Universitá degli Studi di Napoli
Federico II

Ph.D. Thesis

in Statistics

XXVIII Ciclo

Two-Step Reconciliation of Time Series

New Formulation and Validation

Enrico Infante

Dipartimento di Scienze Economiche e Statistiche

Two-Step Reconciliation of Time Series

New Formulation and Validation

Author:

Enrico Infante

Universitá degli Studi di Napoli

Dipartimento di Scienze Economiche e Statistiche

email: enrico.infante@gmail.com

Tutor:

Germana Scepi

April 2017

to my matryoshka

Two-Step Reconciliation of Time Series

New Formulation and Validation

Enrico Infante

Abstract

Two-step reconciliation methods solve the temporal constraint in the first step, while in the second step the contemporaneous constraint is satisfied without altering the temporal constraint. Both in Quenneville and Rancourt and in Di Fonzo and Marini methods, the methodology used applies the Denton benchmarking technique in the first step.

The work done in this study is based on an alternative two-step procedure for the reconciliation of systems of time series, proposing an algorithm which allows to choose one of the two different solutions for the second step, and introduces the possibility of using well-known established techniques in the first step, such a Chow and Lin, Fernández and Litterman. Furthermore, a way of dealing with the reconciliation of hierarchical systems of time series is presented. An innovative test for detecting common seasonal patterns in time series is also presented. Such test could be used for deciding at which level to seasonally adjust an aggregated time series before applying reconciliation.

Moreover, together with a simulation study, several aspects of the validation of a reconciliation technique are shown, including a new methodology for detecting whether the outliers at the end of series are consistent. Two real examples using the European industrial production index and the euro area quarterly sector accounts data will also be presented.

Acknowledgements

The main word characterising all the people which helped me in this work is probably *patience*. For this reason I have to thank Prof. Germana Scepi, who has helped me throughout this project with a lot of *patience*. The number of emails we exchanged in the latest three months is just incredible.

I also have to thank all the other professors in the department, starting from Prof. Carlo Lauro, who has shown a lot of *patience* in not seeing me enough times.

I thank Marco Marini for his comments on this work, as well as Dario Buono and Gian Luigi Mazzi, my fellows in many conferences. The main idea of the work done in this dissertation hit me during a course given by Tommaso Di Fonzo.

Help was also given from my Ph.D. colleagues, in particular Francesco and Rosanna, which I thank for all their advices.

Clearly, any possible errors that remain are my sole responsibility.

Patience was also shown by my colleagues of the sector accounts team in Eurostat, which have seen me discussing about this dissertation for many months, and have replaced me many times in the latest months, while I was taking days off for finalising it. In particular, I would like to thank the Greek crowd: Orestis for his *patience* at introducing me to the wonderful world of LaTeX; and Christos, for his *patience* when I was literally consuming his coffee machine.

Writing a Ph.D thesis and working at the same time is one of the most challenging things I have ever done. It includes long hours during evenings and sunny weekends spent at home in front of a computer. For this reason, I have to thank my lovely wife Josephine for her great *patience* during all these evenings and weekends. On top of that, she also managed to review the English of this dissertation. Thank you for everything you do, Josephine.

I also have to thank my "Eggplant" friends for their *patience* during the latest months. In particular, I have to thank Alessio, the witness, for his *patience* at introducing me to the Java coding.

Finally, I also have to thank my whole family for all the support and their *patience* in dealing with all the bureaucracy involved.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Time series reconciliation techniques are widely used in practice. The most common applications are in official statistics, and in particular in national accounts, or after performing seasonal adjustment. Usually the data obtained by a direct seasonal adjustment procedure do not sum up to the total series and the annual totals are not in line with the non-seasonally adjusted figures. Two data restrictions are encountered in these cases.

The first restriction is referred to as the contemporaneous constraint: the linear, or non-linear, combinations of the variables under examination are to be fulfilled for each observed period. Techniques for solving the contemporaneous constraints are called balancing techniques.

The second restriction is referred to as the temporal constraint: the high frequency time series are to be in line with the low frequency aggregates.

Techniques for solving the temporal constraints are called benchmarking or temporal disaggregation techniques.

This dissertation refers to high frequency (HF) and low frequency (LF) time series, indicating at which frequencies the time series are observed (e.g. HF is quarterly and LF is annual). In other words, HF and LF are used to distinguish between higher and lower frequencies respectively, and do not indicate the very high frequency time series which are dealt with big data analytics. This is in line with the literature in the field (see, for example, Ciammola et al., 2005). Moreover it must be possible to entirely aggregate the HF to the LF (for example the HF and the LF cannot be bi-monthly and quarterly, respectively).

Benchmarking and temporal disaggregation techniques are in theory two different types of techniques, but in practice they often overlap. Benchmarking techniques are designed to solve the temporal constraints starting from preliminary HF estimations. Temporal disaggregation techniques are designed to obtain HF figures starting from LF values. However, in the majority of the cases, the temporal disaggregation techniques use at least one related time series, which could be used to derive preliminary HF estimations or HF benchmarked data directly. In view of this, benchmarking techniques could be considered as a subset of the temporal disaggregation techniques. In this dissertation the two terms will be used synonymously.

Finally, reconciliation techniques are defined as the statistical processes that aim to restore consistency in a system of time series as regards to both contemporaneous and temporal constraints. Therefore, according to this definition, reconciliation techniques embed both balancing and temporal disaggregation methods.

Balancing and temporal disaggregation techniques have both been long discussed in the literature.

A balancing framework, which takes into account the differences in the accuracy of the preliminary estimates of the variables in the system, has been developed by Stone et al. (1942). Bacharach (1970) discussed the RAS method for matrix balancing, which was introduced in the first half of the nineteenth century, also referred to as the bi-proportional adjustment (see also Stone, 1961).

A very well established literature is available for temporal disaggregation methods. Boot et al. (1967) proposed a smoothing technique in order to preserve the trend of the LF series, while Denton (1971) developed a methodology which reallocates the discrepancies of HF series using preliminary estimates. However, the Denton method is often considered as a mathematical (mechanical) method, while Chow and Lin (1971) worked on optimal regression models in the sense of least squares. Starting from the latter, several methodological variants have been developed, Fernández (1981) and Litterman (1983) being the most predominant. Techniques using dynamic models have also been discussed (see, particularly, Santos Silva and Cardoso, 2001).

Reconciliation techniques have been recently discussed by several authors. Simultaneous approaches, as generalisations of either the Denton or the Chow-Lin methodology, have been presented by some authors (see Di Fonzo and Marini, 2003 and Di Fonzo, 1990, respectively). Two-step approaches, which deal first with the temporal constraint and then by the contemporaneous constraint, have been proposed by Quenneville and Rancourt (2005) and by Di Fonzo and Marini (2011b). Both approaches have the first step in common, whereby the Denton methodology is applied. The reason for applying two-step approaches is justified by the authors since the computational burden might be significant using a simultaneous approach, and also because this more simple approach preserves the original movements of the time series in the second steps.

The work done in this dissertation is based on an alternative two-step procedure for the reconciliation of systems of time series, proposing an algorithm, developed with Java, which allows the choice of one of the two different solutions for the second step, and introduces the possibility of using well-known established techniques in the first step. Although the general reconciliation and balancing techniques deal with both linear and non-linear accounting restrictions, the methodology presented here takes into account only the case of linear combinations, avoiding the non-linear case, which for instance arises when the aggregates are expressed in both current and constant prices (Di Fonzo and Marini, 2007). A schematic presentation of the most used methods for benchmarking, balancing and reconciliation is proposed together with a methodology for validating the results of such approaches, which could be used in general for time series. This study also focuses on practical problems encountered by statistical agencies when dealing with a reconciliation problem in the production of official statistics, ranging from timeliness to validation of the results.

This dissertation is divided into six chapters.

In Chapter 2, a critical review of the literature on temporal disaggregation, balancing and reconciliation of time series will be presented, including an innovative schematic presentation of the mostly used methods. Such kind of schematic overview does not exist in the literature, and it will help putting some order in the field.

Chapter 3 will present an innovative two-step approach for the reconciliation of systems of time series, which keeps the second step unchanged, as presented by the above mentioned authors, while an optimal methodology is applied in the first step. The method is also able to deal with multiple systems of time series when they are nested, using a hierarchical dependent reconciliation procedure. The algorithm used will also be presented. Particular attention will be given to the case of the reconciliation of time series after the seasonal adjustment, presenting the statistical test proposed by Infante et al. (2015), which would help the user to obtain preliminary seasonally adjusted series closer to the contemporaneous constraints.

In Chapter 4, the validation and assessment criteria of reconciliation techniques will be presented, starting from a set of measures of the distance between the preliminary and the reconciled series. An innovative methodology for assessing the quality of time series, including results of a reconciliation technique, will be discussed. Finally, a simulation study of the methodology presented in the previous chapter will also be discussed in this section.

Chapter 5 will be dedicated to the application of the methods on real data sets: the European industrial production index, seasonally adjusted according to a geographical direct approach, and the European quarterly sector accounts. These are typical examples of reconciliation problems encountered in official statistics.

Finally, the conclusions will be discussed in Chapter 6.

Chapter 2

Temporal Disaggregation, Balancing and Reconciliation

Techniques for adjusting a matrix of provisional data are widely used for estimation purposes, where the final estimation should be consistent with the marginal totals of the matrix. One of the possible applications is when dealing with time series in official statistics, for example in national accounts (Eurostat, 2013). In this case, the marginal totals are the contemporaneous constraints and the temporal constraints, meaning that each LF period has its marginal total. The most common case is when the temporal marginal total is the annual figure and the HF series are measured at quarterly or monthly level.

As described in the introduction, temporal disaggregation (or benchmarking) techniques are used to solve the temporal constraints, while balancing techniques are used to solve the contemporaneous constraints (or accounting constraints, as they are sometimes referred to in national accounts). Finally, reconciliation techniques are used when both temporal and contemporaneous constraints are to be solved.

A description of some the techniques which will be described here is in Dagum and Cholette (2006).

2.1 Temporal Disaggregation

Temporal disaggregation techniques could be broadly divided into two main categories: techniques which use a related HF indicator (or series), and techniques which only use the original LF series. Other possible ways of classifications could be done. For example, one can classify the techniques which either use statistical models or not. The ones which do not use statistical models, such as the Denton procedure, are sometimes referred to as mathematical methods. However, all methods have certain statistical properties. Thus, categorising the different methodologies according to whether they use or they do not use a related indicator seems to be the best approach.

Related indicators are sometime referred to as preliminary series. A preliminary series is a preliminary estimate of the variable of interest, expressed in the same measurement unit. A related series is a proxy of the variable of interest, possibly not expressed in the same measurement unit (even more than one related series could be used). Since preliminary series could be used as related series, and related series could be used to derive preliminary series, there is no general difference among the two to be considered, and in this study the expressions of preliminary and related series will be considered synonymous, unless otherwise specified.

A description of temporal disaggregation techniques should involve distribution, interpolation and extrapolation. However, according to Chow and Lin (1971), from the theoretical point of view, the distinction between distribution and interpolation is not justified and therefore they demonstrate how to obtain results in the same framework.

Temporal disaggregation techniques are associated with flow or index (average stock) series, whereby the LF data correspond to the sums or averages of the HF data for each LF observation. Since in this case, HF data are obtained from temporal distribution of LF data, this matter is referred to a distribution problem.

An interpolation problem occurs when dealing with end-of-period (EOP) or beginning-of-period (BOP) stock time series, whereby the LF values are equal to that of the last (or first, respectively) HF observations in the LF times (e.g. the yearly value is equal to the fourth quarter).

Finally, extrapolation refers to the generation of values outside the temporal range of the data, and can be both backward or forward. In other words, estimates of HF data are needed when the related LF value is not yet available.

Notation

When dealing with temporal disaggregation techniques, it is important to identify a notation which holds for all the methods described. Fortunately, a common notation has been identified by many authors: Chow and Lin (1971) introduced the matrix $\mathbf{C}$, while Ciammola et al. (2005) or Chamberlin (2010) provide a good explanation.

Given:

- $y_{H,t}$, $t = 1, 2, ..., n$ the HF series.
- $y_{L,T}$, $T = 1, 2, ..., N$ the LF series.
- s the temporal aggregation order (for example if LF is annual and HF is quarterly, it will be $s = 4$).
- $p_{H,t}$, $t = 1, 2, ..., n$ the preliminary series.
- $\mathbf{X}_{H,t}$, $t = 1, 2, ..., n$ matrix of k related indicators.

According to the nature of the series, different temporal aggregation constraints are defined:

- Flow series: $\sum\limits_{t \in T} y_{H,t} = y_{L,T}$.
- Index series: $\frac{1}{s} \sum\limits_{t \in T} y_{H,t} = y_{L,T}$.
- Stock, EOP series: $y_{H,sT} = y_{L,T}$.
- Stock, BOP series: $y_{H,s(T-1)+1} = y_{L,T}$.

Such constraints could be shown as linear combinations, for each period $T = 1, \ldots, N$:

$$y_{L,T} = c_1 y_{H,s(T-1)+1} + \ldots + c_s y_{H,s(T-1)+s} = \sum_{i=1}^{s} c_i y_{H,s(T-1)+i}$$

Where the s $\times$ 1 vector $\mathbf{c}$ assumes different forms according to the nature of the series:

- Flow series: $\mathbf{c} = (1, 1, \ldots, 1, 1)'$.
- Index series: $\mathbf{c} = \left(\frac{1}{s}, \frac{1}{s}, \ldots, \frac{1}{s}, \frac{1}{s}\right)'$.
- Stock, EOP series: $\mathbf{c} = (1, 0, \ldots, 0, 0)'$.
- Stock, BOP series: $\mathbf{c} = (0, 0, \ldots, 0, 1)'$.

Hence it is possible to define the temporal aggregation matrix as:

$$\mathbf{C} = \mathbf{I}_n \otimes \mathbf{c}' \tag{2.1}$$

Where $\otimes$ is the Kronecker product.

The formulation of the temporal aggregation with the $\mathbf{C}$ matrix will therefore change according to the nature of the data:

- Flow series:

$$
\mathbf{C} = \begin{bmatrix}
1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 1 & 1 & \cdots & 1
\end{bmatrix}
$$

- Index series:

$$
\mathbf{C} = \begin{bmatrix}
{}^1/_s & {}^1/_s & \cdots & {}^1/_s & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & {}^1/_s & {}^1/_s & \cdots & {}^1/_s & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & {}^1/_s & {}^1/_s & \cdots & {}^1/_s
\end{bmatrix}
$$

- Stock, EOP series:

$$
\mathbf{C} = \begin{bmatrix}
0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 1
\end{bmatrix}
$$

- Stock, BOP series:

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & 1 & 0 & \cdots & 0 \end{bmatrix}$$

- In case of extrapolation, there is a need of adding extra $n - sN$ columns of zeroes to the matrix:

$$\mathbf{C} = \begin{bmatrix} & \cdots & 0 & 0 & \cdots & 0 \\ & \cdots & 0 & 0 & \cdots & 0 \\ & \ddots & \vdots & \vdots & \ddots & \vdots \\ & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Where the first part of the matrix depends on the nature of the data as described above.

Thus, the problem could be defined as the estimation of a vector $\hat{\mathbf{y}}_H$ such that the following equation holds:

$$\mathbf{C}\hat{\mathbf{y}}_H = \mathbf{y}_L \qquad (2.2)$$

2.1.1 No indicators available

In the event when neither related indicators nor preliminary series are available, there are only two possible ways of dealing with temporal disaggregation: either applying some kind of mathematical smoothing method, or using statistical models.

The mathematical smoothing methods are the most used, and in many cases they are also the most effective. Naïve and related methods belonging to this group are all attempts of interpolating the unknown HF series with the known LF series. The methodology developed by Boot et al. (1967) (hereafter "BFL", as it is generally called by its authors: Boot, Feibes and Lisman), together with its variants, is probably the most effective and it is often considered as the best approach when no preliminary series or indicators are available (Eurostat, 2010).

Statistical methods developed by Al-Osh (1989) and Wei and Stram (1990) are very interesting from the theoretical point of view, but often very difficult to apply in practice.

2.1.1.1 Naïve and related methods

The naïve method is the very basic and simple way of dealing with temporal disaggregation and it is often used as reference criteria when other methods are used and checked (see, for example, Rodríquez Feijoó and Rodríquez Caro, 2000 or Rodríquez Feijoó et al., 2003). It is clearly not a good method as it introduces a constant step in the HF series and its use in practice should be avoided. This method is often referred to as the "*divided by four*" method, since statistical institutes often work with annual flow series which should be disaggregated in quarterly series. However, this name does not include all general situation for flow series and all the cases with index or stock series.

In the case of flow series, the HF values can be estimated as follows:

$$\hat{y}_{H,t} = \frac{1}{s} y_{L,\lfloor (t-1)/s \rfloor + 1} \tag{2.3}$$

While in the case of index and stock series, there is no need to divide the LF value for the temporal aggregation order:

$$\hat{y}_{H,t} = y_{L,\lfloor (t-1)/s \rfloor + 1}$$

For stock series a slightly better way to proceed is to make a linear interpolation. However, in case no extrapolation is done (via a forecast or backcast of the LF series), it is not possible to interpolate the first s observations for EOP series and the last s observations for BOP series.

For EOP series:

$$\hat{y}_{H,t} = \begin{cases} y_{L,\lfloor (t-1)/s \rfloor + 1} & \forall\, t = 1, \ldots, s \\ y_{L,\lfloor (t-1)/s \rfloor} + \dfrac{y_{L,\lfloor (t-1)/s \rfloor + 1} - y_{L,\lfloor (t-1)/s \rfloor}}{s} \times & \forall\, t = s+1, \ldots, n \\ \times\, (t - s(\lfloor (t-1)/s \rfloor - 1)) & \end{cases}$$

For BOP series:

$$\hat{y}_{H,t} = \begin{cases} y_{L,\lfloor (t-1)/s \rfloor +1} + \dfrac{y_{L,\lfloor (t-1)/s \rfloor +2} - y_{L,\lfloor (t-1)/s \rfloor +1}}{s} \times \\ (t-1-s(\lfloor (t-1)/s \rfloor -1)) \quad \forall\, t = 1, \ldots, n-s \\ \\ y_{L,\lfloor (t-1)/s \rfloor +1} \qquad\qquad\qquad \forall\, t = n-s+1, \ldots, n \end{cases}$$

Linear interpolation for flow series cannot be done in a trivial way, as different weights could be assigned to the HF values of a given LF period. The naïve method is a special case of a linear interpolation, where all the weights are equal.

Lisman and Sandee (1964) propose a method of linear interpolation for annual series which links the quarterly series of a given year T to the annual benchmarks of the year before and after. This is done considering that:

1. The year constraint is respected.
2. The results for a given year are symmetric when inverting the year before and the year after.
3. The results follow a linear trend, meaning that if the yearly totals rise by equal steps (i.e. $y_{L,T} - y_{L,T-1} = y_{L,T+1} - y_{L,T}$), the quarterly figures of year T should also rise by equal steps of length $y_{L,T} - y_{L,T-1}/16$.
4. In case $y_{L,T} - y_{L,T-1} = y_{L,T} - y_{L,T+1}$, the quarterly figures of year T should lie on a sinusoid.

The results for year T are unique:

$$\begin{bmatrix} \hat{y}_{H,1} \\ \hat{y}_{H,2} \\ \hat{y}_{H,3} \\ \hat{y}_{H,4} \end{bmatrix} = \begin{bmatrix} 0.073 & 0.198 & -0.021 \\ -0.010 & 0.302 & -0.042 \\ -0.042 & 0.302 & -0.010 \\ -0.021 & 0.198 & 0.073 \end{bmatrix} \begin{bmatrix} y_{L,T-1} \\ y_{L,T} \\ y_{L,T+1} \end{bmatrix} \tag{2.4}$$

As for the interpolation of the stock series, if there is no extrapolation via a forecast and a backcast of the LF series, it is not possible to interpolate the first and the last s observations of the series. Although the methodology is presented with the special case of $s = 4$, it is possible to derive different schemes.

A similar approach is followed by Zani (1970), which proposes a quadratic interpolation for quarterly series, getting the following results:

$$\begin{bmatrix} \hat{y}_{H,1} \\ \hat{y}_{H,2} \\ \hat{y}_{H,3} \\ \hat{y}_{H,4} \end{bmatrix} = \begin{bmatrix} 0.0547 & 0.2344 & -0.0391 \\ 0.0078 & 0.2656 & -0.0234 \\ -0.0234 & 0.2656 & 0.0078 \\ -0.0391 & 0.2344 & 0.0547 \end{bmatrix} \begin{bmatrix} y_{L,T-1} \\ y_{L,T} \\ y_{L,T+1} \end{bmatrix} \tag{2.5}$$

2.1.1.2 Boot, Feibes and Lisman method and further developments

The BFL, proposed for the first time by Boot et al. (1967), is the most known and used method for temporal disaggregation when no indicator is available. This smoothing approach has had different attempts for generalisations to the multivariate case (see, for example, Quenneville et al., 2013).

Boot et al. (1967) describe the methodology for a constraint minimisation of the squared first and second differences. The method will be described from a more general point of view, showing a constraint minimisation of the squared d-*th* differences, as presented also in Jacobs and Wansbeek (1992), with examples for the first differences case. Although the original technique was presented for deriving quarterly series from annual benchmarks, the methodology could easily be adapted for generalising the derivation of HF series from LF benchmarks. This has been shown in Cohen et al. (1971), which, on one hand, they dealt with any pair of possible combinations of HF and LF, and, on the other hand, they considered the minimization of the sum of the squared of the d-*th* differences between successive sub-period values.

The basic idea is very simple. In order to obtain a smooth series, the authors propose to minimise the sum of squares of the differences of the HF values, subject to the temporal aggregation constraint. Thus, mathematically:

$$\min_{y_H} \sum_{t=d+1}^{n} ((1-L)^d y_{H,t})^2$$

$$\text{s.t.} \sum_{t \in T} y_{H,t} = y_{L,T}$$

Which in case of $d = 1$ becomes:

$$\min_{y_H} \sum_{t=2}^{n} (y_{H,t} - y_{H,t-1})^2$$

$$\text{s.t.} \sum_{t \in T} y_{H,t} = y_{L,T}$$

In matrix form:

$$\begin{aligned} &\min_{y_H} \mathbf{y}'_H \mathbf{B} \mathbf{y}_H \\ &\text{s.t. } \mathbf{C}\mathbf{y}_H = \mathbf{y}_L \end{aligned} \tag{2.6}$$

Where $\mathbf{B} = \mathbf{D}'\mathbf{D}$ and $\mathbf{D}$ is the matrix performing the d-*th* difference. Thus for $d = 1$ they assume the following forms:

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & -1 & 1 & \\ & & & & -1 & 1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & -1 & & & & & \\ & -1 & 2 & -1 & & & \\ & & -1 & 2 & -1 & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 1 \end{bmatrix}$$

Which becomes $\mathbf{B} = \mathbf{D}'\mathbf{D}'\mathbf{D}\mathbf{D}$ in case of $d = 2$ and so on.

By applying the Lagrange function, it could be demonstrated that the solution is given by the following linear system:

$$\begin{bmatrix} \mathbf{B} & \mathbf{C}' \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{y}}_H \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{y}_L \end{bmatrix} \tag{2.7}$$

Where the solution is given by:

$$\hat{\mathbf{y}}_H = \mathbf{B}^{-1}\mathbf{C}' \left(\mathbf{C}\mathbf{B}^{-1}\mathbf{C}'\right)^{-1} \mathbf{y}_L \tag{2.8}$$

In general, all the smoothing methods have the possibility to work without an indicator or preliminary estimates available. However, they are normally difficult to implement if there are null values. Since they all need one LF observation more at the end and at the beginning of the series, which often means that forecasts are needed, the border effect problem occurs.

2.1.1.3 Other methods

Other methodologies which do not use any preliminary or related series have been suggested by different authors.

A similar approach to the BFL has been proposed by Marcellino (1999), which minimises a different loss function given by the mean squared disaggregation error:

$$\begin{aligned} &\min_{\hat{y}_H} tr\left(\mathbf{E}\left(\mathbf{y}_H - \hat{\mathbf{y}}_H\right)\left(\mathbf{y}_H - \hat{\mathbf{y}}_H\right)'\right) \\ &\text{s.t. } \mathbf{C}\mathbf{y}_H = \mathbf{y}_L \end{aligned} \tag{2.9}$$

If on one hand this method has the advantage to be able to deal with missing observations and can be extended to the case when an indicator is available, on the other hand it has the big drawback that requires the use of the covariance matrix of $\mathbf{y}_H$, which is normally unknown and should thus be estimated. The author proposes to derive a disaggregated ARIMA process starting from the aggregated process. This approach creates some doubts when considering that very often in official statistics the LF series are annual series and are only available for very few observations, which generates problems in the identification of the aggregated ARIMA process.

A similar problem is present in the methodology proposed by Stram and Wei (1986) and Wei and Stram (1990), which use the residuals of a preliminary OLS estimation of the LF model in order to estimate the parameters of the HF ARIMA model, obtaining the estimation of the covariance matrix. Similarly, Guerrero (1990) proposes to derive a preliminary estimate of the HF target series and to derive the covariance matrix by applying the traditional methodology of Box and Jenkins (1976). Again, the methods proposed can be also used when an indicator is available. However, because it uses only the N residuals of the LF OLS regressions, it is not applicable unless the number of LF periods is enough to fit an

ARIMA model with a reasonable accuracy (Santos Silva and Cardoso, 2001). An extension of the method to the bivariate case is provided by Hodgess and Wei (2000).

Finally, Al-Osh (1989) proposed a dynamic linear model, using an appropriate state space representation of the the HF ARIMA model, estimating the covariance matrix by applying a Kalman filter to the state space representation. Again, the number of observations plays a big role, making the methodology proposed, as all the ones described in this section, very interesting from a statistical point of view, but with a very limited practical use.

2.1.2 Indicator available

Several options are applicable when a preliminary estimate of the target variable or an indicator is available. Amongst them, the movement preservation principle method firstly developed by Denton (1971), and the set of the optimal regression-based techniques firstly proposed by (Chow and Lin, 1971) are the ones which have been mostly used in practice, in particular by statistical agencies. In all the cases which will be described in this section, the overall quality of the final estimates depends on the quality of the indicator (preliminary series) used, and to the variable's relation with the objective variable to estimate.

2.1.2.1 Naïve and related methods

When a preliminary series is available, it is always possible to equally distribute the discrepancies between the LF benchmark and the HF target series.

By defining the discrepancies $d_{L,T}$ as follows:

$$\sum_{t \in T} p_{H,t} - y_{L,T} = d_{L,T}$$

The naïve solution is given by:

$$\hat{y}_{H,t} = p_{H,t} + \frac{1}{s} d_{L,\lfloor (t-1)/s \rfloor + 1} \tag{2.10}$$

Although in practice this solution has to be avoided because the simple equally distribution of the discrepancies among the HF periods creates a step between the estimate of the last HF period of one LF period and first HF period of the next LF period, it still has a statistical meaning since it could be seen as a solution of the following minimisation problem:

$$\begin{aligned} &\min_{y_H} \sum_{t=1}^{n} \left(y_{H,t} - p_{H,t} \right)^2 \\ &\text{s.t.} \sum_{t \in T} y_{H,t} = y_{L,T} \end{aligned} \tag{2.11}$$

A better (less naïve) solution is to proportionally allocate the intra-LF discrepancies of the indicator in each HF period to the target variable:

$$\hat{y}_{H,t} = y_{L,T} \frac{p_{H,t}}{\sum_{t \in T} p_{H,t}} \tag{2.12}$$

However this approach also creates a step problem between the last HF period of one LF period and the first HF period of the next LF period, thus it should not be used (Eurostat, 2013).

2.1.2.2 Denton method and further developments

One of the most widely used methods for temporal benchmarking is the one which was originally developed by Denton (1971). This approach follows the movement preservation principle obtained by minimising a quadratic penalty (loss) function.

In the original method proposed by Denton, two different functions are proposed. The first one is on levels, with the Additive First Differences (AFD), while the second one is on proportional levels, with the Proportional First Differences (PFD).

In the case of the AFD, the problem can be expressed as follows:

$$\begin{aligned} &\min_{y_H} \sum_{t=1}^{n} \left((y_{H,t} - p_{H,t}) - (y_{H,t-1} - p_{H,t-1})\right)^2 ; \; y_{H,0} = p_{H,0} \\ &\text{s.t.} \sum_{t \in T} y_{H,t} = y_{L,T} \end{aligned} \tag{2.13}$$

With this specification, the function includes also the first term, being:

$$\Delta\left(y_{H,1} - p_{H,1}\right) = \left(y_{H,1} - p_{H,1}\right) - \left(y_{H,0} - p_{H,0}\right)$$

Where $y_{H,0}$ and $p_{H,0}$ are outside the range over which the series is to be adjusted, and thus are generally unknown. In order to solve this problem, Denton propose to take $y_{H,0} = p_{H,0}$, so that:

$$\Delta\left(y_{H,1} - p_{H,1}\right) = \left(y_{H,1} - p_{H,1}\right)$$

This solution, however, does not maximise the parallelism between the observed and adjusted series, as shown by Cholette (1984).

The specification given by Denton minimises the size of the first correction $(y_{H,1} - p_{H,1})$, and pulls the correction curve towards zero at the beginning of the series.

The solution proposed is to remove the first term and thus the equality of the period 0, obtaining the following problem (which is often referred to as modified Denton, or Cholette) in case of PFD:

$$\begin{aligned}
&\min_{y_H} \sum_{t=2}^{n} \left(\left(\frac{(y_{H,t} - p_{H,t})}{p_{H,t}} \right) - \left(\frac{(y_{H,t-1} - p_{H,t-1})}{p_{H,t-1}} \right) \right)^2 = \sum_{t=2}^{n} \left(\frac{y_{H,t}}{p_{H,t}} - \frac{y_{H,t-1}}{p_{H,t-1}} \right)^2 \\
&\text{s.t.} \sum_{t \in T} y_{H,t} = y_{L,T}
\end{aligned} \tag{2.14}$$

It is clear that the AFD formulation is very similar to the BFL method. Both the AFD and the PFD variants could be seen as a special case when the preliminary series is equal to 1:

$$\sum_{t=2}^{n}\left((y_{H,t}-1)-(y_{H,t-1}-1)\right)^2 \equiv \sum_{t=2}^{n}\left(\frac{y_{H,t}}{1}-\frac{y_{H,t-1}}{1}\right)^2 = \sum_{t=2}^{n}\left(y_{H,t}-y_{H,t-1}\right)^2$$

Similarly to the BFL method, it is also possible to define different penalty functions by using the Additive Second Differences (ASD) or the Proportional Second Differences (PSD).

In a general framework, the problem is expressed by the following matrix formulation:

$$\begin{aligned} &\min_{y_H}\ (\mathbf{y}_H-\mathbf{p}_H)'\,\mathbf{M}\,(\mathbf{y}_H-\mathbf{p}_H) \\ &\text{s.t. } \mathbf{C}\mathbf{y}_H=\mathbf{y}_L \end{aligned} \tag{2.15}$$

Which is solved by applying the lagrangean:

$$L=(\mathbf{y}_H-\mathbf{p}_H)'\,\mathbf{M}\,(\mathbf{y}_H-\mathbf{p}_H)+2\lambda'\,(\mathbf{C}\mathbf{y}_H-\mathbf{y}_L)$$

The results will be obtained by the solution of the following system:

$$\hat{y}_{H,t}=\begin{cases} \dfrac{\partial L}{\partial \mathbf{y}_H} = 0 \\ \dfrac{\partial L}{\partial \lambda} = 0 \end{cases} \Rightarrow \begin{cases} \mathbf{M}\mathbf{y}_H + \mathbf{C}'\lambda = \mathbf{M}\mathbf{p}_H \\ \mathbf{C}\mathbf{y}_H = \mathbf{y}_L \end{cases}$$

Therefore, in the case where the matrix $\mathbf{M}$ is singular, the solution of the benchmarking problem is part of the solution of the following linear system:

$$\begin{bmatrix} \mathbf{M} & \mathbf{C}' \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{y}}_H \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{M}\mathbf{p}_H \\ \mathbf{y}_L \end{bmatrix} \tag{2.16}$$

That is:

$$\begin{bmatrix} \hat{\mathbf{y}}_H \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \mathbf{C}' \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{M}\mathbf{p}_H \\ \mathbf{y}_L \end{bmatrix}$$

$$\hat{\mathbf{y}}_H = \mathbf{p}_H + \mathbf{M}^{-1}\mathbf{C}' \left(\mathbf{C}\mathbf{M}^{-1}\mathbf{C}'\right)^{-1} (\mathbf{y}_L - \mathbf{C}\mathbf{p}_H)$$

Where different solutions are obtained by changing the nature of $\mathbf{M}$ and $\mathbf{p}_H$:

- If $\mathbf{M} = \mathbf{I}$ and $\mathbf{p}_H = \mathbf{1}$, the solution corresponds to the naïve without indicator.
- If $\mathbf{M} = \mathbf{I}$, the solution corresponds to the naïve with indicator.
- If $\mathbf{M} = \mathbf{D}'\mathbf{D}$ and $\mathbf{p}_H = \mathbf{1}$, the solution corresponds to the BFL.
- If $\mathbf{M} = \mathbf{D}'\mathbf{D}$, the solution correspond to the modified Denton AFD.
- If $\mathbf{M} = \mathbf{P}_H^{-1}\mathbf{D}'\mathbf{D}\mathbf{P}_H^{-1}$, where $\mathbf{P}_H = diag\,(\mathbf{p}_H)$, the solution correspond to the modified Denton PFD.

For BFL and Denton methods, one can also easily to derive the specifications when using differences higher than one.

As for the extrapolation, the form of the matrix $\mathbf{C}$ implies that the HF benchmark-to-indicator ratios $\left(\frac{y_{H,t}}{p_{H,t}}\right)$ for all the extrapolated periods are equal to the benchmark-to-indicator ratios of the last HF observation of the last available LF period. In other words, the growth rates of the extrapolated values estimated by Denton correspond to the growth rates of the preliminary series. It is clear that by using this approach the back data of the preliminary time series are not considered at all for the estimation of the extrapolated values.

