A Multivariate Denton Method for Benchmarking Large Data Sets

Jacco Daalmans[1], Nino Mushkudiani[2]

[1]Affiliation: j.daalmans@cbs.nl

[2]Affiliation: n.mushkudiani@cbs.nl

Abstract

Benchmarking is a process of achieving mathematical consistency between low-frequency (e.g. annual) and high-frequency (e.g. quarterly) data. Statistics Netherlands is planning to apply a new benchmarking method on Dutch national accounts.

To incorporate the economic relations of Dutch national accounts into the model, new methodological features, such as ratio constraints, soft constraints and inequality constraints were needed. In this paper we extended the multivariate Denton method of Bikker and Buijtenhek (2006) with these methodological features. The extended method is formulated as an easy understandable quadratic programming problem.

To apply this method to Dutch national accounts it is crucial that the software can cope with large data sets, i.e. over 100,000 variables. Such large-scale applications of reconciliation methods are rare in the literature. In this paper we show, however, for these large problem sizes we can find solutions, by using a state-of-the-art optimization solver like CPLEX.

To test whether its possible to incorporate the constraints of the Dutch supply and use tables in the proposed model, a simulation experiment was conducted. The results of this experiment are excellent: all comments of national account experts could be adequately incorporated in the model. Also, the comparison of the results with published figures of Dutch national account does not show much (unexplainable) deviation.

Keywords: Benchmarking, Denton method, National accounts

1. Introduction

National Account data have to satisfy certain accounting rules. Since different data sources and different statistical compilation methods are used for annual and quarterly data, violations of the accounting rules occur naturally. Macro-integration is a process of achieving consistency. The first step of this process consists of the correction of errors. In the second step a balancing process is carried out. This paper focuses on the second step only.

The literature on this topic goes back to Stone et al. (1942), who presented a constrained, generalised least squares method, in which variances are used as reliability weights. The size of the adjustments that are made to the data are negatively correlated to the reliability of the data. Several other macro-integration methods are described in Wroe et al. (2004, Annex A).

This paper deals with a specific macro-integration problem called benchmarking. Benchmarking is the process to make high frequency and low frequency data consistent. Without loss of generality, henceforth it is assumed that the time series are quarterly data and the benchmarks are annual figures. Typically, the annual data sources provide the most reliable information about overall levels and the quarterly data sources provide information about short-term changes. Thus, the aim of benchmarking is to make the quarterly data coherent with annual totals, while preserving all quarter-to-quarter changes as much as possible. The latter is called the movement preservation principle (Denton, 1971).

The benchmarking methods in the literature can be broadly classified into purely numerical methods and model-based methods. Bloem et al. (2001, Chapter VI) and Dagum and Cholette (2006) give a comprehensive overview of methods.

In the model-based field we have regression models, for instance Cholette-Dagum (1994), ARIMA model-based methods proposed by Hillmer and Trabelsi (1987) and state space models proposed by Durbin and Quenneville (1997). Closely related to the regression method is the method of Chow and Lin (1971) for the interpolation of time series, i.e. deriving quarterly data from annual data, although this is not a benchmarking method in the strict sense. The Chow and Lin method may suffer from step problems, i.e large gaps between the fourth quarter of one year and the first quarter of the next year.

A classical reference to a method in the numerical field is Denton (1971). The Denton method is a quadratic programming method and was initially proposed for univariate data. As the method is based on the movement-preservation principle, it avoids the step problem. Di Fonzo and Marini (2003) have extended the Denton method for multivariate data. In addition to temporal alignment, multivariate data often also have to satisfy a set of contemporaneous constraints (i.e. restrictions between different time-series in one time-period). Subsequently, Bikker and Buijtenhek (2006) have added reliability weights to the multivariate Denton method.

In order to incorporate all economic relations of the national accounts to the model, new methodological features had to be added to the multivariate Denton method of

Bikker and Buijtenhek (2006). These involve: soft constraints, ratio and inequality constraints. Magnus et al. (2000) already have incorporated these features into a reconciliation method, more general than for benchmarking purposes.

Although it is mentioned by Bloem et al. (2001) that the Denton method is well suited for large scale applications, to the knowledge of the authors, such applications have not been performed at NSI's (National Statistical Institutes). The lack of adequate technology used to be an obstacle (Nicolardi, 2000).

