



Department of Statistical Sciences
University of Padua
Italy

UNIVERSITÀ
DEGLI STUDI
DI PADOVA
DIPARTIMENTO
DI SCIENZE
STATISTICHE

New Procedures for the Reconciliation of Time Series

Tommaso Di Fonzo

Department of Statistical Sciences
University of Padua
Italy

Marco Marini

Statistics Department
International Monetary Fund
USA

Abstract: We propose new simultaneous and two-step procedures for reconciling systems of time series subject to temporal and contemporaneous constraints according to a Growth Rates Preservation (GRP) principle. Two nonlinear optimization algorithms are used: an interior-point method applied to the constrained problem and a Newton's method with Hessian modification applied to a suitably reduced-unconstrained problem. Both techniques exploit the analytic gradient and Hessian of the GRP objective function, making full use of all the derivative information at disposal. We apply the proposed GRP procedures to two large systems of economic series, and compare the results with those of other reconciliation procedures based on the Proportional First Differences (PFD) principle, a linear approximation of the GRP principle widely used by data-producing agencies. Our experiments show that (i) an optimal solution to the nonlinear GRP problem can be efficiently achieved through the proposed Newton's optimization algorithms, and (ii) GRP-based procedures preserve better the growth rates in the system than linear PFD solutions, especially for series with high temporal discrepancy and high volatility.

Keywords: Benchmarking, Reconciliation, Movement preservation, Constrained nonlinear optimization, Newton's method, Interior-point.

JEL classification codes: C22, C61, C82.

Contents

1	Introduction	1
2	Growth Rates Preservation for Benchmarking and Reconciliation	4
3	Optimization algorithms for the GRP problem	7
3.1	An interior-point method for the constrained GRP problem	9
3.2	A Newton's method for the unconstrained GRP problem	11
4	GRP in Two-Step Reconciliation Procedures	13
5	Applications	14
6	Conclusions	19
	Appendix A. Matrix representation of the constraints	21
	Appendix B. Gradient and Hessian of the global GRP criterion	24
	References	25

Department of Statistical Sciences
Via Cesare Battisti, 241
35121 Padova
Italy

tel: +39 049 8274168
fax: +39 049 8274170
<http://www.stat.unipd.it>

Corresponding author:
Tommaso Di Fonzo
tel: +39 049 827 4158
difonzo@stat.unipd.it
<http://www.stat.unipd.it/~difonzo>

New Procedures for the Reconciliation of Time Series

Tommaso Di Fonzo

Department of Statistical Sciences
University of Padua
Italy

Marco Marini

Statistics Department
International Monetary Fund
USA

Abstract: We propose new simultaneous and two-step procedures for reconciling systems of time series subject to temporal and contemporaneous constraints according to a Growth Rates Preservation (GRP) principle. Two nonlinear optimization algorithms are used: an interior-point method applied to the constrained problem and a Newton's method with Hessian modification applied to a suitably reduced-unconstrained problem. Both techniques exploit the analytic gradient and Hessian of the GRP objective function, making full use of all the derivative information at disposal. We apply the proposed GRP procedures to two large systems of economic series, and compare the results with those of other reconciliation procedures based on the Proportional First Differences (PFD) principle, a linear approximation of the GRP principle widely used by data-producing agencies. Our experiments show that (i) an optimal solution to the nonlinear GRP problem can be efficiently achieved through the proposed Newton's optimization algorithms, and (ii) GRP-based procedures preserve better the growth rates in the system than linear PFD solutions, especially for series with high temporal discrepancy and high volatility.

Keywords: Benchmarking, Reconciliation, Movement preservation, Constrained nonlinear optimization, Newton's method, Interior-point.

JEL classification codes: C22, C61, C82.

1 Introduction

The benchmarking problem arises when time series data for the same target variable are measured at different frequencies with different level of accuracy, and there is the need to remove discrepancies between the annual benchmarks and the corresponding aggregates (either sums or averages) of the sub-annual values. For example, the optimal combination of annual levels and quarterly movements requires an adjustment that preserves as much as possible the short-term movements in the preliminary infra-annual figures subject to the restrictions provided by the annual constraints.

The reconciliation problem is commonly known as the adjustment process of a

system of time series where both *temporal aggregation* constraints, for each individual series across the temporal dimension, and *contemporaneous* constraints, for each individual period across the variables of the system, must be satisfied. As for the benchmarking problem, this adjustment should be done according to some movement preservation principle such that the temporal profiles of the original series are preserved to the highest possible degree.

Benchmarking and reconciliation problems are typically faced by statistical agencies in the production of official statistics. Typical examples are the compilation of quarterly supply and use tables (SUT) in national accounts, where the quarterly flows are required to satisfy more comprehensive level estimates from annual SUT and to be in line with the many row and column equalities of the tables; or the production of seasonally adjusted (SA) estimates following a direct approach, when SA series are required to be consistent with their annual unadjusted levels and SA aggregates need to be in line with SA components in any observed period.