For this reason, Bloem et al. (2001) propose to modify the Denton PFD method by applying the so-called enhanced Denton PFD method, which introduce a new formulation of the constraints, with the possibility of adding an explicit forecast of the benchmark-to-indicator ratios for the extrapolated periods. A matrix formulation of the problem is derived by Di Fonzo and Marini (2012b). This kind of approach introduces some control from the user in the extrapolation practice, which will depend from the forecast of the benchmark-to-indicator ratios.

The Denton method is sometimes referred to as a two-step or indirect method. This is because the methodology needs the use of a preliminary series, which is somehow close to satisfying the temporal constraints and is expressed in the same unit measure. When only a related indicator is available, than a procedure to derive a preliminary estimate to benchmark is needed.

This could be done by a simple extrapolation, which assumes that that an available indicator $x_{H,t}$ has the same growth rates of the preliminary estimate of $y_{H,t}$. The preliminary estimates are very often derived according to a linear regression at LF level between the target series and the related indicators (Eurostat, 2013):

$$\mathbf{y}_L = \mathbf{X}_L \beta + \varepsilon_L$$

The OLS estimator $\hat{\beta}$ of β is then used with the HF related indicators in order to derive the preliminary series:

$$\mathbf{p}_H = \mathbf{X}_H \hat{\beta}$$

In this last expression, it is important to correctly deal with the constant term, so that for flow series the first column in $\mathbf{X}_L$ has all values equal to 1, while for index series the first column in $\mathbf{X}_H$ has all values equal to $^1/_s$.

More complex models could be used, for example, the user could use a regression in first differences (dynamic models) or assume that the residual term follows a first order autoregressive model.

2.1.2.3 Regression-based techniques

Optimal regression-based methods for temporal disaggregation have been firstly introduced by Chow and Lin (1971), which also provided a first general formulation of the interpolation, distribution and extrapolation problems. In this class of methods static models are used, in the sense that the dynamics are only present in the disturbances.

It is assumed that the following regression model holds at HF level:

$$\mathbf{y}_H = \mathbf{X}_H \beta + \mathbf{u}_H \tag{2.17}$$

With:

$$\mathbf{E}\left[\mathbf{u}_H \mid \mathbf{X}_H\right] = \mathbf{0}$$
$$\mathbf{E}\left[\mathbf{u}_H \mathbf{u}_H' \mid \mathbf{X}_H\right] = \mathbf{V}_H$$

Where β is a vector of regression coefficients, $\mathbf{u}_H$ is the disturbances series and $\mathbf{V}_H$ is the covariance matrix of the disturbances.

This model is clearly not observable, as $\mathbf{y}_H$ is the target series. However, when pre-multiplying by $\mathbf{C}$, the following is obtained:

$$\begin{aligned}\mathbf{C}\mathbf{y}_H &= \mathbf{C}\mathbf{X}_H \beta + \mathbf{C}\mathbf{u}_H \\ \mathbf{y}_L &= \mathbf{X}_L \beta + \mathbf{u}_L \end{aligned} \tag{2.18}$$

With:

$$\mathbf{E}\left[\mathbf{u}_L \mathbf{u}_L' \mid \mathbf{X}_H\right] = \mathbf{V}_L = \mathbf{C}\mathbf{V}_H \mathbf{C}'$$

This model is observable as it only contains LF variables. The optimal solutions will depend on the hypothesis regarding the disturbances $\mathbf{u}_H$. The matrix $\mathbf{C}$ has basically the role of transforming the variables from HF to LF.

It is worth noting that it must be $n \geq sT$, and in the case when $n > sT$, there is an extrapolation problem as well.

The solution in the BLUE sense is given by:

$$\hat{\beta} = \left(\mathbf{X}_L'\mathbf{V}_L^{-1}\mathbf{X}_L\right)^{-1}\mathbf{X}_L'\mathbf{V}_L^{-1}\mathbf{y}_L \tag{2.19}$$

$$\hat{\mathbf{y}}_H = \mathbf{X}_H\hat{\beta} + \mathbf{V}_H\mathbf{C}'\mathbf{V}_L^{-1}\left(\mathbf{y}_L - \mathbf{X}_L\hat{\beta}\right) \tag{2.20}$$

Considering that:

$$\hat{\mathbf{u}}_L = \mathbf{y}_L - \mathbf{X}_L\hat{\beta}$$

And by setting $\mathbf{L} = \mathbf{V}_H\mathbf{C}'\mathbf{V}_L^{-1}$, solution 2.20 can be written as follows:

$$\hat{\mathbf{y}}_H = \mathbf{X}_H\hat{\beta} + \mathbf{L}\hat{\mathbf{u}}_L \tag{2.21}$$

With:

$$\mathbf{E}\left[(\hat{\mathbf{y}}_H - \mathbf{y}_H)(\hat{\mathbf{y}}_H - \mathbf{y}_H)'\right] = \\ (\mathbf{I}_n - \mathbf{LC})\mathbf{V}_H + (\mathbf{X}_H - \mathbf{LX}_L)\left(\mathbf{X}_L'\mathbf{V}_L^{-1}\mathbf{X}_L\right)^{-1}(\mathbf{X}_H - \mathbf{LX}_L)'$$

Expression 2.21 can thus be seen as the sum of a systematic part, $\mathbf{X}_H\hat{\beta}$, which gives the dynamic profile of the HF related series to the final estimate, and an adjustment part, $\mathbf{L}\hat{\mathbf{u}}_L$, which recovers the temporal constraints (Santos Silva and Cardoso, 2001). This expression encompasses distribution, interpolation and extrapolation, according to the definition of the matrix $\mathbf{C}$.

It is clear that solution 2.20 depends on $\mathbf{V}_H$, which is often unknown and should be identified and estimated.

The simplest case is to assume that the model is an OLS, thus the disturbances are serially uncorrelated with constant variance. So, considering $\mathbf{V}_H = \sigma^2 \mathbf{I}_t$, and since $\mathbf{CC}' = s\mathbf{I}_t$, the solution is:

$$\begin{aligned}\hat{\mathbf{y}}_H &= \mathbf{X}_H \hat{\beta} + \sigma^2 \mathbf{I}_t \mathbf{C}' \left(\sigma^2 \mathbf{CC}'\right)^{-1} \left(\hat{\mathbf{y}}_L - \mathbf{X}_L \hat{\beta}\right) \\ &= \mathbf{X}_H \hat{\beta} + \frac{1}{s} \mathbf{C}' \left(\hat{\mathbf{y}}_L - \mathbf{X}_L \hat{\beta}\right) \end{aligned} \tag{2.22}$$

Which corresponds to the naïve solution. This result is not surprising, as the OLS model does not deal with serial correlation and thus is normally not fit for time series.

Chow and Lin (1971) propose that the residual term $u_{H,t}$ of model 2.17 follows a first order autoregressive process, AR (1):

$$u_{H,t} = \rho u_{H,t-1} \varepsilon_t \tag{2.23}$$

With:

$$\mathbf{E}\left[\varepsilon_t\right] = 0$$
$$\mathbf{E}\left[\varepsilon_t^2\right] = \sigma_\varepsilon^2$$

With this formulation, the covariance matrix assumes the following form:

$$\mathbf{V}_H = \frac{\sigma_\varepsilon^2}{1-\rho^2} \begin{bmatrix} 1 & & & & \\ \rho & 1 & & & \\ \rho^2 & \rho & \ddots & & \\ \vdots & \vdots & \ddots & 1 & \\ \rho^{t-1} & \rho^{t-2} & \dots & \rho & 1 \end{bmatrix} \tag{2.24}$$

This is the GLS model in case ρ is known. It is also important noticing that in case $\rho = 0$, the model becomes the OLS, leading to the naïve solution. If $\rho < 0$, the smoothing might introduce large volatility in the series and alter its temporal profile because of the negative autocorrelation.

However, in most cases, ρ is unknown and should be estimated. Literature on the matter refers to at least three alternative approaches.

The first approach is the one originally proposed by Chow-Lin, which used the idea of Cochrane and Orcutt (1949) starting from the relation between ρ and the first order autocorrelation coefficient of the residual term for the annual model ϕ_L, which in case of interpolation is simply equal to ρ, while in case of distribution is equal to:

$$\phi_L = \frac{\rho\left(\rho+1\right)\left(\rho^2+1\right)^2}{2\left(\rho^2+\rho+2\right)}$$

Therefore, starting from an initial estimate of ϕ_L, obtained by applying the OLS to model 2.18, ρ is iteratively computed by replacing the new values of ϕ_L until convergence.

However, as shown in Bournay and Laroque (1979), this approach is not feasible, as the function ϕ_L is not monotonic in ρ in the interval $[-1, 1]$, since there are two solutions for $-0.13 < \phi_L \leq 0$ and no solutions for $\phi_L < -0.13$. As mentioned by Ciammola et al. (2005), this is because, for example, the aggregation of a quarterly AR(1) process yields to an annual ARMA(1,1) process, so that there is no biunivocal correspondence between ϕ_L and ρ.

A better alternative is the one suggested by Bournay and Laroque (1979), which follow the maximum likelihood (ML) approach assuming the normality of the residuals. Thus, the problem is stated as:

$$\max_{\rho} L\left(\rho, \hat{\beta}\right) = \frac{t}{2}\left(-1 - \log\left(\frac{2\pi}{t}\right)\right) - \frac{t}{2}\log\left(\hat{\mathbf{u}}_L \mathbf{V}_L^{-1} \hat{\mathbf{u}}_L'\right) - \frac{1}{2}\log|\mathbf{V}_L|$$

The authors also demonstrate the existence of a maximum in the interval $]-1, 1[$. In practice, the estimation is performed by calculating $\mathbf{V}_H$, $\hat{\beta}$ and $\hat{\mathbf{u}}_L$ for a grid of values of ρ, and choosing the value $\hat{\rho}$ for which $L\left(\rho, \hat{\beta}\right)$ is a maximum over the grid.

The third approach is the one proposed by Barbone et al. (1981), which estimate the parameter ρ by minimising the sum of squared residuals (SSR), using thus an EGLS estimator. The statement of the problem is:

$$\min_{\rho} SSR\left(\rho, \hat{\beta}\right) = \hat{\mathbf{u}}_L \mathbf{V}_L^{-1} \hat{\mathbf{u}}_L' = \left(\mathbf{y}_L - \mathbf{X}_L \hat{\beta}\right)' \mathbf{V}_L^{-1} \left(\mathbf{y}_L - \mathbf{X}_L \hat{\beta}\right)$$

The estimation is performed by applying the algorithm proposed by Hildreth and Lu (1960), calculating $SSR\left(\rho, \hat{\beta}\right)$ in a initial grid of values for ρ, and continuing iteratively until convergence.

A different solution to avoid the estimation problem is given by Fernández (1981), which proposes a random walk model, ARIMA $(0, 1, 0)$, for the HF noise:

$$u_{H,t} = u_{H,t-1} + \varepsilon_t \tag{2.25}$$

With:

$$u_{H,0} = 0$$

$$\mathbf{E}\left[\varepsilon_t\right] = 0$$

$$\mathbf{E}\left[\varepsilon_t^2\right] = \sigma_\varepsilon^2$$

The covariance matrix can be written:

$$\mathbf{V}_H = \sigma_\varepsilon^2 \left(\tilde{\mathbf{D}}'\tilde{\mathbf{D}}\right)^{-1} = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & 2 & \dots & 2 & 2 \\ 1 & 2 & \dots & 3 & 3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 2 & \dots & t-1 & t \end{bmatrix} \tag{2.26}$$

Where the matrix $\tilde{\mathbf{D}}$ is an approximate first difference matrix, with the following form:

$$\tilde{\mathbf{D}} = \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix}$$

It is important to notice that $\tilde{\mathbf{D}}$ is not to be confused with the matrix $\mathbf{D}$ used so far (for example in the BFL method), which is the exact first difference matrix, while $\tilde{\mathbf{D}}$ is the approximate first difference matrix, changed because of the initial condition imposed by the Fernández model.

The biggest advantage of the Fernández approach is that the covariance matrix is completely known.

A slightly different approach is followed by Di Fonzo (2003b), which builds on the Chow-Lin model and its variants proposing a deltalog model, using a Taylor approximation for the additivity of the variables, getting the following model:

$$\Delta \log \mathbf{y}_H = \Delta \mathbf{X}_H \beta + \varepsilon_H \tag{2.27}$$

The author gives an economic interpretation of the deltalog model when applied using the Fernández model as the target HF variable is estimated so that the rates of change of the target HF variable are approximatively coherent with the LF counterpart.

A different approach is given by Litterman (1983), which suggest that the residual term of model 2.17 follows a random walk Markov model:

$$\begin{aligned} u_{H,t} &= \quad u_{H,t-1} + e_{H,t} \\ e_{H,t} &= \quad \alpha e_{H,t-1} + \varepsilon_t \end{aligned} \tag{2.28}$$

With:

$$u_{H,0} = e_{H,0} = 0$$
$$\mathbf{E}\left[\varepsilon_t\right] = 0$$
$$\mathbf{E}\left[\varepsilon_t^2\right] = \sigma_\varepsilon^2$$

Which basically corresponds to a first differentiation of model 2.17 in order to recover the stationarity of the residual term, using an ARIMA

$(1, 1, 0)$ model.

In this case the covariance matrix is given by:

$$\mathbf{V}_H = \sigma_\varepsilon^2 \left(\tilde{\mathbf{D}}'\mathbf{H}'\mathbf{H}\tilde{\mathbf{D}}\right)^{-1} \tag{2.29}$$

Where:

$$\mathbf{H} = \begin{bmatrix} -\alpha & 1 & & & & \\ & -\alpha & 1 & & & \\ & & \ddots & \ddots & & \\ & & & -\alpha & 1 & \\ & & & & -\alpha & 1 \end{bmatrix}$$

The Fernández approach can also be seen as a particular case of Litterman when $\alpha = 0$. As for the estimation of the parameter α, the same SSR and ML approaches seen in Chow-Lin can be followed.

2.1.2.4 Dynamic models

Chow-Lin's method and related approaches base their methodology on a static model, in the sense that the dynamics are only left in the residual term, and focus on the problem of the estimation of the covariance matrix of the residuals. Building on the work done by Hendry and Mizon (1978), which shows that models with autoregressive residuals can be seen as restricted dynamic models, some authors have provided solutions using dynamic models to the temporal disaggregation problem.

A first attempt to identify a dynamic model has been done by Palm and Nijman (1984). Solutions have been also proposed by Salazar et al. (1997) and Gregoir (2003) which work on the minimisation of a quadratic loss function. However, these do not provide direct estimates of the target variable as the estimation is made conditional to the first observation of the LF variable, obtaining results which depend on unknown initial conditions. Poissonier (2013) focused more on stock variables.

From a practical point of view, Santos Silva and Cardoso (2001), hereafter SSC, provide a more interesting dynamic extension of the Chow-Lin approach.

Starting from model 2.17, the dynamic extension is the following:

$$y_{H,t} = \kappa y_{H,t-1} + \mathbf{x}'_{H,t}\beta + \varepsilon_t \tag{2.30}$$

Where $|\kappa| < 1$ in order to achieve stationarity, and in the special case where $\kappa = 0$, model 2.30 becomes equal to 2.17.

Building on Tserkezos (1991) and Klein (1958), a recursive substitution can be used:

$$y_{H,t} = \left(\sum_{i=0}^{+\infty} \kappa^i \mathbf{x}'_{H,t-i}\right)\beta + \left(\sum_{i=0}^{+\infty} \kappa^i \varepsilon_{t-i}\right)$$

$$y_{H,t} = \left(\sum_{i=0}^{t-1} \kappa^i \mathbf{x}'_{H,t-i}\right)\beta + \kappa^t y_{H,0} + \left(\sum_{i=0}^{t-1} \kappa^i \varepsilon_{t-i}\right)$$

With:

$$y_{H,0} = \left(\sum_{i=0}^{+\infty} \kappa^i \mathbf{x}'_{H,-i} \right) \beta + \left(\sum_{i=0}^{+\infty} \kappa^i \varepsilon_{-i} \right)$$

$$\eta = \mathbf{E}\left[y_{H,0} \mid \mathbf{x}_0, \mathbf{x}_0, \ldots \right] = \left(\sum_{i=0}^{+\infty} \kappa^i \mathbf{x}'_{H,-i} \right)$$

$$_{\kappa}\mathbf{x}'_{H,t} = \left(\sum_{i=0}^{t-1} \kappa^i \mathbf{x}'_{H,t-i} \right)$$

Thus, model 2.30 can be written as:

$$y_{H,t} = \ _{\kappa}\mathbf{x}'_{H,t}\beta + \kappa^t \eta + u_{H,t} \tag{2.31}$$

With:

$$u_{H,t} = \kappa u_{H,t-1} + \varepsilon_t$$

$$u_{H,0} = 0$$

In matrix form it becomes:

$$\mathbf{y}_H = \ _{\kappa}\mathbf{X}_H \beta + \ _{\kappa}\mathbf{q}\eta + \mathbf{u}_H \tag{2.32}$$

Where $_{\kappa}\mathbf{q} = \left(\kappa, \kappa^2, \ldots, \kappa^n\right)'$.

By considering:

$$_{\kappa}\mathbf{D} = \begin{bmatrix} 1 & & & & \\ -\kappa & 1 & & & \\ & -\kappa & 1 & & \\ & & \ddots & \ddots & \\ & & & -\kappa & 1 \end{bmatrix}$$

Model 2.32 can be rewritten as follows:

$$ {}_{\kappa}\mathbf{D}\mathbf{y}_H = \mathbf{X}_H\beta + \mathbf{q}\eta + \varepsilon_H = \mathbf{Z}_H\gamma + \varepsilon_H $$

Where $\mathbf{q} = (\kappa, 0, \ldots, 0)'$, $\mathbf{Z}_H = [\mathbf{X}_H \mid \mathbf{q}]$ and $\gamma = [\beta' \mid \gamma]'$.

Pre-multiplying by ${}_{\kappa}\mathbf{D}^{-1}$:

$$ \mathbf{y}_H = {}_{\kappa}\mathbf{D}^{-1}\mathbf{Z}_H\gamma + {}_{\kappa}\mathbf{D}^{-1}\varepsilon_H = {}_{\kappa}\mathbf{Z}\gamma + \mathbf{u}_H \tag{2.33} $$

With:

$$ \mathbf{E}\left[\mathbf{u}_H \mid \mathbf{Z}_H\right] = \mathbf{0} $$

$$ \mathbf{E}\left[\mathbf{u}_H\mathbf{u}_H' \mid \mathbf{Z}_H\right] = \mathbf{V}_H $$

And the covariance matrix of the residual term is expressed as following:

$$ \mathbf{V}_H = \sigma_\varepsilon^2 \left({}_{\kappa}\mathbf{D}'\,{}_{\kappa}\mathbf{D}\right)^{-1} = \sigma_\varepsilon^2 \begin{bmatrix} 1 & \kappa & \kappa^2 & \cdots & \kappa^{n-2} & \kappa^{n-1} \\ \kappa & 1 & \kappa & \cdots & \kappa^{n-3} & \kappa^{n-2} \\ \kappa^2 & \kappa & 1 & \cdots & \kappa^{n-4} & \kappa^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \kappa^{n-2} & \kappa^{n-3} & \kappa^{n-4} & \cdots & 1 & \kappa \\ \kappa^{n-1} & \kappa^{n-2} & \kappa^{n-3} & \cdots & \kappa & 1 \end{bmatrix} $$

The model observed at LF is given by:

$$ \mathbf{y}_L = \mathbf{C}\,{}_{\kappa}\mathbf{Z}_H\gamma + \mathbf{C}\mathbf{u}_H \tag{2.34} $$

With:

$$
{}_{\kappa}\mathbf{Z}_L = \mathbf{C}\,{}_{\kappa}\mathbf{Z}_H
$$

$$
{}_{\kappa}\mathbf{V}_L = \mathbf{C}\,{}_{\kappa}\mathbf{V}_H\mathbf{C}'
$$

Santos Silva and Cardoso (2001) suggest to estimate the parameter κ using the ML approach. The statement of the problem is:

$$
\max_{\kappa} L\left(\kappa, \hat{\beta}\right) = \frac{T}{2}\log 2\pi - \frac{1}{2}|{}_{\kappa}\mathbf{V}_L| - \frac{1}{2}\left({}_{\kappa}\mathbf{y}_L - {}_{\kappa}\mathbf{Z}_L\beta\right)'\,{}_{\kappa}\mathbf{V}_L^{-1}\left({}_{\kappa}\mathbf{y}_L - {}_{\kappa}\mathbf{Z}_L\beta\right)
$$

In practice, the authors propose to calculate the ML function in a grid of admissible values of κ, and take the value which maximises the ML function. Di Fonzo (2003b) derives a solution in line with the classical Chow-Lin approach:

$$
\begin{aligned}
{}_{\kappa}\hat{\gamma} &= \left({}_{\kappa}\mathbf{Z}_L'\,{}_{\kappa}\mathbf{V}_L^{-1}\,{}_{\kappa}\mathbf{Z}_L\right)'\,{}_{\kappa}\mathbf{Z}_L'\,{}_{\kappa}\mathbf{V}_L^{-1}\mathbf{y}_L \\
\hat{\mathbf{y}}_H &= {}_{\kappa}\mathbf{Z}_H\,{}_{\kappa}\hat{\gamma} + {}_{\kappa}\mathbf{V}_H\mathbf{C}'\,{}_{\kappa}\mathbf{V}_L^{-1}\left(\mathbf{y}_L - {}_{\kappa}\mathbf{Z}_L\,{}_{\kappa}\hat{\gamma}\right)
\end{aligned}
\tag{2.35}
$$

By setting ${}_{\kappa}\mathbf{L} = {}_{\kappa}\mathbf{V}_H\mathbf{C}'\,{}_{\kappa}\mathbf{V}_L^{-1}$, the covariance matrix of the estimated values is given by:

$$
\begin{aligned}
&\mathbf{E}\left[\left(\hat{\mathbf{y}}_H - \mathbf{y}_H\right)\left(\hat{\mathbf{y}}_H - \mathbf{y}_H\right)'\right] = \\
&\left(\mathbf{I}_n - {}_{\kappa}\mathbf{L}\mathbf{C}\right){}_{\kappa}\mathbf{V}_H + \left(\mathbf{X}_H - {}_{\kappa}\mathbf{L}\mathbf{X}_L\right)\left(\mathbf{X}_L'\,{}_{\kappa}\mathbf{V}_L^{-1}\mathbf{X}_L\right)^{-1}\left(\mathbf{X}_H - {}_{\kappa}\mathbf{L}\mathbf{X}_L\right)'
\end{aligned}
$$

In their paper, Santos Silva and Cardoso (2001) mention that the specification of model 2.30 could be done also by including further lags of the dependent variable. However such enriched models have not been extensively discussed in the literature, and the estimation of the parameters could be cumbersome.

A good resume of methods using dynamic models is in Di Fonzo (2003b).

2.1.2.5 Other methods

A very interesting method which minimises a loss function, is the growth rates preservation (GRP) principle (Bozik and Otto, 1988; Causey and Trager, 1982; Trager, 1982). The idea is that the growth rates are a natural measure of the movements of a time series, and thus should be used instead of the movements preservation principle given by the AFD or PFD Denton variants.

The minimisation problem can be expressed as follows:

$$
\begin{aligned}
&\min_{y_H} \sum_{t=2}^{n} \left(\frac{y_{H,t}}{y_{H,t-1}} - \frac{p_{H,t}}{p_{H,t-1}} \right)^2 \\
&\text{s.t.} \sum_{t \in T} y_{H,t} = y_{L,T}
\end{aligned} \tag{2.36}
$$

It is clear that in this case the function to minimise is non-linear, and it is impossible to find an explicit analytical expression for the solution (Di Fonzo and Marini, 2010). In order to find a solution, different minimisation algorithms can be used (Di Fonzo and Marini, 2011a). However, the final estimates will depend in any case from the algorithm used.

Many interesting studies have been conducted using the GRP method (see, *inter alia*, Daalmans and Di Fonzo, 2014; Di Fonzo and Marini, 2013; Hood, 2005; Reber and Park, 2014).

Temurshoev (2012) presents entropy-based versions of the Denton and GRP methods, which ensure that the sign of the estimated values are the same of the corresponding preliminary series.

Finally, in the last class of methods the LF series is considered as the realisation of an ARIMA process, and the HF values are considered as missing observations (Jones, 1980). The original idea was based on the ML approach, were the function was built excluding the prediction errors associated to the missing observations and proposed to use forecasts obtained by applying Kalman filter. Among other extensions, Gomez and Marvall (1994) proposed to use an approach which is able to deal also with non-stationary time series.

Building on the theory of structural time series models (see, for example, Harvey, 1990), some authors provided temporal disaggregation techniques which are basically in the class of missing observations methods (Proietti, 1999). The main advantage of this class of methods is the possibility to perform simultaneously seasonal adjustment and temporal disaggregation, as shown by Moauro and Savio (2005). Proietti (2005) provides a good resume of the main optimal models revisited in the state-space form.

Jun et al. (2016) propose to use an indirect method which extrapolates the preliminary series according to a regression model, and benchmarks the series using a state space model.

2.2 Balancing

Balancing techniques are used in order to realign a set of variables to contemporaneous, or accounting, constraints. They are often used in official statistics, particularly in national accounts, where there are three approaches for measuring the gross domestic product (GDP): the production or output approach (the value of all goods and services produced within the economy less production costs), the expenditure approach (all the expenditure on goods and services which are not used up or transformed in a productive process), and the income approach (the sum of all income generated by production activity).

Normally, there are discrepancies between the three approaches, as they are normally calculated using different sources (Eurostat, 2013). Thus, in order to publish only one figure for the GDP, they have to be balanced for the sake of consistency.

As will be shown in this section, the methodology for balancing has been developed in the first half of the previous century, and the main work has been done by Stone et al. (1942) and Bacharach (1970).

2.2.1 Adjustment schemes

At least three different adjustment schemes can be considered, given $a_{i,t}$, $i = 1, \dots, m$, the provisional values of the m target variables $y_{i,t}$ (in this specific framework there is no need for the variables to be time series),

z_t, a constraint such that $\sum_{i=1}^{m} y_{i,t} = z_t$, and the observed discrepancy $d_t = z_t - \sum_{i=1}^{m} y_{i,t}$.

The first one is the naïve approach, which simply distributes the discrepancies evenly:

$$\hat{y}_{i,t} = a_{i,t} + \frac{1}{m}\left(z_t - \sum_{i=1}^{m} a_{i,t}\right) = a_{i,t} + \frac{1}{m}d_t \tag{2.37}$$

This approach is obviously not a good practice, as it distributes the discrepancies without considering the values (size) of the target variable.

The proportional allocation of the discrepancies (often called pro-rata approach), which was first used by Matuszewski et al. (1964), seems to be a better approach:

$$\hat{y}_{i,t} = a_{i,t} + \frac{a_{i,t}}{\sum_{i=1}^{m} a_{i,t}}\left(z_t - \sum_{i=1}^{m} a_{i,t}\right) = a_{i,t}\frac{z_t}{\sum_{i=1}^{m} a_{i,t}} \tag{2.38}$$

In matrix form, this solution can be written as follows:

$$\hat{\mathbf{Y}} = \mathbf{RA}$$

Where $\mathbf{Y}$ is the matrix with the target variables, $\mathbf{R}$ is a diagonal matrix with the adjustment factors and $\mathbf{A}$ is the matrix with the preliminary series.

It is important to note that solution 2.38 cannot be used if negative values are present. This is why some authors (see, for instance, ABS,

2009) suggested to use a plus-minus proportional adjustment:

$$\hat{y}_{i,t} = a_{i,t} + \frac{|a_{i,t}|}{\sum_{i=1}^{m} |a_{i,t}|} \left(z_t - \sum_{i=1}^{m} a_{i,t} \right)$$

Which brings to different adjustment factors for positive and negative values:

$$\hat{y}^{+}_{i,t} = a_{i,t} \left(1 + \frac{d_t}{\sum_{i=1}^{m} |a_{i,t}|} \right)$$

$$\hat{y}^{-}_{i,t} = a_{i,t} \left(1 - \frac{d_t}{\sum_{i=1}^{m} |a_{i,t}|} \right)$$

Finally, the third approach is the so called proportional squared:

$$\hat{y}_{i,t} = a_{i,t} + \frac{a^2_{i,t}}{\sum_{i=1}^{m} a^2_{i,t}} \left(z_t - \sum_{i=1}^{m} a_{i,t} \right) \tag{2.39}$$

To better show the statistical meaning of the proportional scheme, the general problem could be seen as a least squares adjustment of the data:

$$\min_{y_{i,t}} \sum_{i=1}^{m} \omega_{i,t} \left(y_{i,t} - a_{i,t} \right)$$
$$\text{s.t.} \sum_{i=1}^{m} y_{i,t} = z_t$$

Which is solved using the lagrangean multiplier:

$$L = \sum_{i=1}^{m} \omega_{i,t} \left(y_{i,t} - a_{i,t} \right) - 2\lambda \left(z_t - \sum_{i=1}^{m} y_{i,t} \right)$$

$$\hat{y}_{i,t} = \begin{cases} \dfrac{\partial L}{\partial y_{i,t}} & = & 0 \\ \\ \dfrac{\partial L}{\partial \lambda} & = & 0 \end{cases} \qquad \Rightarrow \; \hat{y}_{i,t} = a_{i,t} + \frac{\dfrac{1}{\omega_{i,t}}}{\sum\limits_{i=1}^{m} \dfrac{1}{\omega_{i,t}}} \left(z_t - \sum_{i=1}^{m} a_{i,t} \right)$$

It can be easily seen that according to the form of $\omega_{i,t}$, the solution of the minimisation problem can lead to the mentioned schemes. In particular:

- If $\omega_{i,t} = 1$, the naïve approach is used.
- IF $\omega_{i,t} = 1/a_{i,t}$, the proportional approach is used.
- If $\omega_{i,t} = 1/a^2_{i,t}$, the proportional squared approach is used.

The weights $\omega_{i,t}$ are often linked to the variability of the variables to be adjusted, as can be seen, for example, in Van Tongeren and Magnus (2011).

In the case of the proportional scheme, this is implicitly done by using the variance, which means that larger variables are considered relatively more reliable than smaller ones. Hence, it seems that the (normalised) Coefficient of Variation (CV) is a better measure of the reliability.

The implied variance in the three approaches is equal to 1, $a_{i,t}$ and $a^2_{i,t}$, respectively, while the implied reliability index (in percentage) is equal to $1/a_{i,t}$, $1/\sqrt{a_{i,t}}$ and 1, respectively. Basically both the naïve and the proportional approaches assume that the bigger the variable, the bigger the reliability, adjusting relatively more the smaller variables. This is not the case when using the proportional squared scheme, which assumes a constant reliability of the variables (in terms of CV).

2.2.2 Standard techniques for matrix balancing

The issue of matrix balancing dates back to the first half of the previous century, even before the very famous paper by Stone et al. (1942). The techniques used can be divided at least in two big groups: bi-proportional methods (see, above all, Stone, 1961 and Bacharach, 1970) and constrained optimisation methods. Among other domains, they are very often used in national accounts in order to balance input-output tables (Eurostat, 2008b).

2.2.2.1 Bi-proportional adjustment: RAS

One of the most widely used techniques for balancing a table with fixed marginal totals is the so called RAS method, which has been introduced in the thirties and has been heavily discussed in literature (Bacharach, 1970; Lahr and De Mesnard, 2004; Stone, 1961). The methodology is largely applied in the balancing of the Input Output tables in national accounts (Eurostat, 2008b).

Given $\mathbf{A}$, a $m \times n$ matrix of preliminary values whose generic element is a_{ij}, $\mathbf{X}$, the objective matrix to be estimated, $\mathbf{u}$, the $m \times 1$ vector of observed row totals, and $\mathbf{v}$, the $n \times 1$ vector of observed column totals such that:

$$u_i = \sum_{j=1}^{n} x_{ij}$$
$$v_j = \sum_{i=1}^{m} x_{ij}$$
$$T = \sum_{i=1}^{m} u_i = \sum_{j=1}^{n} u_j$$

The objective is to find a matrix $\hat{\mathbf{X}}$ such that:

$$\sum_{j=1}^{n} \hat{x}_{ij} = u_i$$
$$\sum_{i=1}^{m} \hat{x}_{ij} = v_j$$

The RAS algorithm proceeds with an iterative calculation of the $\hat{x}_{ij}$, until convergence, in the following way:

Starting value

$\hat{x}_{ij}^{(0)} = a_{ij}$

First iteration

Firstly, consistency is achieved with the $\mathbf{u}$ vector:

$$r_i^{(1)} = \frac{u_i}{\sum_{j=1}^{n} \hat{x}_{ij}^{(0)}} \qquad \hat{x}_{ij}^{(1)} = r_i^{(1)} \hat{x}_{ij}^{(0)}$$

Secondly, consistency is achieved with the $\mathbf{v}$ vector (albeit loosing the consistency achieved with $\mathbf{u}$):

$$s_j^{(1)} = \frac{v_j}{\sum_{i=1}^{m} \hat{x}_{ij}^{(1)}} \qquad \hat{x}_{ij}^{(2)} = s_j^{(1)} \hat{x}_{ij}^{(1)}$$

k-th iteration

$$r_i^{(k)} = \frac{u_i}{\sum_{j=1}^{n} \hat{x}_{ij}^{(2k-2)}} \qquad \hat{x}_{ij}^{(2k-1)} = r_i^{(k)} \hat{x}_{ij}^{(2k-2)}$$

$$s_j^{(k)} = \frac{v_j}{\sum_{i=1}^{m} \hat{x}_{ij}^{(2k-1)}} \qquad \hat{x}_{ij}^{(2k)} = s_j^{(k)} \hat{x}_{ij}^{(2k-1)}$$

End of the procedure

The procedure will be stopped when either:

$$\left| \sum_{i=1}^{m} \hat{x}_{ij}^{(2k-1)} - v_j \right| < \delta$$

Or:

$$\left| \sum_{j=1}^{n} \hat{x}_{ij}^{(2k)} - u_i \right| < \delta$$

For a given small tolerance $\delta > 0$.

It has to be noted that in case of k complete iterations, the generic estimated value is:

$$\hat{x}_{ij} = r_i^{(k)} \ldots r_i^{(1)}\, a_{ij}\, s_j^{(1)} \ldots s_j^{(k)} \tag{2.40}$$

From this last expression it is clear why the RAS method is referred to as a bi-proportional adjustment, since basically it is an extension of the proportional adjustment for two dimensions, with r_i being the proportions for the row and s_j being the proportions for the columns.

