycle, the year-over-year percentage change for unadjusted data does not perform well as an analytical tool. It does not account for the trading day effect, leaving the analyst wondering if the number or distribution of days is affecting the monthly or quarterly movement. Nor does it address the problem of an outlier month distorting the basis of comparison. The Olympic Games held in Vancouver in February is an example of an outlier month as it created a spike in international travel into British Columbia. When calculating year-over-year changes for February 2010, the increase over February 2009 is exaggerated by the strength in the current month and the decline occurring twelve months later is similarly overstated.

Seasonally adjusted data help address the problem of comparing with an outlier month. When comparing seasonally adjusted data, the analyst examines adjacent months. If there is an event in the previous month that created a large irregular movement, the analyst is more likely to observe this irregular component and discount it. If the size of the irregular component cannot be estimated, the outlier month can be ignored and the trend determined from comparing the current month with earlier months.

The 'cash for clunkers' program in the US, which allowed households with older cars to receive money toward a new vehicle purchased in July and August 2009, is one example of how the issue of an outlier month is partially addressed by using seasonally adjusted data. Some analysts speculated that sales would spike during the 'cash for clunkers' program and then sharply retreat as demand dried up. However, as shown in Figure 3, after a small dip in September, sales followed their pre-clunkers trajectory and by late in 2009 had grown above the pre-'clunkers' level. Ignoring July, August, and September therefore reveals the underlying trend of the seasonally adjusted series.

While the seasonally adjusted year-over-year change still offers an outdated view of recent trends relative to the month-to-month change, it does provide context to the monthly analysis by comparing it with the longer-term trend. The year-over-year calculation using seasonally adjusted data is preferable to the same-month year-ago change in unadjusted data, since the trading day adjustments are removed from the data prior to the calculation. As a result, the data are smoother and the trend is clearer, as the year-over-year change in the seasonally adjusted data is not obscured by the presence of trading day effects.

For example, in Figure 4 the decline in retail sales in late 2008 was revealed earlier in the seasonally adjusted year-over-year change and the decline is more gradual than in the unadjusted data. February 2009 was the weakest month for unadjusted retail sales, dropping 9%. In contrast, the year-over-year change in seasonally retail sales showed a decrease of 5% in February and the trend indicated that the rate of decline was slowing from its low of -7% in December 2008. The gradual improvement in year-over-year seasonally adjusted sales in January accords with the increase that began in monthly seasonally adjusted sales. The main reason for the larger decline in the unadjusted data was that 2008 was a leap year; moreover, this extra day was a Friday, which is the second busiest day of the week for retail sales. The seasonally adjusted data accounted for the effects of the leap year and the additional Friday as part of the overall calendar adjustment, which is why the largest drop in the year-over-year seasonally adjusted retail sales occurred in December 2008, not February 2009.

Conclusion

Seasonal patterns affect all of the major economic indicators. Because it removes recurring seasonal patterns, seasonal adjustment facilitates analysis by leaving only the irregular pattern to account for when focusing on the underlying cyclical trend. When the economy is steadily growing or contracting, seasonally adjusted data usually offer an easy-to-read update for analysts. Other times, such as when a major labour disruption occurs or a new holiday is instituted, there will be a large irregular component. Moreover, in turbulent economic times, there will be additional volatility in the trend-cycle, requiring analysts to be innovative.

*Diana Wyman is a senior economic analyst, Current Economic Analysis Group, Statistics Canada.

International Transport Forum

OECD

Redesigned

International Transport Forum statistical website

The following statistical products are available:

Statistics by topic
Quarterly data briefings analysing transport trends
Statistical papers including reports, analysis and studies
Global transports trends
Release calendar

JRC
EUROPEAN COMMISSION

ipsc
Institute for the Protection and Security of the Citizen

Training

Constructing Composite Indicators: From Theory to Practice

Unit of Econometrics and Applied Statistics

20-21 May 2010, Ispra, Italy

Objective and content

To provide participants with general guidelines for constructing composite indicators following the guidelines provided in the OECD (2008) Handbook. Topics covered:

- concept of composite indicator, strengths and weaknesses & domains of frequent use,
- development of the theoretical framework & selection of indicators,
- preparation of the data set (imputation and normalisation),
- multivariate analysis,
- weighting and aggregation methods,
- robustness and sensitivity analysis,
- presentation and dissemination aspects,
- tips and tricks

News in brief

The OECD ORBIS database: responding to the need for firm-level micro-data

Firm-level micro-data have become increasingly important in OECD activities not only for econometric analysis that captures heterogeneity across enterprises but also in order to conduct different and more detailed types of data aggregation (e.g. by geographical unit, firm size, industry).

This OECD working paper describes the development of the OECD ORBIS micro database which includes more than 200 variables providing financial and ownership information for over 44 million companies across the world. The paper provides an overview of the nature of these data, and of the advantages and disadvantages of such administrative micro-data for research. It describes the data included in the OECD ORBIS database and the cleaning procedures that have been undertaken to identify suspicious values.

For further information: http://www.olis.oecd.org/olis/2010doc.nsf/linkTo/std-doc(2010)1

Forthcoming - OECD Factbook 2010

OECD Factbook is the best-selling, annual title from the OECD. It provides a global overview of today's major economic, social and environmental indicators, in a range of user-friendly formats.

More countries than ever are covered in greater detail, enabling direct comparisons for many indicators between OECD Members and Brazil, Russian Federation, India, Indonesia, China and South Africa.

