Statistics Netherlands

Division of Macro-economic Statistics and Dissemination
Development and support department

P.O.Box 4000
2270 JM Voorburg
The Netherlands

Alignment of Quarterly Sector Accounts to annual data

Reinier Bikker and Susanne Buijtenhek

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

Project number: *MOO-204072-02*
BPA number: *2006-24-MOO*
Date: *22 March 2006*

Contents

ALIGNMENT OF QUARTERLY SECTOR ACCOUNTS TO ANNUAL DATA

Summary: This paper presents a benchmarking process centred around an automatic benchmarking procedure. The process roughly consists of three parts: preparatory editing, to remove the largest inconsistencies; automatic smoothing of the remaining inconsistencies; and a plausibility check on the results. The process is iterative: if the plausibility check reveals that the results are not satisfactory, they can be improved upon by fine-tuning the input and again applying the automatic algorithm. In the preparatory editing phase, the largest initial differences are resolved by manually editing the input data. A good identification of the cases that require manual intervention is important to obtain acceptable quality without spending too much time. The automatic part of the process is based on a Denton-type optimization algorithm under restrictions. This model uses reliability weights and allows for missing indicators and exogenous quarterly series. With these ingredients, the method proves to be flexible enough to benchmark structured time series originating from many different sources, widely varying in reliability. A good benchmarking method, however, does not automatically mean plausible results. The automatic procedure can only produce plausible results, if plausible input is fed into the benchmarking model and if the plausibility is explicitly checked by subject matter specialists.

Keywords: Benchmarking, temporal disaggregation, structural time series, multivariate Denton method, Quarterly Sector Accounts.

1. Introduction

The quality of Quarterly Sector Accounts (QSA) depends, among many other aspects, on their connection to the corresponding Annual Sector Accounts. In particular, time series analysis requires consistent data. Usually the data sources of the annual totals are independent from those of the quarterly data. Thus, differences between annual figures and the corresponding quarterly figures occur naturally. As consistency is generally considered an important quality aspect, making the data more useful for economists and policy makers, consistency must be restored.

Due to the sheer amount of data to be processed, and the short time period available for processing, it is not feasible to do all alignment, including (re)balancing, by hand. In some cases, differences are large and resolving them correctly may require attention from specialists. In most cases, differences are small and could be smoothed in a mechanical way. Therefore, a semi-automatic benchmarking method is called for.

This paper describes the results of a project in which we developed such a method. The method applies to both non-financial and financial accounts, including financial balance sheets. There are no essential methodological differences between both cases. The method includes criteria for discriminating between manual and automatic adaptation, and an automatic method to resolve the remaining discrepancies after large differences have been resolved by hand. The automatic part of the method includes a Denton approach. The data passed through the method must satisfy consistency requirements while staying as close as possible to the original short-term movements in the data.

This paper consists of the following parts. Section 2 starts with a general description of the whole benchmarking process. The sections 3 to 5 give a more detailed description of the various parts of this process. Section 3 provides an in-depth description of the benchmarking model that can be used in this semi-automatic alignment procedure. Section 4 describes the input data and the pre-processing phase. Section 5 gives details on the plausibility checks. Our conclusions and recommendations can be found in section 6.

2. A semi-automatic benchmarking process

2.1 The annual update cycle

Each year in July, Statistics Netherlands publishes Annual Sector Accounts, including financial accounts and balance sheets. The Annual Sector Accounts compilation cycle follows a predetermined schedule spanning the period of a year. New and revised Annual Sector Accounts are published for the three most recent years. The Annual Sector Accounts for last year are mostly based on provisional or incomplete data. Before reaching a final status, these accounts are updated twice in subsequent years, while the primary source data situation gradually improves. So each year Statistics Netherlands publishes provisional accounts of last year, revised semi-final accounts of the last but one year and final accounts of the last but two year.

The Dutch Quarterly Sector Accounts, including financial accounts and balance sheets, are compiled 90 days after closing of the quarter. This means that all four quarters of a given year are compiled some months before Annual Sector Accounts for this year become available for the first time.

Both the annual and the quarterly compilation cycles roughly consist of two phases. In the data collection and processing phase, a full set of initial accounts is compiled for each sector in the accounting system. In the balancing phase, these accounts are then manually balanced by subject matter specialists in order to achieve a fully consistent system.

The Quarterly Sector Accounts should be aligned to the most recent published set of Annual Sector Accounts at all time. This means that each year a time series of twelve subsequent quarters must be updated to the most recent annual figures. Moreover, due to the annual compilation and publication strategy each quarter will have to be aligned to updated annual figures precisely three times.

2.2 Process design

The process we describe in this paper centres around an automatic benchmarking method. This automatic method is based on a Denton-type optimisation algorithm under restrictions. Such an automatic benchmarking model is essentially a mechanical tool, which can only be applied with proper pre-processing of the input data. For best results, the larger annual-quarter differences must be identified and resolved (approximately) before applying the automatic model. Afterwards plausibility checks on the results are necessary. If the plausibility checks reveal that results are not satisfactory, they can be improved upon by fine-tuning the input and again applying the algorithm. This step may be repeated as often as necessary. The whole process is designed in such a way that it can be carried out rather efficiently. As the benchmarking algorithm takes away the burden of many small adjustments, it leaves sufficient room for the important task of plausibility checking. Only the largest and most important differences require manual intervention. Figure 1 presents a schematic outline of the benchmarking process.