In matrix form, expression 2.40 assumes the following form:

$$\hat{\mathbf{X}} = \mathbf{RAS} \tag{2.41}$$

R being the diagonal matrix with the r_i values, and **S** the diagonal matrix with the s_j values. From expression 2.41 it is very clear where the method takes its name.

Uribe et al. (1965) and Theil (1967) have shown that the RAS approach generates a solution which could be seen as the same solution of the following minimisation problem:

$$\begin{aligned} \min_{x_{i,j}} \; & \sum_{i=1}^{m}\sum_{j=1}^{n} x_{i,j} \log \frac{x_{i,j}}{a_{i,j}} \\ \text{s.t.} \; & \sum_{j=1}^{n} x_{ij} = u_i \\ & \sum_{i=1}^{m} x_{ij} = v_j \end{aligned}$$

Which is minimising an entropic distance between $x_{i,j}$ and $a_{i,j}$, also defined as surprise (Bacharach, 1970).

The RAS approach has some good properties. The results $\hat{x}_{i,j}$ preserve the zeroes and the positivity of the preliminary values $a_{i,j}$, while it is possible to easily introduce a priori information by fixing any known $\hat{x}_{i,j}$ (Israilevich, 1986). However, it is important to mention that the RAS methodology is not invariant to linear transformation of the matrix.

Finally Gilchrist and St. Louis (1999) have developed a method (TRAS) to use known information beyond that for the column and row totals.

The original RAS methodology is uncapable of dealing with negative values. For this reason ABS (2009) has developed a methodology to avoid this problem, following the same concept of the plus-minus adjustment scheme. In practice, the proportional iterations are done according to the sign of each preliminary value. Therefore, for the k-th iteration, the values to be calculated are the following:

$$
{}^{+}r_i^{(k)} = 1 + \frac{u_i - \sum_{j=1}^{n} \hat{x}_{ij}^{(2k-2)})}{\sum_{j=1}^{n} \left| \hat{x}_{ij}^{(2k-2)} \right|} \qquad {}^{-}r_i^{(k)} = 1 - \frac{u_i - \sum_{j=1}^{n} \hat{x}_{ij}^{(2k-2)}}{\sum_{j=1}^{n} \left| \hat{x}_{ij}^{(2k-2)} \right|}
$$

$$
{}^{+}s_j^{(k)} = 1 + \frac{v_j - \sum_{i=1}^{m} \hat{x}_{ij}^{(2k-1)}}{\sum_{1=1}^{m} \left| \hat{x}_{ij}^{(2k-1)} \right|} \qquad {}^{-}s_j^{(k)} = 1 - \frac{v_j - \sum_{i=1}^{m} \hat{x}_{ij}^{(2k-1)}}{\sum_{1=1}^{m} \left| \hat{x}_{ij}^{(2k-1)} \right|}
$$

Other kind of bi-proportional balancing procedures have been developed. Above all, it is worth mentioning the diagonal similarity scaling algorithm (see for example Eaves et al., 1985), which starts dealing with the element $a_{i,j}$, for which the row sum differs greatly from the column sum. If on one hand this approach is able to handle upper and lower bounds on the margin totals, on the other hand it requires column sums to be equal to row sums (Lahr and De Mesnard, 2004).

2.2.2.2 The approach by Stone and further developments

As an alternative to bi-proportional approaches, the matrix balancing can be performed by constrained optimisation of a function, which is often quadratic. The approach was firstly introduced by Stone et al. (1942) and have been discussed by many authors (see, *inter alia*, Bacharach, 1970, Lahr and De Mesnard, 2004 and Di Fonzo and Marini, 2007).

In general terms, a good feature of the optimisation approach is that it can also deal with *endogenous* constraints, meaning that a preliminary row vector $\mathbf{u}$ is available and has to be adjusted as well. This case is not covered by the RAS method (Di Fonzo and Marini, 2007). On the other hand, this class of methods does not always preserve the positive sign of the preliminary variables.

Many different functions to be minimised have been proposed in the literature. Extensive lists can be found in Lahr and De Mesnard (2004) and Jackson and Murray (2004). Almon (1968) suggests to use the Euclidean distance:

$$\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{i,j}-a_{i,j})^2 \tag{2.42}$$

Which is a particular case of the Hölder norm for $\theta = 2$:

$$\sum_{i=1}^{m}\sum_{j=1}^{n}|x_{i,j}-a_{i,j}|^{\theta}$$

Lahr (2001) suggests to use the weighted absolute differences:

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\left(a_{i,j}\left|x_{i,j}-a_{i,j}\right|\right) \tag{2.43}$$

While Matuszewski et al. (1964) had suggested to use the weighted squared differences:

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\left(a_{i,j}\left(x_{i,j}-a_{i,j}\right)^2\right) \tag{2.44}$$

A different approach is to use the normalised absolute differences:

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{\left|x_{i,j}-a_{i,j}\right|}{a_{i,j}} \tag{2.45}$$

Finally Deming and Stephan (1940) and Friedlander (1961) propose to use the normalised squared difference, which is the χ^2 of Pearson:

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{\left(x_{i,j}-a_{i,j}\right)^2}{a_{i,j}} \tag{2.46}$$

Criteria 2.42 and 2.46 can be seen as quadratic positive definite (QDP) functions (Di Fonzo, 2003b) of the form:

$$\left(\tilde{\mathbf{x}}-\tilde{\mathbf{a}}\right)'\mathbf{Q}^{-1}\left(\tilde{\mathbf{x}}-\tilde{\mathbf{a}}\right) \tag{2.47}$$

Where $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{x}}$ are the vectorised data of the preliminary matrix $\mathbf{A}$ and the objective matrix $\mathbf{X}$, respectively.

In the cases of $\mathbf{Q}=\mathbf{I}_{mn}$ and $\mathbf{Q}=diag\left(\mathbf{p}\right)$ criteria 2.47 becomes equal to criteria 2.42 and 2.46, respectively.

The least square adjustment subject to linear restrictions of Stone et al. (1942) can be expressed by the following simple linear model:

$$\mathbf{a} = \mathbf{x} + \varepsilon \tag{2.48}$$

With:

$$\mathbf{E}\left[\varepsilon\right] = \mathbf{0}$$
$$\mathbf{E}\left[\varepsilon\varepsilon'\right] = \mathbf{V}$$

Where $\mathbf{V}$ is known and $\mathbf{a}$ is a $mn \times 1$ vector of preliminary values which do not fulfil the set of linear constraints:

$$\mathbf{Bx} = \mathbf{b}$$

Where $\mathbf{B}$ is a known matrix of order $k \times mn$ with $k < mn$ and $\mathbf{b}$ is a $k \times 1$ known vector.

Under this constraints, Di Fonzo and Marini (2007) shows that, for $\mathbf{Q} = \mathbf{V}$, $\hat{\mathbf{x}}$ is an efficient estimator of $\mathbf{a}$:

$$\hat{\mathbf{x}} = \mathbf{a} + \mathbf{VB}'\left(\mathbf{BVB}\right)^{-1}\left(\mathbf{b} - \mathbf{Ba}\right) \tag{2.49}$$

With:

$$\mathbf{E}\left[\left(\hat{\mathbf{x}} - \mathbf{x}\right)\left(\hat{\mathbf{x}} - \mathbf{x}\right)'\right] = \mathbf{V} - \mathbf{VA}'\left(\mathbf{AVA}'\right)^{-1}\mathbf{V}$$

Basically, in order to satisfy the constraints, the data are adjusted considering their relative variances. Weale (1988) shows that the estimates are ML in case of normality assumption.

The main drawback of the approach by Stone et al. (1942) is that the covariance matrix $\mathbf{V}$ is assumed as being known and must be somehow specified.

2.2.3 Balancing in practice

While the multivariate proportional adjustment seems to be a *simple and fairly reasonable* approach (Di Fonzo, 2003a), and it is often used by statistical agencies, other more complex and effective methods are also available.

Probably, one of the most used balancing procedures is the so called *ad hoc* balancing, which means that the balancing is done according to known qualitative indicators on the variables to be balanced. A usual example of this approach is the balancing of the preliminary values of the GDP when different results have been obtained from the output and expenditure approaches (Eurostat, 2013), and the discrepancies are often all added to the variable "*Changes in inventories and net acquisition of valuables*", which is considered as the weakest one.

In practice, the *ad hoc* balancing is often used together with other kinds of (statistical) balancing techniques, making the best use of all qualitative and quantitative information available to the user. This happens, for example, when part of the data come from a source which is considered of a higher level (quality), or is given from a different domain and cannot be modified for consistency reasons.

2.3 Reconciliation

The balancing techniques which have been discussed in the previous section, such as the pure bi-proportional approach, are not designed to be applied to time series, thus in most cases, if used on time series, they do not preserve the dynamics of the related indicator(s). It is clear, however, that when dealing with more time series, the contemporaneous constraints show up together with temporal constraints, and both constraints should be handled.

A first attempt to develop a multivariate regression-based temporal disaggregation technique has been done by Rossi (1982), while a complete formulation of the problem is given by Di Fonzo (1990). Some authors have tried to develop a multivariate approach following Denton's approach (Bikker et al., 2010; Di Fonzo and Marini, 2003, 2011b). Finally, two-step approaches have been introduced by Quenneville and Rancourt (2005) and further extended by Di Fonzo and Marini (2011b).

2.3.1 Simultaneous approaches

While the multivariate proportional adjustment described before is able to solve only the contemporaneous constraints, and thus is strictly a balancing procedure, different formulations have been given in order to develop a multivariate approach which deals with both the contemporaneous and temporal constraints.

2.3.1.1 Multivariate Denton

The approach by Denton has been expanded to the multivariate case in order to also deal with the contemporaneous constraints (Di Fonzo and Marini, 2003).

Given $y_{H,j,t}$, $p_{H,j,t}$ and $y_{L,j,T}$, with $j = 1, \ldots, m$, three sets of m time series denoting the objective HF series to be reconciled, the observed HF preliminary series and the LF benchmarks,, respectively and given $z_{H,t}$ the observed HF contemporaneous benchmark (accounting constraint), the multivariate formulation of the Denton AFD problem is the following:

$$\min_{y_{H,j}} \sum_{j=1}^{m} \sum_{t=2}^{n} \left((y_{H,j,t} - p_{H,j,t}) - (y_{H,j,t-1} - p_{H,j,t-1})\right)^2 \tag{2.50}$$

$$\begin{aligned} \text{s.t.} \quad & \sum_{t \in T} y_{H,j,t} = y_{L,j,T} \quad \forall j = 1, \ldots, m \\ & \sum_{j=1}^{m} y_{H,j,t} = z_{H,t} \qquad \forall t = 1, \ldots, n \end{aligned}$$

And the relative PFD problem:

$$\min_{y_{H,j}} \sum_{j=1}^{m}\sum_{t=2}^{n}\left(\left(\frac{(y_{H,j,t}-p_{H,j,t})}{p_{H,j,t}}\right)-\left(\frac{(y_{H,j,t-1}-p_{H,j,t-1})}{p_{H,j,t-1}}\right)\right)^2 \tag{2.51}$$

$$\begin{aligned} \text{s.t.} \sum_{t\in T} y_{H,j,t} &= y_{L,j,T} \quad \forall j=1,\dots,m \\ \sum_{j=1}^{m} y_{H,j,t} &= z_{H,t} \qquad \forall t=1,\dots,n \end{aligned}$$

In order to define the matrix notation of the constraints, let's consider the following quantities:

- $\mathbf{y}_H = (\mathbf{y}_{H,1},\dots,\mathbf{y}_{H,j},\dots,\mathbf{y}_{H,m})'$, the $mn \times 1$ vector with the m series to be reconciled.
- $\mathbf{p}_H = (\mathbf{p}_{H,1},\dots,\mathbf{p}_{H,j},\dots,\mathbf{p}_{H,m})'$, the $mn \times 1$ vector with the m preliminary time series.
- $\mathbf{y}_L = (\mathbf{y}_{L,1},\dots,\mathbf{y}_{L,j},\dots,\mathbf{y}_{L,m})'$, the $mN \times 1$ vector with the m LF temporal benchmark series.
- $\mathbf{z}_H$, the $n \times 1$ time series with the whole set of HF contemporaneous benchmarks (constraints).
- $\mathbf{y}_a = \begin{bmatrix} \mathbf{z}_H \\ \mathbf{y}_L \end{bmatrix}$ the $(n+mN) \times 1$ vector containing the n contemporaneous HF benchmarks (accounting constraints) and the mN temporal LF benchmarks.

- $\mathbf{H} = \begin{bmatrix} \mathbf{1}'_m \otimes \mathbf{I}_n \\ \\ \mathbf{I}_m \otimes \mathbf{C} \end{bmatrix}$ the aggregation matrix, where $\mathbf{1}'_m$ is $m \times 1$ vector of unitary elements.

The whole set of temporal and contemporaneous constraints are expressed in matrix form by the following expression:

$$\mathbf{H}\mathbf{y}_H = \mathbf{y}_a \tag{2.52}$$

So the multivariate version of the Denton problem is expressed by the following:

$$\begin{aligned} &\min_{y_H} \; (\mathbf{y}_H - \mathbf{p}_H)' \, \boldsymbol{\Omega} \, (\mathbf{y}_H - \mathbf{p}_H) \\ &\text{s.t. } \mathbf{H}\mathbf{y}_H = \mathbf{y}_a \end{aligned} \tag{2.53}$$

Which is solved by applying the lagrangean, as for the univariate case:

$$L = (\mathbf{y}_H - \mathbf{p}_H)' \, \boldsymbol{\Omega} \, (\mathbf{y}_H - \mathbf{p}_H) + 2\lambda' \, (\mathbf{H}\mathbf{y}_H - \mathbf{y}_a)$$

$$\hat{y}_{H,t} = \begin{cases} \dfrac{\partial L}{\partial \mathbf{y}_H} & = & 0 \\ \\ \dfrac{\partial L}{\partial \lambda} & = & 0 \end{cases} \Rightarrow \begin{cases} \boldsymbol{\Omega}\mathbf{y}_H & + & \mathbf{H}'\lambda & = & \boldsymbol{\Omega}\mathbf{p}_H \\ \\ \mathbf{H}\mathbf{y}_H & & & = & \mathbf{y}_a \end{cases}$$

$$\begin{bmatrix} \boldsymbol{\Omega} & \mathbf{H}' \\ \\ \mathbf{H} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{y}}_H \\ \\ \lambda \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Omega}\mathbf{p}_H \\ \\ \mathbf{y}_a \end{bmatrix} \tag{2.54}$$

Leading to the following solution:

$$\hat{\mathbf{y}}_H = \mathbf{p}_H + \boldsymbol{\Omega}^{-1}\mathbf{H}'\left(\mathbf{H}\boldsymbol{\Omega}^{-1}\mathbf{H}'\right)^{-}(\mathbf{y}_a - \mathbf{H}\mathbf{p}_H) \tag{2.55}$$

Where $\left(\mathbf{H}\boldsymbol{\Omega}^{-1}\mathbf{H}'\right)^{-}$ is the Moore-Penrose generalised inverse of the matrix $\left(\mathbf{H}\boldsymbol{\Omega}^{-1}\mathbf{H}'\right)$, being not invertible (see Di Fonzo and Marini, 2003 for its derivation).

As for the univariate case, according to the choice of the matrix $\boldsymbol{\Omega}$, different solutions can be found (Di Fonzo, 2003a):

- If $\boldsymbol{\Omega} = \mathbf{I}_m \otimes (\mathbf{D}'\mathbf{D})$, the solution corresponds to the multivariate Denton AFD.
- If $\boldsymbol{\Omega} = \mathbf{I}_m \otimes (\mathbf{D}'\mathbf{D}'\mathbf{D}\mathbf{D})$, the solution corresponds to the multivariate Denton ASD.
- If $\boldsymbol{\Omega} = \mathbf{P}_H^{-1}(\mathbf{I}_m \otimes (\mathbf{D}'\mathbf{D}))\,\mathbf{P}_H^{-1}$, where $\mathbf{P}_H = diag\,(\mathbf{p}_H)$, the solution corresponds to the multivariate Denton PFD.
- If $\boldsymbol{\Omega} = \mathbf{P}_H^{-1}(\mathbf{I}_m \otimes (\mathbf{D}'\mathbf{D}'\mathbf{D}\mathbf{D}))\,\mathbf{P}_H^{-1}$, the solution corresponds to the multivariate Denton PSD.

Di Fonzo and Marini (2003) give also the results for two systems of time series and split the cases into whether the constraints are binding or unbinding (exogenous or endogenous, respectively). However, using partitioned matrices, the calculations given are rather cumbersome and mathematically complex.

A better solution for generalising the problem to more systems of time series and to include the case of endogenous constraint, seems to be the one proposed by Di Fonzo and Marini (2011b). It builds on the general constraints proposed in the problem 2.53, by replacing contemporaneous constraints $\mathbf{z}_H$ with the set of k HF contemporaneous (accounting) constraints for each of the k systems, $\tilde{\mathbf{z}}_H = (\mathbf{z}_{H,1}, \ldots, \mathbf{z}_{H,i}, \ldots, \mathbf{z}_{H,k})'$, and the vector $\mathbf{1}'_m$ of matrix $\mathbf{H}$, with a $k \times m$ matrix $\mathbf{G}$, which specify the k linear constraints between $\mathbf{y}_H$ and $\tilde{\mathbf{z}}_H$, so that:

$$\tilde{\mathbf{y}}_a = \begin{bmatrix} \tilde{\mathbf{z}}_H \\ \mathbf{y}_L \end{bmatrix}$$

$$\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{G} \otimes \mathbf{I}_n \\ \mathbf{I}_m \otimes \mathbf{C} \end{bmatrix}$$

The authors provide some examples on how to build the matrix $\mathbf{G}$ in different practical cases, and specify that normally it contains values equal to 0, 1 and -1. As mentioned, this specification of the whole set of constraints (temporal and contemporaneous) encompasses the cases of a set of k systems of time series with both exogenous and endogenous constraints.

Problem 2.53 can thus be re-written as follows:

$$\begin{aligned} &\min_{y_H} (\mathbf{y}_H - \mathbf{p}_H)' \, \mathbf{\Omega} \, (\mathbf{y}_H - \mathbf{p}_H) \\ &\text{s.t.} \tilde{\mathbf{H}} \mathbf{y}_H = \tilde{\mathbf{y}}_a \end{aligned} \tag{2.56}$$

It is easy to note that in case $k = 1$ and $\mathbf{G} = \mathbf{1}'_m$, problem 2.56 becomes 2.53, being a particular case.

The first element of system 2.54 is symmetric, indefinite, singular sparse and large, making the adoption of the Moore-Penrose generalised inverse difficult from a computational point of view. In order to obtain the direct solution 2.55, Di Fonzo and Marini (2011b) suggest to apply the matrix factorisation algorithm proposed by Duff (2004), reducing the computational time.

Di Fonzo and Marini (2015) propose an alternative simultaneous approach based on a multivariate generalisation of the GRP principle. The authors state that such approach gives the best results for the preservation of growth rates of the preliminary series.

2.3.1.2 Multivariate optimal methods

Some authors have tried to generalise the Chow-Lin temporal disaggregation method to the multivariate case. A first specification of the problem has been done by Rossi (1982), while a complete discussion is in Di Fonzo (1990), Di Fonzo (2003a) and Eurostat (2013).

The set of m HF regression models is given by:

$$\mathbf{y}_{H,j} = \mathbf{X}_{H,j}\beta_j + \mathbf{u}_{H,j} \tag{2.57}$$

With:

$$\mathbf{E}\left[\mathbf{u}_{H,j}\right] = \mathbf{0}$$
$$\mathbf{E}\left[\mathbf{u}_{H,i}\mathbf{u}'_{H,j}\right] = \mathbf{V}_{H,i,j} \qquad \forall i,j = 1,\ldots,m$$

Where $\mathbf{X}_{H,j}$ are the m matrices including the related series of $\mathbf{y}_{H,j}$. Models 2.57 can be grouped and re-written in the following form:

$$\begin{bmatrix} \mathbf{y}_{H,1} \\ \vdots \\ \mathbf{y}_{H,j} \\ \vdots \\ \mathbf{y}_{H,m} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{H,1} & & & & \\ & \ddots & & & \\ & & \mathbf{X}_{H,j} & & \\ & & & \ddots & \\ & & & & \mathbf{X}_{H,m} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_j \\ \vdots \\ \beta_m \end{bmatrix} + \begin{bmatrix} \mathbf{u}_{H,1} \\ \vdots \\ \mathbf{u}_{H,j} \\ \vdots \\ \mathbf{u}_{H,m} \end{bmatrix}$$

Or, more compactly:

$$\mathbf{y}_H = \mathbf{X}_H\beta + \mathbf{u}_H \tag{2.58}$$

Which is not directly observable. Similarly to what was done to model 2.17, it is possible to pre-multiply for $\mathbf{H}$:

$$\mathbf{H}\mathbf{y}_H = \mathbf{H}\mathbf{X}_H\beta + \mathbf{H}\mathbf{u}_H$$
$$\mathbf{y}_L = \mathbf{X}_L\beta + \mathbf{u}_L \tag{2.59}$$

With:

$$\mathbf{E}\left[\mathbf{u}_L\mathbf{u}'_L \mid \mathbf{X}_H\right] = \mathbf{V}_L = \mathbf{H}\mathbf{V}_H\mathbf{H}'$$

The solutions are given:

$$\hat{\beta} = \left(\mathbf{X}_L'\mathbf{V}_L^{-}\mathbf{X}_L\right)^{-1}\mathbf{X}_L'\mathbf{V}_L^{-}\mathbf{y}_L \tag{2.60}$$

$$\hat{\mathbf{y}}_H = \mathbf{X}_H\hat{\beta} + \mathbf{V}_H\mathbf{H}'\mathbf{V}_L^{-}\left(\mathbf{y}_L - \mathbf{X}_L\hat{\beta}\right) \tag{2.61}$$

Where $\mathbf{V}_L^{-}$ is the Moore-Penrose generalised inverse of $\mathbf{V}_L$. By setting $\mathbf{L} = \mathbf{V}_H\mathbf{H}'\mathbf{V}_L^{-}$, and considering that $\hat{\mathbf{u}}_L = \mathbf{y}_L - \mathbf{X}_L\hat{\beta}$, solution 2.61 can be written as follows:

$$\hat{\mathbf{y}}_H = \mathbf{X}_H\hat{\beta} + \mathbf{L}\hat{\mathbf{u}}_L \tag{2.62}$$

Which corresponds to solution 2.21 in the multivariate case and has the same interpretation.

Finally we can express the covariance matrix of the estimated values:

$$\begin{aligned}&\mathbf{E}\left[\left(\hat{\mathbf{y}}_H - \mathbf{y}_H\right)\left(\hat{\mathbf{y}}_H - \mathbf{y}_H\right)'\right] = \\ &\qquad\left(\mathbf{I}_n - \mathbf{L}\mathbf{H}\right)\mathbf{V}_H + \left(\mathbf{X}_H - \mathbf{L}\mathbf{X}_L\right)\left(\mathbf{X}_L'\mathbf{V}_L^{-}\mathbf{X}_L\right)^{-1}\left(\mathbf{X}_H - \mathbf{L}\mathbf{X}_L\right)'\end{aligned}$$

As the matrices $\mathbf{V}_{H,i,j}$ are normally unknown, they have to be estimated by making assumptions on the residuals $\mathbf{u}_H$. Considering also the computational aspects, at least two approaches have been considered in the literature (Di Fonzo, 2003a; Eurostat, 2013):

1. Multivariate white noise.

   In this case the covariances are expressed by:

$$\mathbf{E}\left[\mathbf{u}_{H,i}\mathbf{u}_{H,j}'\right] = \sigma_{i,j}$$

Which is:

$$\mathbf{E}\left[\mathbf{u}_H \mathbf{u}_H'\right] = \mathbf{\Sigma} \otimes \mathbf{I}_n \tag{2.63}$$

Where the elements $\sigma_{i,j}$ of the matrix $\mathbf{\Sigma}$ can be estimated using the OLS residuals $\hat{\mathbf{u}}_L$. Di Fonzo (1990) shows that in this case the inversion of the matrix $\mathbf{V}_L$ is simplified by a suitable partition of $\mathbf{\Sigma}$ obtained by deleting the last row and the last column.

2. Multivariate random walk.

   This is the multivariate generalisation of the method proposed by Fernández (1981):

$$\begin{aligned} u_{H,t} &= u_{H,t-1} + \varepsilon_t \\ u_{H,0} &= 0 \\ \mathbf{E}\left[\varepsilon_t\right] &= 0 \\ \mathbf{E}\left[\varepsilon_r \varepsilon_s'\right] &= \begin{cases} 0 & \text{if } r \neq s \\ \mathbf{\Sigma} & \text{if } r = s \qquad r,s = 1,\dots,n \end{cases} \end{aligned}$$

   Thus:

$$\begin{aligned} \mathbf{E}\left[\mathbf{u}_{H,t}\right] &= 0 \\ \mathbf{E}\left[\mathbf{u}_{H,r} \mathbf{u}_{H,s}'\right] &= \mathbf{\Sigma} \min\left\{r,s\right\} \end{aligned}$$

   Which means:

$$\mathbf{E}\left[\mathbf{u}_H \mathbf{u}_H'\right] = \mathbf{\Sigma} \otimes \mathbf{D}'\mathbf{D} \tag{2.64}$$

   Where the elements $\sigma_{i,j}$ of the matrix $\mathbf{\Sigma}$ are again estimated using the OLS residuals $\hat{\mathbf{u}}_L$.

As for the univariate case, an extrapolation problem is faced when $n > sN$. In this case, the extra k observations for the m time series are the following:

$$\mathbf{y}_{H,e} = \begin{bmatrix} \mathbf{y}_{H,e,sN+1} \\ \vdots \\ \mathbf{y}_{H,e,sN+h} \\ \vdots \\ \mathbf{y}_{H,e,sN+k} \end{bmatrix}$$

Where $sN + k = n$.

Being also $\mathbf{X}_{H,e}$ and $\mathbf{u}_{H,e}$, the correspondent matrix of related series and vector of disturbances, respectively, the model could be expressed as follows:

$$\begin{bmatrix} \mathbf{y}_H \\ \mathbf{y}_{H,e} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_H \\ \mathbf{X}_{H,e} \end{bmatrix} \beta + \begin{bmatrix} \mathbf{u}_H \\ \mathbf{u}_{H,e} \end{bmatrix}$$

Or, in compact form:

$$\tilde{\mathbf{y}}_H = \tilde{\mathbf{X}}_H \beta + \tilde{\mathbf{u}}_H \tag{2.65}$$

With:

$$\mathbf{E}\left[\tilde{\mathbf{u}}_H\right] = 0$$

$$\mathbf{E}\left[\tilde{\mathbf{u}}_H \tilde{\mathbf{u}}'_H\right] = \tilde{\mathbf{V}}_H = \begin{bmatrix} \mathbf{V}_H & \mathbf{\Gamma}' \\ \mathbf{\Gamma} & \mathbf{V}_{H,e} \end{bmatrix}$$

Where $\mathbf{\Gamma} = \mathbf{E}\left[\mathbf{u}_{H,e}\mathbf{u}'_H\right]$ and $\mathbf{V}_{H,e} = \mathbf{E}\left[\mathbf{u}_{H,e}\mathbf{u}'_{H,e}\right]$.

Di Fonzo (1990, 2003a) distinguishes the cases when a vector of contemporaneous constraints is present and also when it is not.

If there is no contemporaneous constraint he refers to *pure extrapolation*. In this case the solution of model 2.65 is given by:

$$\hat{\mathbf{y}}_{H,e} = \mathbf{X}_{H,e}\hat{\beta} + \mathbf{\Gamma}\mathbf{H}'\mathbf{V}_L^{-}\left(\mathbf{y}_L - \mathbf{X}_L\hat{\beta}\right) \tag{2.66}$$

It is obvious that this is not the solution for a reconciliation problem, but it can be considered as a multivariate method for temporal disaggregation.

When the contemporaneous constraint $\mathbf{z}_H$ is present, the author refers to *constrained extrapolation*. In this case, given $\mathbf{z}_{H,e}$, the last k observations of $\mathbf{z}_H$ such that $\mathbf{H}_e\mathbf{y}_{H,e} = \mathbf{z}_{H,e}$, the solution of the reconciliation problem is given by:

$$\tilde{\hat{\beta}} = \left(\tilde{\mathbf{X}}_L'\tilde{\mathbf{V}}_L^{-}\tilde{\mathbf{X}}_L\right)^{-1}\tilde{\mathbf{X}}_L'\tilde{\mathbf{V}}_L^{-}\tilde{\mathbf{y}}_L \tag{2.67}$$

$$\tilde{\hat{y}}_H = \tilde{\mathbf{X}}_H\tilde{\hat{\beta}} + \tilde{\mathbf{V}}_H\tilde{\mathbf{H}}'\tilde{\mathbf{V}}_L^{-}\left(\tilde{\mathbf{y}}_L - \tilde{\mathbf{X}}_L\tilde{\hat{\beta}}\right) \tag{2.68}$$

Where $\tilde{\mathbf{y}}_L = \tilde{\mathbf{H}}\tilde{\mathbf{y}}_H$, $\tilde{\mathbf{X}}_L = \tilde{\mathbf{H}}\tilde{\mathbf{X}}_H$, $\tilde{\mathbf{V}}_L = \tilde{\mathbf{H}}\tilde{\mathbf{V}}_H\tilde{\mathbf{H}}'$ and $\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_e \end{bmatrix}$.

Solution 2.68 encompasses distribution, interpolation and extrapolation, and it is the solution of a reconciliation problem because it solves both temporal and contemporaneous constraints.

2.3.2 Two-Step reconciliation techniques

Two-step reconciliation methods have been firstly introduced by Quenneville and Rancourt (2005), in order to reconcile series which have been seasonally adjusted using a direct approach, and further developed by Di Fonzo and Marini (2011b). This approach has the advantage of being very simple and effective, and does not require big computational problems. The basic idea is to divide the reconciliation problem in two steps:

1. Use the univariate modified PFD Denton technique on each variable of the system, solving the temporal constraint.
2. Balance the system of time series in each LF period, solving the contemporaneous constraint.

The methodology has been developed in order to reconcile a system of time series with binding exogenous constraints.

2.3.2.1 The first step

In the first step (which is basically the same in all the two-step procedures proposed in the literature so far), the univariate modified Dentond PFD is applied to the m series, obtaining the benchmarked (temporally disaggregates) HF series $\mathbf{b}_{H,j}$, such that:

$$\mathbf{C}\mathbf{b}_{H,j} = \mathbf{y}_{L,j}$$

Or, in compact form:

$$(\mathbf{I}_m \otimes \mathbf{C})\,\mathbf{b}_H = \mathbf{y}_L$$

Where $\mathbf{b}_H = (\mathbf{b}_{H,1}, \ldots, \mathbf{b}_{H,j}, \ldots, \mathbf{b}_{H,m})'$.

However, given an already benchmarked contemporaneous constraint $\mathbf{z}_H$, the series $\mathbf{b}_{H,j}$ do not satisfy the contemporaneous constraints:

$$\sum_{j=1}^{m} \mathbf{b}_{H,j} \neq \mathbf{z}_H$$

Or, in compact form:

$$\left(\mathbf{1}'_m \otimes \mathbf{I}_n\right) \mathbf{b}_H \neq \mathbf{z}_H$$

Generally speaking, it is believed that after the first step, the discrepancies between the sum of the $\mathbf{b}_{H,j}$ and the contemporaneous constraints $\mathbf{z}_H$ are reduced. This is because the contemporaneous constraints are already temporally benchmarked and satisfy the contemporaneous constraint at LF level:

$$\mathbf{C}\mathbf{z}_H = \mathbf{z}_L$$

$$\mathbf{C}\mathbf{z}_H = \sum_{j=1}^{m} \mathbf{y}_{L,j}$$

2.3.2.2 The second step

In the second step the balancing procedure is applied to each of the N LF periods, keeping the temporal constraint satisfied. Denoting $\mathbf{b}_{H,j}$, the m reconciled series, both temporal and contemporaneous constraints will be satisfied:

$$\begin{aligned} \mathbf{C}\mathbf{r}_{H,j} &= \mathbf{y}_{L,s} \quad \forall j = 1, \dots, m \\ \sum_{j=1}^{m} \mathbf{r}_{H,j} &= \mathbf{z}_H \end{aligned}$$

Or, in compact matrix form:

$$\mathbf{H}\mathbf{r}_H = \mathbf{y}_a$$

Where $\mathbf{r}_H = (\mathbf{r}_{H,1}, \dots, \mathbf{r}_{H,j}, \dots, \mathbf{r}_{H,m})'$.

In order to solve the contemporaneous constraints, a constrained optimisation of a quadratic function is applied, following the approach by Stone et al. (1942).

Quenneville and Rancourt (2005) propose to solve the following problem:

$$\begin{aligned} &\min_{r_{H,j}} \sum_{j=1}^{m} \sum_{t=(T-1)s+1}^{Ts} \frac{(r_{H,j,t} - b_{H,j,t})^2}{b_{H,j,t}} \quad \forall T = 1, \dots, N \\ &\text{s.t.} \sum_{j=1}^{m} r_{H,j,t} = z_{H,t} \quad \forall t = (T-1)\,s + 1, \dots, Ts \end{aligned} \tag{2.69}$$

Which basically corresponds to balancing the N intra-LF tables according to criteria 2.46.

A slightly different approach has been proposed by Dagum and Cholette (2006), which, building on Beaulieu and Bartelsman (2004), propose the following criteria (Di Fonzo and Marini, 2011b):

$$\begin{aligned}
&\min_{r_{H,j}} \sum_{j=1}^{m} \sum_{t=(T-1)s+1}^{Ts} \frac{\left(r_{H,j,t} - b_{H,j,t}\right)^2}{|b_{H,j,t}|} \quad \forall T = 1, \ldots, N \\
&\text{s.t.} \sum_{j=1}^{m} r_{H,j,t} = z_{H,t} \quad \forall t = (T-1)\, s + 1, \ldots, Ts
\end{aligned} \tag{2.70}$$

This approach has the advantage that can be applied also when one or more $b_{H,j,t}$ are negative, a situation which is often encountered in practice (see for example the balancing done in national accounts Eurostat, 2010).

