

# Statistical benchmarking: the Dutch case



EUROPEAN  
COMMISSION



*Europe Direct is a service to help you find answers to your questions about the European Union*

**New freephone number:**

**00 800 6 7 8 9 10 11**

A great deal of additional information on the European Union is available on the Internet.  
It can be accessed through the Europa server (<http://europa.eu.int>).

Luxembourg: Office for Official Publications of the European Communities, 2005

ISSN 1725-4825

ISBN 92-79-01322-X

© European Communities, 2005



**Statistics Netherlands**

Division of Macro-economic Statistics and Dissemination  
Department of price statistics and short-term indicators

*P.O.Box 4000  
2270 JM Voorburg  
The Netherlands*

---

**Statistical Benchmarking**

The Dutch case.

Paper to be presented at the OECD/Eurostat workshop, April 2005.

Symon Algera

(the author is Head macro-economic short-term statistics ; Statistics Netherlands)

Remarks:

The views expressed in this paper are those of the author and do not necessarily reflect the policies of Statistics Netherlands.

---

*Project number:*

*BPA number:*

*Date:*

*8 March 2005*

## **Introduction**

Keywords: system of statistics, coherence, consistency, revision policy, dissemination policy, quality assurance.

Purpose: to provide a description of benchmarking and related issues from the viewpoint of a statistical bureau.

The paper starts with a description of the System of statistics (section 1), a necessary first step in a description of benchmarking. Issues related to benchmarking are dealt with in sections 2: Quality assurance; 3: Revision policy; and 4: Dissemination policy. These aspects are all described from the viewpoint of a statistical institute, highlighting the specific role of these institutes. Some suggestions are made and some examples are presented for the Dutch case. This paper does not describe *benchmarking techniques* (notwithstanding some experiences in this field, especially smoothing techniques with respect to the “step problem”). It focuses on statistical and organisational aspects. The fact that Statistics Netherlands is a highly centralised statistical institute makes it possible to contribute with respect to these issues.

The annex contains some recent information with respect to benchmarking and labour market statistics (author : Sylvia de Vries)

Note: "Benchmarking is defined as the process of correcting inconsistencies between two sources of data for the same target variable which are produced with different frequencies" (OECD)

## 1 System of statistics

As benchmarking deals with relationships between statistics, the first step in a description of this process is a description of the system of statistics. In the Dutch case this system can be illustrated by a *statistical matrix*.

*Statistical matrix (example for construction statistics)*

<i>Time</i>	<i>Degree of integration</i>		
	<i>Single</i>	<i>Combined</i>	<i>Integrated</i>
Future	Tendency surveys Building permits	Index of producers' confidence	
Month	Turnover	Volume production index	
Quarter			Quarterly national accounts
Year	Structural business statistics	Annual branch Report	National accounts

The *time* column distinguishes statistics relating to future, monthly, quarterly and annual data.

Under *degree of integration*:

- the column headed *single* contains statistics which are obtained from a single survey of statistical units. The data are simply the survey findings;
- in the *combined* column the statistics are the outcome of combining different statistics (and different surveys). They are the result of taking statistics together, but are not yet integrated. An example is the deflation of turnover (from one source) using prices (from another source e.g. price statistics) and weighting schemes (from another source again e.g., national accounts);
- the *integrated* column contains data that are the outcome of an integration process involving detailed checks of all available information. The national accounts and the quarterly national accounts are examples of integrated statistics.

We can use such a matrix, to help explain some relevant aspects of coherence. Reading from top to bottom of each column, the reported results become more reliable and detailed, but less timely. Reading from left to right, the data become more comprehensive and reliable, but as a rule again less timely.

The various aspects of benchmarking (e.g. quality assurance and revision policy) can be illustrated on the basis of such a matrix.