Both benchmarking and reconciliation can be performed setting up constrained minimization problems of some mathematical criterion aimed at preserving at the best the movements in the sub-annual values. It is commonly understood by many authors and practitioners that an ideal movement preservation principle should be formulated as an explicit preservation of the period-to-period rates of change of the preliminary series (Helfand *et al.*, 1977; Bloem *et al.*, 2001), according to which the sum of the squared differences between the growth rates of the target series and the growth rates of the preliminary series is minimized. The *Growth Rates Preservation* (GRP) criterion, however, gives rise to a nonlinear problem (NLP), whose solution can only be achieved by recurring to numerical optimization algorithms.

Denton (1971) proposed alternative movement preservation principles (i.e. objective criteria) for benchmarking, that give rise to quadratic-linear optimization problems in the target values. The benchmarked values can thus be found using an explicit formula involving simple matrix operations. In particular, the *Proportionate First Differences* (PFD) criterion, one of the variant proposed by Denton, has been very successful in practical benchmarking applications. The PFD criterion looks for benchmarked estimates aimed at minimizing the sum of squared proportional differences between the target values and the unbenchmarking values.

Di Fonzo and Marini (2011a) extended the PFD criterion to the reconciliation of a system of time series subject to both temporal and contemporaneous constraints. A simultaneous solution to the problem was proposed, which exploits the sparsity of the linear system to be solved. Furthermore, a two-step reconciliation strategy was recommended to reduce the complexity of the problem in the case of large systems: a benchmarking procedure is applied for each series at the first step, and then, the benchmarked series are reconciled year-by-year using a least squares balancing procedure. The work demonstrated empirically that a two-step procedure with a least squares adjustment proportional to the squared level of the benchmarked series at the second step results in a close approximation of the Denton PFD simultaneous solution.

At the individual series level (i.e. for univariate benchmarking), the PFD criterion, with the modification for the starting condition due to Cholette (1984), has often been claimed to be a close approximation of the GRP principle. Thanks to

the satisfactory results obtained from the modified Denton's PFD technique and from other related linear solutions (e.g. Dagum and Cholette, 2006), few significant progress has been achieved towards the development of an efficient and robust optimization algorithm to minimize the nonlinear GRP function for benchmarking and reconciliation problems. A benchmarking procedure based on the GRP criterion was first implemented by Causey and Trager (1981; see also Trager, 1982, and Bozik and Otto, 1988). To solve the NLP defined by the GRP criterion, Causey and Trager developed a steepest descent (SD) algorithm based on first-derivative information (i.e. the gradient). However, using only first-derivative information may result in poorly efficient procedures, characterized by slow convergence and possible troubles in finding actual minima of the objective function. More recently, Brown (2010, 2012) proposed a gradient-based procedure that uses a Conjugate Gradient (CG) algorithm, but the results were broadly in line with the Causey and Trager's SD procedure in terms of efficiency and robustness of the solutions.

Our interest in developing GRP-based benchmarking and reconciliation procedures has been motivated by two reasons. First, we think that recent advances in optimization algorithms along with the huge increase in computational power of computers make it possible to solve the nonlinear GRP problem nowadays more accurately and more rapidly than in the past. Second, once a more efficient optimization algorithm is available, we aim at assessing empirically how true is the supposed approximation of the optimal, nonlinear GRP objective function by the linear PFD solution in both benchmarking and reconciliation problems.

In our recent works we have found that massive improvements in both efficiency and robustness of the benchmarking results can be obtained exploiting both the analytical gradient vector and Hessian matrix of the GRP function. First, Di Fonzo and Marini (2010) found that an interior-point method (Nocedal and Wright, 2006), which uses second-order derivative information, provides more accurate and faster solutions compared to other gradient-based optimization procedures. Second, Di Fonzo and Marini (2011b) proposed a Newton's method with Hessian modification that can be applied after the original constrained benchmarking problem is transformed into an unconstrained problem. This Newton's method is particularly appealing for practitioners because it is easy to implement, computationally robust and time-efficient.

With these effective implementations of the GRP benchmarking procedure we have been able to clarify the nature of the PFD approximation. Using a simulation exercise, we showed the conditions under which the PFD solution provides a close approximation to GRP (Di Fonzo and Marini, 2010). We also found that the approximation works particularly well when the movements in the preliminary series are smooth (low variance in the growth rates) and the relationship with the (annual) target series is relatively stable (no changes in the ratio between their levels), but deteriorates as soon as the preliminary series is lumpy (or affected by strong seasonal effects) or presents sudden level shifts compared to the target series.

In this paper we continue our research on the GRP by extending our proposed Newton's method to solve reconciliation problems subject to both temporal and contemporaneous constraints. Similarly to the benchmarking problem, we transform the original constrained nonlinear reconciliation problem into an equivalent uncon-

strained nonlinear problem in order to apply the Newton's method. We show that the extension is straightforward, but there are complications in the derivation of the unconstrained system due to the presence of constraints having different nature and the larger dimensions of the system. On the contrary, the interior-point method applied to the constrained system is shown to be a fast and feasible procedure even for very large systems. We apply the Newton's method and the interior-point method to reconciling two real-life systems of time series, using both a simultaneous approach and a two-step strategy. We compare the results with the PFD-based reconciliation procedures presented in Di Fonzo and Marini (2011a) in order to demonstrate the effectiveness of the GRP adjustment in terms of both computational efforts and quality of the results.