Finally, in order to preserve the reliability of the variables, Di Fonzo and Marini (2011b), building on Stuckey et al. (2004), consider the following problem:

$$\begin{aligned}
&\min_{r_{H,j}} \sum_{j=1}^{m} \sum_{t=(T-1)s+1}^{Ts} \left(\frac{r_{H,j,t} - b_{H,j,t}}{b_{H,j,t}}\right)^2 \quad \forall T = 1, \ldots, N \\
&\text{s.t.} \sum_{j=1}^{m} r_{H,j,t} = z_{H,t} \quad \forall t = (T-1)\, s + 1, \ldots, Ts
\end{aligned} \tag{2.71}$$

In matrix form, the problem can be written as follows:

$$\begin{aligned}
&\min_{r_{H,T}} \left(\mathbf{r}_{H,T} - \mathbf{b}_{H,t}\right)' \boldsymbol{\Omega} \left(\mathbf{r}_{H,T} - \mathbf{b}_{H,T}\right) \quad \forall T = 1, \ldots, N \\
&\text{s.t. } \mathbf{H}\mathbf{r}_{H,T} = \mathbf{y}_{a,T}
\end{aligned} \tag{2.72}$$

Where:

- $\mathbf{r}_{H,T} = (\mathbf{r}_{H,T,1}, \ldots, \mathbf{r}_{H,T,j}, \ldots, \mathbf{r}_{H,T,m})'$, $\forall T = 1, \ldots, N$, is the $ms \times 1$ vector including the values of the m reconciled variables for the LF value T.

- $\mathbf{b}_{H,T} = (\mathbf{b}_{H,T,1}, \ldots, \mathbf{b}_{H,T,j}, \ldots, \mathbf{b}_{H,T,m})'$, $\forall T = 1, \ldots, N$, is the $ms \times 1$ vector including the values of the m balanced variables for the LF value T, as obtained after the first step.

- $\mathbf{y}_{a,T} = \begin{bmatrix} z_{H,T} \\ \mathbf{y}_{L,T} \end{bmatrix}$, with $\mathbf{y}_{L,T} = (y_{L,T,1}, \ldots, y_{L,T,j}, \ldots, y_{L,T,m})'$, is the $(m+1) \times 1$ vector containing all the the contemporaneous and temporal constraints of the LF value T.

- $\mathbf{H}$ is built considering only one LF period.

The $ms \times ms$ matrix $\boldsymbol{\Omega}$ is chosen according to the approach applied in the second step. For problems 2.69, 2.70 and 2.71, $\boldsymbol{\Omega}$ is a diagonal matrix with non-zero entries equal to $\dfrac{1}{b_{H,T,j}}$, $\dfrac{1}{|b_{H,T,j}|}$ and $\dfrac{1}{b^2_{H,T,j}}$ respectively.

The final solution is then given by the following expression:

$$\hat{\mathbf{r}}_{H,T} = \mathbf{b}_{H,T} + \boldsymbol{\Omega}^{-1}\mathbf{H}' \left(\mathbf{H}\boldsymbol{\Omega}^{-1}\mathbf{H}'\right)^{-1} (\mathbf{y}_{a,T} - \mathbf{H}\mathbf{b}_{H,T}) \tag{2.73}$$

Di Fonzo and Marini (2011b) suggest using covariance matrices of the variables to be balanced or alterability coefficients defined by the users for each series, in a subjective way.

However, this kind of information is normally unavailable. Hence, in order to decide which two-step approach to follow, the implicit assumption of approaches 2.69 and 2.70 is that smaller series are of lower quality than larger ones. This is because smaller series are adjusted relatively more than bigger ones, independently from their relative reliabilities. If this assumption is not true, it might result in an alteration of the temporal profile of the smaller series moving between different LF values which causes a step problem.

If the same reliability (in terms of coefficient of variations) is assumed for all the series of the system, the second step of the reconciliation should be performed according to approach 2.71.

The authors also specify that approach 2.71 gives results which are very close to the multivariate Denton method.

Di Fonzo and Marini (2012a, 2015) extended the two-step reconciliation methods by applying the GRP method in the first step. Such approach gives results which are very similar to the multivariate version of the GRP method.

2.4 Conclusions

Some concluding remarks could be done regarding the techniques available in the literature for temporal disaggregation, balancing and reconciliation. If from one hand, according to the different situations, one should search for the best technique to be used from a statistical point of view, it is also important to mention practical problems which statistical agencies face during the production of official statistics.

Temporal disaggregation and benchmarking techniques have been nowadays widely discussed in the literature and they are ordinarily used in practice. Still some remarks are to be done, in order not to make the mistake of using the wrong method in a certain situation.

The approaches by Fernández and Litterman imply that $u_{H,t}$ is non-stationary, so that the estimation is basically performed using differenced series. In this case, the assumption is that there is no long term relationship between the LF series and the relative indicator(s), making the two variants unsuitable for stationary or co-integrated time series. Without considering the co-integration, OLS or GLS techniques could lead to spurious regressions. Although for domains like national accounts this could be a small problem, because the indicators used normally approximate the variables to be estimated, it cannot be ignored. The problem is principally related to the extrapolation part, where the absence of co-integration when using a non-differentiate model (such as Chow-Lin), may lead to large revisions when the extrapolated values are later replaced by the LF series.

Thus, for stationary or co-integrates series, the original Chow-Lin approach is to be used (Santos Silva and Cardoso, 2001; Sax and Steiner, 2013). Alternatively, a good solution is to use the dynamic model, which could reduce the probability of misspecification.

Regarding the revisions which occur when a new LF benchmark is available, all the methods described here produce a completely revised HF series. A solution which is sometimes used when applying the Denton methodology, is to apply the procedure on a moving window so that past data are not revised at all. If from one point of view this is an effective solution, it is not completely correct from a statistical point of view. It is a matter of fact that revisions to the past data are normally very close to zero and, moreover, the availability of new information can only generate better results.

It is also very important to mention that Denton's approach follows a movement preservation principle of the related indicator (preliminary series), generating an estimate of the target series which is similar to the indicator even if this indicator is not correlated with the target HF series.

When seen as an indirect approach, the Denton method could lead to good results, and various approaches could be envisaged (see, *inter alia*, the French approach described in Eurostat, 2013). However, this normally foresees the application of a regression model at LF level. Thus, in this case it seems better to proceed directly with a regression expressed at HF level by using an optimal regression-based approach.

As regards to the extrapolation, it is quite clear that the Denton method has drawbacks, since the extrapolation is performed using only the latest LF and related HF periods. Whilst the enhanced Denton methodology could be used to bypass this problem, the issue of how to estimate the benchmark-to-ratios remains. On the other hand, a regression-based approach might perform very well, but generates unsatisfying results when the model is misspecified.

As a conclusion, the optimal methods seem to be more appropriate from a statistical point of view, giving results according to the correlation between the indicator and the final estimate. However, it must be stressed that, in general, optimal methods would require more attention from the user, who has to evaluate the model estimated, and they also need a sufficient number of observed periods in order to be able to estimate a model. This problem is not encountered when applying the Denton method, as the results can be obtained also when the number of available LF periods is very low.

Balancing is a very known issue, and the techniques which are still used basically are related to the work done by Stone et al. (1942). Indeed the papers by Sir Richard Stone are big milestones in the field of balancing (for more information about the contributions of the author in the field, see Marangoni and Rossignoli, 2014).

Statistical and mathematical balancing techniques are probably the best choice when qualitative information is unavailable, and both the general approach proposed by Stone et al. (1942) and the multivariate proportional adjustment seem to perform well. The *ad hoc* balancing is however often chosen at least to solve part of the balancing problem, and this is the best choice given that the qualitative information is correct.

Reconciliation techniques are definitely less used in practice, or done only by those statistical agencies which have the luck of having some expert in the field. Simultaneous approaches have been developed since the early nineties, but their use in practice has been limited due to the computational issue, while two-step reconciliation techniques have been only recently developed and seem to be a very good alternative.

All two-step reconciliation methods have the initial assumption that there is no need to look at the dynamics in the second step of the procedure because the temporal profiles are preserved in the first step and not altered in the second step. In their paper, Di Fonzo and Marini (2011b) show that this is generally true in practical problems. Moreover, the authors state that approach 2.71 gives results which are very close to the simultaneous Denton method and that *"very good performances have been registered"*.

Thus, two-step reconciliation practices appear to be very convenient because of their simplicity and low computational time. It is a matter of fact that statistical agencies have to also deal with a third constraint: the *time constraint*. Very often, preliminary estimates (such as seasonally adjusted data directly obtained) have to be reconciled and validated within hours (maybe due to a specific regulation, see for example Eurostat, 2010), creating big challenges for the users.

Finally, it must be stressed that all benchmarking, balancing and reconciliation techniques are designed to adjust series (variables) which are in any case close to the constraints to be fulfilled. In other words, the quality of the results are inversely proportioned to the discrepancies between the preliminary series and the constraints.

2.4.1 A schematic resume

In the literature, some kind of classifications have been done for temporal disaggregation and balancing techniques (see, for example, Marcellino, 1999 and Lahr and De Mesnard, 2004). An interesting survey of all the methods developed for temporal disaggregation can be found in Pavía-Miralles (2010), which, however, does not present a schematic view of the available methods.

Here an innovative schematic resume of the techniques for temporal disaggregation, balancing and reconciliation is presented. If, on one hand, it does not differ too much from the classifications already done, on the other hand for the first time it represents the methods in a schematic way.

Table 2.1 presents the techniques for temporal disaggregation and benchmarking, classifying them according to the use or not of a related indicator (or preliminary series). Table 2.2 presents the available methodology for balancing, and finally the techniques for reconciliation are classified in table 2.3.

Table 2.1: Techniques for temporal disaggregation

Methodology for Temporal Disaggregation	
Without indicators	**With indicators**
Naïve	**Naïve**
	Pro-rata
Interpolation – Lisman-Sandee – Zani	**Min. of a loss function** – Denton – Cholette (modified Denton) – GRP
Min. of a loss function – BFL – Marcellino	**Regression-based models** **Static** – Chow-Lin – Fernández – Litterman **Dynamic** – Salazar – Gregoir – SSC
Time series methods – Wei-Stram – Al-Osh – Guerrero	**Missing observations** – Jones – Gomez-Maravall
	Structural models – Moauro-Savio – Proietti

Table 2.2: Techniques for balancing

Methodology for Balancing
Naïve
Proportional (including plus-minus)
Proportional squared
Bi-proportional methods – RAS (including plus-minus) – TRAS
Minimisation of a loss function: Stone Approaches by: – Almon – Lahr – Matuszewski – Deming – Dagum-Cholette – Stuckey
***Ad hoc* balancing**

Table 2.3: Techniques for reconciliation

Methodology for Reconciliation	
Simultaneous	**Two-Step**
Multivariate Denton – Di Fonzo-Marini – Bikker *et. al*	**First step** – Modified Denton PFD – GRP
Multivariate BLUE – Rossi – Di Fonzo	**Second Step** – Quenneville-Rancourt – Dagum-Cholette – Di Fonzo-Marini

Chapter 3

An alternative two-step reconciliation method

Two-step reconciliation methods described in Chapter 2 seem to be very promising thanks to their simplicity and efficiency, and have no problems with the computational burden. On the other hand, the techniques presented so far in the literature are not very flexible.

In this chapter, an alternative methodology for the two-step reconciliation methods is presented. This methodology could be seen as a generalisation of all the methods described in the literature, which are extended to the possibility of using different techniques in both the first and the second steps, adding a clear flexibility to the two-step reconciliation techniques. Such methodology has been implemented in Java, as a plug-in of JDemetra+ (Grudkowska, 2015), the official European tool designed for seasonal adjustment.[1]

Finally, the focus will be given to the case when reconciliation is applied after seasonal adjustment, presenting a statistical test for identifying common seasonal patterns between different series, which could be used to determine at which level the series need to be adjusted, and thus determine the system(s) of series which will be reconciled.

A short description of the topics presented in this chapter can be found in Infante and Scepi (2017).

3.1 Regression-based two-step reconciliation

When developing the two-step reconciliation methods, all the literature to date focuses on the second step. They all use the univariate modified Denton approach in the first step (and in particular they work with the PFD variant), with the exception of Di Fonzo and Marini (2012a), which propose to use the GRP approach in the first step.

However, as seen in Chapter 2, other temporal disaggregation methods could be used in the first step, according to the different situations encountered in practice. In particular, the movement preservation principle followed by the Denton methodology might not be appropriate when the

[1] JDemetra+ has been officially recommended, since 2 February 2015, to the members of the European Statistical System (ESS) and the European System of Central Banks (ESCB) as software for seasonal and calendar adjustment of official statistics. More details regarding the tool and how to download and install are on the CROS portal (https://ec.europa.eu/eurostat/cros/).

user is unsure about the correlation between the target unknown series and the relative indicator, or when there are many HF periods to extrapolate. Thus, in these cases, a methodology which also detects the degree of correlation should be used, like the regression-based optimal approaches.

From this perspective, two-step approaches are very flexible, making possible the application of whatever temporal disaggregation (benchmarking) technique in the first step.

Given a set of m HF time series which need be to be reconciled to their LF counterparts $\mathbf{y}_{L,j}$ and to the accounting constraint $\mathbf{z}_H$, it is possible to apply a univariate regression-based technique in order to estimate the benchmarked HF series:

$$\mathbf{y}_{H,j} = \mathbf{X}_{H,j}\beta + \mathbf{u}_{H,j} \quad \forall j = 1, \ldots, m \tag{3.1}$$

Obtaining the following results:

$$\hat{\beta}_j = \left(\mathbf{X}'_{L,j}\mathbf{V}^{-1}_{L,j}\mathbf{X}_{L,j}\right)^{-1}\mathbf{X}'_{L,j}\mathbf{V}^{-1}_{L,j}\mathbf{y}_{L,j} \tag{3.2}$$

$$\hat{\mathbf{y}}_{H,j} = \mathbf{b}_{H,j} = \mathbf{X}_{H,j}\hat{\beta}_j + \mathbf{V}_{H,j}\mathbf{C}'\mathbf{V}^{-1}_{L,j}\left(\mathbf{y}_{L,j} - \mathbf{X}_{L,j}\hat{\beta}_j\right) \tag{3.3}$$

Where the covariance matrices $\mathbf{V}_{H,j}$ are estimated according to the Chow-Lin, Fernández or Litterman solutions.

The m solutions are grouped in the vector:

$$\hat{\mathbf{y}}_H = \mathbf{b}_H = (\mathbf{b}_{H,1}, \ldots, \mathbf{b}_{H,j}, \ldots, \mathbf{b}_{H,m})'$$

Such vector $\mathbf{b}_H$ is then balanced for each LF period T:

$$\min_{r_{H,T}} \left(\mathbf{r}_{H,T} - \mathbf{b}_{H,t}\right)' \boldsymbol{\Omega} \left(\mathbf{r}_{H,T} - \mathbf{b}_{H,T}\right) \quad \forall T = 1, \ldots, N$$

$$\text{s.t.} \mathbf{H}\mathbf{r}_{H,T} = \mathbf{y}_{a,T}$$

Obtaining the following results:

$$\hat{\mathbf{r}}_{H,T} = \mathbf{b}_{H,T} + \boldsymbol{\Omega}^{-1}\mathbf{H}' \left(\mathbf{H}\boldsymbol{\Omega}^{-1}\mathbf{H}'\right)^{-1} \left(\mathbf{y}_{a,T} - \mathbf{H}\mathbf{b}_{H,T}\right)$$

Where $\boldsymbol{\Omega}$ is, for example, the diagonal matrix with generic term $\frac{1}{b^2_{H,T,j}}$, according to the approach suggested by Di Fonzo and Marini (2011b). Alternatively, the approaches by Quenneville and Rancourt (2005) or Dagum and Cholette (2006) could be used.

Such way of dealing with reconciliation of the time series allows the user to use qualitative information by partially applying an *ad hoc* balancing in the second step, leaving only a sub-set of the variables to be finally reconciled and changing the contemporaneous constraints accordingly.

Although not expressively foreseen by Di Fonzo and Marini (2011b), the contemporaneous constraints $\mathbf{z}_H$ could also be the result of a temporal disaggregation technique. Thus, given the LF series of the contemporaneous constraints $\mathbf{z}_L$, and a set (usually one) of related indicators which are included in $\mathbf{X}_{H,z}$, it is easy to estimate $\mathbf{z}_H$ by applying, for example, a regression-based technique:

$$\begin{aligned}
\mathbf{z}_H &= \mathbf{X}_{H,z}\beta + \mathbf{u}_{H,z} \\
\hat{\beta}_z &= \left(\mathbf{X}'_{L,z}\mathbf{V}_{L,z}^{-1}\mathbf{X}_{L,z}\right)^{-1}\mathbf{X}'_{L,z}\mathbf{V}_{L,z}^{-1}\mathbf{y}_{L,z} \\
\hat{\mathbf{z}}_H &= \mathbf{X}_{H,z}\hat{\beta}_z + \mathbf{V}_{H,z}\mathbf{C}'\mathbf{V}_{L,z}^{-1}\left(\mathbf{y}_{L,z} - \mathbf{X}_{L,z}\hat{\beta}_z\right)
\end{aligned}$$

Following this method, it is possible to implement a cascade approach for reconciling more systems of time series when they are nested.

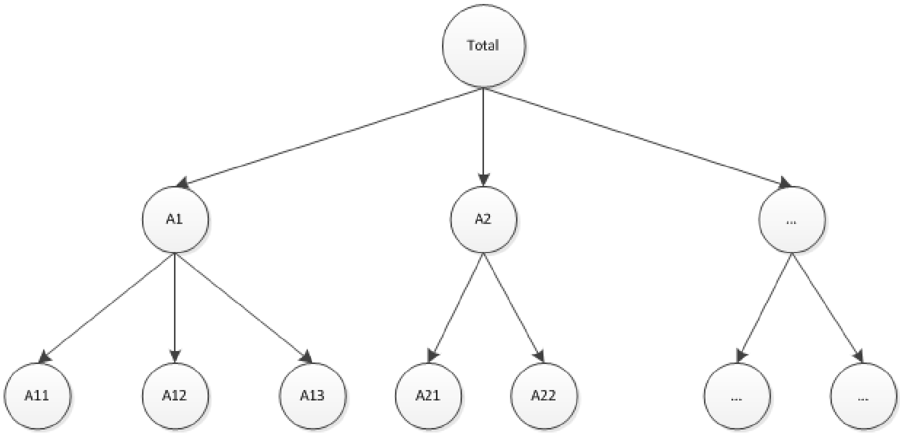
Given that $\hat{\mathbf{z}}_H = \sum_{j=1}^{m} \mathbf{y}_{H,j}$, the m series $\mathbf{y}_{H,j}$ might be the contemporaneous constraints of m systems of time series (or less), so that:

$$\mathbf{y}_{H,j} = \sum_{h \in j} \mathbf{a}_{H,j,h}$$

Where $\mathbf{y}_{H,j}$ are the time series of the first layer (total, one system of time series), and $\mathbf{a}_{H,j,h}$ are the time series of the second layer (a_h, m systems of time series). A generic example of a hierarchical chart is presented in Figure 3.1.

It is clear that the results obtained in the second layer of the systems will be dependent from the results obtained on the temporal disaggregation of the contemporaneous constraint in the first layer. However this is a drawback only in theory, while it could actually be a good solution when the overall quality of the time series in the second layer is low, which is often the case in official statistics, especially when going to very detailed series.

Figure 3.1: Generic hierarchical chart



More details regarding hierarchical time series are in Hyndman et al. (2011) and Taieb et al. (2017).

While here only a formal presentation of the case of two layers has been given, the procedure can be applied to whatever number of layers, keeping in mind that the results obtained in the lower layers will be dependent on all the results obtained by the temporal disaggregation of the contemporaneous constraints of all the higher layers.

Two practical examples are presented in Chapter 5: a small scale case after seasonal adjustment (monthly industrial production index), and a medium-large scale case for the aggregation of euro area quarterly sector accounts data.

3.2 Algorithm and software

In recent years, Eurostat developed a tool for temporal disaggregation, JEcotrim, composed of a set of plug-ins of the official tool for seasonal adjustment, JDemetra+, version 1.2.0. Different plug-ins have been developed in order to perform each available method: modified Denton, Chow-Lin, Fernández, Litterman, RAS-PM and two-step reconciliation (always using the modified PFD Denton approach in the first step). An additional plug-in has also been developed in order to perform batch jobs of the available methods. In this case, after the job is initiated by the user, the program is run in the background without interaction by the user. Such way of working with statistical methods, which is also available for the standard seasonal adjustment functionalities of JDemetra+, is very useful in practice when the user has to run a big number of series with very limited time (the so called time constraint), but could also bring to non-satisfactory results when the initial specifications are not correct for specific time series.

In order to be useful in an official statistics' framework, where the time constraint is often a real problem, it is very important that a good tool for two-step reconciliation is able to:

1. Handle more systems of time series at the same time.
2. Be flexible enough, so that different choices are available on both the first and the second steps.

While the first point is handled by the original JEcotrim solution (only in batch mode), for the first step of the reconciliation procedure, the tool obliges the user to use a modified Denton PFD technique. In order to solve this problem, a new, more flexible, version of JEcotrim has been developed.

The first very important improvement is on the compatibility of JEcotrim with the latest version of JDemetra+, 2.1.0. The original JEcotrim was only able to work with version 1.2.0, a beta version which was not yet recommended as official tool by Eurostat.

The original code is based on a general model, which is called as shown in the box below, where "*phi1*" and "*phi2*" correspond to the Chow-Lin's ρ and Litterman's α, respectively.

```
public abstract class GlobalMethod extends TemporalDisaggregationMethod {
    ...
    @Override
    public TemporalDisaggregationMethodResult process() throws Exception {
        compute();

        // Estimation scanning phi, else phi is given by the user
        double phi = 0.0;
        ScanningResult scanning = new ScanningResult();
        if (arflag) {
          scanning = Scanning.scanning(y0, X, C, em, nbStep, phi1, phi2,
               method);
          phi = scanning.getPhi();
        } else {
          phi = arfix;
          scanning.setPhi(phi);
        }

        // Getting HF and LF covariance matrix, annualised preliminary
        V(phi);
        X0();
        V0();
        GeneralLeastSquaresResult gls = GeneralLeastSquares.gls(y0, X0, V0);
```

```
        // GLS final estimates of the parameters conditional to the estimated
            phi
        TemporalDisaggregationMethodResult eco = finalized(gls.getBeta(),gls.
            getSsr());
        eco.setGls(gls);
        eco.setScanning(scanning);

        return eco;
    }
    ...
}
```

The method to be applied is than recalled according to different functions. The box below shows the case of Chow-Lin.

```
public class ChowlinMethod extends GlobalMethod {
    ...
    @Override
    protected void V(double phi) {

        // Compute Chow-Lin covariance matrix
        V = null;
        {
            HDPMatrix sequentialMat = new HDPMatrix(n, 1);
            sequentialMat.seqm(1.0, phi);
            V = HDPMatrix.Convert2Toeplitz(sequentialMat.internalStorage(),
                false);
            V.mul(1.0 / (1.0 - Math.pow(phi, 2)));
        }
    }

    @Override
    protected void X() {

        // Compute X matrix with or without intercept
        if (hfm == HighFreqDisturb.WITH) {
            Matrix o = new Matrix(n, 1);
            o.set(1.0);

            X = Functions.concatenationHorizontale(o, X);
        }
    }
}
```

A similar code is provided for the Litterman method, while for the

Fernández method the general model is replaced since the parameters of the residual term are fixed and thus do not need estimation.

In this study, three new plug-ins have been created, in order to allow the user to choose between Chow-Lin, Fernández and Litterman in the first step, without forcing the use of modified PFD Denton, as described in the previous section. The Java code which recalls the Chow-Lin method in the first step is shown in the box below.

```
// Step 1
private MMatrix compute(MMatrix LF, MMatrix HF, FreqAggrOrder s) throws
    Exception {
        Matrix mat = null;

        for (int i = 0; i < LF.getColumnsCount(); i++) {
            // INPUT (LF and HF series)
            Matrix m_lf = LF.getSubMatrixFullRows(i, i + 1);
            Matrix m_hf = HF.getSubMatrixFullRows(i, i + 1);

            // Chow-Lin process for each series
            PrintMatrix.ACTIVE_LOG = false;
            ChowlinMethod chowlin = new ChowlinMethod(m_lf, m_hf,
                HighFreqDisturb.WITH, s, taggr, EstimMethod.ML, 100, 0.00,
                0.99, true, 0.0);
            TemporalDisaggregationMethodResult r = chowlin.process();
            PrintMatrix.ACTIVE_LOG = true;
            if (mat == null) {
                mat = r.getYdisag();
            } else {
                mat = Functions.concatenationHorizontale(mat, r.getYdisag());
            }
        }

        // Results first step using Chow-Lin
        MMatrix mmat = new MMatrix(mat, null);
        return mmat;
    }
```

As it is, the process estimates the Chow-Lin model with intercept, using the maximum likelihood approach, scanning in 100 equally distant possible values of ρ between 0 and 0.99.

Similar coding has been developed for the Fernández and Litterman methods, by changing the process function as shown in the boxes below. In the case of Litterman, there is no scanning as the parameter of the residual AR process is fixed to be equal to 1.

```
FernandezMethod fernandez = new FernandezMethod(m_lf, m_hf,
    HighFreqDisturb.WITH, s, taggr);
TemporalDisaggregationMethodResult r = fernandez.process();
```

```
LittermanMethod litterman = new LittermanMethod(m_lf, m_hf,
    HighFreqDisturb.WITH, s, taggr, EstimMethod.ML, 100, 0.00,
    0.99, true, 0.0);
TemporalDisaggregationMethodResult r = litterman.process();
```

3.3 Seasonal adjustment before reconciliation

Benchmarking techniques are often used in order to transform the results of a seasonal adjustment procedure so that the annual totals for seasonally adjusted and row series are equal. When applicable, the annual totals of seasonally adjusted estimates are benchmarked to the annual totals of the calendar adjusted series, leaving discrepancies between the calendar adjusted and the row series. In the event where more series linked by an accounting constraint are seasonally adjusted, the benchmarking problem becomes a reconciliation problem if the constraint is also the result of a seasonal adjustment procedure (direct approach).

A generic aggregated time series y_t can be expressed as follows:

$$y_t = f\left(x_{1,t}, \ldots, x_{k,t}, \ldots, x_{S,t}\right) \tag{3.4}$$

A special case is when the function $f(\cdot)$ is additive, which can be generalized as follows:

$$y_t = \omega_1 x_{1,t} + \ldots + \omega_k x_{k,t} + \ldots + \omega_S x_{S,t} = \sum_{k=1}^{S} \omega_k x_{k,t} \tag{3.5}$$

Where $\omega_1, \ldots, \omega_S$ are general weights.

A very simple example of this kind of aggregate is the European Union GDP, which is the sum of the GDPs of the 28 EU countries (in this case the weights ω_k are all equal to 1).

Seasonal adjustment is a well-known topic that has been studied by many authors (see, for example, Granger, 1978). Amongst others, two classes of methods are systematically used in many statistical agencies (Eurostat, 2015): the model-based approach (TRAMO/SEATS, see Gómez and Maravall, 2001; Maravall and Pérez, 2011) and the filter-based approach (X11 family, see, for instance, Findley and Hood, 2000 or Findley, 2005). For practical analysis, consult Buono (2004) or Gysels and Osborn (2001).

In this study a new test is proposed. It is based on a three-way ANOVA model, which aims at identifying whether disparate series have a common seasonal pattern. The main advantage of this test is that it gives information about which series have a common seasonal pattern before seasonally adjusting them, so that it can be considered as an a priori method. A first elaboration of this idea is in Buono and Infante (2012), while a more complete formulation is in Infante et al. (2015). The need of such kind of test is also stated in Cristadoro and Sabbatini (2000).

3.3.1 The innovative test

The classical test for moving seasonality (Higginson, 1975) is based on a two-way ANOVA model, where the two factors are the time frequency (usually months or quarters) and the years, respectively. A Bartlett-type test for moving seasonality has been proposed by Surtradhar and Dagum (1998). A test based on a three-way ANOVA model (see Cohen, 2007) is presented in this study, in order to test the presence of a moving seasonality between different series, and not between the years of the same series, as established by the classic moving seasonality test. The three factors are the time frequency, the years and the series.

The tested variable in the classical test for moving seasonality is the final estimation of the unmodified Seasonal-Irregular differences absolute value if the decomposition model is an additive one, or the Seasonal-Irregular ratio minus one absolute value, if the decomposition model is a multiplicative one. The series of the Seasonal-Irregular ratios, using the tool X-13 ARIMA, is presented in table D8, see Ladiray and Quenneville (2001) for a detailed explanation of the X-13 tables.

As the test needs to be performed a priori (e.g. before running a seasonal adjustment procedure), it is impossible to use the Seasonal-Irregular differences (or ratios) as used in the test for moving seasonality. Thus, for creating the trend series T_{kt}^{HP}, a Hodrick-Prescott filter is applied to each series x_{kt}. Such filter is widely used, especially for macroeconomic series, and it seems to be the most appropriate trend estimation when dealing with these kinds of series (Harvey and Trimbur, 2008; Hodrick and Prescott, 1997). Other trend estimation methods may be applied for different types of series.

Thus, it is possible to calculate the variable obtained by subtracting the trend series from the original one:

$$SI_{ijk} = x_{ijk} - T_{ijk}^{HP} \tag{3.6}$$

The notation SI is kept in order to remark the fact that it is a de-trended series. As such, the tested variable is a three-dimensional array (cube), where in the rows there is the i-th time frequency, in the columns there is the j-th year, and in the depth there is the j-th series. As the series involved in the test can be added up before or after the seasonal adjustment procedure, it is evident that they must have the same scale. The test is performed only on the part of the time series that covers all the observations of entire years.

The model is specified as follows:

$$SI_{ijk} = a_i + b_j + c_k + e_{ijk} \tag{3.7}$$

This equation implies that the value SI_{ijk} represents the sum of:

- A term a_i, $i = 1, \ldots, M$, representing the numerical contribution due to the effect of the i-th time frequency (usually $M = 12$, for monthly series, or $M = 4$, for quarterly series).

- A term b_j, $j = 1, \ldots, N$, representing the numerical contribution due to the effect of the j-th year.

- A term c_k, $k = 1, \ldots, S$, representing the numerical contribution due to the effect of the k-th series of the aggregate.

- A residual component term e_{ijk}, which is assumed to be normally distributed with zero mean, constant variance and zero covariance. It represents the effect, on the values of the SI, of the whole set of factors not explicitly taken into account in the model.

The test is based on the decomposition of the variance of the observations:

$$S^2 = S_M^2 + S_N^2 + S_S^2 + S_R^2 \tag{3.8}$$

Denoting:

- The general mean: $\bar{x} = \dfrac{1}{MNS} \sum_{i=1}^{M} \sum_{j=1}^{N} \sum_{k=1}^{S} |SI_{ijk}|$.

- The M time frequency means: $\bar{x}_{i\bullet\bullet} = \dfrac{1}{NS} \sum_{j=1}^{N} \sum_{k=1}^{S} |SI_{ijk}|$.

- The N yearly means: $\bar{x}_{\bullet j\bullet} = \dfrac{1}{MS} \sum_{i=1}^{M} \sum_{k=1}^{S} |SI_{ijk}|$.

- The S series means: $\bar{x}_{\bullet\bullet k} = \dfrac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} |SI_{ijk}|$.

Then it is possible to compute the following quantities:

- $S_M^2 = \dfrac{NS}{M-1}\sum_{i=1}^{M}(\bar{x}_{i\bullet\bullet} - \bar{x})^2$ is the between time frequencies variance. It is the effect that measures the magnitude of the seasonality.
- $S_N^2 = \dfrac{MS}{N-1}\sum_{j=1}^{N}(\bar{x}_{\bullet j\bullet} - \bar{x})^2$ is the between years variance. It is the effect that measures the movement of the seasonality in the same series.
- $S_S^2 = \dfrac{MN}{S-1}\sum_{k=1}^{S}(\bar{x}_{\bullet\bullet k} - \bar{x})^2$ is the between series variance. It is the effect that measures the movement of the seasonality between different series.
- $S_R^2 = \dfrac{\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{S}(|SI_{ijk}| - \bar{x}_{i\bullet\bullet} - \bar{x}_{\bullet j\bullet} - \bar{x}_{\bullet\bullet k} + 2\bar{x})^2}{(MNS-1)-(M-1)-(N-1)-(S-1)}$ is the residual variance.

Hence, the null hypothesis is the following:

$$H_0: \; c_1 = c_2 = \ldots = c_S \tag{3.9}$$

When H_0 is not rejected, it implies that there is no change in the seasonality over the series, i.e. it is not possible to exclude that the series have a common similar seasonal pattern.

If the null hypothesis is true, the relative test statistics is required to follow a Fisher-Snedecor distribution with $(S-1)$ and $(MNS-1)-(M-1)-(N-1)-(S-1)$ degrees of freedom and can be written as:

$$F = \frac{S_S^2}{S_R^2} \tag{3.10}$$

In the case that the null hypothesis is rejected, an option could be to run the test on sub-groups of the series, in order to discover which ones present similar seasonal movements. The use of a cluster analysis could also be explored in order to group these series which present common seasonal patterns. In any case, the number of series tested at the same time should not be too high.

The test power analysis, together with a simulation study and a practical application, is presented in Infante et al. (2015).

3.3.2 The use and possible improvements

One possible practical application of the test is to choose which approach to use for seasonal adjustment.

To obtain seasonally adjusted figures, at least two different approaches can be applied (for more details see Astolfi et al., 2001a,b):

- Direct Approach: the seasonally adjusted data are computed directly by seasonally adjusting the aggregate y_t.
- Indirect Approach: the seasonally adjusted data are computed indirectly by seasonally adjusting data per each series x_{kt}. The seasonally adjusted y_t is then given by the sum of the seasonally adjusted components.

A third option could be the mixed approach. If it is possible to define a criterion in order to separate the series in groups, creating sub-aggregates (e.g. these series have common seasonal pattern), then it is possible to compute the seasonally adjusted figures by summing the seasonally adjusted data of these sub-aggregates.

The direct and the indirect approaches have been discussed for many years, and there is no consensus on which is the best approach (see, for instance, Maravall, 2006 or Hood and Findley, 2001).