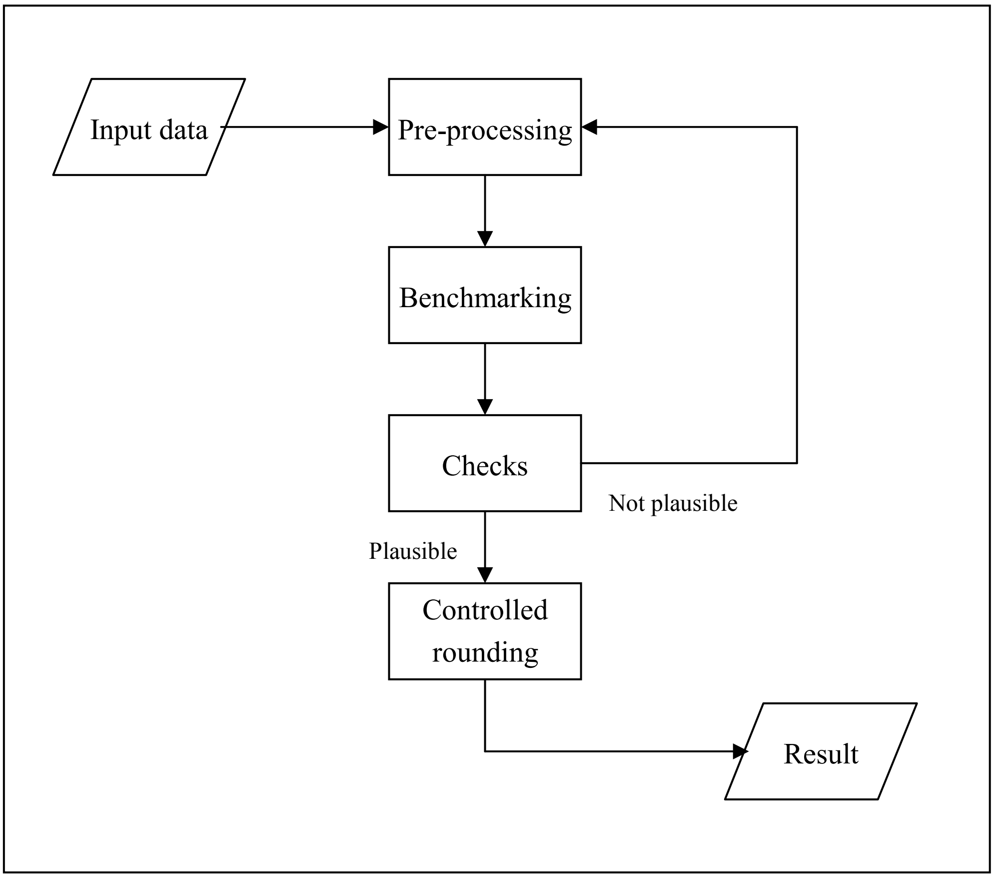

Figure 1. Schematic outline of the benchmarking process

The automatic benchmarking model is the core of the process. Its main characteristic is that quarter-to-quarter movements are preserved as much as possible, while enforcing annual alignment. Simultaneously, all accounting rules applicable in each quarter are satisfied. The benchmarking model is based on the “movement preservation principle” formulated by Denton (1971). Additionally, the algorithm uses reliability weights in order to deal with quality differences of input data. Furthermore, in some cases quarterly data should not be changed at all. Hence, the algorithm allows for exogenous quarterly series. The automatic model is described in detail in Section 3.

Successful application of the automatic core of the benchmark process relies on suitable input data. Differences that are too large, e.g., due to imperfections in the quarterly source data, cannot always be satisfactorily resolved by an automatic method. In order to deal with such cases, a pre-processing phase has been introduced. Pre-processing of the data starts with analysing the differences between quarterly and annual figures. According to the importance of the variable, a threshold value for allowed initial differences can be specified, above which manual intervention is required. The intervention amounts to manually attributing (a sufficiently large part of) the differences to the correct quarters. When the remaining initial differences are small, they can be smoothed in a mechanical way. The details of pre-processing are described in section 0.

After the automatic benchmarking is applied, the quality of its output must be judged by subject matter specialists. For this, they can be aided by visualisation tools and several kinds of analysing tools. Section 5 will describe this part of the process in more detail.

When the results are deemed acceptable, there is one final step to be taken. This step is the controlled rounding process, where all results are rounded to integer values, under the same constraints used in the benchmarking process. The controlled rounding process will be briefly visited in section 3.3.

3. The benchmarking model

3.1 Aligning high and low frequency time series

Statistical data can be compiled at different frequencies. Most usually, data is published each month, each quarter or each year. When dealing with high and low frequency data on the same variable, one often encounters consistency problems, e.g. the quarterly data does not add up to the annual figures. High and low frequency data are typically obtained independently and as both may be subject to all kinds of statistical disturbance, inconsistencies occur naturally. In another setting consistency is present originally, but it is disturbed by statistical operations. In these settings, the process of restoring consistency is usually called *benchmarking*. The same consistency problems arise when statisticians or economists try to create high frequency data from low frequency data, with the help of related indicators. This process is usually called *temporal disaggregation*. Benchmarking and temporal disaggregation are essentially the same problem. In both cases high frequency input data must be made to fit the low frequency data. The only difference, usually, is the quality of the high frequency indicators.

A method for restoring consistency can be based on several different statistical models. In the univariate field we have the Chow-Lin method (1971), which applies a least squares estimator to estimate a high frequency time series from one or more related series, in such a way that the high frequency data are connected to the low frequency data. This method finds the solution closest to a set of indicators, but may create step problems when the regression relations change quickly over time. Another classical reference is Denton (1971). The Denton method[1] is based on the "movement preservation principle", which aims to avoid step problems. The Denton method can be used with or without a related time series. Without related series the method creates a smooth interpolation or distribution of the low frequency data.

It is also possible to benchmark structural time series. Apart from temporal alignment, structural time series must also satisfy a set of contemporaneous constraints. Di Fonzo and Marini (2003) extended the Denton method to perform multivariate benchmarking. More sophisticated multivariate approaches include the state space approach (e.g. Harvey (1990)) and multivariate Bayesian benchmarking (e.g. Broemeling (1985)). We did not explore these types of models as we felt the more basic approach provided by the Denton method would be best suited for a first large scale practical application.

3.2 The multivariate Denton method

3.2.1 Outline

In this section we will present a benchmarking method based on a Denton-type optimization algorithm under restrictions. The main characteristic of this algorithm is that quarter-to-quarter movements are preserved as much as possible while quarterly-annual alignment is achieved. The resulting quarterly series can be seen as composed of the seasonal components of the original quarterly data, superimposed on the (interpolated) trend-cycles, which are obtained from the annual figures.

Simultaneously to quarterly-annual alignment, all accounting rules applicable are satisfied. The algorithm uses reliability weights in order to deal with quality differences of input data. Furthermore, in some cases quarterly data should not be changed at all. Hence, the algorithm allows for exogenous quarterly series.