Statistical Netherlands is planning to implement an extended version of the multivariate benchmarking method of Bikker and Buijtenhek (2006) in its production process of national accounts in the near future. This application requires software that can cope with large data sets, namely over 10,000 time series. The software will be based on a state-of-the-art, commercial quadratic programming (QP) solver. The Bureau of Economic Analysis (BEA) also used similar software for the implementation of a reconciliation method (Chen, 2006), but their method is not aimed at benchmarking.

A simulation experiment was conducted on data of Dutch supply and use tables to find out whether all economic constraints that are used in the current benchmarking process can be included in the model. The experiment was successful and now Statistics Netherlands is in the process of implementing the method into its production system.

In Section 2 we describe the univariate and a multivariate Denton method and its extensions. Section 3 deals with the current and future benchmarking process of Statistics Netherlands, the software implementation and the experiment on Dutch national accounts.

2. The model

2.1 The univariate model

In this section we discuss the classical Denton method. The aim of this method is to find a benchmarked time-series $\hat{x}_t$, $t = 1,\ldots,T$, that preserves as much as possible all quarter-to-quarter changes of the original series x_t, subject to annual benchmarks.

Denton proposed several measures to define the quarter-to-quarter changes. We consider the proportional first-order function and the additive first-order function. The additive function keeps additive corrections ($\hat{x}_t - x_t$) as constant as possible over all periods. The proportional function is designed to preserve the (proportional) growth rates of x_t and therefore keeps the relative corrections ($\hat{x}_t / x_t$) as constant as possible over all periods.

In mathematical terms the additive Denton model is defined as follows:

$$\min_{\hat{x}} \sum_{t=2}^{T} \left((\hat{x}_t - x_t) - (\hat{x}_{t-1} - x_{t-1}) \right)^2 \tag{2.1}$$

$$\text{such that} \sum_{t=4(j-1)+1}^{t=4(j-1)+4} \hat{x}_t = y_j \qquad j = 1, \ldots, T/4 \tag{2.2}$$

and the proportional Denton model is given by

$$\min_{\hat{x}} \sum_{t=2}^{T} \left((\hat{x}_t / x_t) - (\hat{x}_{t-1} / x_{t-1}) \right)^2 \tag{2.3}$$

$$\text{such that} \sum_{t=4(j-1)+1}^{t=4(j-1)+4} \hat{x}_t = y_j \qquad j = 1, \ldots, T/4 \tag{2.4}$$

where j is an index of the year and y_j is an annual value. The set of restrictions expresses the alignment of four quarters to annual totals.

2.2 The multivariate case

The extension of the univariate Denton model to the multivariate case of Bikker and Buijtenhek (2006) is straightforward. The multivariate, additive model is given by

$$\min_{\hat{x}} \sum_{i=1}^{N} \sum_{t=2}^{T} \frac{1}{\left(w_{it}^{A}\right)^2} \left((\hat{x}_{it} - x_{it}) - (\hat{x}_{it-1} - x_{it-1}) \right)^2 \tag{2.5}$$

$$\text{such that } C^H \hat{x} = b^H, \tag{2.6}$$

where i is the index for the time-series, N denotes the total number of time-series, and w_{it}^{A} is a reliability weight of the i'th time-series. The set of restrictions involves both contemporary (i.e. over different time-series and one period) and intertemporal (over different periods and one time-series) restrictions.

The extension for the proportional model is similar. In Bikker and Buijtenhek (2006) the proportional and the additive models are combined. The user has to specify for each time-series whether the model is proportional, additive or fixed, that last one is discussed in the following subsection.

2.3 Soft constraints, fixed quarters, ratio constraints, and inequalities

In this subsection we propose extensions to the multivariate Denton method.

At Statistics Netherlands a lot of subject matter knowledge is used in the reconciliation of national accounts. For instance: for some perishable goods the value of the change of stocks, summed over the four quarters of one year, should not differ much from 0. In order to include such knowledge into the model soft constraints are needed. A set of soft constraints is given by

$$C^S \hat{x} \sim (b^S, (w^C)^2). \tag{2.7}$$

where b^S is a vector with target values, and w^C is a vector of weights. Similar notation will be used throughout this section.