To date, many authors presented a posteriori analysis on the results of the different approaches (Bušs, 2009; Geweke, 1979; Hindrayanto, 2004; Otranto and Triacca, 2000). For an overview of seasonality tests, refer to Busetti and Harvey (2003) and Rau (2006). As seasonal adjustment deals with unobserved components, the evaluation criteria of an a posteriori analysis depends on many factors (e.g. the method used) and could be a bit weak.

The main drawback to be considered as regards to the direct approach is that there is no accounting consistency between the aggregate and individual series. Another drawback of the direct approach is the directional inconsistency, as for some periods it could be that the components move in one direction while the aggregate moves in the opposite one. A controversial point with the direct approach is the so called cancel-out effect. If there are two series with opposite patterns of seasonality, then the aggregated series will possibly show no seasonality. For example, the aggregated series can show no seasonality even if all the individual series have seasonality. According to Maravall (2006), this is not a drawback.

On the other hand, the indirect approach also has some drawbacks. First of all, the presence of residual seasonality should always be carefully checked in all of the indirectly seasonally adjusted aggregates. In that case, applying an indirect approach means working with a larger number of series, and therefore the calculation burden could be quite big.

The numerical results obtained by performing the different approaches are usually close in terms of medium and long term evolution, but they can still diverge in terms of signs of the growth rates in the short term period. They are likely to coincide if the aggregate is an algebraic sum, the decomposition model is additive, there are no outliers and the filters used is the same for all the series. These conditions are rarely met in a real data set.

According to the ESS guidelines on seasonal adjustment (Eurostat, 2015), the indirect adjustment is preferred if the series x_{kt} do not show similar seasonal patterns. Otherwise, the direct approach is preferred if the series show common seasonal patterns and approximately the same timing in their peaks and troughs. In this case, the aggregation will produce a smoother series with no loss of information on the seasonal patterns. The direct approach is preferred for transparency and accuracy, while the indirect approach is preferred for consistency.

Rejecting the null hypothesis means that the direct approach should be avoided, and the indirect one should be taken in consideration. Once the sub-groups of the series with common similar seasonal pattern are determined, a mixed approach can be used.

As mentioned, the decision between direct and indirect (or mixed) approach is relevant in order to understand whether the series could be reconciled or only benchmarked. By using any kind of direct or mixed approach, the reconciliation problem arises. It is clear that the user may decide not to reconcile the system, as suggested by some authors (Maravall, 2006), but it is sometimes the recommended policy, especially in certain domains, like national accounts or unemployment (Eurostat, 2015).

In this case, the test presented in this section could be used in order to determine whether the series have to be seasonally adjusted directly (and thus reconciled) or indirectly (and hence there is no reconciliation problem). By applying a mixed approach, the series have to still be reconciled in each of the subgroups identified.

As it stands, the test could still be improved by carrying out some more in-depth analysis, which would ideally include:

- Seasonal co-movements test (Centoni and Cubadda, 2011): the seasonal co-movements test could be used in order to assess that the test presented here is well-detecting the seasonal movements of the different time series.
- The use of a different filter for trend estimation. In particular the selection of the parameter λ could follow different methodologies (see for example Maravall and Del Rio, 2001).
- The testing of the assumptions made on the residual term of model

3.7.

- Outliers: a detailed study on how the presence of outliers impacts on the test performance. In seasonal adjustment, usually three different kinds of outliers are considered: the Additive Outlier (AO), the Transitory Change (TC) and the Level Shift (LS). It may be interesting to see how the different outliers (and their combinations) would impact the results of the test.

3.4 Conclusions

Two-step reconciliation practices are getting more and more applied by different users, especially in the world of official statistics.

As several well-established methodologies for temporal disaggregation are considered to be good practices (minimisation of a loss function, regression-based models, etc.), a two-step reconciliation method should allow the user to choose amongst them in the first step.

The methodology presented in this chapter not only permits the application of different methods in both the first and the second steps of the procedure, but also allows the reconciliation of more systems of time series in the case they are nested. This cascade approach is justified from a practical expectation, which is that the lower series are normally of a lower quality.

Moreover, the tool JEcotrim has been modified in order to allow the user to chose between different methods in the first step. In particular, three new plug-ins have been developed, which perform the Chow-Lin, Fernández and Litterman methods respectively.

An important remark on direct and indirect approach for seasonal adjustment has to be noted. Without delving into the philosophical debate on which of the two approaches is the best in absolute terms, it is important to note that the results of a seasonal adjustment procedure will in any case depend on the quality of the time series to be adjusted. When trying to adjust component series which are of too low quality, it is obvious that summing them up would not lead to satisfactory results for the aggregates series since there could be a cancel-out effect (seasonality is present in the component series but not in the aggregate, or vice-versa), or simply the seasonality in the component series is not regular. Hence, a direct approach is to be followed in these cases, creating a reconciliation problem.

From this point of view, the test for common seasonal patterns presented here is a useful tool in order to decide at which level the seasonal adjustment is to be performed, helping the practitioner to get a reconciliation problem where the discrepancies to be adjusted are smaller.

Chapter 4

Validation

The results of a reconciliation procedure should be assessed according to given criteria. This is basically done by taking in consideration two main aspects: the distance and the differences in the dynamics between the preliminary series and the reconciled ones.

Taking into examination these two aspects, several summary indices are presented in this chapter, as well as an innovative technique for assessing the impact of possible outliers in the system of time series.

The validation of the methodology proposed in Chapter 3 is an important part of the research performed. For this reason, a simulation study is presented in this chapter.

4.1 Validation and assessment criteria

Many assessment criteria could be followed in order to evaluate the performances of reconciliation techniques. According to Stuckey et al. (2004) and Di Fonzo and Marini (2011b), three main principles should be followed: 1. The levels of the reconciled series should be as close as possible to the levels of the preliminary series; 2. The movements of the reconciled series should be as close as possible to the movements of the preliminary series; 3. Highly volatile series are altered more than less volatile series.

While it is evident that the distance between the preliminary (related) series and the reconciled series should be small, it is important to use adequate measures of such distance. Amongst other possibilities, this study follows the approaches suggested by Ladiray and Mazzi (2003) and Di Fonzo and Marini (2011b).

As regarding to the levels, the absolute percentage differences (APD) and the squared percentage differences (SPD) could be easily computed. Given:

$$APD_{j,t} = \left| \frac{r_{H,j,t}}{p_{H,j,t}} - 1 \right|$$

For each of the m time series, the mean and maximum absolute percentage difference, as well as the mean squared percentage difference are, respectively:

$$meanAPD_j = \frac{1}{n}\sum_{t=1}^{n} APD_{j,t} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{r_{H,j,t}}{p_{H,j,t}} - 1\right| \tag{4.1}$$

$$maxAPD_j = \max_t APD_{j,t} = \max_t \left|\frac{r_{H,j,t}}{p_{H,j,t}} - 1\right| \tag{4.2}$$

$$meanSPD_j = \sqrt{\frac{1}{n}\sum_{t=1}^{n} APD_{j,t}^2} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\frac{r_{H,j,t}}{p_{H,j,t}} - 1\right)^2} \tag{4.3}$$

The mean absolute percentage difference and the mean squared percentage difference could also be computed for the whole system of series:

$$meanAPD = \frac{1}{mn}\sum_{j=1}^{m}\sum_{t=1}^{n}\left|\frac{r_{H,j,t}}{p_{H,j,t}} - 1\right| \tag{4.4}$$

$$meanSPD = \sqrt{\frac{1}{mn}\sum_{j=1}^{m}\sum_{t=1}^{n}\left(\frac{r_{H,j,t}}{p_{H,j,t}} - 1\right)^2} \tag{4.5}$$

If more systems are reconciled together (for example after applying a cascade approach) it would be very useful to compute these indices for the whole set of systems, given that the unit measure is the same.

As regards the movements, the mean and maximum absolute percentage difference of the growth rates (APDG), as well as the mean squared absolute percentage difference of the growth rates (SPDG) and the concordance of growth rates (C1) could be calculated. Hence, starting from the absolute percentage difference of the growth rates:

$$APDG_{j,t} = \left| \frac{r_{H,j,t}}{r_{H,j,t-1}} - \frac{p_{H,j,t}}{p_{H,j,t-1}} \right|$$

The indices are computed as follows:

$$meanAPDG_j = \frac{1}{n-1} \sum_{t=2}^{n} APDG_{j,t} = \frac{1}{n-1} \sum_{t=2}^{n} \left| \frac{r_{H,j,t}}{r_{H,j,t-1}} - \frac{p_{H,j,t}}{p_{H,j,t-1}} \right|$$

$$maxAPDG_j = \max_t APDG_{j,t} = \max_t \left| \frac{r_{H,j,t}}{r_{H,j,t-1}} - \frac{p_{H,j,t}}{p_{H,j,t-1}} \right|$$

$$meanSPDG_j = \sqrt{\frac{1}{n-1} \sum_{t=2}^{n} APDG_{j,t}^2} = \sqrt{\frac{1}{n-1} \sum_{t=2}^{n} \left(\frac{r_{H,j,t}}{r_{H,j,t-1}} - \frac{p_{H,j,t}}{p_{H,j,t-1}} \right)^2}$$

$$C1_j = \frac{1}{n-1} \frac{\sum_{t=2}^{n} \left| sign\left(\frac{r_{H,j,t}}{r_{H,j,t-1}} - 1 \right) + sign\left(\frac{p_{H,j,t}}{p_{H,j,t-1}} - 1 \right) \right|}{2}$$

And the relative indices for the whole system of series are given by:

$$meanAPDG = \frac{1}{m(n-1)} \sum_{j=1}^{m} \sum_{t=2}^{n} \left| \frac{r_{H,j,t}}{r_{H,j,t-1}} - \frac{p_{H,j,t}}{p_{H,j,t-1}} \right|$$

$$meanSPDG = \sqrt{\frac{1}{m(n-1)} \sum_{j=1}^{m} \sum_{t=2}^{n} \left(\frac{r_{H,j,t}}{r_{H,j,t-1}} - \frac{p_{H,j,t}}{p_{H,j,t-1}} \right)^2}$$

$$C1 = \frac{1}{m(n-1)} \frac{\sum_{j=1}^{m} \sum_{t=2}^{n} \left| sign\left(\frac{r_{H,j,t}}{r_{H,j,t-1}} - 1 \right) + sign\left(\frac{p_{H,j,t}}{p_{H,j,t-1}} - 1 \right) \right|}{2}$$

Regarding the movements, a very important point is that there is common seasonality between each preliminary indicator and the correspondent reconciled series. In other words, there should be no residual seasonality in the reconciled seasonally adjusted series (see Evans, 2004) and the seasonality should remain the same when reconciliation is performed on non-seasonally adjusted series.

Finally, taking into consideration the larger components of the systems, it could be interesting to compute a weighted version of the absolute and squared percentage differences of the series and of the growth rates, where the weights $\omega_{j,t}$ are given by the shares of the values of the series in the grand totals (contemporaneous constraints). In this case the indices become:

$$meanWAPD = \frac{1}{n}\sum_{j=1}^{m}\sum_{t=1}^{n}\omega_{j,t}\left|\frac{r_{H,j,t}}{p_{H,j,t}} - 1\right| \tag{4.6}$$

$$meanWSPD = \sqrt{\frac{1}{n}\sum_{j=1}^{m}\sum_{t=1}^{n}\omega_{j,t}\left(\frac{r_{H,j,t}}{p_{H,j,t}} - 1\right)^2} \tag{4.7}$$

$$meanWAPDG = \frac{1}{n-1}\sum_{j=1}^{m}\sum_{t=2}^{n}\omega_{j,t}\left|\frac{r_{H,j,t}}{r_{H,j,t-1}} - \frac{p_{H,j,t}}{p_{H,j,t-1}}\right| \tag{4.8}$$

$$meanWSPDG = \sqrt{\frac{1}{n-1}\sum_{j=1}^{m}\sum_{t=2}^{n}\omega_{j,t}\left(\frac{r_{H,j,t}}{r_{H,j,t-1}} - \frac{p_{H,j,t}}{p_{H,j,t-1}}\right)^2} \tag{4.9}$$

These kind of indices could only be correctly computed if the contemporaneous constraints are given by a simple summation of the series.

4.2 Outlier identification as validation criteria

In general terms, the user would expect that the outliers identified in the preliminary series are unchanged after applying a reconciliation technique (especially in the last observations), otherwise important information would be lost.

In this section, a methodology for identifying outliers at the end of the series is presented. Such methodology has been firstly discussed by Infante and Buono (2013) and Buono et al. (2016), which developed it in order to identify commodity risk in price statistics. Similar ideas have been discussed by Maravall and Caporello (2003); Revilla and Rey del Castillo (1999, 2000).

When analysing market price data (Eurostat, 2008a), it is important to identify the factors causing their fluctuations. Such fluctuations are often due to the market price risk, which can be broadly defined as the threat of losses due to changes in market parameters. Market risk can be affected by the commodity risk, defined as the threat that a change in the price of a production input will adversely impact a producer who uses that input. Commodity production inputs are usually raw materials (cotton, corn, wheat, copper, etc.). Factors that can affect commodity prices include political and regulatory changes, seasonal variations, technology and market conditions. These factors have an impact on the volatility of the data, and may affect the predictability, generating uncertainty. For more details regarding commodity and market risk see, one ca refer to Dusak (1973) or Giot and Laurent (2003).

With a special focus on the end-series observations, a new technique is presented in the following section, in order to face such uncertainty by detecting the degree of the possibility of having a commodity risk occurring in the series. The underlining idea is that when the observed data differs considerably from the expected forecasted trend, the commodity risk may be present. In the same way, it is also possible to detect potential outliers within end series observations.

While the original application was done on price statistics, nothing prevents the use of this technique as outlier detection method in any other domain.

4.2.1 Proposed methodology

Given a time series x_t, the procedure follows three main steps: identification of the seasonal ARIMA model; estimating forecast intervals; and detecting the volatility degree. The idea is to fit a seasonal ARIMA model to the series, where the last r observations are removed within the sample, estimating the forecast confidence intervals on those values not considered in the first step, and checking whether the observed values fall inside or outside the confidence bands.

1. **Identification of the model**:

   The first step of the procedure is to model the series without the last r observations. Here a seasonal ARIMA$(p, d, q)(P, D, Q)_s$ model is used:

$$\phi_p(B)\,\Phi_P(B^s)\,(1-B)^d\,(1-B^s)^D\,x_{t^\star}=\theta_q(B)\,\Theta_P(B^s)\,\varepsilon_{t^\star} \tag{4.10}$$

Where $t^\star = t - r$ and B is the lag operator.

All the validation checks should be performed as per usual practice. If calendar effects and outliers are present in the series $x_{t^\star}$, a RegARIMA model could be used.

2. **Estimating forecast intervals**:

   In the second step, the seasonal ARIMA forecast intervals are computed for r observations which were not taken into account during the first step:

   $$\hat{x}_{t^\star}(h) \pm z_{\alpha/2}\sqrt{VAR\left[e_{t^\star}(h)\right]} \tag{4.11}$$

   where $\hat{x}_{t^\star}$ is the punctual forecast at time $t^\star + h$, $z_{\alpha/2}$ is the percentile of a standardized normal distribution and $e_{t^\star}$ is the forecast error at time $t^\star + h$. Commonly, the interval is taken at 95% level. A detailed analysis of such kind of forecast intervals is in Chatfield (2001).

   The intervals are computed for each $h = 1, \ldots, r$, in order to obtain r intervals, for all observations not considered in the model during the first step.

3. **Detecting the volatility degree**:

   In the third step the observed values at time $t^\star + h$, with $h = 1, \ldots, r$, are compared with the forecast intervals computed during the second step. If the observed value at time $t^\star + h$ is not inside the forecast interval at time $t^\star + h$, then the commodity risk (outlier) is detected due to the volatility of the series.

4.2.2 Some considerations

In a general framework, the number r should not be too high, otherwise the forecasted observations would be too far from the ones used for modelling the series. On the other hand, it should neither be too low, given the need to analyse as much information possible. From this perspective, in the original paper, the authors suggest to consider $r = 3$ in the case of monthly series.

When applied to the preliminary series and to the correspondent results of a reconciliation technique, it appears natural to consider r equal to the number of periods which are extrapolated.

When a given observed value falls outside the interval, it may be classified as an outlier. The idea is that, as the dynamics and the outliers of the preliminary series should be the same, this methodology is applied to both the preliminary and the reconciled estimates, and the results are checked whether they are the same. If one or more outliers are observed in the preliminary series, but not in the reconciled series (or vice versa), then the reconciliation performed is not considered as being satisfactory.

Information about the type of the identified outlier is given by looking at all the r intervals together.

Three different cases, regarding additive outliers, transitory changes and level shifts are considered in this study:

- If in a given period $t^\star + h^\star$, the observed value is outside the interval, while all the other $r - 1$ observed values are inside the respective intervals, then the observed value at time $t^\star + h^\star$ is considered as an additive outlier.
- If in the given periods $t^\star + h^\star$ and $t^\star + h^\star + 1$, the observed values are outside the respective intervals on the same side, but the observed value at time $t^\star + h^\star + 2$ is inside its respective the interval, then the observed value at time $t^\star + h^\star$ is classified as a transitory change.
- If, starting from a given period $t^\star + h^\star$, all the observed values are outside the respective intervals on the same side, then the observed value at time $t^\star + h^\star$ is classified as a level shift.

Further improvements to the methodology may include the use of a multivariate model for a group of series and changing the way of setting up the forecast intervals, which could be done by considering just one observation ahead, and thus changing the number of observations in the model for each h.

4.3 Simulations

A simulation can be broadly defined as an imitation of a system (Robinson, 2014). A simulation predicts the performance of an operations system under a specific set of inputs. As such, simulation is an experimental approach to modelling, that is, a "what-if" analysis tool, and should be seen as a form of decision support system with the aim of finding an optimum scenario.

Simulation models are able to explicitly represent the variability, interconnectedness and complexity of a system, such as a reconciliation problem. As a result, with a simulation it is possible to predict a system performance, to compare alternative system designs and to determine the effect of alternative policies on the system performance.

In this section, simulations are carried out in order to explore the behaviours of two-step reconciliation practices, and especially to evaluate their performances when changing the temporal disaggregation technique applied in the first step.

4.3.1 Scheme

Five different schemes have been created for simulating different cases of reconciliation. In particular, amongst all possible dimensions which could be taken into account, two aspects of the problem are here considered: the size of the series, and the size of the discrepancies between the preliminary series and the constraints.

The procedure for simulating the reconciliation problem follows four steps:

1. Creating sets of four monthly series which follow a seasonal ARIMA $(1,1,0)(0,1,0)_{12}$ model, using random significant coefficients and relative sizes changing according to the schemes. The series span is 15 years and 9 months, in order to allow analysis for the extrapolation.

2. Aggregating the time series created in the first step, to obtain the temporal constraints.

3. Calculating the total series by summing up the series created in the first step, in order to obtain the contemporaneous constraints.

4. Creating preliminary time series by artificially increasing the series created in the first step. The increasing factor is chosen randomly in a relatively small interval for all observations of each time series. The size of the factor changes according to the schemes.

The procedure, done using R (R Core Team, 2017), has been repeated in order to create 100 systems, of four time series each, to be reconciled. Once created, the systems are than reconciled in JEcotrim.

The simulation schemes are presented in Table 4.1. The discrepancies inserted have to always be considered as averages, as they are randomly fluctuating between plus and minus one percentage extra point.

The R code used for generating the series for scheme 1A is shown in the box below. A very similar code is used for the other schemes.

```
## Calling the library
library(forecast)
library(timeSeries)

## Setting the seed for replicating the random numbers
set.seed(101)

## Initiating the matrix with the quarterly series
syst <- 100
nseries <- 4*syst
Xquart <- matrix(nrow=189,ncol=nseries)
Xprel <- matrix(nrow=189,ncol=nseries)

## Creating the vector with the ARIMA coefficients, in intervals [a,b] or [c,d
    ]
ncoeff <- 4*syst
par <- numeric(ncoeff)
a = -0.9
b = -0.5
c = 0.5
d = 0.9
for (i in 1:ncoeff){
 g <- runif(1, 0, b-a+d-c)
 if( g < (b-a) ){
    par[i] <- a + g
 }else{
    par[i] <- c + g - (b-a)
 }
}

## Initiating the size of the series of the system for scheme 1A
siz <- numeric(nseries)
for (i in 1:nseries){
siz[4*i-3] <- 5000
siz[4*i-2] <- 5000
siz[4*i-1] <- 5000
siz[4*i] <- 5000
}

## Creating the system of time series using airline models
for (i in 1:nseries){
model1 <- Arima(ts(rnorm(189)+siz[i],freq=12), order=c(1,1,0), seasonal=c
    (0,1,0), fixed=c(phi=par[i]))
model1[2] <- 400
Xquart[,i] <- simulate(model1, nsim=189)
}
Xquart = ts(Xquart,frequency=12,start=2001)

## Calculating the annual constraints
```

```
Y0 <- aggregate(Xquart, nfrequency=1)

## Calculating the contemporaneous constraints
Zquart <- matrix(nrow=189,ncol=syst)
for (i in 1:syst){
Zquart[,i] <- Xquart[,4*i-3]+Xquart[,4*i-2]+Xquart[,4*i-1]+Xquart[,4*i]
}
Zquart = ts(Zquart,frequency=12,start=2001)

## Simulating the preliminary series for scheme 1A
rand <- matrix(nrow=189,ncol=nseries)
rand[,seq(1, nseries, 4)] <- runif(189*(nseries/4), 1.09, 1.11)
rand[,seq(2, nseries, 4)] <- runif(189*(nseries/4), 1.09, 1.11)
rand[,seq(3, nseries, 4)] <- runif(189*(nseries/4), 1.09, 1.11)
rand[,seq(4, nseries, 4)] <- runif(189*(nseries/4), 1.09, 1.11)
Xprel <- Xquart * rand
```

Table 4.1: Simulation schemes

Schemes	***S1***	***S2***	***S3***	***S4***
Size				
A1	25%	25%	25%	25%
A2	25%	25%	25%	25%
B1	40%	40%	10%	10%
B2	40%	40%	10%	10%
B3	40%	40%	10%	10%
Discrepancies				
A1	9-11%	9-11%	9-11%	9-11%
A2	9-11%	9-11%	1-3%	1-3%
B1	9-11%	9-11%	9-11%	9-11%
B2	1-3%	1-3%	9-11%	9-11%
B3	9-11%	9-11%	1-3%	1-3%

4.3.2 Results

In order to read the results of the simulations, it is important to clarify that according to the way the schemes are built, by using discrepancies with a fixed average, they should generate results which are very similar in all the four methods applied. From this perspective, even very small differences in the results are important for determining which approach generates best results.

The results for the five schemes are presented in Tables 4.2, 4.3, 4.4, 4.5 and 4.6 for the whole exercise and for the extrapolations only. Very extreme results have been eliminated from the analysis because they were simply generated by values very close to zero, which were exploiting the ratios.

Although it is clear that in all the cases the four different combinations of methods for the first and the second steps are performing well, hence are all reliable approaches, some differences are still observable.

In terms of growth rates, the Quenneville-Rancourt (QR) approach is always performing worse than the Di Fonzo-Marini (FM) approach. This is particularly true for schemes 2A, 2B and 2C, where the sizes of the series are different and the QR approach has more difficulties to keep the size of the growth rates for the smaller series. This is evident since the method is the worst performing in terms of signs of the growth rates (C1) for scheme 2B, where the discrepancies assigned to the small series are big, while the discrepancies in the big series are small.

Regarding the differences between the Chow-Lin (CL) and the modified Denton (PFD) methods, it is clear that by applying the CL method, the results of the extrapolations are always better. This is again particularly true for schemes 2A, 2B and 2C, as the sizes of the series are not the same, hence the CL method manages to get better extrapolated values (i.e. closer to fulfil the contemporaneous constraints).

Similar conclusions could be made by looking at the boxplots of the results, which are shown in Chart 4.1 for scheme 1A, in Charts 4.2 and 4.3 for scheme 1B, in Chart 4.4 for scheme 2A, in Charts 4.5 and 4.6 for scheme 2B and in Charts 4.7 and 4.8 for scheme 2C. Two different charts are needed if the sizes of the discrepancies amongst the four series of the system have been assigned differently.

The results obtained by using the Chow-Lin approach in the first step are slightly better in terms of volatility, and the lower parts of the boxes are always smaller than the higher parts (apart from the small discrepancies of scheme 1B). Additionally, the Denton method shows more extreme values than the Chow-Lin method, which is therefore more stable. No big differences are visible in the comparison between the results of the Quenneville-Rancourt and the Di Fonzo-Marini methods.

Table 4.2: Simulations results: scheme 1A

Scheme 1A	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
Complete				
MeanAPD	9.08%	9.08%	9.09%	9.09%
MeanSPD	9.09%	9.09%	9.09%	9.09%
MeanAPDG	0.33%	0.29%	0.32%	0.28%
MeanSPDG	0.49%	0.46%	0.41%	0.36%
C1	92.24%	92.30%	92.33%	92.41%
Extrapolation				
MeanAPD	9.10%	9.10%	9.16%	9.13%
MeanSPD	9.11%	9.10%	9.18%	9.13%
MeanAPDG	0.37%	0.28%	0.36%	0.27%
MeanSPDG	0.45%	0.38%	0.45%	0.38%
C1	93.83%	93.83%	93.52%	93.52%

Table 4.3: Simulations results: scheme 1B

Scheme 1B	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
Complete				
MeanAPD	5.52%	5.52%	5.52%	5.52%
MeanSPD	5.53%	5.53%	5.54%	5.53%
MeanAPDG	0.34%	0.30%	0.33%	0.29%
MeanSPDG	0.51%	0.49%	0.42%	0.37%
C1	92.12%	92.24%	92.17%	92.27%
Extrapolation				
MeanAPD	5.54%	5.53%	5.60%	5.57%
MeanSPD	6.59%	6.58%	6.67%	6.61%
MeanAPDG	0.38%	0.29%	0.38%	0.28%
MeanSPDG	0.47%	0.40%	0.47%	0.40%
C1	92.90%	93.21%	93.21%	93.21%

Table 4.4: Simulations results: scheme 2A

Scheme 2A	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
Complete				
MeanAPD	9.08%	9.08%	9.09%	9.09%
MeanSPD	9.09%	9.09%	9.09%	9.09%
MeanAPDG	0.36%	0.27%	0.35%	0.26%
MeanSPDG	0.52%	0.43%	0.44%	0.32%
C1	87.13%	88.50%	87.22%	88.45%
Extrapolation				
MeanAPD	9.09%	9.08%	9.17%	9.17%
MeanSPD	9.09%	9.09%	9.18%	9.18%
MeanAPDG	0.38%	0.28%	0.38%	0.27%
MeanSPDG	0.46%	0.39%	0.46%	0.39%
C1	90.12%	90.12%	90.12%	89.81%

Table 4.5: Simulations results: scheme 2B

Scheme 2B	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
Complete				
MeanAPD	5.52%	5.52%	5.52%	5.52%
MeanSPD	5.53%	5.54%	5.54%	5.53%
MeanAPDG	0.39%	0.28%	0.38%	0.27%
MeanSPDG	0.55%	0.45%	0.48%	0.34%
C1	86.60%	87.99%	86.47%	87.96%
Extrapolation				
MeanAPD	5.52%	5.52%	5.61%	5.61%
MeanSPD	6.58%	6.58%	6.69%	6.68%
MeanAPDG	0.41%	0.30%	0.40%	0.29%
MeanSPDG	0.49%	0.41%	0.49%	0.42%
C1	89.20%	89.20%	88.27%	88.27%

Table 4.6: Simulations results: scheme 2C

Scheme 2C	*Chow-Lin*		*Denton PFD*	
	QR	***FM***	***QR***	***FM***
Complete				
MeanAPD	5.52%	5.52%	5.52%	5.52%
MeanSPD	5.53%	5.53%	5.54%	5.53%
MeanAPDG	0.36%	0.27%	0.36%	0.26%
MeanSPDG	0.53%	0.44%	0.44%	0.32%
C1	87.16%	88.34%	87.19%	88.46%
Extrapolation				
MeanAPD	5.52%	5.52%	5.61%	5.61%
MeanSPD	6.58%	6.58%	6.64%	6.65%
MeanAPDG	0.38%	0.28%	0.38%	0.28%
MeanSPDG	0.46%	0.39%	0.46%	0.39%
C1	89.51%	89.20%	89.20%	88.89%

Figure 4.1: APD: boxplots for scheme 1A

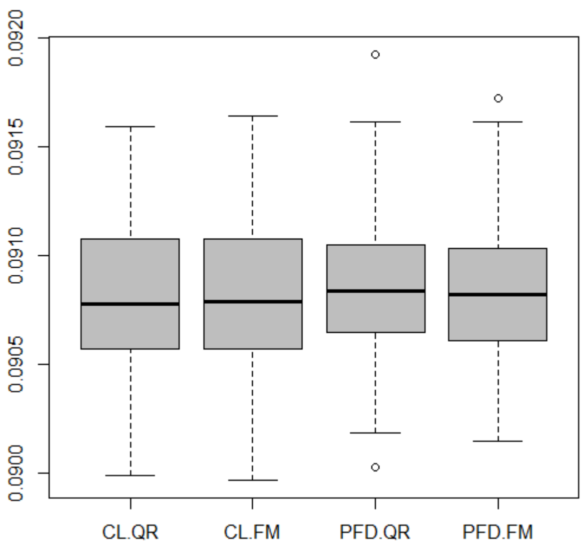


Figure 4.2: APD: boxplots for scheme 1B, small discrepancies

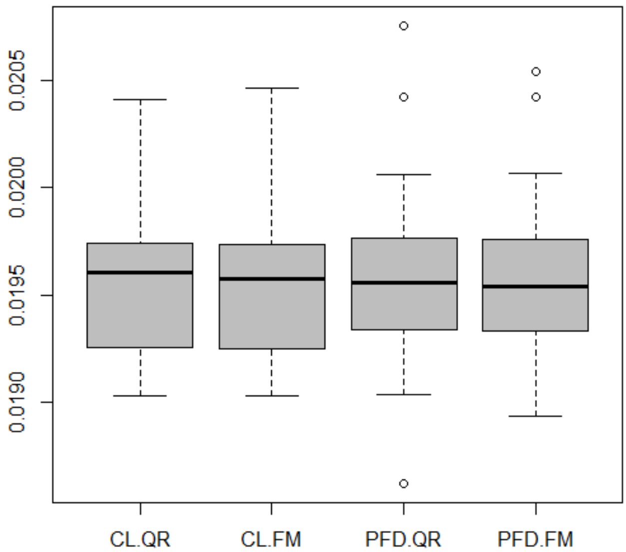


Figure 4.3: APD: boxplots for scheme 1B, big discrepancies

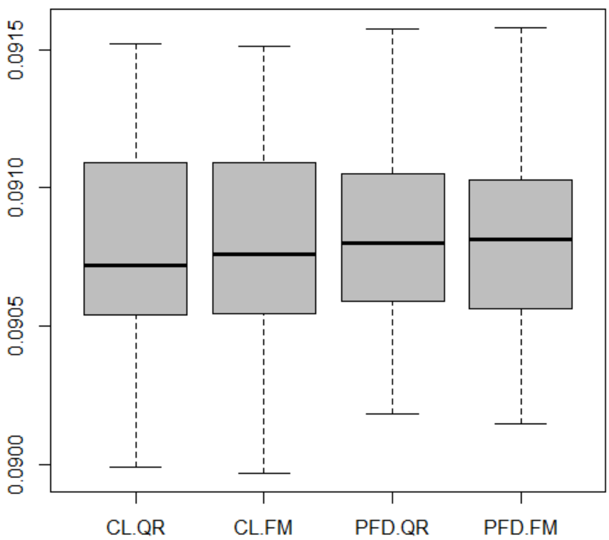


Figure 4.4: APD: boxplots for scheme 2A

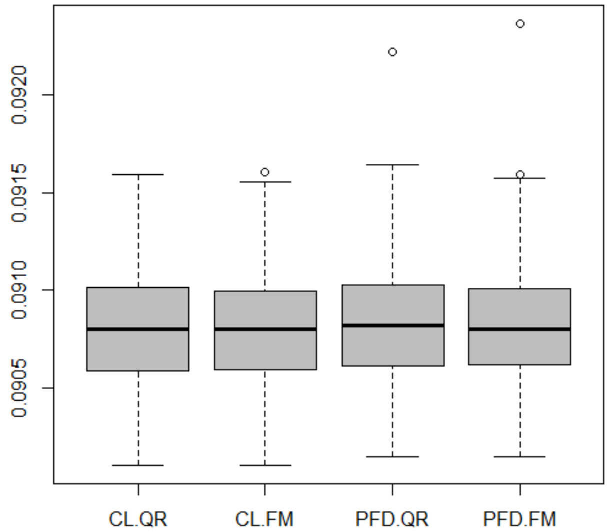


Figure 4.5: APD: boxplots for scheme 2B, small discrepancies

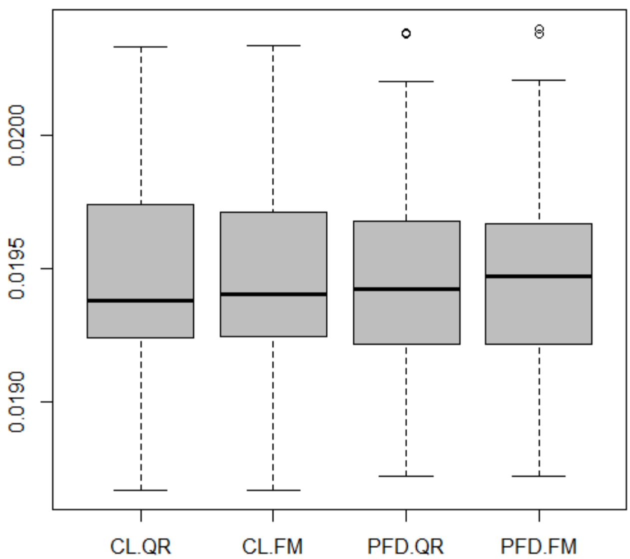


Figure 4.6: APD: boxplots for scheme 2B, big discrepancies

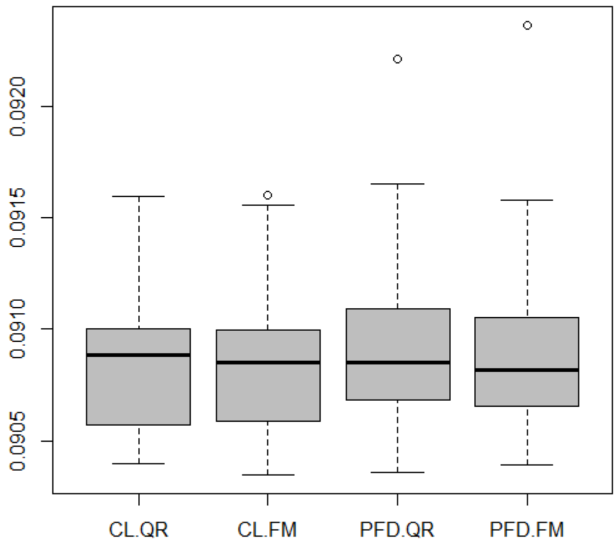


Figure 4.7: APD: boxplots for scheme 2C, small discrepancies

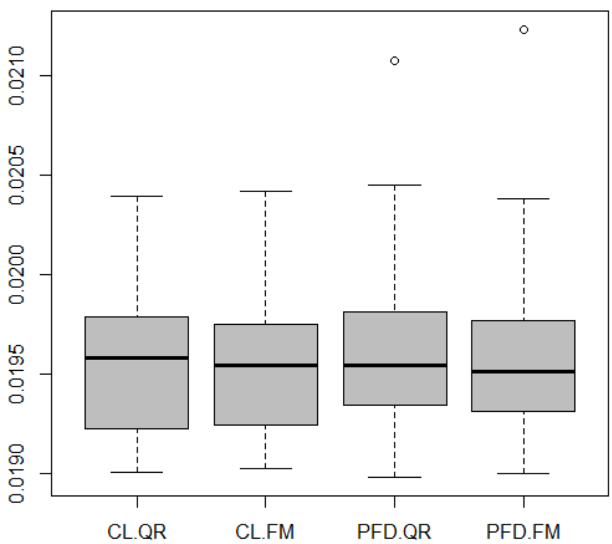
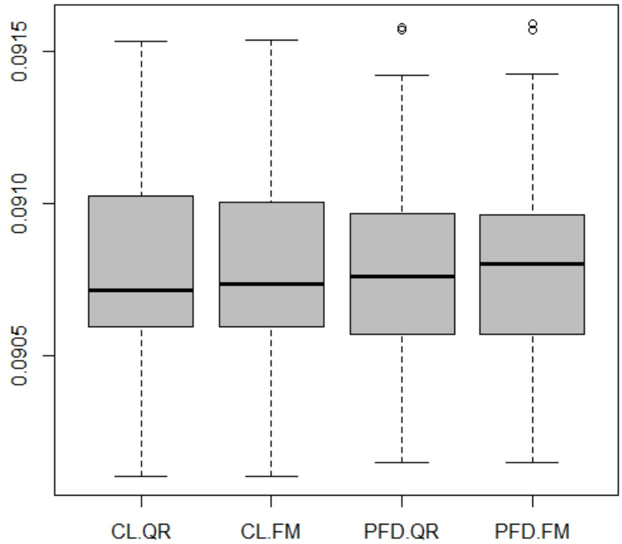


Figure 4.8: APD: boxplots for scheme 2C, big discrepancies



4.4 Conclusions

Validation and assessment criteria are of great importance when evaluating the performances of a given methodology and when comparing different methodologies. This is also true for reconciliation techniques.