A distinction can be made between *vertical benchmarking* (adjustments in the column, from monthly statistics to quarterly and annual statistics) and *horizontal benchmarking* (adjustments in the row from single, to partially and fully integrated statistics). The combination of these two approaches results in *diagonal benchmarking*. In statistics practice, most international experiences and literature are related to vertical benchmarking. This paper will illustrate that SN has some experiences with horizontal and diagonal benchmarking.

The following examples will be elaborated: monthly construction indicators and quarterly national accounts.

#### *Monthly production index for construction .*

The place of the monthly production indicator for construction in the system of statistics is illustrated in the example. It is a monthly *combined* and *not-integrated* indicator. Building blocks (sources) for the production index are the various *single* statistics. The indicator is published prior to the quarterly national accounts (QNA), which describe construction in a fully integrated way using all existing quarterly and monthly sources. The matrix shows the relationships.

The production process of the indicator can be described very shortly as follows. The main sources are: monthly turnover figures (value), price indices, building permits and working days. With the aid of these sources a monthly production indicator is compiled. Production is defined as *value added in constant prices*. This means that the indicator can be considered as a predictor of the *value added of construction* as described more reliably in the system of quarterly and annual accounts. The metadata are identical (see section 3). As a consequence of this relationship the compilation processes of the monthly production indicator and the QNA are linked as follows. When the production indicator is compiled for the third month of a quarter, the resulting quarterly indicator (the sum of the three months) is used as input for the compilation of the QNA (which is carried out with the help of supply and use tables, see below). Usually the balancing process leads to adjustment of the original input, in this case the quarterly index of production (as the sum of the three production indicators), and consequently to adjustments of the three underlying monthly indicators. This means that with the first release of a quarter in the QNA the two (already published) monthly production indicators are adjusted and the production of the third month is adjusted before publishing. At SN this production scheme (which is roughly also used for the monthly production index of manufacturing) is considered to be the optimal procedure using all the existing information and providing coherent and consistent monthly and quarterly information concerning construction.

This procedure can be considered as an example of *statistical benchmarking* as part of the compilation process (horizontal/diagonal benchmarking).

### *Quarterly National Accounts (QNA)*

The place in the system of statistics is clear. It is the integrated framework of short-term statistics. The reliability of short term statistics can be assessed by comparing the results of these statistics with the results of later, more integrated statistics. It is then relatively simple to obtain quantitative measures of the quality dimension of the statistical output. Furthermore, the benefits of coherence can be used for benchmarking and possibly for adjustments in the monthly and quarterly statistics which have served as building blocks for the QNA.

The QNA are compiled from quarterly supply and use tables. These tables are based on the available monthly and quarterly data and a set of assumptions. Using this information pre-balanced supply and use tables are constructed. In the first instance, the tables are rendered consistent with the aid of a discrepancy column. The balancing process consists of eliminating this discrepancy column. The values in the discrepancy column are a good indicator of the quality of the underlying statistics. Moreover, further analyses may lead to feedback to and adjustment of the source statistics.

This, too, is an example of statistical benchmarking (horizontal/diagonal benchmarking).

These examples show that short-term indicators can benefit from benchmarking, not only benchmarking based on a comparison of monthly statistics with quarterly and annual statistics, but also benchmarking based on a comparison of single with integrated statistics.

The following sections examine the related issues of quality assurance, revision policy and dissemination policy.

## **2 Quality assurance**

Benchmarking as described in the previous section has a clear relationship with quality assurance. As illustrated in the examples, short-term statistics can benefit from the output of annual statistics and integrated statistics. Another related issue is quantitative quality measurement.

There is a clear coherence between statistics: firstly in terms of predictive power (short-term statistics predict annual statistics), and secondly in terms of consistency (e.g. between structural business statistics and national accounts, or between statistics which are part of accounting systems).

Assuming that this statistical coherence exists, the reliability of statistics can be assessed by comparing the results of the underlying statistics with the results of later and more integrated data. It is then relatively simple to obtain quantitative measures of this quality dimension of statistical output.