In addition, StatLinks below every table and graph contain further data in Excel™ about the other countries currently going through the OECD accession process – Chile, Estonia, Israel and Slovenia.

Each edition of OECD Factbook contains a special chapter - this year the focus is on The Crisis and Beyond.

Available mid-May at www.oecd.org/bookshop

Out soon

Are the New Millennium Learners Making the Grade?

Technology Use and Educational Performance in PISA 2006

Using data from PISA 2006, this book analyzes to what extent investments in technology enhance educational outcomes. One of the most striking findings is that the digital divide in education goes beyond the issue of access to technology. A new second form of digital divide has been identified: the one existing between those who have the right competencies to benefit from computer use, and those who do not. These competencies and skills are closely linked to the economic, cultural and social capital of the student. This finding has important implications for policy and practice. Governments should make an effort to clearly convey the message that computer use matters for the education of young people and do their best to engage teachers and schools in raising the frequency of computer use to a level that becomes relevant. If schools and teachers are really committed to the development of 21st century competencies, such an increase will happen naturally. And only in these circumstances will clear correlations between technology use and educational performance emerge.

TALIS 2008 Technical Report

The OECD's new Teaching and Learning International Survey (TALIS) has been designed to provide data and analyses on the conditions needed for effective teaching and learning in schools. As the first international survey with this focus, it seeks to fill important information gaps that have been identified at the national and international levels of education systems.

This TALIS Technical Report describes the development of the TALIS instruments and methods used in sampling, data collection, scaling and data analysis phases of the first round of the survey. It also explains the rigorous quality control programme that operated during the survey process, which included numerous partners and external experts from around the world.

Recent publications

All OECD publications can be ordered on line at: www.oecd.org/bookshop

Structural and Demographic Business Statistics 2009

This publication provides a summary of the information available in the OECD Structural and Demographic Business Statistics (SDBS) database. Key economic variables are presented at the International Standard of Industrial Classification, Revision 3 (ISIC) 2-digit level, for all industries and OECD member countries.

It provides a wealth of information at a very detailed sectoral level including: turnover, value-added, production, operating surplus, employment, labour costs and investment. The breakdown by industrial sector, including services, is supplemented by a further breakdown into business size classes, further enhancing the analytical capabilities of the database. The database also includes business demography statistics such as business birth, death and survival rates as well as the number of high-growth enterprises.

As such, the database and the publication will help answer such questions as: Which sectors have experienced positive/negative growth in recent years? What contribution do small businesses make to economic activity? How does the structure of businesses vary across OECD countries? Which industrial sectors have the highest labour productivity? How does labour productivity vary by business size? Are small and medium enterprises (SMEs) more or less profitable per employee than large businesses? Which sectors invest most?

Geographical Distribution of Financial Flows to Developing Countries 2010: Disbursements, Commitments, Country Indicators

This publication provides comprehensive data on the volume, origin and types of aid and other resource flows to around 150 developing countries for the period 2004-2008. The data show each country's intake of official development assistance and well as other official and private funds from members of the OECD's Development Assistance Committee, multilateral agencies and other key donors. Key development indicators are given for reference.

Atlas of Gender and Development How Social Norms Affect Gender Equality in non-OECD Countries

Illustrated with graphics and maps, the Atlas of Gender and Development gives readers a unique insight into the impact of social institutions - traditions, social norms and cultural practices - on gender equality in 124 non-OECD countries.

The Atlas of Gender and Development is an indispensable tool for development practitioners, policy makers, academics and the wider public. It provides detailed country notes, maps and graphics describing the situation of women in 124 developing and transition countries using a new composite measure of gender inequality - the Social Institutions and Gender Index (SIGI) - developed by the OECD Development Centre.

AGENDA

FORTHCOMING OECD MEETINGS

Date	Meeting
26-29 APRIL 2010	Meeting on the Management of Statistical Information Systems, Daejeon, Korea, Joint United Nations Economic Commission for Europe (UNECE)/Eurostat/Organization for Economic Cooperation and Development (OECD) (http://www.unece.org/stats/documents/2010.04.msis.htm)
7-8 JUNE 2010	Committee on Statistics (CSTAT), Statistics Directorate, Paris, France

Unless otherwise indicated attendance at OECD meetings and working parties is by invitation only.

OTHER STATISTICS MEETINGS

Date	Meeting
26-28 APRIL 2010	Work Session on Gender Statistics, Geneva, Switzerland, United Nations Economic Commission for Europe (UNECE) (http://www.unece.org/stats/documents/2010.04.gender.htm)
28-30 APRIL 2010	Joint Eurostat/UNECE Work Session on Demographic Projections, Lisbon, Portugal (http://www.unece.org/stats/documents/2010.04.projections.htm)
8-10 JUNE 2010	58th plenary session of the Conference of European Statisticians, Paris, France

European Conference on Quality in Official Statistics Q2010

Organised by Statistics Finland and Eurostat

Venue: Finlandia Hall, Helsinki, Finland, on 4-6 May 2010

For more information:

http://q2010.stat.fi/index/

Similarly to its predecessors, the Q2010 Conference will offer participants a platform for high-level discussions on various types of quality issues and methodological aspects of statistics production. The Q2010 will in particular focus on meeting current and future challenges for reengineering the statistical business processes of the European Statistical System.

The Statistics Newsletter
for the extended OECD Statistical Network
Issue No. 48 - April 2010