In this chapter, a number of indices which measure the distance between the preliminary and the reconciled series have been presented, taking into account different aspects such as levels, movements and volatility of the series. Apart from computing these indices on single series and on a single system of time series, it is also possible to do so on a set of systems, giving a single evaluation value (for a given aspect) to the whole problem. These indices are fundamental tools for evaluating the results of the reconciliation techniques.

An innovative validation assessment has also been presented, which could be seen as a generic tool for detecting outliers in the last observations of the time series. Such tool is particularly helpful in order to evaluate the performances of the extrapolations. Outliers should not be present in the extrapolated reconciled results, unless they have already been observed in the related (preliminary) series.

The simulations performed on the different two-step reconciliation practices have shown that the possibility to choose among different methods in both the first and the second step could help the user to obtain better results. In particular, in the first step, the CL method is generating better results for the extrapolations than the PFD method, while in the second step the FM method performs slightly better than the QR method in terms of growth rates. From this point of view, it seems that the two-step reconciliation procedure which is performing better is the one using the CL method in the first step and the FM method in the second step.

However, this is not always true, and it should always depend on the situation and on the choice of the user. For example, in the case of estimating missing back data at HF level, the user might not need to extrapolate any data, and thus could prefer to use the PFD method in the first step.

In addition, when only few LF values are available, the performance of the CL model could be poor (or it could even be impossible to estimate the values), which would increase the risk of generating big revisions each time a new LF observation is available.

Finally, the user might need to focus mainly on the biggest items of the system to be reconciled, leaving the smaller variables with worse results than the bigger ones, and thus preferring the use of the QR method in the second step.

Chapter 5

Empirical applications

The innovative two-step reconciliation technique described in Chapter 3 is here applied in two different empirical case studies, and it is compared to the original two-step methods.

The first empirical application is performed on a relatively small sized system of time series, which is the directly seasonally adjusted European Union industrial production index. This index series is the aggregation of the 28 indices of the European Union's member states. This is a typical example of a reconciliation problem, where the directly adjusted benchmarked aggregated series is the contemporaneous constraint, and the annual totals of the original member states series are the temporal constraints.

The second application is medium-large sized. The euro area quarterly sector accounts are calculated in such a way that the aggregation process is not a simple sum of the national components. Many inconsistencies arise after such a process, meaning that accounting relationships between the variables are not fulfilled. Moreover, in many cases, the quarterly data are not in line with the corresponding annual figures, generating a reconciliation problem. As more nested contemporaneous constraints are present, the cascade approach will be used. This is a complex reconciliation problem which is treated with a partial *ad hoc* balancing of the accounts.

5.1 Industrial production index

The industrial production index (IPI, sometimes also called industrial output index or industrial volume index) is a business cycle indicator which measures monthly changes in the price-adjusted output of industry (mining and quarrying; manufacturing; electricity, gas, steam and air conditioning supply).

It is one of the most important short-term statistics indicators. It is used to identify turning points in the economic development in the early stages, and to assess the future development of GDP. For this purpose, at the European level it is available on a monthly basis, in a detailed activity breakdown and with a rather short delay (1 month and 10 days).

The objective of the production index is to measure changes in the volume of output at close and regular intervals, normally monthly. It provides a measure of the volume trend in value added over a given reference period.

The production index is a theoretical measure that must be approximated by practical measures. Value added at basic prices can be calculated from turnover (excluding VAT and other similar deductible taxes directly linked to turnover), adding capitalised production, other operating income and the changes in stocks, subtracting purchases of goods and services and the difference between taxes on products which are linked to turnover but not deductible and any subsidies on products received.

More details about how the IPI is compiled in the European Union countries are available on the Commission Regulation 1503/2006.[2]

5.1.1 Description of the problem

Eurostat requires European Union member states to transmit calendar adjusted data for the IPI. Additionally, member states are encouraged to transmit seasonally adjusted indices. If they do not, Eurostat calculates the seasonally adjusted indices using TRAMO/SEATS method in JDemetra+ v. 2.0.0 software for the individual member states.

In 2012, the method for seasonal adjustment of European Union (EU28) IPI data was changed from a direct to an indirect approach.

Therefore, the European indices are currently calculated from national indices, taking into account the relative share of each member state in the appropriate geographical aggregate, for the gross and calendar adjusted forms. This is done at each level of the activity classification.

[2] http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32006R1503.

Table 5.1: European Union weights for aggregating the IPI

Country	***Weight***	***Country***	***Weight***	***Country***	***Weight***
Belgium	2.92	Croatia	0.39	Poland	3.63
Bulgaria	0.32	Italy	12.39	Portugal	1.19
Czech Rep.	1.94	Cyprus	0.08	Romania	1.01
Denmark	1.88	Latvia	0.11	Slovenia	0.37
Germany	27.18	Lithuania	0.16	Slovakia	0.65
Estonia	0.13	Luxembourg	0.14	Finland	1.63
Ireland	1.77	Hungary	1.06	Sweden	3.09
Greece	1.08	Malta	0.04	UK	11.88
Spain	6.66	Netherlands	3.97	EA19	74.80
France	11.61	Austria	2.72	EU28	100.00

In other words, seasonally adjusted series for the European aggregates are calculated from corresponding national series (geographically indirect seasonal adjustment).

Each member state's share in the European aggregates (European Union or euro area), in terms of weights, is given for each activity or group of activities. The weights by country for the total industry are given in Table 5.1.

The main reason for which the European statistical office decided to move towards an indirect geographical approach is to achieve consistency. However, without benchmarking the results obtained at countries level, the consistency is only achieved according to the contemporaneous constraint and not according to the temporal constraints. Moreover, by doing this Eurostat gives more importance to the data adjusted at country level then to the aggregates data.

Basically, the aggregates will be different from the direct adjusted series. From this point of view, the adjustment is not done at the level of the aggregated series, but at the level of the countries' data.

A different approach is possible. The European IPI series could be adjusted directly and benchmarked according to either the modified Denton approach, or to a regression-based technique. Subsequently, the system of time series could be reconciled by applying a two-step method, choosing between the modified Denton approach and a regression-based approach in the first step, and between the Quenneville-Rancourt method and the Di Fonzo-Marini method in the second step.

The index variables are compiled at monthly level, and comprehensive data are available from January 2000 to November 2016.[3] A first important consideration is that due to the fact that data are available till November, a big number of monthly observations (11) is to be extrapolated in the first step of the procedure. The complete set of dimensions for the IPI reconciliation problem is expressed in Table 5.2.

For the purpose of this exercise, the calendar adjusted series published by Eurostat have been seasonally adjusted, instead of using the seasonally adjusted series published by the European statistical office. This has been done using TRAMO/SEATS as available on the JDemetra+ 2.1.0. A complete automatic modelling has been chosen (without calendar effects, as the series are already calendar adjusted), with manual intervention for the cases where the results were not satisfactory.

[3]Industrial production data have been downloaded in January 2017 from the Eurostat database (http://ec.europa.eu/eurostat/data/database).

Table 5.2: Dimensions of the IPI reconciliation

Dimension	***Notation***	***Value***
Type of variables	*type*	INDEX
Number of LF observations	N	16
Temporal aggregation order	s	12
Number of HF observations	n	203
Number of extrapolated observations	$n - sN$	11
Number of variables	m	28
Number of accounting relationships	k	1

5.1.2 Results

Different two-step reconciliation methods have been applied to the IPI case, varying the methods used in both the first and the second step. In particular, four different combinations have been performed by using the modified Denton PFD (PFD) or the Chow-Lin (CL) approaches in the first step, and the Quenneville-Rancourt (QR) or the Di Fonzo-Marini (FM) methods in the second step.

The first important aspect of such reconciliation problem is that the discrepancies to be distributed are very small, which is usual when reconciling seasonally adjusted data.

Detailed results of the mean squared percentage differences by country are shown in Table 5.3. From the table it is evident that the series which have been adjusted the most are the indices of Germany, Italy, UK and France, which are the four countries with the highest weights in the system.

While it is clear that there are almost no differences between the four methods used, it is also important to note that Chow-Lin always performs equal or very slightly better than Denton. The only case where this is not true is for the Netherlands, where if the approach QR is applied in the second step, the Denton method performs slightly better than the Chow-Lin method. Moreover, for Ireland and Slovakia, the Chow-Lin method performs definitely better than the Denton method.

No real differences have been observed between the Quenneville-Rancourt and the Di Fonzo-Marini methods. This is mainly because the benchmarked series obtained after the first step are already very close to satisfying also the contemporaneous constraints, resulting in very small adjustments in the second step.

The statistics for the whole system are presented in Table 5.4.

Once again, it is clear that the results are very similar for all combinations of the methods in the two steps. However, the regression-based approach has always equal or slightly better results than the Denton approach.

The weighted statistics are much higher than the non-weighted ones. This is because, as shown in Table 5.3, the countries with higher weights have been adjusted more than the countries with lower weights.

Chart 5.1 shows the mean absolute percentage differences of the total EU28, obtained by using the the Di Fonzo-Marini approach in the second step and Chow-Lin (CL-FM) or modified Denton PFD (PFD-FM) methods in the first step. Similar results have been obtained by applying the Quenneville-Rancourt method in the second step. It is clear from the chart that the differences between the two approaches are very small, and that they are basically concentrated on the last periods of the series, where the figures are extrapolated.

In order to compare the performances of the extrapolation for the various methods, the same statistics are also reported in Table 5.5 for the extrapolated data only. Once again all the statistics are equal or better in the case of the Chow-Lin approach. However, although they are very low, on the levels, the results obtained using the Chow-Lin method are clearly better than the ones obtained by using the Denton method. Again, no real differences have been observed between the Quenneville-Rancourt and the Di Fonzo-Marini methods.

Finally, a methodology for outlier detection has been applied as described in Chapter 4, in order to check whether the outliers detected in the extrapolation span of the preliminary series are the same as those detected in the reconciled series. Table 5.6 presents the detailed results obtained for those observations which are flagged as outliers in the preliminary series but not in the reconciled estimates, or vice-versa. Such inconsistency has been noted in three observations when using the Chow-Lin method, and in four observations when using the Denton method.

Table 5.3: Mean SPD by country

Country	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
Belgium	0.07%	0.07%	0.07%	0.07%
Bulgaria	0.12%	0.12%	0.12%	0.12%
Czech Republic	0.06%	0.06%	0.10%	0.10%
Denmark	0.06%	0.07%	0.07%	0.07%
Germany	0.60%	0.59%	0.61%	0.60%
Estonia	0.06%	0.06%	0.07%	0.07%
Ireland	0.10%	0.10%	0.26%	0.26%
Greece	0.05%	0.06%	0.05%	0.06%
Spain	0.19%	0.20%	0.19%	0.20%
France	0.26%	0.27%	0.27%	0.27%
Croatia	0.06%	0.06%	0.07%	0.07%
Italy	0.37%	0.38%	0.38%	0.38%
Cyprus	0.12%	0.12%	0.12%	0.12%
Latvia	0.17%	0.17%	0.17%	0.17%
Lithuania	0.03%	0.03%	0.06%	0.06%
Luxembourg	0.07%	0.07%	0.07%	0.07%
Hungary	0.09%	0.09%	0.10%	0.10%
Malta	0.12%	0.12%	0.13%	0.13%
Netherlands	0.14%	0.13%	0.13%	0.13%
Austria	0.09%	0.09%	0.09%	0.09%
Poland	0.09%	0.09%	0.10%	0.09%
Portugal	0.12%	0.12%	0.12%	0.12%
Romania	0.11%	0.11%	0.12%	0.12%
Slovenia	0.14%	0.14%	0.14%	0.14%
Slovakia	0.20%	0.20%	0.41%	0.41%
Finland	0.22%	0.22%	0.22%	0.22%
Sweden	0.12%	0.12%	0.12%	0.12%
United Kingdom	0.26%	0.26%	0.27%	0.27%

Table 5.4: Statistics of the IPI problem

Index	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
MeanAPD	0.12%	0.12%	0.12%	0.12%
MeanSPD	0.19%	0.19%	0.21%	0.21%
MeanAPDG	0.06%	0.06%	0.06%	0.06%
MeanSPDG	0.14%	0.14%	0.14%	0.14%
C1	98.51%	98.53%	98.51%	98.46%
MeanWAPD	0.84%	0.83%	0.85%	0.84%
MeanWSPD	0.70%	0.70%	0.71%	0.71%
MeanWAPDG	0.66%	0.66%	0.67%	0.67%
MeanWSPDG	0.61%	0.61%	0.62%	0.61%

Table 5.5: Statistics of the IPI extrapolations

Index	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
MeanAPD	0.11%	0.11%	0.23%	0.23%
MeanSPD	0.17%	0.17%	0.40%	0.40%
MeanAPDG	0.07%	0.07%	0.09%	0.08%
MeanSPDG	0.14%	0.13%	0.15%	0.15%
C1	97.73%	97.73%	97.73%	97.73%
MeanWAPD	0.78%	0.78%	0.94%	0.94%
MeanWSPD	0.60%	0.61%	0.73%	0.73%
MeanWAPDG	0.85%	0.85%	0.89%	0.89%
MeanWSPDG	0.60%	0.61%	0.62%	0.63%

Figure 5.1: Mean APD for the total EU28

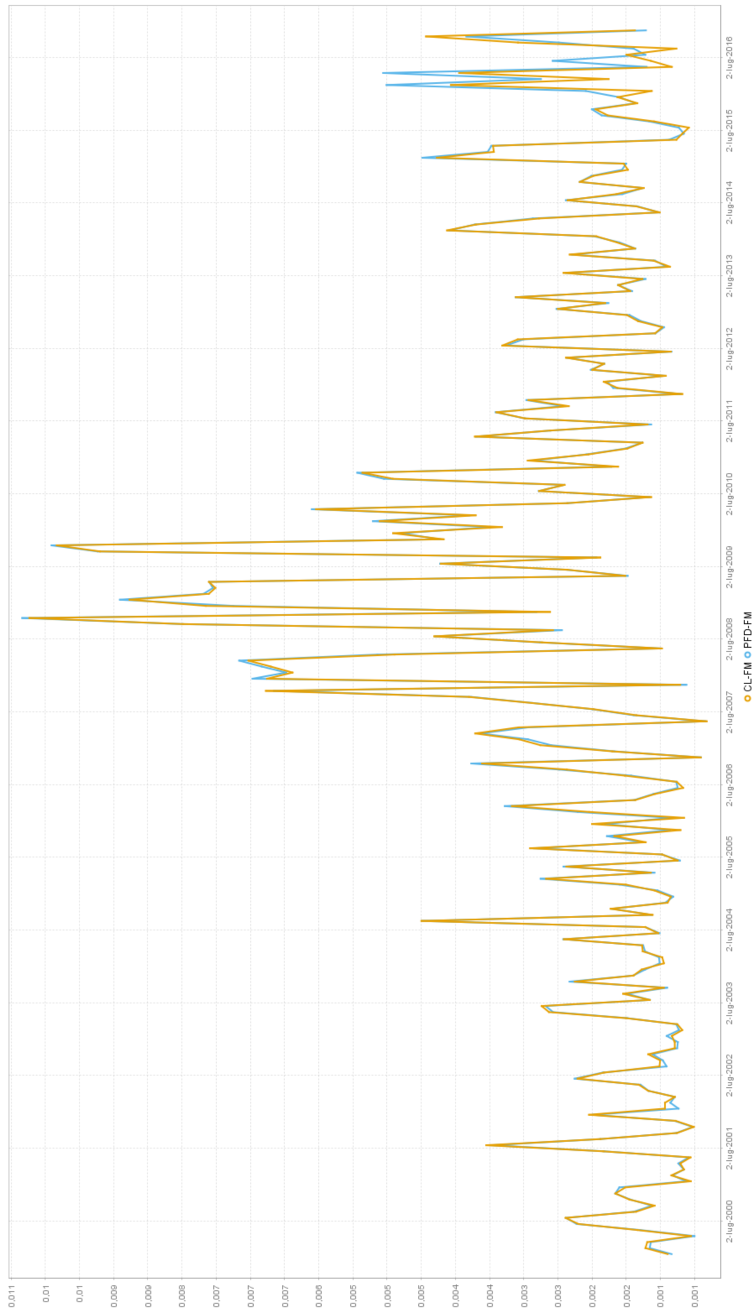


Table 5.6: Outlier consistency for the IPI extrapolations

Observation	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
IE M5	✓	✓	✗	✗
HR M11	✗	✗	✗	✗
UK M6	✗	✗	✗	✗
UK M8	✗	✗	✗	✗

5.1.3 An alternative approach

In order to obtain the reconciled figures of the European Union, a different solution can be the use of a mixed approach for the seasonal adjustment of the 28 member states, choosing the countries to be adjusted together according to the reccomendations of Eurostat (2015). This can be easily done by applying the test proposed in Chapter 3, so that groups of series with the same seasonal patterns are selected.

Among all the possible groups which could have been selected by applying the test iteratively on all possible combinations of countries, in this case the combinations which minimise the number of groups have been chosen. The results have generated a total of seven groups, four of which are composed of single series, since their seasonal patterns are different from all the other series: Italy, Netherlands, Poland and the United Kingdom. The other three groups are composed as follows (the total weights of the groups are indicated in parenthesis):

1. AGG1 (35.12%): Belgium, Bulgaria, Czech Republic, Germany,

Greece, Croatia, Latvia, Lithuania and Romania.

2. AGG2 (19.78%): Denmark, Estonia, France, Luxembourg, Hungary, Malta, Austria, Portugal, Slovenia and Slovakia.

3. AGG3 (13.23%): Ireland, Spain, Cyprus, Finland and Sweden.

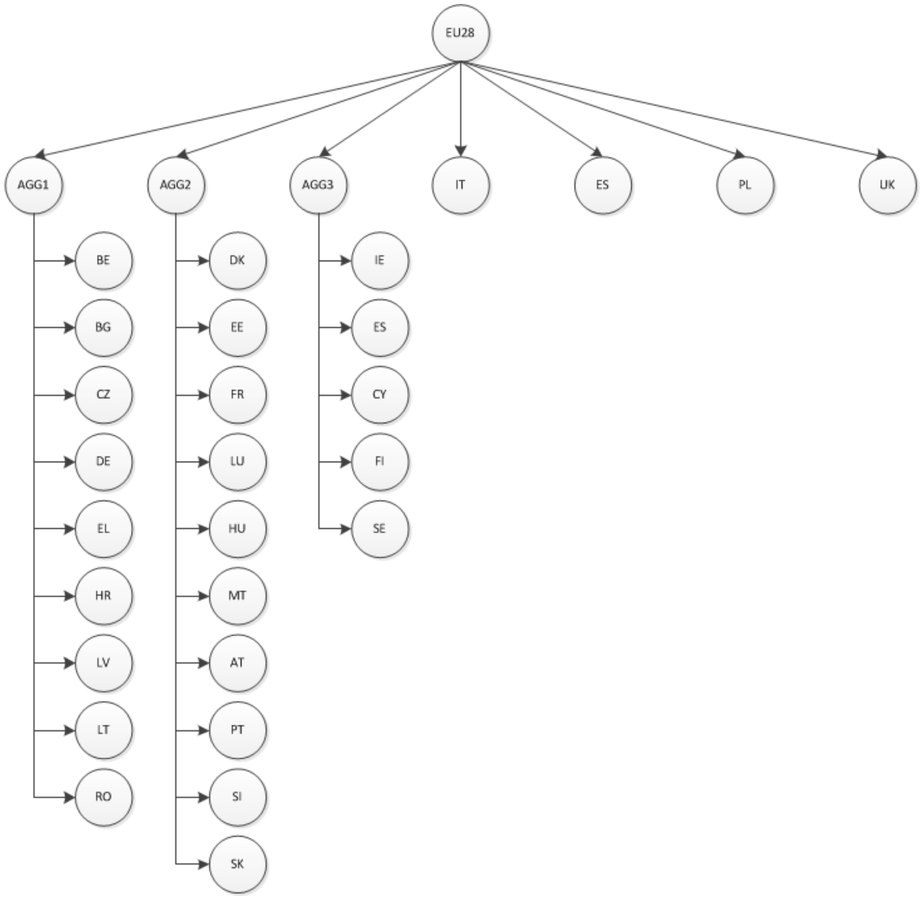
As all countries have to be adjusted, the problem becomes a hierarchical reconciliation problem, whereby the three reconciled series of each group are the results of the first layer, and thus, following the approach described in Chapter 3, the contemporaneous constraints of the series in the second layer. The overall hierarchical chart is shown in Chart 5.2.

The results obtained by country, including the ones obtained on the three groups, are listed in Table 5.7. While it is evident that in some cases the results are different from the ones obtained by using the indirect method for seasonal adjustment, it is still evident that the results obtained by using the Chow-Lin approach in the first step are always better than the ones obtained when using the Denton method.

Finally, the overall statistics of the problem are shown in Table 5.8, while the statistics of the extrapolations are shown in Table 5.9. In order to read the tables it is important to note that the weighted statistics are not directly comparable to the ones derived when the indirect approach was used, since the number of variables of the reconciliation problem is different.

Overall, the results obtained by applying the mixed approach to the seasonal adjustment process are worse than the results obtained by applying

Figure 5.2: Hierarchical chart of the IPI (mixed approach)



the indirect approach. It is also evident that the results obtained by using the Chow-Lin method in the first step are better than the ones obtained when using the Denton method. No clear distinction can be done for the methodology applied in the second step. However, the discrepancies to the preliminary series are still very small, and thus, in general, the methods applied produce good results.

Table 5.7: Mean SPD by country (mixed approach)

Country	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
Belgium	0.07%	0.07%	0.08%	0.07%
Bulgaria	0.12%	0.12%	0.12%	0.12%
Czech Republic	0.06%	0.06%	0.11%	0.11%
Denmark	0.09%	0.09%	0.09%	0.09%
Germany	0.64%	0.63%	0.65%	0.64%
Estonia	0.05%	0.05%	0.07%	0.07%
Ireland	0.15%	0.17%	0.26%	0.26%
Greece	0.05%	0.05%	0.06%	0.06%
Spain	0.49%	0.49%	0.50%	0.50%
France	0.44%	0.44%	0.44%	0.45%
Croatia	0.06%	0.06%	0.07%	0.07%
Italy	0.31%	0.31%	0.31%	0.31%
Cyprus	0.12%	0.12%	0.12%	0.12%
Latvia	0.16%	0.16%	0.17%	0.17%
Lithuania	0.03%	0.03%	0.06%	0.06%
Luxembourg	0.06%	0.06%	0.07%	0.07%
Hungary	0.09%	0.09%	0.10%	0.10%
Malta	0.12%	0.12%	0.13%	0.13%
Netherlands	0.11%	0.11%	0.12%	0.12%
Austria	0.12%	0.12%	0.12%	0.12%
Poland	0.07%	0.07%	0.08%	0.07%
Portugal	0.12%	0.12%	0.12%	0.13%
Romania	0.11%	0.11%	0.12%	0.12%
Slovenia	0.13%	0.13%	0.14%	0.14%
Slovakia	0.20%	0.20%	0.41%	0.41%
Finland	0.24%	0.24%	0.25%	0.25%
Sweden	0.24%	0.23%	0.24%	0.24%
United Kingdom	0.17%	0.17%	0.17%	0.17%
AGG1	0.50%	0.49%	0.50%	0.49%
AGG2	0.29%	0.29%	0.29%	0.29%
AGG3	0.19%	0.20%	0.19%	0.20%

Table 5.8: Statistics of the IPI problem (mixed approach)

Index	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
MeanAPD	0.17%	0.17%	0.18%	0.18%
MeanSPD	0.27%	0.27%	0.27%	0.27%
MeanAPDG	0.12%	0.13%	0.12%	0.12%
MeanSPDG	0.21%	0.21%	0.21%	0.21%
C1	96.53%	96.39%	96.53%	96.32%
MeanWAPD	3.35%	3.34%	3.37%	3.36%
MeanWSPD	1.33%	1.32%	1.34%	1.33%
MeanWAPDG	2.64%	2.63%	2.63%	2.62%
MeanWSPDG	1.08%	1.08%	1.08%	1.07%

By looking to the results on the extrapolations, it is clear that they are worse when applying the Chow-Lin approach in the first step than when using the mixed approach On the other hand, when applying the Denton approach, the results are better, and therefore the gap between the two methods is reduced.

The results are also confirmed by Chart 5.3, which shows the mean APD for the total EU28 in the case the mixed approach has been applied for the seasonal adjustment procedure.

Figure 5.3: Mean APD for the total EU28 (mixed approach)

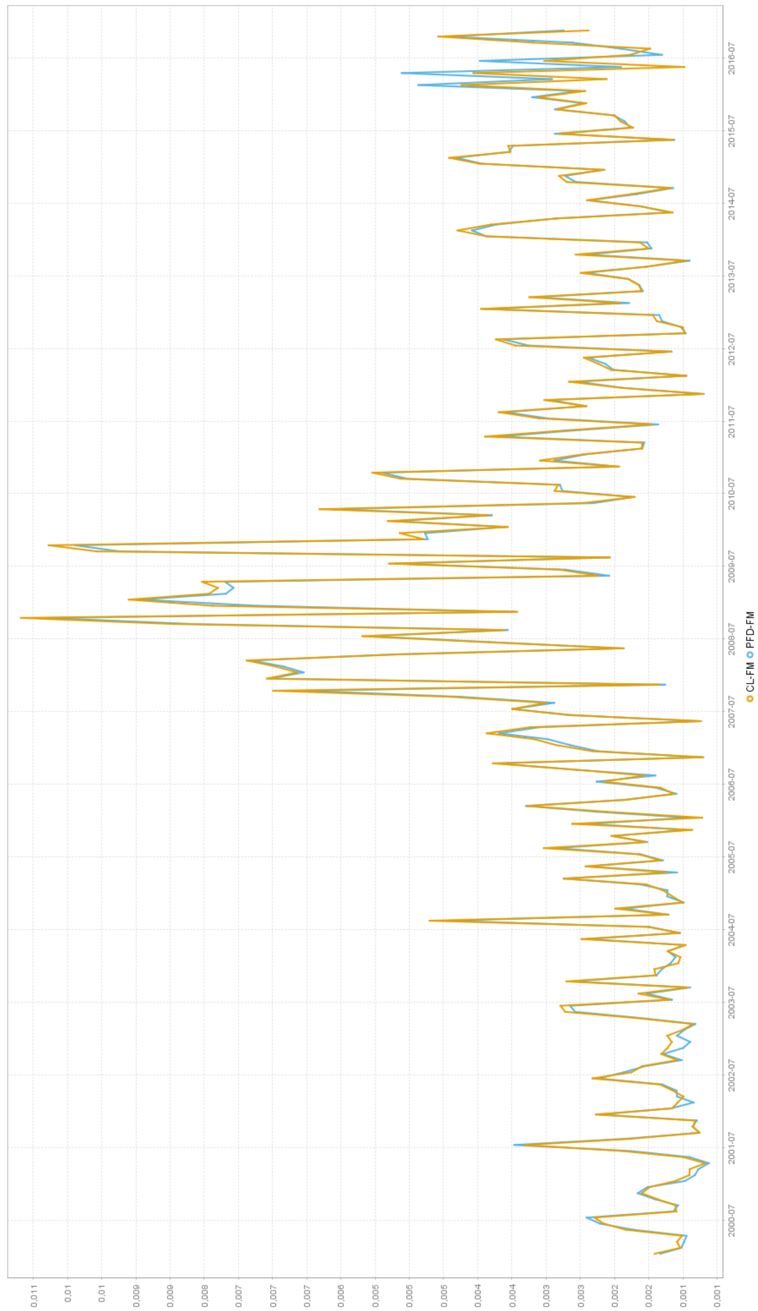


Table 5.9: Statistics of the IPI extrapolations (mixed approach)

Index	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
MeanAPD	0.14%	0.14%	0.18%	0.17%
MeanSPD	0.21%	0.21%	0.23%	0.23%
MeanAPDG	0.15%	0.15%	0.15%	0.15%
MeanSPDG	0.21%	0.21%	0.21%	0.21%
C1	97.40%	97.40%	97.40%	97.40%
MeanWAPD	2.90%	2.90%	3.12%	3.12%
MeanWSPD	1.07%	1.08%	1.12%	1.14%
MeanWAPDG	3.27%	3.26%	3.30%	3.29%
MeanWSPDG	1.07%	1.08%	1.09%	1.10%

5.2 The European sector accounts

European sector accounts group together economic subjects with similar behaviour into institutional sectors, such as:

- Non-financial corporations (S11).
- Financial corporations (S12).
- Government (S13).
- Households and non-profit institutions serving households (S1M).
- Rest of the world (S2).

Additionally, there are flows belonging to the so called unspecified total economy (S1N), while the total economy (S1) is normally given by the sum of the other domestic sectors. Grouping economic subjects in this way greatly helps to understand the functioning of the economy.

European sector accounts present a complete and consistent set of data for all the resident sectors. Apart from providing comprehensive information on the economic activities of the institutional sectors, they also provide information on the interactions between these sectors and the rest of the world.

The institutional sectors combine institutional units with broadly similar characteristics and behaviours (Eurostat, 2010): households and non-profit institutions serving households (NPISHs), non-financial corporations, financial corporations, and the government. Transactions with non-residents and the financial claims of residents on non-residents, or vice versa, are recorded in the rest of the world account.

The households sector comprises of all households, and includes household firms. These cover sole proprietorships and most partnerships that do not have an independent legal status. Hence, the households sector, in addition to consumption, also generates output and entrepreneurial income. In the European accounts, NPISHs, such as charities and trade unions, are grouped with households. Their economic weight is relatively limited. A detailed discussion about the split between the household and the NPISH sectors is in Gregorini et al. (2016).

The non-financial corporations sector comprises of all the private and public corporate enterprises that produce goods or provide non-financial services to the market.

The government sector excludes public enterprises and comprises central, state (regional) and local government and social security funds.

The financial corporations sector comprises of all the private and public entities engaged in financial intermediation, such as monetary financial institutions (broadly equivalent to banks), investment funds, insurance corporations and pension funds.

Complete and consistent quarterly rest of the world accounts for the euro area and the European Union (EU) are compiled. This means that cross-border transactions and financial claims amongst the euro area/EU member states have been removed from the rest of the world accounts and, in particular, the asymmetries in the bilateral trade statistics have been eliminated. Consequently, imports and exports are much smaller than they would have been if a simple aggregation of the national data had been used. About half of the external trade of the individual member states is within the euro area/EU.

5.2.1 Description of the problem

As already mentioned, the European Quarterly Sector Accounts (QSA) are compiled in such a way that they are not a simple sum of the countries' data. This is because of different reasons:

1. **Estimations**:

   According to the European regulation (Eurostat, 2010), the member states whose GDP at current prices represents less than 1% of the corresponding Union total are not obliged to transmit the whole set of QSA data to Eurostat. Thus, these missing figures have to be estimated in order to compile the aggregates.

   Moreover, not all member states are 100% compliant with the transmission regulation, which in facts results in more figures to be estimated for the compilation of the aggregates.

2. **General government sector**:

   The data for the general government sector (S13) are replaced by the data obtained from the short-term public finance statistics (STPFS) collection at aggregated level, as the countries' data follow a more strict validation process and are considered of higher quality.