Remark. The remainder of this section (3.2.2–3.2.6) is very technical and may be skipped on first reading.

---

[1] The Denton approach will be explained in some detail in the next section.

3.2.2 Notation

Our problem deals with a set of quarterly sources x_{ik}, where $i = 1,...,M$ is a label to distinguish different series, and index $k = 1,...,N$ enumerates the quarters. The sources are generally independent of the corresponding annual sources and independent among themselves. Because of the independency of the sources, the consistency is not automatically guaranteed. Any kind of statistical disturbance, like sample errors and non sample errors, can generally be found in the sources. We will assume that these statistical disturbances are not correlated between different series. We will also assume that there the statistical disturbances are not correlated in time. Both these assumptions may in fact not always be true, yet we seldomly know how variables are correlated. If the correlations were known they could be modelled explicitly.

In the following sections we will represent the quarterly sources as a single vector x, which contains the source data for all M variables during N quarters. The order in which we vectorize the source data is of no importance, but for clarity let us make the choice

$$x = (x_{11},...,x_{1N}, x_{21},...,x_{2N},...,x_{M1},...,x_{MN})' . \tag{1}$$

Analogously, the vector y contains the corresponding annual totals. The benchmarked results are represented by the vector $\hat{x}$.

3.2.3 Consistency

There are two types of consistency within our system, balancing and connection to the annual totals. Balancing is based on a set of contemporaneous linear restrictions, which is actually the definition of the accounting framework. Connection to annual totals means either that the sum of the quarterly flows adds up to the corresponding annual total or that the fourth quarter final stock equals the annual final stock. These are also linear restrictions, be it temporal ones, as they relate variables in different quarters. Figure 2 below gives a graphical representation of a simplified example.

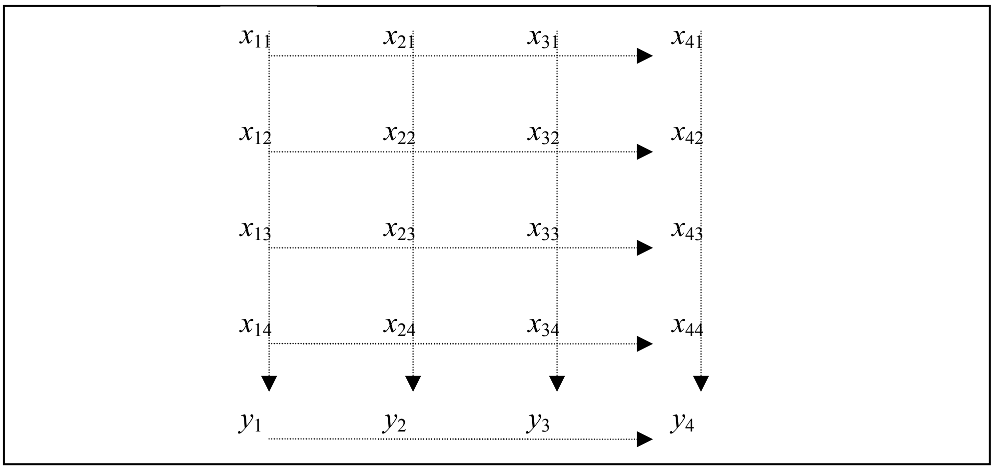

Figure 2. An example of temporal and contemporaneous linear restrictions

Figure 2 shows four different transactions, x_{1q}, x_{2q}, x_{3q} and x_{4q}, in four subsequent quarters, labelled with subscript q, together with their four annual totals, denoted y_i. The horizontal arrows in figure 2 represents a contemporaneous restriction within each quarter. In each quarter the equality $x_{1q} + x_{2q} + x_{3q} = x_{4q}$ must hold. The vertical arrows represent the temporal alignment. If transaction x_{1q} contains flow data, the temporal restriction for connection to annual totals would be $x_{11} + x_{12} + x_{13} + x_{14} = y_1$. Annual connection for stock data would be represented by $x_{14} = y_1$.

Contemporaneous and temporal restrictions could be treated independently, e.g. in an iterative and hierarchical way[2]. It is important to realize that from a statistical point of view the optimal solution can only be found if they are treated simultaneously. Together they form one set of linear restrictions, which the benchmarked quarterly figures must satisfy,

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \hat{x} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \tag{2}$$

where C_1 and C_2 represent the constraints for balancing and connection to annual totals, respectively. Vector b_1 contains either the values of exogenous variables or zeros and vector b_2 contains the annual totals.

We will only deal with linear restrictions. There are other types of consistency restrictions. For example, most stocks cannot be negative. Inequalities however, are much more difficult to deal with mathematically and numerically. Another type of relations are non-linear constraints, like ratios and products. If a non-linear relation has an economic meaning, it could be desirable to include it in a benchmarking procedure. Non-linear constraints are also difficult to deal with mathematically. However, non-linear constraints can be linearized and can be included in our model as such. We will not discuss non-linear constraints explicitly. Apart from binding constraints, the model can also deal with non-binding constraints, where $C_{\mathrm{NB}}\hat{x} \approx b_{\mathrm{NB}}$. For this the non-binding restriction must be rewritten into a binding one by introducing an error term, such as $C_{\mathrm{NB}}\hat{x} + \varepsilon = b_{\mathrm{NB}}$ with $\varepsilon \sim N(0,\sigma)$.

3.2.4 Minimizing adaptations

The set of restrictions alone does not lead to a unique solution, as this set generally is not complete. Several degrees of freedom remain, from which a solution must be chosen. The intuitive idea is that, somehow, the benchmarked results must be "as close as possible" to the original sources. Therefore we must introduce a so called *objective function,* to measure the overall difference between the benchmarked data and the original data. According to the least squares principle the objective function must be a quadratic function of the differences between benchmarked and original values, given by

[2] For an example see Laniel and Fyfe (1989, pp. 464-465).

$$(\hat{x}-x)'\Omega^{-1}(\hat{x}-x), \tag{3}$$

where the matrix Ω must be designed in such a way that the intended objectives are met.