The paper is organized as follows. In section 2 the optimization problem in terms of GRP is discussed and compared to the classical benchmarking procedure by Denton (1971), modified by Cholette (1984). The extension to reconciliation with both temporal and contemporaneous constraints is then presented. In section 3 we illustrate the two algorithms used to minimize the GRP function: an interior-point method and a Newton's method with Hessian modification. Both algorithms exploit first and second-order derivatives, whose analytic expressions are derived in Appendix B. In section 4 we discuss two-step reconciliation procedures based on the GRP criterion. In order to analyze the distinctive features of the proposed procedures, section 5 presents applications of the Newton's method and the interior-point method for benchmarking and reconciling real-life systems of series, namely 175 quarterly series from the EU Quarterly Sector Accounts (EUQSA), and 236 monthly series from the Canadian Monthly Retail Trade Survey (MRTS). Section 6 presents some final remarks and conclusions, and draws future research lines.

2 Growth Rates Preservation for Benchmarking and Reconciliation

Benchmarking and reconciliation problems are solved through the minimization of an objective function of the unobserved values of one or more target series, which must satisfy given temporal and contemporaneous aggregation constraints. Let us denote the target series by $y_{j,t}$, where the two sub-indices indicate the cross-sectional dimension and the temporal dimension, respectively: $j = 1, \dots, m$ and $t = 1, \dots, n$, with m the number of series considered (if $m = 1$, we have a univariate benchmarking problem) and n the number of high-frequency periods. Each series is denoted in vector form as $\mathbf{y}_j = [y_{j,1}, y_{j,2}, \dots, y_{j,n}]'$. The whole system of series can be conveniently stacked in a single vector as $\mathbf{y} = [\mathbf{y}'_1, \dots, \mathbf{y}'_m]'$. In the following, we assume that for benchmarking and reconciliation problems the system constraints assume the general form

$$\mathbf{A}\mathbf{y} = \mathbf{b} \tag{1}$$

where \mathbf{A} is a $(r \times mn)$ matrix of any given real numbers defining the relationships between the mn observations and \mathbf{b} is the r -dimensional vector with the known quantities (benchmarks) of the system. Appendix A shows some examples of \mathbf{A} and \mathbf{b} for the most common benchmarking and reconciliation cases.

The objective function is usually set as a distance metric between the unobserved target series $y_{j,t}$ and some preliminary series $p_{j,t}$ observed at the same frequency. For economic series, especially those observed at the infra-annual level, it is sensible to define this metric in terms of movements of the series: the user of such statistics is generally much more interested in the dynamic of a monthly or quarterly series (e.g. how much it has grown since the last month or quarter) rather than in its level (e.g. how much is the level in the month or the quarter). For this reason, objective functions for benchmarking and reconciliation problems are most commonly known as *movement preservation principles*.

Causey and Trager (1981; see also Monsour and Trager, 1979, and Trager, 1982) consider a minimization problem for benchmarking, in which the criterion to be minimized is explicitly related to the growth rate, which is a natural measure of the movement of a time series:

$$\min_{y_t} \sum_{t=2}^n \left(\frac{y_t}{y_{t-1}} - \frac{p_t}{p_{t-1}} \right)^2. \quad (2)$$

The benchmarked values y_t^* , $t = 1, \dots, n$, minimize the criterion (2) subject to the aggregation constraints $\sum_{t \in T} y_t = Y_T$, $T = 1, \dots, N$, where index T denotes the low-frequency period (e.g., the year). In other words, the benchmarked series is estimated in such a way that its temporal dynamics, as expressed by the growth rates $\left(\frac{y_t^* - y_{t-1}^*}{y_{t-1}^*} \right)$, $t = 2, \dots, n$, be ‘as close as possible’ to the temporal dynamics of the preliminary series, where the ‘distance’ from the preliminary growth rates $\left(\frac{p_t - p_{t-1}}{p_{t-1}} \right)$ is given by the sum of the squared differences.

Two observations in order. Looking at the criterion to be minimized in (2), it clearly appears that it is grounded on an “ideal” movement preservation principle, “formulated as an explicit preservation of the period-to-period rate of change” of the preliminary series (Bloem *et al.*, 2001, p. 100). Second, the constrained minimization of the objective function (2) is nonlinear in the target values y_t , has not linear first-order conditions for a stationary point, and thus an explicit, analytic expression for the solution cannot be found. A solution can only be found using nonlinear optimization algorithms. We think that this complication has made the GRP benchmarking approach less appealing for data-producing agencies, and limited its use in practical applications.

Denton (1971) proposed a benchmarking procedure grounded on the *Proportionate First Differences* (PFD) between the target and the preliminary series¹:

$$\min_{y_t} \sum_{t=2}^n \left(\frac{y_t}{p_t} - \frac{y_{t-1}}{p_{t-1}} \right)^2. \quad (3)$$

Since the PFD criterion (3) is a quadratic function of the target values, the benchmarked series \mathbf{y}^* can be expressed in closed form and its values found with simple matrix operations (Dagum and Cholette, 2006; Di Fonzo and Marini, 2010).

¹Criterion (3) is expressed according to Cholette (1984), who modified the PFD criterion proposed by Denton, in order to correctly deal with the starting conditions of the problem.