3. **Intra flows**:

   It is clear that when aggregating variables related to the rest of the world sector, a simple sum of the countries' data would not give the correct figures, as the definition of the rest of the world sector changes according to the geographical area. Thus, the figures for S2 have to be replaced by an estimation of the extra flows.

4. **European institutions**:

   Apart from the countries, the European aggregates also include the European institutions, which are international organisations belonging to the European Union. The euro area aggregates include only two institutions: the European Central Bank (ECB) and the European Stability Mechanism (ESM), which have a relatively small impact on the total.

Preliminary QSA estimates are hence obtained by summing up the countries' data, including the estimations, replacing the S13 sector with the STPFS aggregates, and replacing the S2 sector with a proportional allocation of intra and extra flow according to the Balance of Payments (BoP) statistics to the rest of the world total. In this case, only the euro area is considered. The estimation has been performed according to seasonal ARIMA models, when possible. ECB and ESM accounts have not been included to the preliminary estimates due to their confidentiality and their small weight in the euro area total.

Annual Sector Accounts (ASA) data are in principle available for all the member states, although some estimations for missing data are also needed. They could be used for benchmarking the QSA data. Because of various reasons, like discrepancies between QSA and ASA data at countries level, or estimations of missing data, the differences between the annualised preliminary QSA estimates and the ASA benchmarks could be very large.

Tables 5.10 and 5.11 contain the list of the existing QSA production variables for each sector and both uses (paid) and resources (received) sides, while Tables 5.12 and 5.13 contain the list of the distributive variables. The cells shaded in gray indicate when the variables do not exist, while the check (✓) and x (✗) symbols denote variables for which the European aggregates are compiled or not, respectively. In few cases, the variables are constrained to be equal to zero.

The first important differentiation to be done is between the production (P) and distributive (D) transactions.

Distributive transactions need to be *balanced*, which means that the total uses (pay) must be equal to the total resources (rec). For all distributive transactions, it should be:

$$\mathrm{S1}pay + \mathrm{S2}pay = \mathrm{S1}rec + \mathrm{S2}rec \tag{5.1}$$

The total S1 is given by the sum of the sectors:

$$\begin{aligned} \mathrm{S1}pay &= \mathrm{S1N}pay + \mathrm{S11}pay + \mathrm{S12}pay + \mathrm{S13}pay + \mathrm{S1M}pay \\ \mathrm{S1}rec &= \mathrm{S1N}rec + \mathrm{S11}rec + \mathrm{S12}rec + \mathrm{S13}rec + \mathrm{S1M}rec \end{aligned} \tag{5.2}$$

Thus, equation 5.1 can be written as:

$$\begin{aligned} &\mathrm{S1N}pay + \mathrm{S11}pay + \mathrm{S12}pay + \mathrm{S13}pay + \mathrm{S1M}pay + \mathrm{S2}pay = \\ &= \mathrm{S1N}rec + \mathrm{S11}rec + \mathrm{S12}rec + \mathrm{S13}rec + \mathrm{S1M}rec + \mathrm{S2}rec \end{aligned}$$

Or:

$$\begin{aligned} &\mathrm{S1N}pay + \mathrm{S11}pay + \mathrm{S12}pay + \mathrm{S1M}pay + \\ &- \mathrm{S1N}rec - \mathrm{S11}rec - \mathrm{S12}rec - \mathrm{S1M}rec = \\ &= \mathrm{S13}rec + \mathrm{S2}rec - \mathrm{S13}pay - \mathrm{S2}pay \end{aligned} \tag{5.3}$$

The right elements of this last expression are actually known (as they cannot be modified), and thus they can be used as contemporaneous constraints. From this point of view, the balancing of the QSA data is done by partially using an *ad hoc* balancing procedure.

Transactions P51C and NP are here considered together with the distributive transactions for completeness, but are actually not reconciled. P51C has no discrepancies at all, while NP has no annual constraint.

It should be noted that equation 5.3 represents the most generic situation, when all the variables exist for both uses and resources. Actually, in most of the cases, equation 5.3 has less elements, as shown in Table 5.12.

Moreover, this approach is followed only for the one digit series, while the lower digits series will be reconciled according to the hierarchical approach shown in Chapter 3, leaving sector S12 as a residual. This is because for non-financial sector accounts, sector S12 is often the less interesting from an economic point of view, while more interesting indicators, such as the households saving rate or the non-financial corporations investment rate, could be computed on other sectors.

As regards to production transactions, the main equation to be considered is the so called goods and services balancing for the total economy:

$$\mathrm{P2} + \mathrm{P3} + \mathrm{P5} + \mathrm{P6} = \mathrm{P1} + \mathrm{P7} + \mathrm{D21}pay - \mathrm{D31}rec \qquad (5.4)$$

The gross value added (B1G) is obtained from the countries' data, and it is used to derive P1 and P2:

$$\mathrm{B1G} = \mathrm{P1} - \mathrm{P2}$$

So, equation 5.4 could be written as:

$$\mathrm{P3} + \mathrm{P5} = \mathrm{P1} - \mathrm{P2} + \mathrm{P7} - \mathrm{P6} + \mathrm{D21}pay - \mathrm{D31}rec \tag{5.5}$$

Where the right part of this expression is known, and thus could be used as contemporaneous constraints.

Once P5 is known for the total economy, the figures for the sectors are derived by considering the values for S1 as the contemporaneous constraints. Once this is done, for each sector other than S13 (and S12, which is again derived as a residual), a two-step reconciliation is applied to the lower digit transactions:

$$\mathrm{P5} = \mathrm{P51G} + \mathrm{P5M}$$

The complete hierarchical chart for the QSA problem shown in 5.4.

Finally, it is worth mentioning that in this exercise no balancing transactions have been considered. This is because the balancing transactions can always be calculated starting from the distributive and production transactions. Of course, different ways of proceeding are possible. For example, the user could fix the balancing transactions and reconcile the other variables to them.

The QSA variables, expressed in millions of euros, are compiled at quarterly level, with data available from the first quarter of 1999 to the third quarter of 2016, while ASA data are available until year 2015.[4] The complete set of dimensions for the QSA reconciliation problem is expressed in Table 5.14.

5.2.2 Results

As done for the IPI problem, four different two-step reconciliation methods have been applied to the QSA case, varying the methods used in both the first step (modified Denton PFD or Chow-Lin) and the second step (Quenneville-Rancourt or Di Fonzo-Marini). Table 5.15 shows the detailed mean squared percentage differences for each of the system considered, where it is clear that different systems bring to different results. Generally (but not always), the Di Fonzo-Marini method performs better than the Quenneville-Rancourt one, while nothing generic could be said regarding the differences between the Chow-Lin and Denton methods.

It is also worth to note that while all discrepancies are in general terms small, contrary to what has happened to the IPI problem (which was one single system of time series), here the situation varies from system to system, with discrepancies between the preliminary series and the annual constraints which in same cases reach 15%. Moreover, it should be noted that for some specific series, the first step of the procedure does

[4] QSA and ASA data have been downloaded in January 2017 from the Eurostat database (http://ec.europa.eu/eurostat/data/database).

Table 5.10: List of the QSA production variables

Code	*Uses*						*Resources*					
	S1N	S11	S12	S13	S1M	S2	S1N	S11	S12	S13	S1M	S2
P1								✗	✗	✗	✗	
P1O										✗		
P2		✗	✗	✗	✗							
P3				✓	✓							
P31				✓	✓							
P32				✓								
P5		✓	✓	✓	✓							
P51G		✓	✓	✓	✓							
P5M		✓	✓	✓	✓							
P6						✓						
P61						✓						
P62						✓						
P7												✓
P71												✓
P72												✓

not have to be performed since the series are already benchmarked to the annual totals.

A very interesting example, which clearly shows how the different methods perform, is given by the reconciliation of the system for the breakdown of P5 in P51G and P5M for sector S11, whose results are given in Table 5.16, while the charts of the preliminary variables are shown in Charts 5.5 and 5.6.

In order to interpret the results, it is important to note that the two variables of the system have quite different sizes, with P51G being quite big, and P5M being smaller, and it alternates negative and positive values. Moreover, the discrepancies between the quarterly preliminary series and the annual constraints are smaller before 2009 and start increasing in the last years.

Figure 5.4: Hierarchical chart of the QSA problem

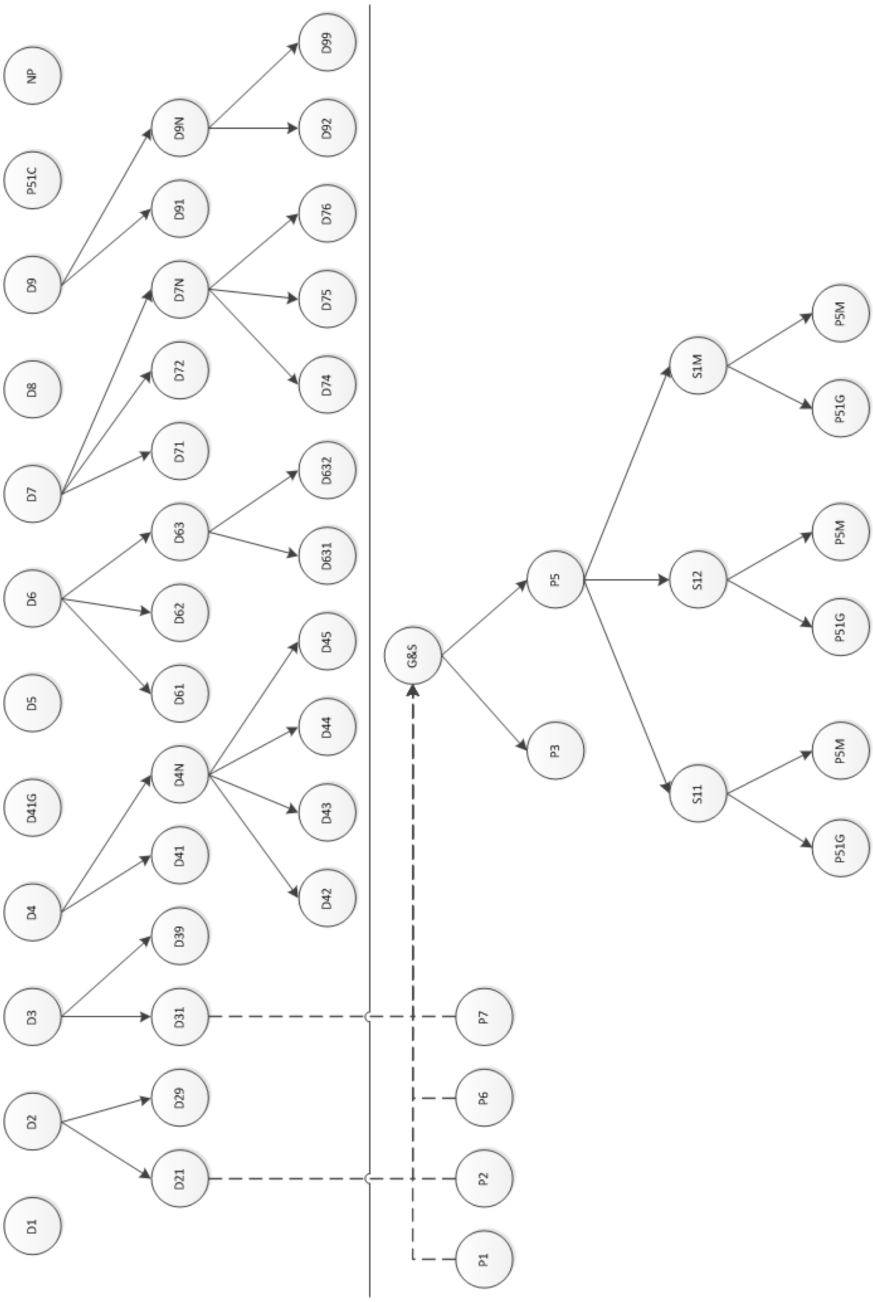


Table 5.11: Codes of QSA production variables

Code	***Description***
P1	Intermediate consumption
P1	Output
P1O	Market output, output for final use and payments for other non-market output
P2	Intermediate consumption
P3	Final consumption expenditure
P31	Individual consumption expenditure
P32	Collective consumption expenditure
P5	Gross capital formation
P51G	Gross fixed capital formation
P5M	Changes in inventories and acquisitions less disposals of valuables
P6	Exports of goods and services
P61	Exports of goods
P62	Exports of services
P7	Imports of goods and services
P71	Imports of goods
P72	Imports of services

Charts 5.7 and 5.8 show the absolute percentage differences of the reconciled results of P51G obtained using the Quenneville-Rancourt and the Di Fonzo-Marini techniques, respectively. The two charts clearly show that the results obtained are much better in terms of stability, when using the Chow-Lin technique in the first step. Moreover, the results of the Denton approach deteriorate in the latest years, when the differences between the preliminary series and the annual constraints are bigger, creating also a problem in the extrapolated values. Finally, it could be noted that the stability is actually higher for the results obtained using the Di Fonzo-Marini approach in the second step.

While for P51G the results are presented according to the method applied in the second step, the results obtained for P5M are instead reported

Table 5.12: List of the QSA distributive variables

Code	*Uses*						*Resources*					
	S1N	S11	S12	S13	S1M	S2	S1N	S11	S12	S13	S1M	S2
D1		✓	✓	✓	✓	✓					✓	✓
D2	✓	✓	✓	✓	✓					✓		✓
D21	✓									✓		✓
D211										✗		
D29		✓	✓	✓	✓					✓		✓
D3				✓		✓	✓	✓	✓	✓	✓	
D31				✓		✓	✓					
D39				✓		✓		✓	✓	✓	✓	
D4		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
D41		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
D4N		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
D42		✓	✓			✓		✓	✓	✓	✓	✓
D43		✓	✓			✓		✓	✓	= 0	= 0	✓
D44		= 0	✓	= 0		✓		✓	✓	✓	✓	✓
D45		✓	✓	✓	✓			✓	= 0	✓	✓	
D41G		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
D5		✓	✓	✓	✓	✓				✓		✓
D6		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
D61					✓	✓		✓	✓	✓	✓	✓
D62		✓	✓	✓	✓	✓					✓	✓
D63				✓	✓						✓	
D631				✗	✗							
D632				✗	✗							
D7		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
D71		✓	✓	✓	✓	✓			✓	= 0		✓
D72			✓	= 0		✓		✓	✓	✓	✓	✓
D7N		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
D74				✗		✗				✗		✗
D75		✗	✗	✗	✗	✗		✗	✗	✗	✗	✗
D76				✗								✗
D8		✓	✓	✓	✓	✓					✓	✓
D9		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
D91		✓	✓		✓	✓				✓		✓
D9N		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
D92				✗		✗		✗	✗	✗	✗	✗
D99		✗	✗	✗	✗	✗		✗	✗	✗	✗	✗
P51C		✓	✓	✓	✓			✓	✓	✓	✓	
NP		✓	✓	✓	✓	✓						

Table 5.13: Codes of QSA distributive variables

Code	***Description***
D1	Compensation of employees
D2	Taxes on production and imports
D21	Taxes on products
D211	Value added type taxes (VAT)
D29	Other taxes on production
D3	Subsidies
D31	Subsidies on products
D39	Other subsidies on production
D4	Property income
D41	Interest
D4N	Property income other than interest
D42	Distributed income of corporations
D43	Reinvested earnings on direct foreign investment
D44	Property income attributed to insurance policy holders
D45	Rents
D41G	Total interest before FISIM allocation
D5	Current taxes on income, wealth, etc.
D6	Social contributions and benefits
D61	Net social contributions
D62	Social benefits other than social transfers in kind
D63	Social transfers in kind
D631	Social transfers in kind - non-market production
D632	Social transfers in kind - purchased market production
D7	Other current transfers
D71	Net non-life insurance premiums
D72	Non-life insurance claims
D7N	Other current transfers (excl. transfers within general government)
D74	Current international cooperation
D75	Miscellaneous current transfers
D76	VAT and GNI - based EU own resources
D8	Adjustment for the change in pension entitlements
D9	Capital transfers
D91	Capital taxes
D9N	Investment grants and other capital transfers
D92	Investment grants
D99	Other capital transfers
P51C	Consumption of fixed capital
NP	Acquisitions less disposals of non-financial non-produced assets

Table 5.14: Dimensions of the QSA reconciliation

Dimension	***Notation***	***Value***
Type of variables	*type*	FLOW
Number of LF observations	N	17
Temporal aggregation order	s	4
Number of HF observations	n	71
Number of extrapolated observations	$n - sN$	3
Number of variables	m	89
Number of accounting relationships	k	28

according to the approach followed in the first step, using the Chow-Lin method for Chart 5.9 and the modified Denton PFD method for Chart 5.10. In this case, it is very clear that by applying the Chow-Lin method, the results obtained are much better than when the modified Denton PFD method is used. Moreover, by using the PFD method, the results are again worsening in the latest periods.

These results are somehow not surprising. As the preliminary series are not much in line with the annual benchmarks in the latest periods, the results obtained by using the modified Denton PFD method are not very good. This is because the Chow-Lin approach is able to capture the autocorrelation of the series, while the Denton approach simply mathematically redistributes the discrepancies according to the minimisation function. This creates results which are more stable (less volatile) when using the Chow-Lin technique. Moreover, the Denton approach has a negative effect on the extrapolations, since it does not perform well if there is any shock in the series.

Regarding the second step, the main noticeable thing is that the quality of the results obtained by using the Di Fonzo-Marini method are correlated negatively with the size of the variables. This is because the implied reliability of the variables is exactly the same when using the Di Fonzo-Marini technique, while it depends on the size of the variables when using the Quenneville-Rancourt method. Consequently, by using the Di Fonzo-Marini approach, the smaller variables will be adjusted less than when using the Quenneville-Rancourt one.

For this reason, the user has to make a choice. If the interpretation and the quality of all the series of the system have the same importance, then the Di Fonzo-Marini approach should be used. Otherwise, by applying the Quenneville-Rancourt method, the results obtained on the smaller series would be of a lower quality than the results obtained on bigger series.

Finally, by looking at the complete results in Table 5.17, it seems clear that the method performing better is the one using the Chow-Lin technique in the first step and the Di Fonzo-Marini technique in the second step.

These results are even more clear in Table 5.18, where the statistics are calculated on extrapolated values only. Once again, the Chow-Lin method is better indicated for dealing with the extrapolation of the data.

Table 5.15: Mean squared percentage differences by system

System	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
D1	0.02%	0.02%	0.02%	0.02%
D2	1.64%	1.67%	1.37%	1.37%
D3	7.99%	6.63%	7.92%	6.58%
D4	0.64%	0.62%	0.64%	0.61%
D4_S11pay	0.88%	0.83%	0.84%	0.81%
D4_S1Mpay	0.86%	0.74%	0.86%	0.74%
D4_S11rec	0.98%	0.94%	0.90%	0.84%
D4_S1Mrec	0.87%	0.79%	0.84%	0.79%
D4N_S11pay	0.89%	0.79%	0.86%	0.78%
D4N_S11rec	0.84%	0.77%	0.78%	0.69%
D4N_S1Mrec	0.96%	0.87%	0.92%	0.86%
D41G	1.00%	0.99%	1.00%	0.99%
D5	0.95%	0.92%	0.95%	0.93%
D6	0.07%	0.07%	0.07%	0.07%
D6_S1Mrec	0.18%	0.19%	0.19%	0.19%
D7	2.35%	2.34%	2.30%	2.29%
D7_S11pay	3.18%	3.18%	3.10%	3.10%
D7_S1Mpay	3.10%	3.10%	3.10%	3.10%
D7_S11rec	4.11%	4.07%	3.93%	3.91%
D7_S1Mrec	3.95%	3.96%	3.93%	3.94%
D8	0.21%	0.21%	0.22%	0.21%
D9	3.88%	3.59%	3.79%	3.50%
D9_S11pay	0.37%	1.00%	0.37%	1.00%
D9_S1Mpay	2.22%	1.38%	2.24%	1.41%
G&S	0.89%	0.88%	0.89%	0.88%
P5	1.27%	1.24%	1.26%	1.24%
P5_S11	1.58%	1.56%	8.21%	8.54%
P5_S1M	0.25%	0.04%	0.25%	0.04%

Figure 5.5: P51G for S11, preliminary series

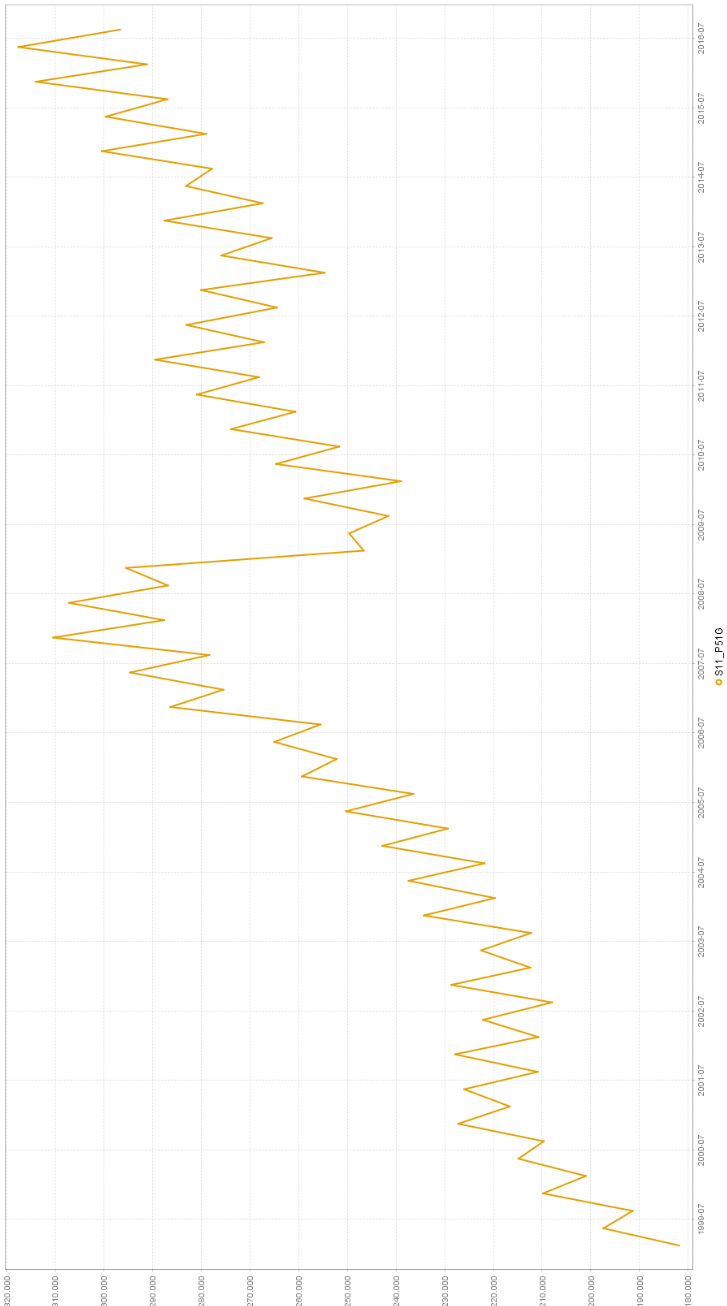


Figure 5.6: P5M for S11, preliminary series

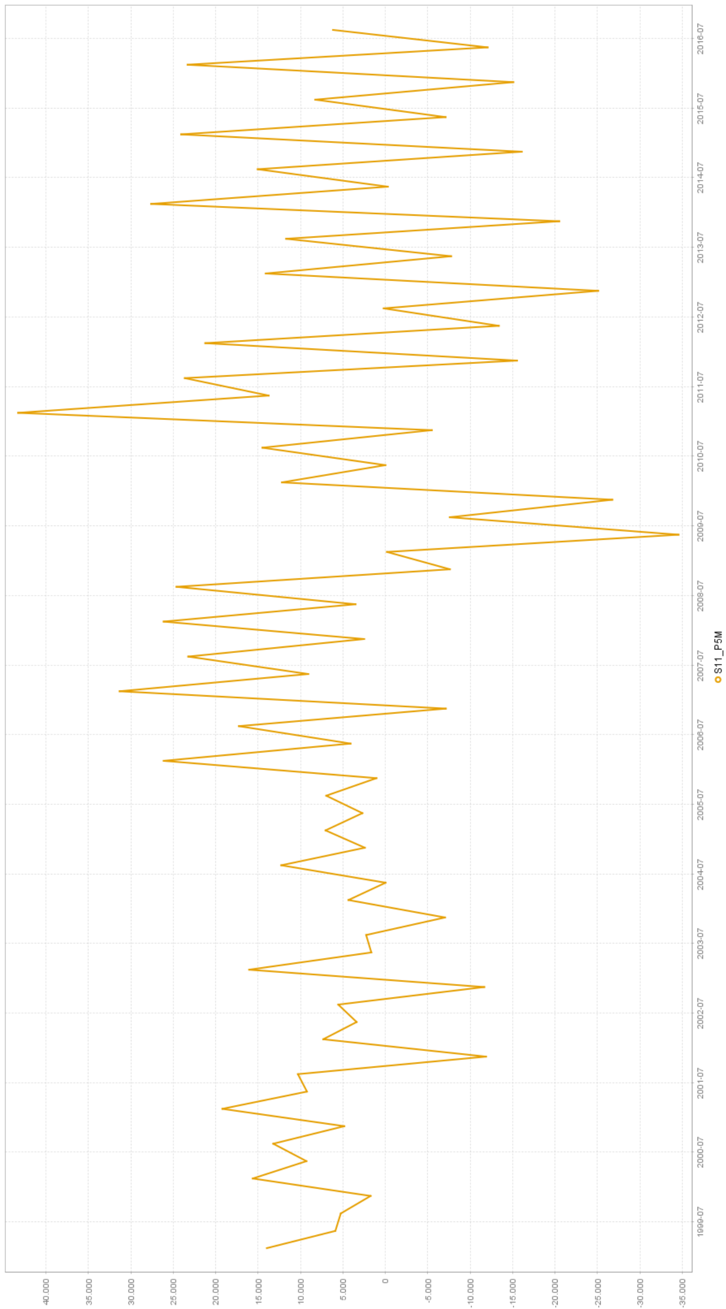


Figure 5.7: APD of reconciled P51G for S11, Quenneville-Rancourt

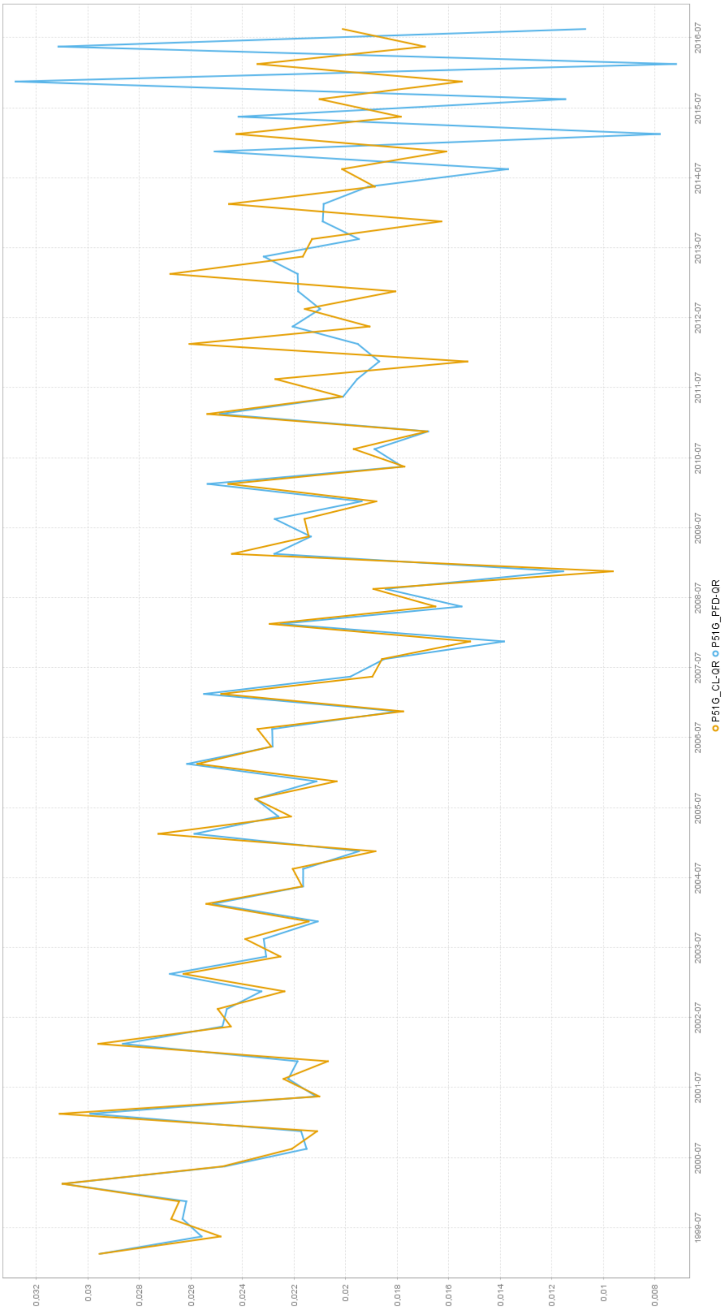


Figure 5.8: APD of reconciled P5M for S11, Di Fonzo-Marini

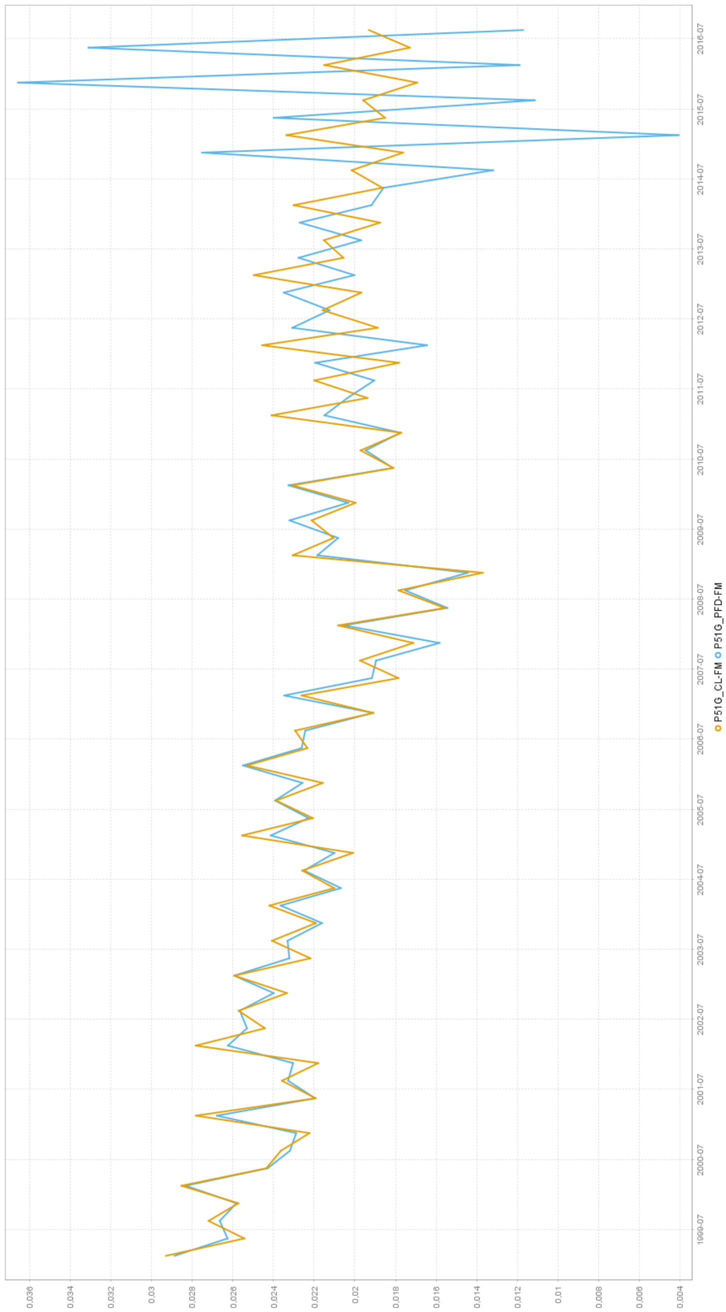


Figure 5.9: APD of reconciled P5M for S11, Chow-Lin

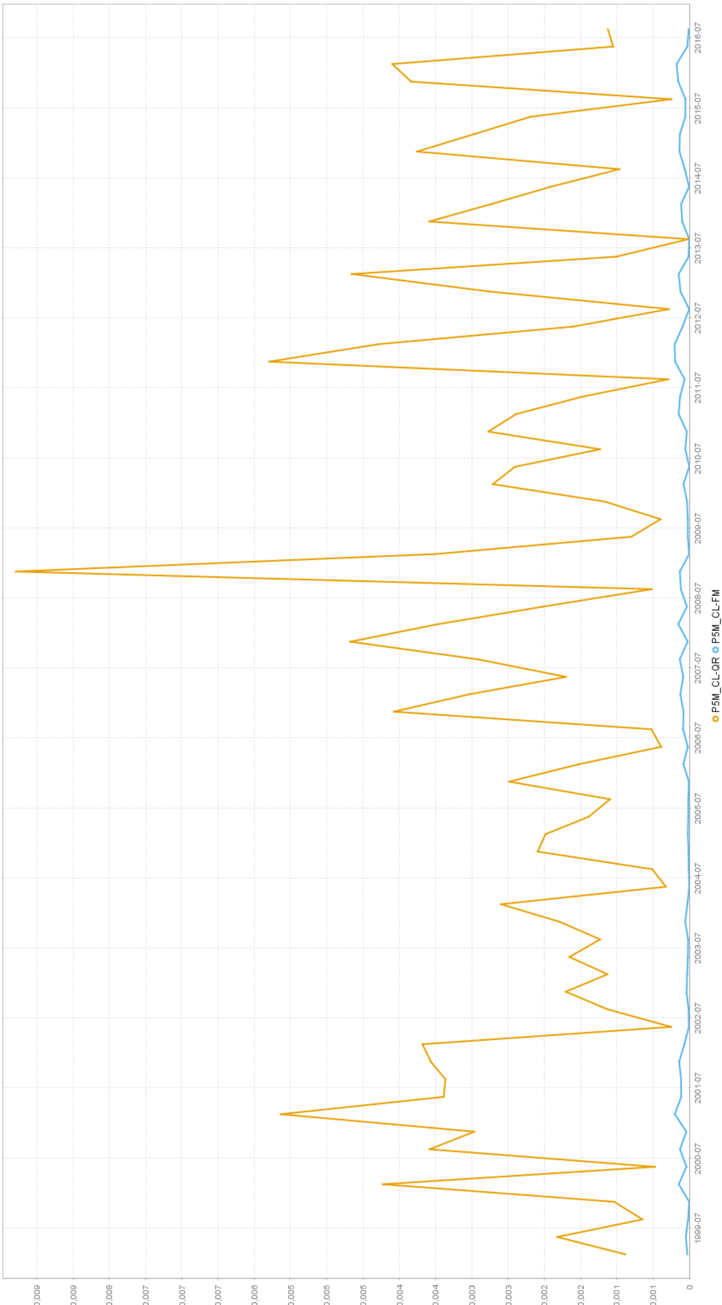


Figure 5.10: APD of reconciled P5M for S11, PFD

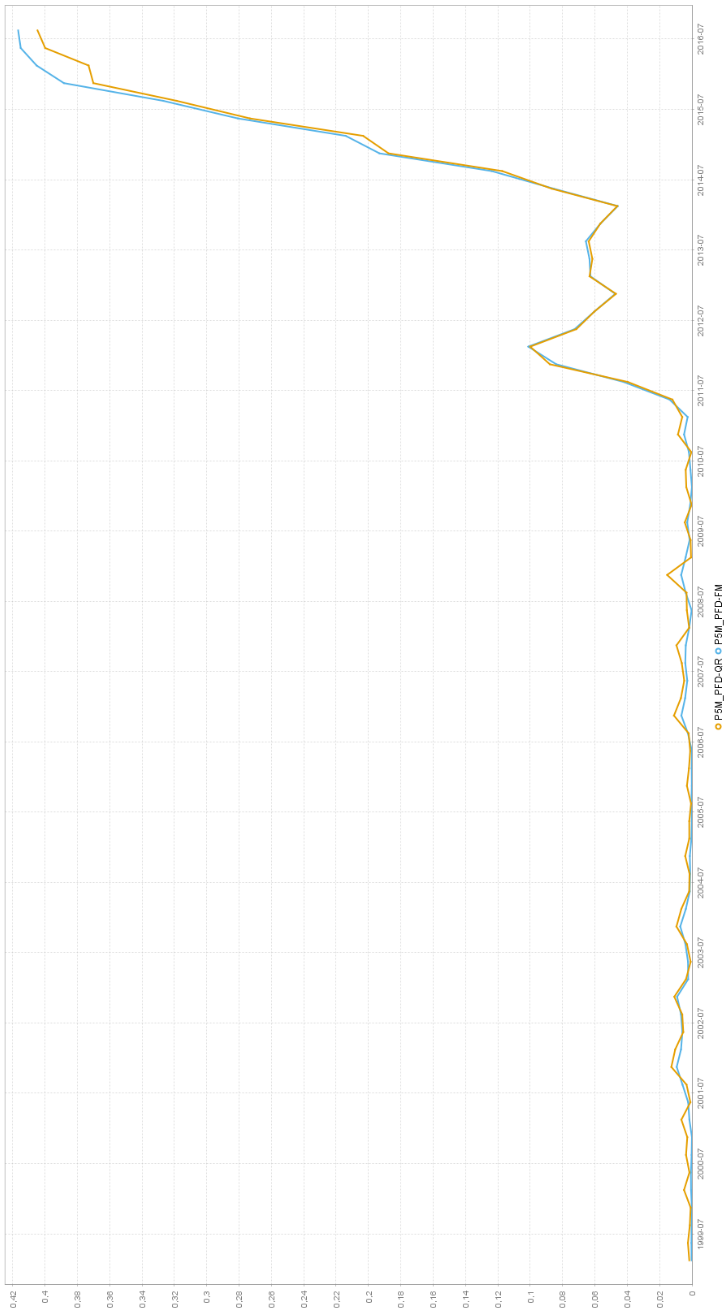


Table 5.16: Statistics of the P5 reconciliation in S11

Index	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
P51G				
MeanAPD	2.19%	2.18%	2.16%	2.17%
MaxAPD	3.11%	2.93%	3.28%	3.66%
MeanSPD	2.22%	2.20%	2.22%	2.22%
MeanAPDG	0.45%	0.32%	0.54%	0.52%
MaxAPDG	1.22%	0.77%	4.21%	4.97%
MeanSPDG	0.53%	0.36%	0.92%	1.02%
C1	100.00 %	100.00 %	100.00 %	100.00 %
P5M				
MeanAPD	0.22%	0.01%	5.14%	5.20%
MaxAPD	0.93%	0.02%	40.46%	41.66%
MeanSPD	0.28%	0.01%	11.40%	11.87%
MeanAPDG	27.39%	0.43%	32.21%	14.43%
MaxAPDG	1766.22 %	26.11%	1710.09 %	461.74 %
MeanSPDG	211.26 %	3.13%	206.73 %	63.95%
C1	100.00 %	100.00 %	100.00 %	100.00 %