Our problem is now formulated as a standard quadratic optimization problem, where we want to find the minimum of the objective function (3), subject to the constraints $Cx = b$. Yet, first we must choose the form of the matrix Ω. For instance, if Ω is set equal to the covariance matrix, we obtain a model that resembles the well-known model of Stone et all (1942), except that in this model x contains data from different periods, linked by temporal constraints. This benchmarking model would indeed generate benchmarked results as close as possible to the original quarterly values. Applying this model for our problem would not be a good idea, because the extent of the *initial differences* between the quarters and their annual totals may vary greatly from year to year. If the changes made to the quarter-to-quarter growth rates are not considered during the benchmarking process, a discontinuity may be introduced between the last quarter of one year and the first quarter of the next year. This is usually called the *step problem*. Staying as close as possible to the original quarter-to-quarter movements is a much better governing principle for finding an optimal solution. This idea, first explained by Denton (1971), is called the *movement preservation principle*.

An important aspect of the problem is that the sources of the quarterly data are not all equally reliable. Some data sources are sample survey results and therefore suffer from both sampling errors and non-sampling errors. Other data sources are not supposed to change during the reconciliation process, because they originate from government registers or simply because they have been published earlier. Yet another class of data sources consists of cases where quarterly data is not available but can be estimated using related indicators. Some data may be missing altogether and must be estimated as a result during the reconciliation process. The overall reliability of the data sources can therefore range from virtually zero to infinite. The reconciliation procedure must be able to accommodate all these types of data input if it is to be applicable in practice.

For implementing the Denton principle, Ω is constructed as follows[3]. Let

$$X = \mathrm{diag}(x) \tag{4}$$

and let the vector v contain the coefficients of variation or relative standard errors of x_{ik}, $v_{ik} = \mathrm{SE}(x_{ik})/x_{ik}$, vectorized in the same way as x. Now, let

$$V = \mathrm{diag}(v), \tag{5}$$

so $X'V'VX$ is the variance matrix. The movement preservation principle can then be implemented by setting Ω equal to one of two possible expressions,

[3] Much of this can be found in greater detail in Di Fonzo and Marini (2003).

$$\Omega = \begin{cases} \left(D'\left(X'V'VX\right)^{-1}D\right)^{-1} & \text{additive} \\ X'\left(D'\left(V'V\right)^{-1}D\right)^{-1}X & \text{proportional,} \end{cases} \tag{6}$$

where the matrix $D = I_M \otimes D_N$, while the $N \times N$ matrix

$$D_N = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 & 0 \\ 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix}. \tag{7}$$

The bottom $N-1$ rows of this operator calculate the first differences of a single variable, while the first row represents the fixed quarter as explained in Denton (1971). We need this extra row in order to let D_N be of full row rank. Fixing the first quarter is equivalent to the assumption that there is a zeroth quarter preceding the first quarter of the series, which is not changed in the benchmarking process, $\hat{x}_0 = x_0$. The choice of fixing the first quarter is not strictly necessary.

The variance matrix of vector x is given by $X'V'VX$. Matrix V can be used to define the quality differences between the sources. Because the objective function can be multiplied by a scalar without changing the minimum, only relative coefficients of variance need to be specified.

The proportional model is suitable for indicators of which the relative values of subsequent quarters are considered representative for the transaction. This is the case when an indicator is based on a representative sample of the population it is supposed to describe. The proportional model will distribute the initial differences between quarters and annual totals proportionally to the original quarterly values.

For other sources the additive model is more suitable. This is the case if a variable can have both positive and negative values or if the level of the sources is very different from the level of the annual totals. In both these cases the quarter-to-quarter movement in the sources can be greatly enhanced or diminished in the results. Most often, this is undesirable. In general, if the indicator only describes a part of the transaction and is not considered representative for the missing part, the additive model is a better choice.

In equation (6) the additive and proportional models are given separately. For practical purposes the model must be able to handle both types of sources simultaneously. To use the additive and proportional model simultaneously, the original matrix X must be factored into matrices X_{PROP} and X_{ADD} where X_{PROP} contains the real values for the proportional variables and value "1" for the additive variables and X_{ADD} contains the real values for the additive variables and ones for the proportional variables. It is clear that $X_{\text{PROP}}X_{\text{ADD}} = X$. Using this, Ω can be expressed as a single expression for a mixed system of additive and proportional variables,

$$\Omega = X'_{\text{PROP}} \left(D' \left(X'_{\text{ADD}} V'VX_{\text{ADD}} \right)^{-1} D \right)^{-1} X_{\text{PROP}} . \tag{8}$$

Using a well known[4] matrix result, the solution to the problem of minimizing expression (3), subject to the constraints (2) is given by

$$\hat{x} = x + \Omega C (C \Omega C')^{-1} (b - Cx) . \tag{9}$$

In equation (8) the matrix $\Omega C (C \Omega C')^{-1}$ distributes the initial discrepancies $b - Cx$ over the original values in x, in such a way that all discrepancies are resolved.

The matrix C must be of full row rank, or the inverse $(C \Omega C')^{-1}$ will not exist. This means that there cannot be any redundant constraints. Yet, combining temporal and contemporaneous constraints always leads to redundancy. This can easily be seen by starting with M unrelated series that must be benchmarked. For each year we have M restrictions describing their annual alignment. Yet, adding one contemporaneous constraint, makes one of the annual alignment restrictions redundant: if the annual alignment holds for the first $M - 1$ series, and the contemporaneous constraint holds in all quarters, then, as the contemporaneous constraint also holds for the annual totals, the annual alignment automatically holds for the *M*-th series[5]. So, for every contemporaneous constraint that is added, one annual alignment constraint is redundant and must be dropped. Of course, the set of contemporaneous constraints may also contain redundant constraints within itself. All redundancy must be removed, either manually or automatically.

3.2.5 Missing sources

The model in principle requires quarterly input data for all variables. The value of most variables is usually known from previously compiled quarterly data. However, in some special circumstances it may be necessary to estimate quarterly variables from related indicators. It is possible to do this by hand, but it is more efficient to extend the model. Therefore it is desirable that the model can obtain information about the missing data by using the available sources present.

Let M again be the number of variables in each quarter. However, there are only sources for p variables, where $p < M$. The source information for these p variables is given in vector d .