In the literature (Cholette, 1984, Bloem *et al.*, 2001, Dagum and Cholette, 2006) it is often claimed that the PFD procedure produces results very close to the GRP benchmarking. Indeed, the GRP criterion (2) and the PFD criterion (3) are very close to each other. After a bit of algebra, we can write (U.S. Bureau of the Census, 2009, p. 96):

$$\sum_{t=2}^n \left(\frac{y_t}{p_t} - \frac{y_{t-1}}{p_{t-1}} \right)^2 = \sum_{t=2}^n \left[\frac{y_{t-1}}{p_t} \left(\frac{y_t}{y_{t-1}} - \frac{p_t}{p_{t-1}} \right) \right]^2. \quad (4)$$

The term in parentheses on the right-hand side is the difference between the growth rates of the target series and those of the preliminary series, namely the addendum of (2). In the PFD criterion these terms are weighted by the ratio between the target series at $t-1$ and the preliminary series at t . When these ratios are relatively stable over time, which is the case when the ‘benchmark-to-indicator ratio’ $\frac{Y_T}{\sum_{t \in T} p_t}$, $T = 1, \dots, N$ (Bloem *et al.*, 2001), is a smooth series, PFD and GRP are very close to each other. On the contrary, when the ratios $\frac{y_{t-1}}{p_t}$ behave differently, each term in the summation is over-(under-)weighted according to the specific relationship between target and preliminary series in that period. For example, sudden breaks in the movements of y_{t-1}/p_t might arise in case of large differences between the annual benchmarks and the annually aggregated preliminary series. Di Fonzo and Marini (2010) verified empirically this relationship, showing that PFD and GRP benchmarked estimates are close when the variability of the preliminary series and/or its bias are low with respect to the target variable. When this is not the case (e.g. preliminary series with large growth rates and/or bias), the GRP and PFD results diverge.

Di Fonzo and Marini (2010, 2011b) showed that a GRP-benchmarking problem can be solved very efficiently by using Newton’s optimization methods, which exploit the analytical expressions of the gradient and the Hessian of the objective function. First, the class of methods known as interior-point (IP) methods (also referred to as barrier methods, Nocedal and Wright, 2006), which has proved to be fast and accurate for many nonlinear constrained optimization problems, has been considered. Second, a Newton’s method with Hessian modification applied to a suitably reduced-unconstrained problem has been developed.

For reconciliation problems, both (2) and (3) can be extended to consider all the m series in the system. The global GRP criterion is defined as

$$\min_{y_{j,t}} \sum_{j=1}^m \sum_{t=2}^n \left(\frac{y_{j,t}}{y_{j,t-1}} - \frac{p_{j,t}}{p_{j,t-1}} \right)^2, \quad (5)$$

whereas the global PFD function is

$$\min_{y_{j,t}} \sum_{j=1}^m \sum_{t=2}^n \left(\frac{y_{j,t}}{p_{j,t}} - \frac{y_{j,t-1}}{p_{j,t-1}} \right)^2. \quad (6)$$

Since there is no cross-sectional interaction between different series in (5) and

(6), the relationship between the global criteria follows straightforward from (4):

$$\sum_{j=1}^m \sum_{t=2}^n \left(\frac{y_{j,t}}{p_{j,t}} - \frac{y_{j,t-1}}{p_{j,t-1}} \right)^2 = \sum_{j=1}^m \sum_{t=2}^n \left[\frac{y_{j,t-1}}{p_{j,t}} \left(\frac{y_{j,t}}{y_{j,t-1}} - \frac{p_{j,t}}{p_{j,t-1}} \right) \right]^2. \quad (7)$$

A reconciliation procedure based on the global PFD criterion (6) was presented by Di Fonzo and Marini (2011a). Likewise the benchmarked estimates, the reconciled estimates can be derived as (part of) the solution of a linear system. Compared to the benchmarking case, however, there are two major complications: (i) the size of the system matrix tends to be large, increasing exponentially with the number of variables and the number of periods, and (ii) the presence of constraints having different nature causes rank deficiency of the system matrix. Using very efficient sparse algorithms available in MATLAB[®], Di Fonzo and Marini (2011a) showed that a simultaneous adjustment of all the variables in the system is still feasible even when the system is very large. Two-step reconciliation procedures were also considered, because they are computationally less demanding if sparse matrices facilities are not available (we will consider them again in Section 4).

To our knowledge, a reconciliation procedure minimizing the global GRP criterion (5) has never been attempted. The two complications mentioned before – large size and rank deficiency of the system matrix, are likely to make even more difficult, if not impossible, the application of nonlinear optimization algorithms. However, the Newton’s algorithms considered for the GRP benchmarking case have proved to be very efficient and robust for a single series and look promising for dealing with many variables and many constraints at the same time. Similarly to our previous works on benchmarking, our main interests in this paper are:

- to verify if reconciliation problems (possibly large and complex) based on the minimization of the global GRP criterion (5) can effectively be solved using Newton’s methods exploiting second-order information;
- once an effective GRP reconciliation procedure is developed, to determine how close is the presumed approximation of the PFD reconciled estimates using practical applications.