This quantitative information on the reliability of short-term indicators is indispensable for users, policymakers and statisticians, if only to choose between timeliness and reliability.

Recent policy at Statistics Netherlands puts more emphasis on this kind of quantitative quality measurement. This approach emphasises the *output* of the statistical production process, in addition to quality control of the *process* itself. Of course, a good output quality goes hand in hand with quality management of the process of compiling statistics.

### 3 Revision policy

Benchmarking may entail revisions. When does, and when doesn't it? What is the revision policy? This section elaborates some of these issues further, focusing first on policy (for the future), and subsequently on present practice (and quick wins).

#### *Policy*

SN is currently developing a consistency policy for its central database *StatLine*. The matrix presented above will be used as a starting point and framework for a systematic approach of the decision-making process. A set of criteria will be developed to decide whether comparable indicators will be revised or not. Important criteria are *statistical feasibility* and *user demands*. Let us start with the latter.

#### *Users*

Although revision of statistics has advantages, it also has disadvantages. Many users dislike revision over a period of years, they want calm and stable figures. Moreover in practice after revisions there is often a break between aggregates and details, which users also dislike. This results in the *consistency* versus *stability* dilemma. In reality there is not just one user, but many more and each has his own demands. A recent development in the Netherlands is that an increasing number of scientific users who want consistent figures also want access to the original figures. This is connected with the tendency to perform more *real-time* analyses (i.e. with data as available at a specific moment) instead of *out of sample* analyses with the definite (*ex post*) figures.

*Intermezzo. In some cases the national statistical institutes (NSI's) play a crucial role. For some statistics it is essential that an independent body determines the outcome (e.g. government deficit, HICP, GDP). NSI's usually do this. It seems desirable to decide for which variables statistical offices should play this umpire role and which revision policy should be applied to these statistics. Obviously, for price statistics stability is more important (because, for instance, they are used in contracts) than for other statistics.*

*This issue of professional independence is closely related to recent developments with respect to the European Statistics Code of Practice. The issue of the Greek data*



*for the European deficit procedures induced the ECOFIN Council (Council of Ministers of Finance in the EU) to demand a code of conduct in order to establish Europe-wide minimum standards in the domain of statistics, which reinforce the independence, integrity and accountability of NSI's and of Eurostat. A small taskforce drew up a proposal for discussion. The proposed code is based on 15 principles and consists of three parts: institutional environment, statistical processes and statistical output. Governance authorities and statistical authorities are committed to adhering to the principles and to reviewing its implementation periodically by the use of indicators of good conduct for each of the 15 principles, which are to be used as references.*

*As a member of the taskforce on the Statistical Law (2004) and the subsequent taskforce on the Code of Conduct (2005), Statistics Netherlands has worked on the development of a more sound judicial framework for the ESS. Looking at the European Statistics Code of Practice as a precursor of a more formal regulation, we are strongly in favour continuing the discussions on the revision of the StatLaw.*

*The first principle in this proposal is professional independence. This principle implies that for crucial statistics, benchmarking/adjusting is the preserve of the responsible statistical authorities.*

#### *Statistical feasibility.*

It is easy to replace one aggregate figure by another aggregate figure. The problem in practice is that moving towards the source (building blocks) of the statistics (in the matrix from right to left) the figures become much more detailed and it becomes less realistic and feasible to adjust the original figures. The most extreme consequence of a fully consistent database would be to change the original answers of respondents if inconsistencies appear between integrated and not integrated statistics in the balancing process of the national accounts (and discrepancies have been eliminated on a certain level of aggregation). Although in theory this approach has some advantages (all statistics can be compiled from this fully consistent database, with an optimal micro-macro link) at SN it is considered not realistic or feasible.

#### *SN policy*

SN is currently developing a revision/adjustment policy. A keyword in this revision policy will be *metadata* (e.g. definitions, classifications). Adjustments to later statistics will only be applied systematically if the metadata of the variables in the various statistics are identical. Because the metadata in the columns of the matrix are more identical than in the rows, adjustments will be carried out more often (but not exclusively) in the case of vertical benchmarking than in the case of horizontal benchmarking.