The relation between d and x is given by matrix C_d . Now it is possible to define x as

---

[4] See for instance Magnus and Neudecker (1988)

[5] Also see figure 2 for an example.

$$x = C_d d \, . \tag{10}$$

Matrix C_d is composed from some of the constraints in C_1. Intuitively, it is clear that for each of the $M - p$ missing values, a constraint is needed. Let v_d be the vector with the coefficients of variation related to vector d and let

$$V_d = \text{diag}(v_d) \, . \tag{11}$$

Matrix V can now be written as

$$V = C_d V_d C_d' \, . \tag{12}$$

Using V and x in our basic model gives the results for all M variables.

3.2.6 Exogenous variables

For policy reasons it is required that some variables can be treated as exogenous variables. Although these variables themselves may not be changed, we still need them in the model as they play a role in one or more constraints. Of course, if they are not to be changed at all, exogenous variables must already have a correct connection to the annual totals and they must be consistent among themselves. Furthermore, if the results will be submitted to a controlled rounding algorithm, the exogenous variables must already be correctly rounded.

A convenient way to implement this, is simply to move the values of all exogenous variables appearing in left hand side of eqn. (2) to the right hand side. So, if in the next example *D* must be treated as an exogenous variable, the restriction

$$\begin{pmatrix} 1 & 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} = 0 \tag{13}$$

becomes

$$\begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} = D \, . \tag{14}$$

We see that in this step the exogenous variables must be removed from the vectors *x* (and *v*), and the corresponding columns in C_1 and C_2 must be removed as well. The values in vector Cx corresponding to exogenous variables are subtracted from vector b.

Note that in this step the constraints on annual connection are removed completely for exogenous variables, leaving empty rows in matrix C_2. Other constraints may have become redundant in this step. Empty rows and redundant constraints must be removed, in order to obtain a new constraints matrix with full row rank. The corresponding values in vector b must also be removed.

3.3 Controlled rounding of the results

The benchmarking model described here yields results that are real values, whereas the requirement of Eurostat is that the values be rounded to integer values, typically representing millions of euros. The rounded results must of course satisfy the same constraints. Intuitively it is clear that the results must also be “as close as possible” to the original results. Controlled rounding can be expressed as an integer linear optimisation problem. One of the possible solutions for this problem is a branch and cut algorithm, for which standard software packages is available.

3.4 Implementation

At Statistics Netherlands we have a working implementation of the benchmarking model described in this chapter. This implementation was successfully applied for benchmarking a time series of Quarterly Sector Accounts ranging from 1998 to 2004. This prototype will be adapted for application in regular production as described in section 2.

The model is implemented using the Excel spreadsheet programme, first to gather all the necessary input, and again to organize and visualize the resulting quarterly series. The actual multivariate Denton model itself was implemented in Matlab. This choice was mainly motivated by Matlab's ability to handle large sparse matrices efficiently. The controlled rounding process is implemented in Express, which is a package that is able to solve integer linear optimisation problems. All input data (both annual and quarterly) are obtained from a dedicated Sector Accounts compilation and database system, developed in-house at Statistics Netherlands. Output data are loaded back into the same system.

4. Input data and preparatory editing

This section is concerned with the input of the automatic benchmarking method, and how the (quarterly) input must be prepared for optimal results. We will first describe the input itself, and then move on to the preparatory editing of the data.

4.1 Input data

The input of the model can be divided in different parts, namely:

- annual totals,
- quarterly indicators,
- reliability weights
- auxiliary information.

About the latter we will be brief. The auxiliary information specifies administrative information about the indicators. They consists of three lists, specifying whether the transaction contains stock data or flow data, whether an indicator must be treated additive, proportional or exogenous and whether the quarterly indicator is missing or not. In the remainder of this section we will describe the first three items somewhat more elaborate.

4.1.1 Annual totals

Each year Statistics Netherlands publishes provisional accounts of last year, revised semi-final accounts of the last but one year and final accounts of the last but two year. The differences between the first provisional publication and the revised semi-final can be substantial. The differences between the updated and the final version are usually minor.

Application of the automatic benchmarking process demands that for any variable, a corresponding series of the latest annual totals are available. The annual totals must satisfy some obvious conditions. First, it is only possible to fulfil both contemporaneous and temporal requirements for the quarterly data if the same contemporaneous requirements are met by the annual totals. Moreover, the quarterly figures must be rounded to integer values after benchmarking. Therefore, the annual totals must already be correctly rounded. If these requirements are not met, the benchmarking problem is ill-posed and does not have a suitable solution.

4.1.2 Quarterly indicators

A consequence of the annual update cycle is that each quarter will be aligned three times to updated annual figures. The first time the quarter will be aligned to the provisional annual figures, the second time to revised annual figures and the third time to final annual figures. This allows for some freedom to choose which sets of quarterly data should be used as input to the model.

In order to preserve the relation with the original data sources one should always use the original quarterly indicators as input and not the benchmarked results of last year. When the initial differences are large, the benchmarking process can make a very noticeable change in the trend of the quarterly series. When using such results as input again, the automatic benchmarking model will treat such benchmarking effects as original quarterly movements and will try to preserve them. Manual changes made in the preparatory editing phase of earlier benchmarking procedures, should of course be included in the quarterly input .

4.1.3 Reliability weights

As the quarterly indicators may be based on many different types of sources, they can differ greatly in reliability. Some may be based on *direct source data*, while others may be obtained from *indirect or related sources.* Some data may be lacking a source at all, and will be estimated during the benchmarking process. In order to find a solution that reflects the origins of the data with some credibility, a measure of the reliability must be specified in the input. As explained in Section 3, the model needs weights that are proportional to the standard deviations expressed as a fraction of the indicator value.

In practice it is nearly always impossible to find reasonable estimates of the standard deviations of all indicators. It is possible however, to use a subjective assessment of the relative reliability of the indicators. The compilers of the original data can be asked to determine the relative reliability on an ordinal scale (ranging from “very poor” to “very good”) within their own sector, because they often have a good idea about the difference in reliability between their own indicators. Next a correction factor is used that should both correct for subjectivity and weigh the average quality of the indicators within one sector. With this method, each indicator obtains two reliability measures. They are combined into a single ordinal measure in the way shown in figure 3 (simplified).