In the literature variances are used to define the weights, but in practice these variances are often unavailable. In our model $(w^C)^2$ can be chosen at will. Only the relative values of the weights determine the outcome of optimization procedure.

The superscript S denotes that the constraints are soft (where the superscript H in (2.6) indicates that a constraint is hard). These constraints are included in the model by adding the following penalization terms to the objective function (2.5).

$$+\sum_{r=1}^{R^S} \frac{1}{(w_r^C)^2} \left(b_r^S - \sum_{t=1}^{T} \sum_{i=1}^{N} c_{rit}^S \hat{x}_{it} \right)^2, \tag{2.8}$$

where r is the index for the soft constraints and R^S denotes the total number of constraints.

Although the Denton method focuses on the preservation of quarter-to-quarter, changes, in some cases it may be necessary to stick to a quarterly value, for instance if that value already published. Hard and soft fixed quarters are respectively denoted by

$$\hat{x}_{it.} = x_{it} \text{ and } \hat{x}_{it} \sim (x_{it}, (w_{it}^F)^2). \tag{2.9}$$

Hard fixed quarters are linear equalities and are added to the set of hard constraints (2.6). Soft, fixed quarters are incorporated in the model, by adding the following term to the objective function (2.5)

$$+\sum_{i,t \in F} \left(\frac{\hat{x}_{it} - x_{it}}{w_{it}^F} \right)^2. \tag{2.10}$$

where F is the set of indices of all fixed quarters.

Many economic indicators are ratios of national accounts variables, for example the relation between value added and the output of an industry. To describe these relations hard and soft ratio constraints are added to the model. These constraints are given by

$$\hat{x}_{nt} / \hat{x}_{dt} = v_{ndt} \text{ and } \hat{x}_{nt} / \hat{x}_{dt} \sim (v_{ndt}, (w_{ndt}^R)^2) \tag{2.11}$$

where x_{nt} denotes the numerator time-series, x_{dt} denotes the denominator time series and v_{ndt} is some predetermined value. Following the approach of Magnus et al. (2000), ratio constraints are linearised in the following way

$$\hat{x}_{nt} - v_{ndt}\hat{x}_{dt} = 0 \text{ and } \hat{x}_{nt} - v_{ndt}\hat{x}_{dt} \sim (0,(w_{ndt}^{R'})^2). \qquad (2.12)$$

These linearised ratios are added to (2.6), and to (2.7) respectively.

Note that, essentially, there is no difference between hard linear constraints and hard linearised ratio's, but there is a difference in the soft case. For soft linearised ratios, the value of v_{ndt} is uncertain. But in the case of a linear constraint the randomness is in the values of b_r^S. Due to this difference the weights of soft ratio's are computed in a different way than the weights of soft, linear constraints.

Most economic variables are not allowed to have negative signs. To incorporate this (and other) requirement(s) in the model some inequality constraints are needed. A set of inequalities

$$A\hat{x} \le z. \qquad (2.13)$$

can be easily added to the model.

The problem, defined by (2.5), (2.6), (2.8), (2.10) and (2.13) is a standard (convex) quadratic programming (QP) problem. The problem is well known in the literature, and many efficient solving techniques are available.

2.4 Weights

In the objective function of the aforementioned model several kind of weights are used (weights of additive and proportional changes, soft fixed quarters, soft ratio's and soft, linear constraints). The weights have to be determined, such that the model meets certain properties, see Öhlén (2006).

Expressions for all kind of weights will be proposed in an upcoming paper. These expressions focus on two model properties:

1) Invariance: If all input data are multiplied by the same nonnegative scalar, the outcomes must also be changed by this factor;

2) Ratio symmetry (mentioned by Magnus et al. 2000): The outcome will not change if a soft ratio in the benchmarking model is replaced by its reciproke, i.e. if the restriction $x / y \sim (r,(w_{x/y}^R)^2)$ is replaced by $y / x \sim (1/r,(w_{y/x}^R)^2)$.

3. The application of the model

3.1 Current and intended benchmarking process

In this subsection we explain that the current benchmarking method of the Dutch supply and use tables can be improved by implementing the model described in Section 2.