The first rough contours of a solution of the *consistency* versus *stability* dilemma are:

1. adjust and make consistent the integrated statistics column (this is already practice in most European countries and also at SN);,
2. adjust and make consistent all macro/meso–economic combined short-term statistics with the later integrated statistics (this is practice at SN, but not in many other countries);
3. with respect to the column single statistic: the output of monthly , quarterly and annual statistics should be mutually consistent within certain margins; if the discrepancy is outside these margins adjustments are necessary;
4. do not adjust in all other cases except in the case of obvious mistakes;
5. if changes are made to the original statistics, always keep the original figures and make it possible to show the original figures (this is not practice at SN but will be the policy for the future and it will be made feasible in the new database *StatLine 4*; see section 4).

### ***Present practice***

Apart from developing a policy for the future on the issue of revisions, SN considers it just as important to be more transparent on the present policy (even if this policy is not consistent for all statistics). For this reason a distinction can be made between the various types of revision and an indication can be given of which type of revision in principle applies to a certain series of statistics.

A quick elaboration for the Dutch case gives the following picture

Revisions can be prompted by:

1. A higher response rate.
- 2 . Adjustment to *later* statistics. For example: adjustment of a monthly series to a quarterly series, or of quarterly statistics to annual statistics. Note: in this context we do not consider annual publications (statistics) which are an aggregation of monthly or quarterly statistics as *annual* statistics
3. Adjustment to *integrated* statistics, such as national accounts. For example: adjustment of the index of monthly household consumption to the household consumption estimated within a consistent system of quarterly national accounts.
4. More and improved information from other countries. For example: adjustment of national external trade statistics to Eurozone external trade statistics based on the *asymmetry exercise*.
5. Major (five or ten-yearly) revisions based on new definitions and classifications.

Revisions under item 2 apply to the columns in the statistical matrix, and revisions under item 3 to the rows. Revisions under item 4 apply to the international dimension of the matrix (which is not elaborated) and relate to external trade and balance of payment statistics.

Each series of statistics at SN can be labelled with its revision policy.

For instance:

- Monthly manufacturing index: (1), (3) and (5);
- Monthly manufacturing turnover index: (1);
- Monthly retail sales index: (1);
- Quarterly national accounts (1), (3) and (5).

It would seem useful to provide metadata of this kind on the revision policies for the various statistics.

#### **4 Dissemination policy**

As a consequence of benchmarking , statistics may be revised (for policy: see above). How are revised statistics published? Are the original statistics still available? Which metadata are provided? What is the status of the statistics? This section contains some answers to these questions. A distinction is made between *statistics* and *metadata*.

##### *Statistics*

A consequence of the revision policy may be that the initial statistics are revised several times. Many users want the “definite” figures, but others also want access to the original figures. In the Netherlands there is an increasing demand for original figures as economic analyses are tending to use the *real time* data instead of the *ex post* data. We have not yet decided how to deal with this issue. A promising option is to provide the “definite” figures in the central database (they are the best and completely integrated with other statistics), and to lead users to the original (not yet adjusted) statistics via a system of hyperlinks.

##### *Metadata: Status of statistics*

Another problem is the lack of systematic meta-information with respect to the *status* of short-term indicators. Are they part of an accounting system? Which revision policy applies to them and are they adjusted to other statistics? This kind of information seems indispensable for users. It is also important for country comparisons, as the status may differ between countries.