3 Optimization algorithms for the GRP problem

The GRP criterion considered in the minimization problem (2) is a nonlinear function of the target values. More precisely, it can be shown that it is a non-convex function (Di Fonzo and Marini, 2011b). Differently from the PFD case, the constrained minimization problem based on the GRP function does not have linear first-order conditions for a stationary point, and therefore it is not possible to find an explicit, analytic expression for the solution. On the other hand, provided that both p_t and y_t , $t = 1, \dots, n - 1$, be different from zero, the GRP criterion is a twice continuously differentiable function, making it possible the use of several iterative minimization algorithms (Nocedal and Wright, 2006).

The first implementation of an algorithm for the GRP benchmarking problem was given by Causey and Trager (1981), who used a constrained steepest descent algorithm based on first-derivates information (i.e. the gradient) of the GRP function. The minimization problem is solved in the original variables y_t by using a *feasible direction method*, according to which at each iteration the unconstrained search direction is projected onto the feasible set of solutions defined by the constraints of the problem. The Causey and Trager's procedure is implemented in the DOS-executable programme **BMK1**, which has been used extensively by the US Census Bureau. Brown (2010, 2012) has recently proposed a similar feasible direction method based on the conjugate gradient algorithm and implemented in SAS. However, according to the results this implementation offers little improvements over the original procedure by Causey and Trager.

As a matter of fact, gradient-based algorithms may result in poorly efficient procedures, characterized by slow convergence and possible troubles in finding actual minima of the objective function. Improvements in both efficiency and robustness may be obtained by considering second-order information from the objective function, i.e. the Hessian matrix.

Di Fonzo and Marini (2011b) developed an efficient Newton's method with Hessian modification to solve the GRP benchmarking problem. The algorithm consists of the following steps. First, the analytical expression of the Hessian of the GRP function is derived; second, the original constrained (benchmarking) problem is transformed into an equivalent unconstrained problem; third, a Newton's method that allows a modification of the Hessian in order to preserve positive definiteness is applied. The Newton's method was compared with different gradient-based procedures (including the feasible steepest descent algorithm implemented in **BMK1**) through a benchmarking exercise of hundreds of series. In all the cases considered, the Newton's method significantly outperformed gradient-based methods in terms of both accuracy of the solution and convergence rates.

To our knowledge, the simultaneous reconciliation of time series subject to both temporal and contemporaneous constraints according to the global GRP criterion has never been considered in the literature. Due to the large and sparse nature of the constrained optimization problem (5), we only consider Newton's methods exploiting second-order derivative information, which are by far the most powerful algorithms for these problems.

Let us denote the global GRP function in (5) with $f(\mathbf{y})$. The gradient is the $(mn \times 1)$ vector²

$$\nabla f(\mathbf{y}) = \mathbf{g}(\mathbf{y}) = \{g_i\}_{i=1}^{mn},$$

while the Hessian matrix is defined as

$$\nabla^2 f(\mathbf{y}) = \mathbf{H}(\mathbf{y}),$$

with elements

$$h_{rs} = \frac{\partial^2 f(\mathbf{y})}{\partial y_r \partial y_s} = \frac{\partial g_r}{\partial y_s}, \quad r, s = 1, \dots, mn.$$

²To simplify, we use here a single-index notation for the temporal and cross-sectional dimensions.

The derivation of $\mathbf{g}(\mathbf{y})$ and $\mathbf{H}(\mathbf{y})$ (see Appendix B) is a straightforward extension of the results for the univariate GRP criterion shown by Di Fonzo and Marini (2011b). It is worth noting here that the Hessian is a sparse symmetric matrix with a tri-diagonal structure. We will note in the application section that the sparsity pattern of the Hessian is a key requirement for solving the global GRP problem (5).

We consider the two following optimization algorithms:

- an interior-point method, which is a powerful algorithm for solving large-scale nonlinear constrained problems (Nocedal and Wright, 2006);
- a Newton’s method with Hessian modification applied to a suitably reduced unconstrained transformation of the original constrained reconciliation problem, similar to the procedure developed for the benchmarking problem by Di Fonzo and Marini (2011b).

Both the interior-point method and the Newton’s method used in this work are line-search algorithms. They minimize transformations of the original GRP function

$$\min_{\mathbf{y}} \tilde{f}(\mathbf{y}), \quad (8)$$

which incorporate the constraints, following the same line-search approach:

1. Specify some initial guess of the solution: \mathbf{y}_0 ;
2. Perform an optimality test on \mathbf{y}_0 . If optimal, then stop. Otherwise, for $k = 1, 2, \dots$
 - (a) determine a *search direction* \mathbf{d}_k ;
 - (b) determine a *step length* α_k that leads to an improved estimate of the solution

$$\mathbf{y}_{k+1} = \mathbf{y}_k + \alpha_k \mathbf{d}_k. \quad (9)$$

It is typically required that the search direction \mathbf{d}_k be a *descent direction* for the function \tilde{f} at the point \mathbf{y}_k . This means that for “small” steps taken along \mathbf{d}_k , the function value (or the *merit* function value, in the case of constrained minimization) is guaranteed to decrease:

$$\tilde{f}(\mathbf{y}_k + \alpha \mathbf{d}_k) < \tilde{f}(\mathbf{y}_k) \quad \text{for} \quad 0 < \alpha \leq \epsilon$$

for some $\epsilon > 0$. Line-search algorithms differ by three major steps - the optimality test, computation of \mathbf{d}_k , and computation of α_k . The two methods, briefly described below, differ also in that they minimize different transformations of the original GRP function.