	Reliability per sector		
	High	**Medium**	**Low**
Level 1	High		
Level 2	Medium	High	
Level 3	Low	Medium	High
Level 4		Low	Medium
Level 5			Low

Reliability per transaction

Figure 3. A simplified table for combining "per transaction" and "per sector" reliabilities

Figure 3 shows five overall levels of reliability, from level 1 to level 5. A transaction with medium reliability within a sector with a high reliability obtains level 2. A transaction within a sector with low reliability can at most reach level 3.

The ordinal scale can easily be transformed into an interval scale, which is needed for the model. We can for example transform reliability level 1 to a standard deviation of 1 percent. Reliability level 2 can equally be transformed to a standard deviation of 2 percent, and so on: each additional level adds a factor of 2 to the estimated coefficient of variance.

4.2 Preparatory editing

As the annual figures are updated based upon improved information, so should the quarterly indicators be. Yet, as the automatic benchmarking procedure can efficiently smooth all small inconsistencies, it is only worthwhile to focus on the large initial differences for manual editing. Therefore, the quarterly input consists of the original data, with subsequent corrections for large updates in the annual figures.

The amount of time spent in the pre-processing phase and plausibility check are related. Consider the following two extreme cases.

If enough information is available to reduce all large initial differences, we can put very much effort in preparatory editing. If this process works well, we can assume that all remaining initial differences are small. They can be automatically smoothed with the automatic procedure, without the need for an extensive plausibility check. A drawback of this scenario is that perhaps to much time is spent manually editing the input, as automatically smoothing may deliver equivalent results.

On the other hand, if little time is spent on preparatory editing, the automatic benchmarking process will occasionally have to deal with large differences. In this case a more extensive plausibility check is called for, since the model may not produce an acceptable solution. When the plausibility checks finds unsatisfactory results the process must return to the pre-processing phase, the automatic procedure must be repeated, and the results must then be checked again. This could lead to a very time consuming procedure.

We see from this that the amount of time spent in the pre-processing phase and plausibility check are related. A well focussed pre-processing reduces the need for extensive checks afterwards. In order to find the optimal balance between preparatory editing and plausibility checks, two things are important:

1. a good identification of the cases that require manual intervention;
2. information to solve the inconsistencies.

For each constraint a threshold value can be determined, above which manual editing is required. This threshold value depends on the (relative) size of the inconsistency, the quality of the data sources and the achieved quality of the results. If however, no information is available to solve the inconsistency, the automatic benchmarking procedure has same or even better chances to produce an acceptable solution.

5. Plausibility checks

The model described in section 3 aligns series of quarterly figures to their corresponding annual totals. Three situations with respect to the initial differences between the quarterly and annual figures of a cluster[6] can occur, namely

1. differences are small,
2. differences are large and can to a sufficiently large extent be allocated to one or more specific quarters,
3. differences are large and cannot be allocated.

[6] A cluster is a group of variables that are related by one or more restrictions. A change made to one of those variables can have impact on all other variables in a cluster.

The second situation may arise when additional quarterly information has become available after the regular compilation of the Quarterly Sector Accounts. During the preparatory editing, described in section 4.2, this extra information will be added to the input for the automatic benchmarking model. The benchmarking model can usually smooth the remaining small differences over all quarters involved in a satisfactory way.

If the preparatory editing process cannot make sure that all remaining differences are either in category 1 or 2, the results produced by the model may not be satisfactory due to the large differences that must be smoothed. The reason is that the large initial differences allow for more freedom for the model in allocating these differences, which may give rise to implausible results. So subject matter specialist should focus their plausibility checks on the third category.

The following sections further discuss different ways to check the plausibility of the results.

5.1 Graphical representations

An important tool for evaluating plausibility of time series is a simple graphical representation. The power of a graphical representation lies in the fact that it makes it possible to roughly decompose the series in trend-cycle, seasonal component and incidental effects, without help of any other tools. Several series can be combined in one plot, making it possible to compare series along any conceivable cross section of the data. For obvious reasons, cross sections that group variables, which are related by one or more of the contemporaneous constraints in the benchmarking model, are the most useful. E.g., the evolution of aggregation relations can be made visible in this way, or all different sectors that contribute to the transactions of one single financial instrument.

A particularly useful tool is a time series representation showing both the original series and the benchmarked results. However, it is often difficult to see directly whether an adaptation on a variable is made due to differences between its quarterly values and the corresponding annual total or due to other related variables. Because of that, it is also useful to use a graph of the differences[7] between input and output, in other words, the adaptations made by the model.

Figures 4–6 below illustrate the use of graphical representations on a fictitious example. For simplicity we only consider a single variable. Assume the annual total for the most recent year amounts to three times its ‘usual’ value due to a single huge transaction[8]. Suppose that this transaction has not been covered by the quarterly source data, such that there is a ‘gap’ between annual and quarterly data of approximately the amount of this transaction. If the gap has not been identified and dealt with in the input pre-processing phase, it will be resolved by the automatic

---

[7] Absolute differences for all additive variables and relative differences for all proportional variables.

benchmarking model. Figure 4 shows the results. In order to accommodate the annual-quarterly discrepancy according to Denton's Movement Preservation Principle, the model allocates it to all four quarters of the year such that the quarter-to-quarter movements are respected as good as possible.