*Intermezzo (1). An example of two extremes in this context: the monthly index of industrial production in the US (as compiled by the FED), is not used at all in the compilation of quarterly GDP (as compiled by BEA) , so there is no numerical relationship of this index with industrial production in the quarterly accounts. There are no revisions of the monthly production index when new information on value added of industry becomes available from quarterly or national accounts. So in the US, the index of industrial production is a quite stable and transparent figure, but inconsistent with the quarterly and national accounts. The monthly index of industrial production in the Netherlands, on the other hand, is completely integrated in the quarterly and national accounts. This monthly index is adjusted when new and more accurate quarterly and annual information becomes available. The monthly figures and the quarterly and national accounts are fully consistent. This is what users want for analytical purposes. On the other hand, the monthly figures on industrial production are then adjusted a number of times. So in the Netherlands the index of industrial production is a consistent but rather unstable figure (this kind of background information may be important for the users of the figures). Moreover: a possible bias in the monthly figures has been made visible.*

*Intermezzo (2) This example also shows that the organisation of the compilation of statistics is connected to the benchmarking issue. Statistical benchmarking can be applied more exhaustively and systematically and on a larger scale in countries where the compilation process of statistics is centralised and coordinated.*

In its *StatLine* data base, SN plans to show the *status* of the statistics with respect to which revision policies applies to that specific statistic, and with respect to which stage the revision is in. Take, for example, the monthly production index before adjustment to the quarterly national accounts but after a revision to a higher response rate. The metadata in the database will show that (in principle) revision policies (1), (3) and (5) apply to the index (see previous section) and that the specific figure is in revision stage (1) (and as a consequence that at least two stages will follow).

This approach would also have advantages for the databases of international organisations (e.g. Eurostat, OECD). First it provides an insight into the background of the figures of the various countries (see the example for the production indices of USA and the Netherlands); and secondly it provides the possibility for international organisations to place comparable figures of the various countries in the same database. An example for the monthly production index of SN: it is (in theory) possible for an international organisation to place in its international database the index of industrial production for the Netherlands before adjustment to the QNA (revision stage 1), because the indices of the most other countries have the same status; but they can make clear with metadata that SN also has adjusted and more integrated indices. This also means that European indices, for example, can be made definite much earlier than at present. At the moment the Dutch industrial production index is definite after two and a half years, and as a consequence the Eurostat index

may also be revised over a period of years (which Eurostat dislikes, calling it *rewriting history*).

*The new improved statistical database: StatLine 4.*

The new version of Statistics Netherlands' statistical database, *StatLine 4*, also makes it possible in a more technical sense to introduce the plans described above.

*StatLine 4* will standardly show tables which contain the most *definite* outcome of the statistics, but it will give users access to former estimates as well. Furthermore, one of the features in *StatLine 4* is that users can save the table they have selected as a link; they can choose whether they want to link to the real-time data or to the latest update of the data.

At the moment SN is implementing a uniform system for presenting the dissemination policy for all statistics. In *StatLine 4* the dissemination policy per statistic will be made available.

## Annex : **(Un)employment as an example**

(author : Sylvia de Vries)

(the author is Head of “Consumption, households and labour” (which includes the labour accounts) within the National accounts department; Statistics Netherlands)

The labour market is an interesting area in terms of benchmarking. Labour is both a complex and a dynamic phenomenon, as it measured through various surveys and viewed from different perspectives. Labour data are obtained from both household surveys (labour force survey), establishment surveys (employment, earnings and labour cost surveys, mainly using salary administrations) and register data from social security institutions. And labour can be considered done from an economic (as a production factor) as well as a social perspective (socio-economic status).

It is important that the labour statistics are placed within a solidly defined framework, just like the national accounts, so that the NSI's in Europe can develop a consistent statistical system for the labour market. This will produce internationally comparable figures that will be able to play a pivotal role in European labour policy. In the Netherlands we compile labour statistics using an integrated *labour accounts* approach. The labour accounts have been part of the national accounts since the ESA 1995 revision; they present one figure on employment and compensation of employees. This has not only improved the linkage of economic to social statistics for users, but has also expanded data quality check possibilities for both sides.