3.1 An interior-point method for the constrained GRP problem

Interior-point (IP) methods (also referred to as barrier methods) are a powerful class of algorithms to solve linear and nonlinear optimization problems. Early IP methods were originally developed in the 1960’s to solve nonlinear programming problems.

However, these methods were soon abandoned due to computational difficulties. IP methods were rediscovered after the appearance of Karmarkar's seminal paper (1984). Karmarkar proposed a new polynomial algorithm for linear programming, based on the idea that iterates should be calculated not on the boundary, but in the interior of the feasible region. Its application to linear programming programs showed huge increase of performance compared to the Dantzig's simplex method, which was the standard algorithm for linear programming at the time. Karmarkar's work started an explosion in research activity on IP methods in both linear and nonlinear programming applications (for more details on the development of interior-point methods, see Lesaja, 2009).

IP methods are nowadays considered the most powerful algorithms for large-scale nonlinear programming. At each step of IP methods, it is necessary to solve linear systems that are usually sparse to some extent. Recent progress in sparse matrix algorithms makes it possible to perform this step efficiently and accurately, as shown by Di Fonzo and Marini (2011a) for the simultaneous reconciliation according to the global PFD criterion. In a sense, IP methods represent the most advanced methods for solving large and sparse reconciliation problems based on the global GRP nonlinear function.

In this work we use the IP method available in the `fmincon` function of the Matlab[®] Optimization Toolbox³. This implementation is based on the algorithm proposed by Byrd *et al.* (1999, 2000). A description of the IP algorithm in `fmincon` is given by Waltz *et al.* (2006)⁴.

The IP approach to constrained minimization is to solve a sequence of approximate minimization problems that are easier to solve than the original constrained problem (which may also contain inequalities). These approximate problems are called barrier models. To solve each barrier problem, the algorithm uses one of two main types of steps at each iteration:

- a Newton step, attempting to solve the linear system defined by the Karush-Kuhn-Tucker (KKT) conditions⁵ for the approximate problem using direct linear algebra. In particular, the algorithm makes an LDL factorization of the system matrix;
- a conjugate gradient step, minimizing a quadratic approximation to the barrier problem using a trust region.

By default, the algorithm first attempts to take a direct step. One result from the LDL factorization is a determination of whether the projected Hessian is positive definite or not. If not, the algorithm attempts the conjugate gradient step. This mechanism of switching between a line search step and a trust region step guarantees to make progress toward feasibility and optimality. We note, however, that the

³The IP method is included as an option of `fmincon` from version 4.0 (R2008a) Optimization Toolbox.

⁴This work actually describes the implementation of the interior-point method in the KNITRO software (Ziena Optimization, 2011), which is similar to the one available in `fmincon` in every aspect.

⁵The KKT conditions are the first-order necessary conditions for a local optimum in constrained problems (Nocedal and Wright, 2006).

algorithm never switches to the conjugate gradient step in our applications. For this reason, the IP method applied here is a pure line-search algorithm.

One appealing feature of IP methods is that they are “infeasible” algorithms; that is, it is not required that the method start from a point in the feasible region. This feature is quite relevant in solving reconciliation problems, because it allows the user to start the algorithm by using the preliminary series $p_{j,t}$. In contrast, feasible methods, like the Newton’s method presented below, need to start from a feasible point and then require a set of reconciled series as input. Given that deriving (preliminary) reconciled estimates to start the algorithm may be an additional complication for practitioners, we view this feature as a clear advantage of the IP method.

3.2 A Newton’s method for the unconstrained GRP problem

In this section we describe the main steps of the Newton’s method with Hessian modification proposed in Di Fonzo and Marini (2011b) and how we have extended it to solve reconciliation problems. As shown in our applications, we note that the presence of both temporal and contemporaneous constraints – especially when the number of variables to be reconciled is very large, may make the procedure less accurate, considerably slower, or even unfeasible.

The main steps of the Newton’s method for benchmarking proposed by Di Fonzo and Marini (2011b) are outlined below:

1. the linear constraints of the system (to preserve temporal aggregation) are eliminated using a procedure based on standard linear algebra operations; the original constrained problem is transformed into an unconstrained one. The transformation is based on the null-space of the $(N \times n)$ temporal aggregation matrix \mathbf{C} (see Appendix A), which permits to eliminate N variables. The problem is then expressed on the basis of $n - N$ ‘free’ variables \mathbf{y}_Z . It is shown that a null-space matrix \mathbf{Z} can be found by a QR factorization of matrix \mathbf{C}^T . However, the pattern of the constraint matrix \mathbf{C} for the benchmarking problem permits to compute the QR factorization in compact form and once for all. No computational effort is therefore required to perform the QR factorization.
2. the optimization problem with equality constraints is transformed into an equivalent unconstrained problem by incorporating the constraints into the objective function. The reduced problem (variables, gradient, Hessian) is derived by transforming the original problem by means of the null-basis matrix \mathbf{Z} . The reduced problem is a function of the unrestricted variables \mathbf{y}_Z only.
3. a Newton’s method is applied to calculate the search direction \mathbf{d}_k , using a modified Cholesky factorization of the Hessian in order to have a positive definite matrix in the Newton equations. Positive definiteness of the Hessian is required to obtain a descent direction. Then, a step length α_k satisfying the Wolfe conditions⁶ is calculated. Iterations stop if $\|\nabla \tilde{f}(\mathbf{y}_{Z,k})\|_1 < \epsilon$.