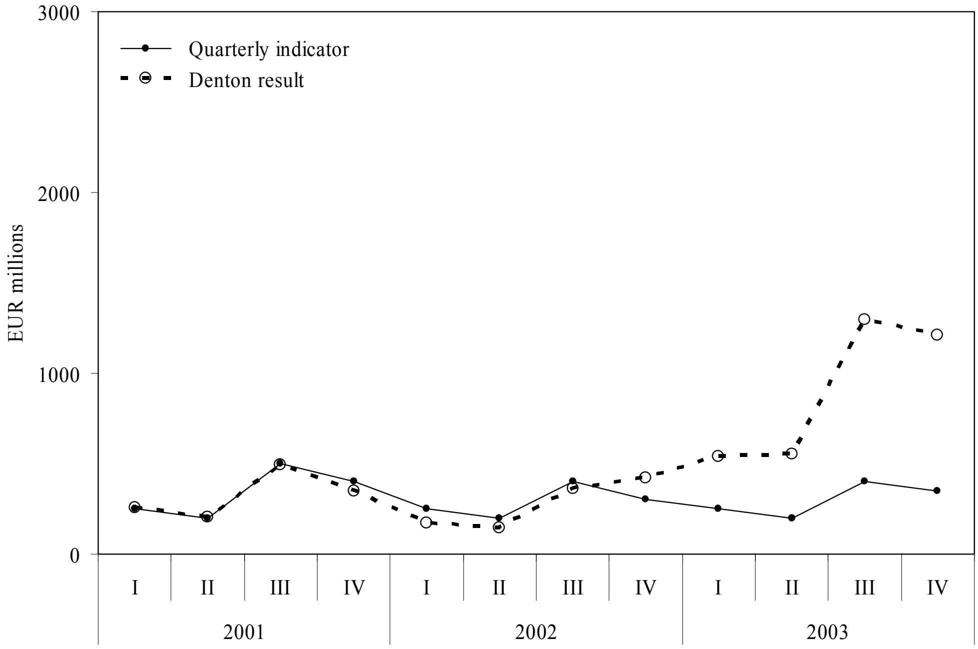

Figure 4. Graphical representation of original (without correction) and bench-marked values of a fictitious quarterly time series

The effects are even more clearly visible when looking at the proportional differences in Figure 5 (open bullets, left scale). It appears that all quarterly amounts for the previous year have been significantly changed, as well as the fourth quarter of the year before that.

Both graphical presentations clearly show that something is wrong, and further inspection is needed. After the missing transaction is identified and the quarterly input series is modified to take it into account, the benchmarking model can be applied again. Figure 6 shows the new result, which obviously is much more satisfactory. This is also clear from the proportional differences in Figure 5 (closed bullets, right scale). It appears that now all quarterly amounts for the previous year as well as the fourth quarter of the year before that, have been changed by only a few percent. Part of these changes, moreover, originate from changes in the annual totals for these years themselves.

---

[8] One can think of, e.g., the impact of the UMTS frequencies sales on the ESA95 category K.2, acquisitions less disposals of non-financial non-produced assets.

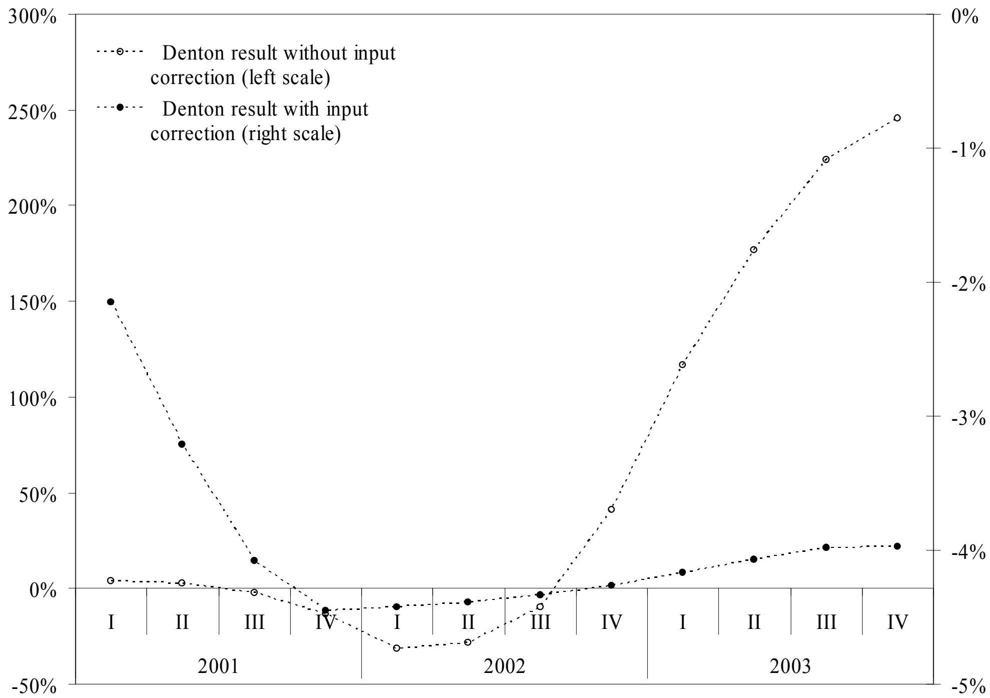

Figure 5. Graphical representation of proportional differences between benchmarked values and original values (with and without correction) of a fictitious quarterly time series

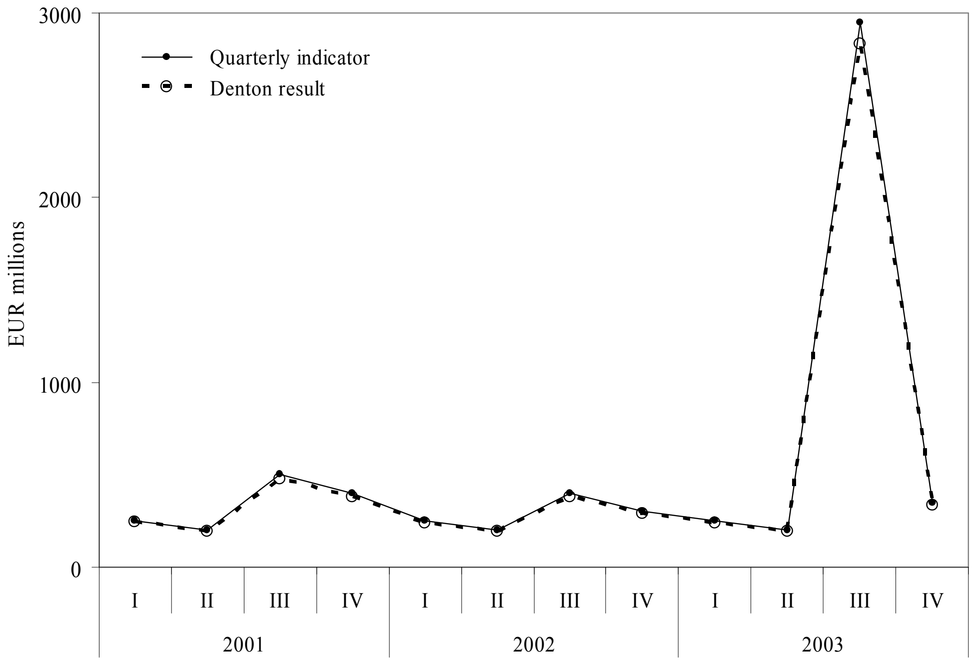

Figure 6. Graphical representation of original (with correction) and benchmarked values of a fictitious quarterly time series