The Dutch case for (un)employment is illustrated in the *statistical matrix* below. Labour accounts provide information on employed persons and jobs, by means of fully integrated statistics. A description of the integration procedure for the annual labour market data is given in SN 1999. Statistics Netherlands uses the number of jobs of employees as an indicator for employment (in the future integrated short-term information on self-employed persons will also become available), publishing a flash and a regular estimate of this statistic for the sake of timeliness. The flash estimate is based on changes for a small set of selected enterprises from the quarterly survey. The regular estimate uses the full source information. Information on the labour force on the other hand is placed in the combined section (due to the fact that statistics are combined, but not yet integrated). Source data come from the labour force survey (LFS), a continuous (household) survey of persons resident in the Netherlands. Monthly figures from both CAPI (computer-assisted personal interviewing) and CATI (computer-assisted telephone interviewing) surveys are used to compile three-month moving averages for the

employed and unemployed labour force. Because of substantial differences between CAPI and CATI adjustments are made using moving ratios.

*Statistical Matrix for employment en unemployment*

<i>Time</i>	<i>Degree of integration</i>		
	<i>Single</i>	<i>Combined</i>	<i>Integrated</i>
<i>Month</i>	Labour Force Survey	3-month moving average (un)employed labour force	
<i>Quarter</i>	Quarterly statistics on employment		Labour accounts: jobs of employees
<i>Year</i>	Annual survey on employment and earnings		Labour accounts: employed persons, jobs and labour input

National accounts measurements of employment differ from the LFS measurements because of differences in definitions. Also differences in population coverage exist, because of place of residence for example (national versus domestic concept), and age. On top of these, on the labour market non-registered paid labour (black market) is not accounted for in the source statistics. An example is domestic cleaners, who mostly receive their money under the counter. Consistency between the employment indicators is provided by taking these differences into account. As a consequence, when the employment figures are published, the integrated figures are leading. Single statistics provide more detailed information and are presented as underlying figures.

Because of the definition used, unemployment (labour force) is only measured through the LFS. As this is a sample survey, it contains noise and there are sporadic outliers. Registered unemployment benefit data are available and can be used to detect these outliers. On the basis of the LFS and register data a model estimate is made for the expected result. Any survey result falling outside the 95% confidence margin surrounding the model estimate will be treated as an outlier. Further analysis should reveal whether a plausible cause can be found, if any irregularities occurred or the figure is a survey outlier. In the latter two cases the figure(s) concerned will be adjusted. The figures of the previous month are subjected to this method once more.

This example describes the benchmarking process at the moment. In the future the *labour market accounts* should provide integrated information for the (un)employed labour force.

## References

Algera, S. (2001); *Short-term indicators; their role in the statistical system and recent developments; the Dutch case*. Background paper for OECD meeting, Paris, June 2002

Algera, S. and S. Keuning (2001). *Some elements of a quality framework for CMFB statistics*.

Algera, S. and S. Keuning (2002). *Quality of statistics: a sketch of some remaining issues* Note for presentation at the ECB-seminar “Eurozone statistics: Challenges for the future”, Frankfurt

Booleman, M., et al. (1999), *Attention to quality within Statistics Netherlands: quantifiable quality characteristics*, Statistics Netherlands.

Booleman, M. (2004), *The Dutch metadata model: introduction of conceptual metadata, process metadata and implementation strategy*, Statistics Investment in the future, Prague, 2004

Booleman, M. (2004), *Conceptual metadata and process metadata: Key elements to measure quality components of the statistical system*, Q2004, Mainz

Kazemier, B., H. Nijmeijer and R. Hijman ( 2004), *How to judge the reliability of provisional National Accounts*, Paper prepared for the European Conference on Quality and Methodology in Official Statistics, Mainz, May 24-26 2004.

Leunis, Wim P. (2000), *Linking Social and Economic Statistics through the 1995 Revision of National Accounts and Labour Accounts* (Statistics Netherlands NA-093)

Statistics Netherlands (1999), *Labour accounts in theory and practice, the Dutch experience* (Statistics Netherlands,).