The current compilation of supply and use tables is carried out as follows. Each year t, annual national accounts are compiled of the years $t - 1$, $t - 2$ and t – 3. After each compilation, quarterly accounts are benchmarked to the new annual accounts. Currently, the benchmarking process is done for each estimate of the annual accounts independently (i.e. three times a year). As a consequence spurious gaps can arise between the last quarter of one year and the first quarter of the next year. (the step-problem).

In the future it is intended to apply the benchmarking process to the three years $t - 1$, $t - 2$ and $t - 3$ simultaneously, by using the multivariate Denton method. Thus, the results will not suffer from the step-problem (at least within the period of three years).

Moreover the application of the Denton model to Dutch national account will enable Statistics Netherlands to improve the efficiency of compiling national accounts. The balancing process of the annual accounts of the years $t - 1$ and $t - 2$ will be skipped. Hence, at the start of the benchmarking process, the annual data of $t - 1$ and $t - 2$ will not satisfy the accounting rules within each year. Both the consistency between annual and quarterly accounts and the consistency within each time period will be achieved by using the multivariate Denton method.

An outline of the new process is as follows: The quarterly values of $t - 3$ will be aligned to the reconciled, annual accounts of $t - 3$. The benchmarked quarterly values will be extrapolated to the quarterly values of $t - 1$ and $t - 2$, by using the initial quarter-to-quarter changes and the annual values of these years. The annual accounts of $t - 1$ and $t - 2$ will be defined by the sum of the underlying quarterly accounts. By applying the multivariate Denton method each of these steps are performed simultaneously.

3.2 Software

In order to be useful for practical implementation at Statistics Netherlands, the benchmarking software should be able to handle very large data sets.

Eurostat developed ECOTRIM (Barcellan, 2004), a software tool which supplies several univariate and multivariate benchmarking and temporal disaggregation techniques, including the multivariate Denton method proposed by Di Fonzo and Marini (2003). Although ECOTRIM is a useful tool, it does not satisfy the requirements imposed by Statistics Netherlands for benchmarking national account data. The reasons for this are that ECOTRIM is not designed for dealing with thousands of time-series, it does not include features like weights, soft constraints, and the possibility to combine the proportional and additive methods of benchmarking into one model.

Statistics Netherlands has built a prototype of benchmarking software, using CPLEX (ILOG, 2008) as solver. This state-of-the-art, commercial optimization solver is able to cope with very large data sets. A benchmarking problem with 13.790 time series, each consisting of up to 3 annual, and 12 quarterly values, was translated into a quadratic optimization problem with approximately 160,000 free variables and 75,000 constraints. By using CPLEX on a 2.00 GHZ, Xeon E5335 with 1024 MB Ram, the optimal solution was found in 45 seconds.

However, in order to conduct the intended benchmarking of Dutch supply and use tables, an optimization problem will have to be solved, with approximately 500,000 free variables and 120,000 constraints. We also performed tests with the simulated data of these large sizes. The prototype software found the optimal solutions. However the constraints for these tests were different than the ones of the supply and use tables.

3.3 A first simulation experiment

A simulation exercise was conducted on the Dutch supply and use tables of 2004-2006. The aim of the experiment was to find out whether all requirements of statistics Netherlands can be incorporated in the model. There was no intention to obtain results that fully satisfy all economic relations. Thus, we did not use exactly the same data as in the reconciliation process of Statistics Netherlands, and we did not incorporate all relations among the variables of the model. For some classes of similar constraints we only worked out examples.

An important feature of the supply and use tables is that they involve fixed price data and current price data. Both kinds of data were used in the experiment. By definition, ratios of current price values and fixed price values are price changes. The values of these price changes had to be preserved as much as possible.

The initial model parameters (like weights and the choice whether a proportional or additive or fixed model should be applied to some time-series) were determined together with experts of the national account department. Parts of the supply and use tables were aggregated, to keep the model as small as possible. As we have mentioned before this model still involves 13.790 time series.

The outcomes of the model were reviewed by subject matter specialists. Their comments were translated into changes of the model set up, for instance different weights, other and or more constraints.