5.2 Balancing items and ratios as plausibility checks

From a macro-economic point of view, balancing items and ratios, derived from underlying variables of revenue and expenditure, are among the most important results of the Quarterly Sector Accounts. Examples of important balancing items are the Gross Domestic Product and net lending/borrowing of the government. An example of an important ratio is the savings ratio of households. Even if the underlying variables seem plausible, the resulting aggregates may be implausible. Therefore, it is important to analyze the plausibility of balancing items and ratios. The graphical tools described above can easily be applied for this goal.

In addition to analysis of direct variables, balancing items and ratios, important information about the economy’s behaviour can be found in quarter-to-quarter growth rates Similarly, it might be interesting to examine the growth between corresponding quarters in subsequent years (in the latter case, seasonal patterns are ignored).

5.3 Statistical discrepancies

The quarterly non-financial accounts and the quarterly financial accounts are partly based on different sources. Moreover, they are presently benchmarked individually, meaning that the budget identity is not enforced by the benchmarking model. This means that statistical discrepancies between net lending/net borrowing measured from the non-financial accounts (B.9) and net lending/net borrowing measured from the financial accounts (B.9F) occur naturally.

The size of statistical discrepancies is a good indicator of the presence of problems in the results and is therefore important to examine. Obviously, annual statistical discrepancy must be taken into account during this plausibility check due to the restriction of annual connection.

5.4 Logical inconsistencies

Earlier, in section 3.2.3, we mentioned the automatic benchmarking procedures cannot handle all types of consistency restrictions. E.g., the benchmarking model cannot assure that stocks are non-negative and also does not handle non-linear constraints. Non-binding restrictions are another example. These are restrictions that hold only approximately. Although the last class of restrictions can be implemented in the model, we did not yet do that. Any restrictions that are not included in the model specification must be checked in the plausibility assessment.

6. Conclusions

This paper presents a benchmarking process centred around an automatic benchmarking procedure. The process roughly consists of three parts:

1. Preparatory editing, to remove the largest inconsistencies.

2. Automatic smoothing of the remaining inconsistencies.
3. A plausibility check on the results.

The process is iterative: if the plausibility check reveals that the results are not satisfactory, they can be improved upon by fine-tuning the input and again applying the automatic algorithm.

In the preparatory editing phase the largest initial differences are resolved by manually editing the input data. The automatic part of the process is based on a Denton-type optimization algorithm under restrictions. This model uses reliability weights and allows for missing indicators and exogenous quarterly series. The plausibility of the benchmarked must be checked afterwards, where the primary focus lies with the variables that have been changed most. Several checks can be performed on the results. The most powerful tools are graphical representations of well chosen cross sections of the data. In addition, economically meaningful aggregates and ratios provide much insight in the plausibility of the underlying results.

In order to be a practical implementation a method must meet some considerations. Statistical data in an accounting framework is usually based on data originating from many different sources, widely varying in reliability. So first of all the method must be flexible enough to deal with all this variety. The multivariate Denton model described in this paper, has proved to be flexible enough to benchmark a time series of Quarterly Sector Accounts ranging from 1998 to 2004. The most important ingredients facilitating this flexibility are the reliability weights, the tolerance for missing indicators and the possibility to specify exogenous quarterly series.

Next, the method must of course produce plausible results. A good benchmarking method does not automatically mean plausible results. The automatic procedure can only produce plausible results, if plausible input is fed into the benchmarking model and if the plausibility is checked by subject matter specialists. The preparatory editing of the data, guided by indicators of the largest differences, and the explicit plausibility checks after benchmarking, provide for this. If the cases that require manual intervention are identified well, the process leads to optimal quality without spending to much time.

Finally, the method must be usable and maintainable in a user friendly way. It is not enough when researchers can operate and maintain the benchmarking model, statisticians and economists without a very profound mathematical background must be able to use it. The benchmarking model must therefore be accompanied with an interface that hides the mathematical intricacies from users with a non-technical background. Ideally, it should be integrated within the sector accounts compilation system. This aspect, however, will be left as a future concern: before taking any definitive implementation steps, we have to gain sufficient practical experience with the benchmarking approach.

References

Broemeling, L.D. (1985), *Bayesian Analysis of Linear Models*, Library of congress cataloguing in publication data, New York

Chow, G.C. and A. Lin (1971), *Best Linear Unbiased Interpolation, and Extrapolation of Time Series by Related Series*, Rev. Economics and Statistics, Vol. 53, Issue 4, pp. 372-375

Denton, F.T. (1971), *Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization*, Journal of the American Statistical Association, Vol. 66, No 333, pp 99-102

Di Fonzo, T., and M. Marini (2003), *Benchmarking systems of seasonally adjusted time series according to Denton's movement principle*, University of Padova

Di Fonzo, T., and M. Marini (2005), *Benchmarking a system of time series: Denton's movement preservation principle vs. data based procedure*, University of Padova

European Commission (1996), *European System of Accounts 1995 (ESA 1995)*, Office for Official Publications of the European Communities, Luxembourg

Harvey, A.C. (1990), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press

Laniel N. and K. Fyfe (1989), *Benchmarking of economic time series*, American Statistical Association, Proceedings of the Business and Economic Statistics Section

Magnus J.R. and H. Neudecker (1988), *Matrix differential calculus with applications in statistics and econometrics*, New York, Wiley.

Stone, J.R.N., D.G. Champernowne and J.E. Meade, (1942), *The Precision of National Income Estimates.* Reviews of Economic Studies, vol. 9, pp. 111-135.