⁶The Wolfe conditions define a set of sufficient conditions for α_k in order to have adequate reduction in the objective function at a minimal cost (Nocedal and Wright, 2006).

4. the solution of the original constrained problem \mathbf{y}^* can be derived by transforming back the solution of the reduced problem \mathbf{y}_Z^* through \mathbf{Z} .

Reconciliation problems are minimization problems with equality constraints, then the same procedure potentially applies. Yet, they are generally large, involving hundreds of variables at a time, and include contemporaneous constraints of different nature in addition to those preserving temporal annual consistency. Due to this additional complexity, we have introduced two changes in the method in order to make the algorithm computable. As explained in section 5, despite these changes we have not been able to apply this procedure to one of the two reconciliation problems faced in this work.

The first change introduced concerns the calculation of the null-space matrix \mathbf{Z} . Due to the addition of contemporaneous constraints, we have not been able to find an ad-hoc expression for \mathbf{Z} as we did for the benchmarking problem. For we have to recur to a QR factorization of \mathbf{A}^T , as suggested by Nocedal and Wright (2006, p. 433) and discussed in Di Fonzo and Marini (2011b). We use the multifrontal sparse QR factorization method implemented in the `spqr` function of the SuiteSparseQR toolbox for Matlab (Davis, 2011). This function has proved to be efficient and reliable in the calculation of a null-space matrix \mathbf{Z} , exploiting the sparsity structure of \mathbf{A}^T . Matrix \mathbf{Z} , however, is no longer sparse as it was in the benchmarking case; it has instead a full lower-triangular form. This complicates matters a lot because \mathbf{Z} is used to pre- and post-multiply the original Hessian for the constrained problem, which is a large and sparse tri-diagonal matrix, to derive the Hessian for the unrestricted problem, which becomes a full matrix with no structure to be exploited through sparse algorithms⁷.

The second change is related to the Hessian modification in the calculation of the step direction. For this step we use the Matlab function `minFunc` (Schmidt, 2006), which is a free analogous of the function `fminunc` of the Optimization Toolbox of Matlab (The MathWorks, 2009). Different options are available for the Hessian modification. For benchmarking problems, a modified Cholesky factorization was found to work very well. Evidently, the reduced (sparse) Hessian (or its transformation) is always positive definite through the iterations. On the other hand, it was not possible to apply the Cholesky factorization to the reduced Hessian of a reconciliation problem. We needed to switch to another `minFunc` option available for symmetric indefinite matrices, based on an LDL factorization. We noted that, given that the reduced Hessian is no longer sparse, the LDL factorization takes much more time when the system is large, or even causes out of memory error.

Differently from the IP method, the Newton's method is a feasible algorithm. Therefore, it requires a feasible point to begin the iterations. For the benchmarking problem, we use the Denton PFD benchmarked series as this should be close to the GRP optimum in most cases. Consistently, we use the simultaneous Denton PFD reconciled series as the starting point for reconciliation problems. This procedure, however, may itself take a while and necessitate of advanced sparse algorithms to be

⁷Boyd and Vandenberghe (2004, p. 143) note that when the system is large eliminating the equality constraints may destroy sparsity of some other useful structure of the problem. This seems to be quite the case for the problem in hand.

solved, as discussed by Di Fonzo and Marini (2011a). An approximation of the simultaneous PFD solution could be derived using two-step reconciliation procedures, which divide and solve the problem in smaller pieces and require less computational time and resources than a simultaneous solution. Two-step procedures are described in the following section.

4 GRP in Two-Step Reconciliation Procedures

When the system of time series is very large, a simultaneous solution can be operationally difficult to apply, mostly if the practitioner either does not intend to or cannot use sparse matrices computation facilities. Simplified solutions are however possible, based on a generalization of the two-step approach proposed by Quenneville and Rancourt (2005) for restoring the additivity of a system of SA time series such that their sum is in line with directly-derived SA totals: firstly, a univariate benchmarking procedure (e.g., the modified Denton PFD benchmarking procedure or the more general regression based benchmarking procedure by Cholette and Dagum, 1994) is used to restore the temporal additivity of every SA series; in the second step, the SA component series are reconciled *one year at a time* using a least squares balancing procedure.