Table 5.17: Statistics of the QSA problem

Index	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
MeanAPD	1.17%	1.00%	1.20%	1.04%
MeanSPD	2.44%	2.20%	2.68%	2.49%
MeanAPDG	1.38%	0.66%	1.29%	0.73%
MeanSPDG	22.92%	3.62%	22.23%	7.78%
C1	97.61%	98.51%	97.61%	98.56%

Table 5.18: Statistics of the QSA extrapolations

Index	***Chow-Lin***		***Denton PFD***	
	QR	***FM***	***QR***	***FM***
MeanAPD	1.21%	0.92%	1.62%	1.34%
MeanSPD	2.06%	1.73%	4.57%	4.62%
MeanAPDG	1.12%	0.65%	1.14%	0.65%
MeanSPDG	2.43%	1.49%	2.66%	1.54%
C1	97.80%	98.53%	98.53%	98.90%

5.3 Conclusions

In this chapter, two applications on real data sets have been presented: a relatively small sized reconciliation problem, as well as a complex medium-large sized one, have been handled with the use of different two-step techniques.

A large emphasis was given to the differences in the results when a regression-based temporal disaggregation method has been applied instead of the modified Denton one. Extra focus has been given to the effects of the extrapolation on the final results.

In the case of the industrial production index, the results have clearly shown that the performances of the two-step reconciliation techniques using a regression-based method in the first step, are better than when the modified Denton method is applied, especially for the extrapolations. These results are not entirely surprising, since many observations are extrapolated and the developments in the last available year have been particularly different than the ones from the previous years, and from the current year from which the data are extrapolated.

Similar results have been obtained by performing the reconciliation techniques after using a mixed approach for seasonal adjustment.

As for the QSA case, the results differ internally to the problem, as its complexity is quite larger than the IPI case. However, once again the best results have been obtained by using the Chow-Lin method in the first step and the the Di Fonzo-Marini one in the second step. Moreover, if the Denton approach is applied in the first step, there is a risk of being unable to capture the correct movements in the extrapolation part.

Finally, the choice between the two approaches in the second step should be done according to the reliability of the variables. Thus, at the decision-making stage, different aspects should probably be considered by the user, including the needs of the data, the importance and the quality of the series to be reconciled.

Chapter 6

Conclusions

The most used techniques for temporal disaggregation, balancing and reconciliation have been described in detail in Chapter 2. The main frameworks for temporal disaggregation have been developed in the seventies, in the well known papers by Chow and Lin (1971) and Denton (1971), and were improved mainly in the eighties, especially by Fernández (1981), Litterman (1983) and Cholette (1984). These techniques are currently widely used at least across statistical offices, mainly due to their simplicity and efficiency.

Balancing techniques date back to the first half of the last century, mainly due to the work done by Stone et al. (1942), who initiated a framework which is still valid and widely used. The approach to balancing, including bi-proportional matrix balancing, such as the RAS methodology, has than been completely formulated by Bacharach (1970). Although it is not a statistical approach, *ad hoc* balancing is often a good solution, which could help the user in deciding on qualitative basis where discrepancies should be allocated, and could also be used in combination with quantitative statistical techniques.

The work done by the above mentioned authors is of extreme importance and validity, and it represents the milestones of the temporal disaggregation, benchmarking and balancing techniques. Although not very recent, these methods are still amongst the most valid and used ones. However, they only consider one dimension, which is either the time frequencies (temporal disaggregation) or the variables (balancing).

Reconciliation techniques manage to solve both the temporal and the contemporaneous constraints, and therefore are to be considered as multidimensional methods. The work done in the nineties by Di Fonzo (1990) helped to formulate a framework for reconciliation in the regression based approach, while a more complete framework for simultaneous methods is in Di Fonzo (2003a) and Di Fonzo and Marini (2011b). Moreover, two-step reconciliation methods have been lately developed by Quenneville and Rancourt (2005), followed by Di Fonzo and Marini (2011b).

In Chapter 2, a new classification of the methods described is provided. Such innovative classification is done for temporal disaggregation, balancing and reconciliation techniques, and includes all the principal methods developed so far.

It was only in recent years that two-step reconciliation methods have started to be used in the production of official statistics by national statistical institutes, and the interest keeps growing. This group of methods are indeed a valid alternative, and could bring value added to statistical agencies.

However, the applications have been limited to the use of a modified Denton method in the first step, and in particular, the first proportionate differences version. If on one hand, the Denton methodology is a well established practice for dealing with benchmarking problems, on the other hand, it has been highlighted that it has some drawbacks, especially in the performances of the extrapolations, as well as when the results are obtained from problematic series. This has also been verified in the simulation exercise performed in Chapter 4 and in the two empirical applications presented in Chapter 5. For this reason, a valid alternative is the use of regression-based techniques in the first step, and it is actually preferable in certain conditions, specifically when the user is not sure about the relationship between the related and the target reconciled series and when there are many observations to be extrapolated.

Such a two-step reconciliation technique, which allows the user to decide which technique to use in both the first and the second step, has been described in Chapter 3, together with the possibility of adjusting nested systems of time series. If on one hand, the results of the lower layers will depend on the results obtained at higher layers, on the other hand, the use of such cascade approach is justified by the fact that in official statistics the series in the lower layers are often also of a lower quality.

From a practical point of view, three new plug-ins have been developed in the JEcotrim tool, dealing with Chow-Lin, Fernández and Litterman, respectively, in the first step, and allowing the user to choose between the approaches by Quenneville-Rancourt and Di Fonzo-Marini in the second step. The algorithms implemented in the tool are very user friendly and therefore could help all kind of users in applying such techniques.

In order to reduce the discrepancies before the reconciliation of a system obtained after a seasonal adjustment process, an innovative test has also been presented. This test has been designed in order to identify common seasonal patterns in a set of time series, so that the user can have an indication regarding at which level the seasonal adjustment is to be performed.

Regarding the validation of the reconciliation techniques, several ways for assessing the quality of the results have been presented in Chapter 4.

Together with a wide set of important statistics for measuring the sizes of the distance between the preliminary and the reconciled series, a new methodology for detecting outliers at the end of the series has been introduced. Such methodology can also be seen as a validation criteria for reconciliation techniques, since the outliers identified in the final series should be the same as the one identified in the preliminary or related series, in order to preserve the movements.

Moreover, a simulation study has been presented in order to verify the validity of four possible combinations of methods between the first and the second step, varying the Chow-Lin and the modified Denton PFD methods in the first step and the Quenneville-Rancourt and the Di Fonzo-Marini methods in the second step. The results clearly showed that all the methods are actually valid and that the possibility of choosing amongst them in different situations would definitely help the user in getting better results.

Finally, two empirical applications have been performed on a relatively small sized system (the industrial production index) and on a more complex medium-large sized set of systems (the quarterly sector accounts). In particular, the results obtained in both cases have shown that from a practical point of view, applying the Chow-Lin method in the first step would improve the extrapolation results. Also, it seems that in all the cases where the size of the series is not similar across variables, applying the Di Fonzo-Marini method would lead to better overall results. However, it should still be the user to decide whether obtaining better results for the small series is a good compromise, or if it would be enough to focus on the big series of the systems, which might be of greater importance. In the latter case the Quenneville-Rancourt approach has to be preferred.

Certain areas mentioned in this study could be analysed further. In particular, a comparison of the results after the first step with the structural models methods could be envisaged, as well as regard to the simultaneous approaches for reconciliation. Furthermore, the simulations could be expanded to analyse how the methods behave with the variations of different dimensions such as the temporal aggregation order or the length of the series. Moreover, different ways of producing the preliminary series could be explored. Finally, the work done could be extended to the case of non-linearity of the contemporaneous constraints.

The tools developed are very practical to use and could be quickly adopted for the production of official statistics by statistical agencies, which too often lack the methodological competence for applying more complicate methods. Moreover, this kind of approach would definitely be useful in tackling what can be considered as a third constraint: the time constraint. As timeliness is one of the dimensions of quality in official statistics (European Statistical System, 2011), it is very common that there is literally no time for in-depth analysis when producing these statistics (the reconciliation part of the production is often completed in few hours), and a technique such as the two-step reconciliation approach presented here is a very good alternative, because it could also help in reducing the processing time, when compared to techniques such as the simultaneous approach.

In conclusion, reconciliation techniques are of a great interest in the world of official statistics, especially in domains such as national accounts or unemployment, where the reconciliation of the data is normally requested. In these cases, flexible two-step reconciliation methods can be easily implemented in the production flow, improving the quality of the data in terms of reducing the distance between the preliminary and the reconciled series, and helping in improving the timeliness of the data. In particular, using the Chow-Lin method in the first step, would improve the quality of the results in many practical circumstances.

Bibliography

ABS (2009). *Population Estimates: Concepts, Sources and Methods.* Australian Bureau of Statistics.

Al-Osh, M. (1989). A dynamic linear model approach for disaggregating time series data. *Journal of Forecasting*, 8(2):85–96.

Almon, C. (1968). Recent methodological advances in input–output in the United States and Canada. Unpublished paper presented at the Fourth International Conference on Input–Output Techniques, Geneva.

Astolfi, R., Ladiray, D., and Mazzi, G. (2001a). Business cycle extraction of Euro-zone GDP: direct versus indirect approach. *Working Papers and Studies.* European Commission.

Astolfi, R., Ladiray, D., and Mazzi, G. (2001b). Seasonal adjustment of European aggregates: Direct versus indirect approach. *Working Papers and Studies.* European Commission.

Bacharach, M. (1970). *Biproportional matrices and input-output change.* Cambridge.

Barbone, L., Bodo, G., and Visco, I. (1981). Costi e profitti in senso stretto: un'analisi su serie trimestrali, 1970-1980. *Bollettino della Banca d'Italia*, 36:465–510.

Beaulieu, J. and Bartelsman, E. (2004). Integrating expenditure and income data: What to do with the statistical discrepancy? *Tinbergen Institute Discussion Paper*, 078(3).

Bikker, R., Daalmans, J., and Mushkudiani, N. (2010). A multivariate Denton method for benchmarking large data sets. Technical report, Statistics Netherlands.

Bloem, A., Dippelsman, R., and Maehle, N. (2001). Quarterly national accounts manual concepts, data sources, and compilation. In *Manual on QNA*. International Monetary Fund.

Boot, J., Feibes, W., and Lisman, J. (1967). Further methods of derivation of quarterly figures from annual data. *Applied Statistics*, 16(1):65–75.

Bournay, J. and Laroque, G. (1979). Réflexions sur la méthode d'élaboration des comptes trimestriels. *Annales de l'inséé*, 36:3–30.

Box, G. and Jenkins, G. (1976). *Time Series analysis Forecasting and Control*. Holden-Day, San Francisco.

Bozik, J. and Otto, M. (1988). Benchmarking: Evaluating methods that preserve month-to-month changes. Technical report, U.S. Bureau of the Census.

Buono, D. (2004). Statistical analysis of NMS and CCs GDPs: outlier detection, seasonal adjustment and cycle extraction. *International journal of Applied Econometrics and Quantitative Studies, Euro-American Association of Economic Development*, 1-1.

Buono, D., Gras, F., Infante, E., and Scepi, G. (2016). The saving rate of a clean household sector. Unpublished manuscript presented at the DSSR2016 conference, Naples.

Buono, D. and Infante, E. (2012). New innovative 3-way ANOVA a-priori test for direcct vs. indirect approach in seasonal adjustment. *Working Papers and Studies*. European Commission.

Busetti, F. and Harvey, A. (2003). Seasonality tests. *Journal of Business and Economic Statistics*, 21(3):420–436.

Bušs, G. (2009). Comparing forecasts of Latvia's GDP using simple seasonal ARIMA models and direct versus indirect approach. In *Munich Personal RePEc Archive*.

Causey, B. and Trager, M. (1982). Derivation of solution to the benchmarking problem: Trend revision. Unpublised research notes, U.S. Bureau of the Census.

Centoni, M. and Cubadda, G. (2011). Modelling comovements of economic time series: a selective survey. *Statistica (Bologna)*, 71(2):267–294.

Chamberlin, G. (2010). Temporal disaggregation. *Economic & Labour Market Review - Methods Explained*, pages 106–121.

Chatfield, C. (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners*, chapter Prediction Intervals. J. Scott Armstrong, Kluwer Academic.

Cholette, P. (1984). Adjusting sub-annual series to yearly benchmarks. *Survey Metholody*, 10(1):35–49.

Chow, G. and Lin, A. (1971). Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The Review of Economics and Statistics*, 53(4):372–375.

Ciammola, A., Di Palma, F., and Marini, M. (2005). Temporal disaggregation techniques of time series by related series: a comparison by a Monte Carlo experiment. *Working Papers and Studies.* European Commission.

Cochrane, D. and Orcutt, G. (1949). Application of least squares regression to relationships containing auto-correlated error terms. *Journal of the American Statistical Association*, 44(245):32–61.

Cohen, B. (2007). *Explaining Psychological Statistics*, chapter Three-way ANOVA, pages 688–746. John Wiley & Sons, New York.

Cohen, K., Muller, W., and Padberg, M. (1971). Autoregressive approaches to disaggregation of time series by related series data. *Applied Statistics*, 20:119–129.

Cristadoro, R. and Sabbatini, R. (2000). The seasonal adjustment of the harmonised index of consumer prices for the euro area: a comparison of direct and indirect method. *Banca d'Italia Econocmi Working Paper*, 371.

Daalmans, J. and Di Fonzo, T. (2014). Denton PFD and GRP benchmarking are friends. An empirical evaluation on Dutch supply and use tables. Unpublished paper presented at the 22nd International Input-Output conference.

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

Deming, E. and Stephan, F. (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathermatical Statistics*, 11(4):427–444.

Denton, F. (1971). Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization. *Journal of the American Statistical Association*, 66(99-102).

Di Fonzo, T. (1990). The estimation of m disaggregate time series when contemporaneous and temporal aggregates are known. *The Review of Economics and Statistics*, 72(1):178–182.

Di Fonzo, T. (2003a). Temporal disaggregation of a system of time series when the aggregate is known: Optimal vs. adjustment methods. In *Workshop on Quarterly National Accounts*, pages 63–77.

Di Fonzo, T. (2003b). Temporal disaggregation of economic time series:

towards a dynamic extension. *Working Papers and Studies*. European Commission.

Di Fonzo, T. and Marini, M. (2003). Benchmarking systems of seasonally adjusted time series according to Denton's movement preservation principle. *Working paper 2003.9*. Dipartimento di Scienze Statistiche, Università di Padova.

Di Fonzo, T. and Marini, M. (2007). Statistical reconcilation of time series. Movement preservation vs. a data based procedure. *Working paper 2007.10*. Dipartimento di Scienze Statistiche, Università di Padova.

Di Fonzo, T. and Marini, M. (2010). Benchmarking and movement preservation. Evidences from real-life and simulated series. *Working Paper 2010.14*. Dipartimento di Scienze Statistiche, Università di Padova.

Di Fonzo, T. and Marini, M. (2011a). A Newton's method for benchmarking time series according to a growth rates preservation principle. *IMF Working Papers*, 179(11).

Di Fonzo, T. and Marini, M. (2011b). Simultaneous and two-step reconciliation of systems of time series: methodological and practical issues. *Journal of the Royal Statistical Society*, 60(2):143–164.

Di Fonzo, T. and Marini, M. (2012a). New procedures for the reconcilation of time series. *Working Paper 2012.6*. Dipartimento di Scienze Statistiche, Università di Padova.

Di Fonzo, T. and Marini, M. (2012b). On the extrapolation with the

Denton proportional benchmarking method. *IMF Working Papers*, 12(169).

Di Fonzo, T. and Marini, M. (2013). *Advances in theoretical and applied statistics*, chapter Benchmarking and movement preservation. Evidences from real-life and simulation series.

Di Fonzo, T. and Marini, M. (2015). Reconciliation of systems of time series according to a growth rates preservation principle. *M. Stat Methods Appl.*

Duff, I. (2004). MA57 - a code for the solution of sparse symmetric definite and indefinite systems. *ACM Transactions on Mathematical Software*, 30(2):118–144.

Dusak, K. (1973). Futures trading and investor returns: An investigation of commodity market risk premiums. *Journal of Political Economy*, 81(6):1387–1406.

Eaves, B., Hoffman, A., Rothblum, U., and Schneider, H. (1985). Line-sum-symmetric scalings of square nonnegative matrices. *Mathematical Programming Study*, 25:124–141.

European Statistical System, C. (2011). European statistics code of practice.

Eurostat (2008a). *European Price Statistics: an Overview.* Eurostat, European Commission.

Eurostat (2008b). *Eurostat Manual of Supply, Use and Input-Output Tables.* Eurostat, European Commission.

Eurostat (2010). *European System of Accounts: ESA 2010*. Eurostat, European Commission.

Eurostat (2013). *Handbook on quarterly national accounts*. Eurostat, European Commission.

Eurostat (2015). *ESS guidelines on seasonal adjustment*. Eurostat, European Commission.

Evans, T. D. (2004). Analysis of raking on seasonally adjusted household gross flows data. *American Statistic Association conference*, pages 1166–1173.

Fernández, R. (1981). A methodological note on the estimation of time series. *The Review of Economics and Statistics*, 63(3):471–476.

Findley, D. (2005). Some recent developments and directors in seasonal adjustment. *Journal of Official Statistics*, 21(2):343–365.

Findley, D. F. and Hood, C. C. (2000). X-12-arima and its application to some italian indicator series. In *In Seasonal Adjustment Procedures - Experiences and Perspectives, Annali di Statistica, Serie X, n. 20*, pages 231–251.

Friedlander, D. (1961). A technique for estimating contingency tables, given marginal totals and some supplemental data. *Journal of the Royal Statistical Society, Series A*, 124:412–420.

Geweke, J. (1979). The temporal and sectoral aggregation of seasonally adjusted time series. In Zellner, editor, *Seasonal analysis of economic time series*. US Bureau of the Census.

Gilchrist, D. and St. Louis, L. (1999). Completing input-output tables using partial information with an application to Canadian data. *Economic Systems Research*, 11:185–193.

Giot, P. and Laurent, S. (2003). Market risk in commodity markets: a VaR approach. *Energy Economics*, 25(5):435–457.

Gómez, V. and Maravall, A. (2001). *A Course in Time Series Analysis*, chapter Seasonal Adjustment and Signal Extraction in Economic Time Series. J. Wiley & Sons, New York.

Gomez, V. and Marvall, A. (1994). Estimation, prediction, and interpolation for nonstationary series with the Kalman filter. *Journal of American Statistical Association*, pages 611–624.

Granger, C. W. J. (1978). *Seasonality: Causation, Interpretation, and Implications*, pages 33–56. NBER.

Gregoir, S. (2003). Propositions pour une désagrégation temporelle basée sur des modèles dynamiques simples. In *Workshop on Quarterly National Accounts*, pages 141–166.

Gregorini, F., Infante, E., and Kozina, M. (2016). Detecting outliers at the end of the series using forecast intervals. Unpublished manuscript presented at the IARIW2016 conference, Dresden.

Grudkowska, S. (2015). *JDemetra+ Reference Manual.* Narodowy Bank Polski, 1.1 edition.

Guerrero, V. (1990). Temporal disaggregation of time series: an ARIMA-based approach. *International Statistical Review*, 58(1):29–46.

Gysels, E. and Osborn, D. (2001). The econometrics analysis of seasonal time series. Cambridge University Press.

Harvey, A. (1990). *The econometric analysis of time series.* Pearson Education.

Harvey, A. and Trimbur, T. (2008). Trend estimation and the Hodrick-Prescott filter. *Journal Japanese Statistical Society*, 38(1):41–49.

Hendry, D. and Mizon, G. (1978). Serial correlation as a convenient simplification, not a nuisance: a comment on a study of the demand for money by the Bank of England. *Econonic Journal*, 88(351):549–563.

Higginson, J. (1975). An f test for the presence of moving seasonality when using census method ii-x-11 variant. Technical report, Statistics Canada.

Hildreth, C. and Lu, J. (1960). *Demand Relations With Autocorrelated Disturbances.* Michigan Agri. Exp. Station Technical Bulletin.

Hindrayanto, I. (2004). Seasonal adjustment: direct, indirect or multivariate method? *Aenorm*, 43.

Hodgess, E. and Wei, W. (2000). Temporal disaggregation of stationary bivariate time series. *Linear Algebra and Its Applications*, 321:175–196.

Hodrick, R. and Prescott, E. (1997). Postwat U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 24:1–16.

Hood, C. (2005). An empirical comparison of methods for benchmarking seasonally adjusted series to annual tools. *Working Papers and Studies*. European Commission.

Hood, C. and Findley, D. (2001). Comparing direct and indirect seasonal adjustment of aggregate series. In *ASA proceedings of the business and economic statistics section*.

Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., and Shang, H. L. (2011). Optimal combination forecasts for hierarchical time series. *Computational Statistics & Data Analysis*, 55(9):2579–2589.

Infante, E. and Buono, D. (2013). New technique for predictability, uncertainty, implied volatility and statistical analysis of market risk using SARIMA forecasts intervals. Unpublished manuscript presented at the NTTS2013 conference, Brussels.

Infante, E., Buono, D., and Buono, A. (2015). A 3-way anova a priori test for common seasonal patterns and its application to direct versus indirect methods. *EURONA - Eurostat Review on National Accounts and Macroeconomic Indicators*, pages 67–77.

Infante, E. and Scepi, G. (2017). Alternative two-step reconciliation practices. Unpublished manuscript presented at the NTTS2017 conference, Brussels.

Israilevich, P. (1986). Biproportional forecasting of input-output tables. Unpublished PhD dissertation.

Jackson, R. and Murray, A. (2004). Alternative input-output matrix updating formulations. *Economic Systems Research*, 16:135–148.

Jacobs, J.and Kroonenberg, S. and Wansbeek, T. (1992). Dividing by 4: An efficient algorithm for the optimal disaggregation of annual data into quarterly data. *CCSO Series*, 7. Center for cyclical and structural research.

Jones, R. (1980). Maximum likelihood fitting of ARIMA models to time series with missing observations. *Technometrics*, 22(3):389–395.

Jun, D. B., Moon, J., and Park, S. (2016). Temporal disaggregation: Methods, information loss, and diagnostics. *Journal of Business & Economic Statistics*, 34(1):53–61.

Klein, L. (1958). The estimation of distribution lags. *Econometrica*, 26:553–565.

Ladiray, D. and Mazzi, G. (2003). Seasonal adjustment of European aggregates: Direct versus indirect. In Manna, M. and Peronaci, R., editors, *Seasonal Adjustment*, pages 37–66. European Central Bank.

Ladiray, D. and Quenneville, B. (2001). *Seasonal Adjustment with the X-11 Method*. Springer.

Lahr, M. (2001). Reconciling domestication techniques, the notion of re-exports, and some comments on regional accounting. *Economic Systems Research*, 13:166–179.

Lahr, M. and De Mesnard, L. (2004). Biproportional techniques in input-output analysis: Table updating and structural analysis. *Economic Systems Research*, 16(2):115–134.

Lisman, J. and Sandee, J. (1964). Derivation of quarterly figures from annual data. *Applied Statistics*, 13(2).

Litterman, R. (1983). A random walk, Markov model for the distribution of time series. *Journal of Business & Economic Statistics*, 1(2).

Marangoni, G. and Rossignoli, D. (2014). Richard Stone's contributions to input-output analysis. *Papers in political economy*, 2(71).

Maravall, A. (2006). An application of the TRAMO-SEATS automatic procedure; direct versus indirect approach. *Computational Statistics & Data Analysis*, 50:2167–2190.

Maravall, A. and Caporello, G. (2003). Program TERROR for quality control of time series data. *Occasional Paper*, (1). Research Department, Bank of Spain.

Maravall, A. and Del Rio, A. (2001). Time aggregation and the Hodrick-Prescott filter. In *Documentos de Trabajo*, volume 0108. Banco de España.

Maravall, A. and Pérez, D. (2011). Applying and interpreting model-based seasonal adjustment: The euro-area industrial production series. In *Documentos de Trabajo*, volume 1116. Banco de España.

Marcellino, M. (1999). Temporal disaggregation, missing observations, outliers and forecasting: A unifying non-model-based procedure.

Matuszewski, T., Pitts, P., and Sawyer, J. (1964). Linear programming estimates of changes in input-coefficients. *Canadian Journal of Economics and Political Science*, pages 202–210.

Moauro, F. and Savio, G. (2005). Temporal disaggregation using multivariate structural time series models. *The Econometrics Journal*, 8(2):214–234.

Otranto, E. and Triacca, U. (2000). A distance-based method for the choice of direct or indirect seasonal adjustment. *Contributi ISTAT*, (1).

Palm, F. and Nijman, T. (1984). Missing observations in the dynamic regression model. *Econometrica*, 52:1415–1435.

Pavía-Miralles, J. (2010). A survey of methods to interpolate, distribute and extrapolate time series. *J. Service Science & Management*, pages 449–463.

Poissonier, A. (2013). Temporal disaggregation of stock variables. the Chow-Lin method extended to dynamic models. Technical report, Institut National de la Statistique et des Études Économiques.

Proietti, T. (1999). Distribution and interpolation revisted: a structural approach. *Statistica*, 58(411-432).

Proietti, T. (2005). Temporal disaggregation by state space methods: Dynamic regression methods revisted. *Working Papers and Studies.* European Commission.

Quenneville, B., Picard, F., and Fortier, S. (2013). Calendarization with interpolating splines and state space models. *of the Royal Statistical Society: Series C (Applied Statistics)*, 62(3):371–99.

Quenneville, B. and Rancourt, E. (2005). Simple methods to studies. *Working Papers and Studies*. European Commission.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rau, R. (2006). A generalized non-parametric seasonality test. In *European Population Conference, Liverpool*.

Reber, R. and Park, S. (2014). Methods of temporal disaggregation for estimating output of the insurance industry. *Bureau of Economic Analysis, Working Paper 0115*.

Revilla, P. and Rey del Castillo, P. (1999). Selection editing methods based on time series modelling. In *UN/UCE Work Session on Statistical Data Editing*.

Revilla, P. and Rey del Castillo, P. (2000). Analysis and quality control from ARIMA modelling. In *UN/UCE Work Session on Statistical Data Editing*.

Robinson, S. (2014). *Simulation: The Practice of Model Development and Use*. John Wiley & Sons.

Rodríquez Feijoó, S. and Rodríquez Caro, A. (2000). Los métodos de trimestralización sin indicador. un estudio comparativo mediante simulación. Technical report, Universidad de Las Palmas de Gran Canaria.

Rodríquez Feijoó, S., Rodríquez Caro, A., and Quintana, D. (2003). Methods for quartely disaggregation without indicators; a comparative study using simulation. *Computational Statistics & Data Analysis*, 43:63–78.

Rossi, N. (1982). A note on the estimation of disaggregate time series when the aggregate is known is known. *The Review of Economics and Statistics*, 64:695–696.

Salazar, E., Smith, R., Weale, M., and Wright, S. (1997). A monthly indicator of GDP. *National Institute Economic Review*, pages 84–90.

Santos Silva, J. and Cardoso, F. (2001). The Chow-Lin method using dynamic models. *Economic Modelling*, 18:269–280.

Sax, c. and Steiner, P. (2013). Temporal disaggregation of time series. *The R Journal*, 5(2):80–87.

Stone, R. (1961). Input-output and national accounts. Technical report, OECD.

Stone, R., Champernowne, D., and Meade, J. (1942). The precision of national income estimates. *The Review of Economic Studies*, 9(2):111–125.

Stram, D. and Wei, W. (1986). A methodological note on the disaggregation of time series totals. *Journal of Time Series Analysis*, 7:293–302.

Stuckey, A., Zhang, X. M., and McLaren, C. H. (2004). Aggregation of seasonally adjusted estimates by a post-adjustment. Research Paper. University of Wollongong, Wollongong.

Surtradhar, B. and Dagum, E. (1998). Bartlett-type modified test for moving seasonality with applications. *The Statistician*, 47(1).

Taieb, S. B., Taylor, J. W., and Hyndman, R. J. (2017). Coherent probabilistic forecasts for hierarchical time series. In *Proceedings of the 34 th International Conference on Machine Learning, Sydney, Australia.*

Temurshoev, U. (2012). Entropy-based benchmarking methods. In *GGDC Working Papers*, volume 122. Groningen: GGDC.

Theil, H. (1967). *Economics and information theory.* North-Holland, Amsterdam.

Trager, M. (1982). Derivation of solution to the benchmarking problem: Relative revision. Unpublished research notes, U.S. Bureau of the Census.

Tserkezos, D. (1991). A distributed lag model for quarterly disaggregation of annual personal disposable income of the Greek economy. *Economic Modelling*, 8(4):528–536.

Uribe, P., de Leeuw, C., and Theil, H. (1965). The information theoretic approach to the prediction of interregional trade flows.

Van Tongeren, J. and Magnus, J. (2011). Bayesian integration of large scale sna data frameworks with an application to guatemala. Technical report, Tilburg University.

Weale, M. (1988). The reconciliation of values, volumes and prices in the national accounts. *Journal of the Royal Statistical Society A*, 151:211–221.

Wei, W. and Stram, D. (1990). Disaggregation of time series models. *Journal of the Royal Statistical Society*, Series B (Methodological), 52(3):453–467.

Zani, S. (1970). Sui criteri di calcolo dei valori trimestrali di tendenza degli aggregate della contabilità nazionale. *Studi e Ricerche, Facoltà di Economia e Commercio, Università degli Studi di Parma VII*, pages 287–349.