The results of this experiment are excellent. All remarks of subject-matter specialists could be properly translated into changes of the model set-up. However, this was not always trivial. For instance, one expert argued that the values of one ratio of two variables should be about the same as the value of an other ratio. Since the Denton method, presented in Section 2, is not able to cope with this type of constraint, ad-hoc solutions were implemented, namely including different constraints, that have a similar effect on the outcomes of the model.

Although only the most important economic relations were incorporated in the model, most of the outcomes closely resemble the published figures of the supply and use table: the difference was smaller than 0.3%-point for most of the main economic indicators.

An explanation for this good result is that the initial discrepancies of most constraints were small. As the benchmark process starts with compiled systems of national accounts, there were no violations of the constraints within the quarters and years. The only discrepancies were between quarterly and annual figures, but these were mainly small. However, in some exceptional cases the initial discrepancies of the constraints were large. The information on how to deal with these inconsistencies are not included in the model. This lead to undesirable results. Since large initial discrepancies in the data are rare, and the desired solution is different for each problem, it is not feasible to adapt the model for each case, beforehand.

Therefore, in the future application, large discrepancies of constraints will be detected and reconciled by hand, before the benchmarking is carried out.

References

Barcellan, R. (1994) ECOTRIM: a Program for Temporal Disaggregation of Time Series, Paper presented at the joint INSEE and Eurostat Quarterly National Accounts Workshop in Paris on December 5th and 6th, 1994

Beaulieu J.J., Bartelsman E.J. (2004**)** Integrating Expenditure and Income Data: What to Do with the Statistical Discrepancy? Finance and Economics Discussion Series 2004-39, Board of Governors of the Federal Reserve System (U.S.)

Bikker R.P., Buijtenhek S. (2006) Alignment of Quarterly Sector Accounts to annual data, CBS Voorburg, http://www.cbs.nl/NR/rdonlyres/D918B487-45C7-4C3C-ACD0 oE1C86E6CAFA/0/Benchmarking_QSA.pdf

Bloem A.M., Dippelsman R.J., Maehle N.O. (2001) Quarterly National Accounts, Manual–Concepts, Data sources and Compilation, IMF, Washington, D.C

Chen, W. (2006) Balanced System of Industry Accounts for the U.S. and Structural Distribution of Statistical Discrepancy, Paper presented at the NBER-CRIW Summer workshop, July 17 – 19 2006, Cambridge

Chow G.C., Lin, A.L. (1971) Best linear unbiased interpolation, distribution and extrapolation of time series by related series, *The Review of Economics and Statistics*, 53, 372–375

Cholette P.A., Dagum E.B. (1994) Benchmarking time series with autocorrelated survey errors. *International Statistical Review*, 62, 365–377

Dagum E.B., Cholette P.A. (2006) Benchmarking, Temporal Distribution and Reconciliation Methods for Time Series, Springer, New York

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*, 66 (333), 99–102

Di Fonzo T., Marini M. (2003) *Benchmarking systems of seasonally adjusted time series according to Denton's movement preservation principle*. University of Padova, Working Paper 2003.09 http://www.oecd.org/dataoecd/59/19/21778574-.pdf

Durbin J., Quenneville B. (1997) Benchmarking by state space models. *International Statistical Review*, 65, 23–48

Hillmer S. C., Trabelsi A. (1987) Benchmarking of economic time series, *Journal of the American Statistical Association*, 82, 1064–1071

ILOG, S.A. (2008) CPLEX 11.0 Reference manual and User's Manual, 2008

Magnus J.R., Tongeren J.W. van, Vos A.F. de (2000) National Accounts Estimation Using Indicators Analysis, *Review of Income and Wealth*, (46), pp. 329–350

Magnus J.R., Danilov D. (2008) On the estimation of a large sparse Bayesian system: The Snaer program, *Computational Statistics & Data Analysis*, 52, pp. 4203–4224

Nicolardi, V. (2000) Balancing Large Accounting Systems: An Application to the 1992 Italian I-O Table, paper presented at the 13th International Conference on Input-Output Techniques, University of Macerata, Italy

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

Wroe D., Kenny P., Rizki U., Weerakkody I. (1998) Reliability and quality indicators for National Accounts Aggregates. ONS (London)