Di Fonzo and Marini (2011a) discussed and applied to real-life systems of time series two-step procedures based on the PFD criterion, and compared them with the simultaneous PFD reconciliation solution based on (6). In the first step, the modified Denton PFD benchmarking technique is used. In the second step, instead, two alternative least-squares adjustments of the benchmarked estimates are employed to reconcile them with the contemporaneous constraints of the system one year at a time. Denoting with $y_{j,t}$ the reconciled series and with $x_{j,t}^{PFD}$ the benchmarked series obtained with the modified Denton PFD method, the two adjustments are based on the following objective criteria:

$$F_T^{PFD-BB} = \sum_{j=1}^m \sum_{t=(T-1)s+1}^{Ts} \frac{\left(y_{j,t} - x_{j,t}^{PFD}\right)^2}{|x_{j,t}^{PFD}|} \quad (10)$$

$$F_T^{PFD-ST} = \sum_{j=1}^m \sum_{t=(T-1)s+1}^{Ts} \left(\frac{y_{j,t} - x_{j,t}^{PFD}}{x_{j,t}^{PFD}}\right)^2 \quad (11)$$

where suffix T , $T = 1, \dots, N$, denotes that the optimization is performed for each low-frequency period separately. Criterion F_T^{PFD-BB} is an adaptation of the original Quenneville and Rancourt's procedure, presenting the absolute value of $x_{j,t}^{PFD}$ at the denominator to allow for the adjustment of possibly negative numbers (Beaulieu and Bartelsman 2004); criterion F_T^{PFD-ST} assumes the squared temporal benchmarked series as normalizing factor, as proposed by Round (2003) and Stuckey *et al.* (2004).

The comparison showed that a two-step procedure with the modified Denton PFD at the first step and a least-squares adjustment based on criterion (11) at the second step is a close approximation of the simultaneous PFD solution. Instead, when criterion (10) is used in lieu of (11) the two-step procedure tends to

over-adjust the dynamic profile of smaller series and under-adjust the movements of larger series. Di Fonzo and Marini (2011a) explain this different behavior in terms of the reliability coefficients implied by the two criteria F_T^{PFD-BB} and F_T^{PFD-ST} . Assuming the coefficient of variability (CV , the ratio between standard deviation and mean) as a reliability indicator⁸, the choice of $|x_{j,t}^{PFD}|$ as denominator in expression (10) implicitly involves $CV_{j,t} = 1/\sqrt{|x_{j,t}^{PFD}|}$, that is: (i) different reliabilities for all variables are considered in the least-squares adjustment, and (ii) large variables are considered relatively more reliable, and thus they are touched relatively less by the reconciliation procedure than small variables. Conversely, the coefficient of variations (i.e. relative reliabilities) are equal for all variables and all times with criterion F_T^{PFD-ST} .

As shown by Di Fonzo and Marini (2011a), the choice of the criterion on which the second step is grounded may have a clear impact on the temporal dynamics of the reconciled series, the adjustments to the preliminary growth rates due to criterion (10) being generally larger than those produced by criterion (11).

However, in our previous work we considered only the modified Denton PFD technique at the first step. We did not investigate other options because our main concern was to compare two-step procedures with the simultaneous PFD reconciliation procedure. Now we are instead interested in two-step procedures having the GRP benchmarking technique at the first step, and a least-squares adjustment at the second step alternatively based on the following two criteria:

$$F_T^{GRP-BB} = \sum_{j=1}^m \sum_{t=(T-1)s+1}^{Ts} \frac{\left(y_{j,t} - x_{j,t}^{GRP}\right)^2}{|x_{j,t}^{GRP}|} \quad (12)$$

$$F_T^{GRP-ST} = \sum_{j=1}^m \sum_{t=(T-1)s+1}^{Ts} \left(\frac{y_{j,t} - x_{j,t}^{GRP}}{x_{j,t}^{GRP}}\right)^2. \quad (13)$$

As for the PFD case, we wish (i) to assess how much alternative choices of the criterion adopted in the second step do affect the temporal dynamics of the reconciled series as compared to the preliminary ones, and (ii) to have empirical confirmation to the expectation that the two-step procedure based on the GRP benchmarking method at the first step and a least-squares adjustment based on (13) at the second step is a close approximation to the simultaneous GRP reconciliation procedure.

5 Applications

In this section we consider the reconciliation of two systems of time series: (i) the EUQSA system, 175 quarterly variables of the European Union's quarterly national accounts (28 quarters) to be reconciled with known annual totals and 30 accounting constraints, and (ii) the MRTS system, 236 monthly series of Canadian seasonally

⁸Higher CVs commonly signal variables of comparatively worse quality: see, for example, Chen (2006), Danilov and Magnus (2008).

adjusted retail trade by provinces (156 months) to be reconciled with annual unadjusted totals and 32 contemporaneous constraints for geographical aggregations (for more details on the two datasets, see Di Fonzo and Marini, 2011a).

Aggregate statistics on the discrepancy in the two systems are presented in table 1. The table presents averages of (in order) mean, absolute mean, standard deviation, minimum, maximum, and range of the discrepancy for both temporal and contemporaneous constraints. The average discrepancy is expressed in % of the benchmark values⁹. It can be noted that EUQSA presents much higher discrepancies than MRTS in both temporal and contemporaneous constraints. Discrepancies in EUQSA are mostly generated by the fact that quarterly preliminary figures are available for just a few member states and not all the countries of the European Union (*data-driven* discrepancy); on the contrary, MRTS discrepancies come from the direct application of seasonal adjustment to each component in the system (*procedure-driven* discrepancy).

Table 1: Average statistics on discrepancy (in % of benchmark values) in the EUQSA and MRTS systems

	n. of series/ constraints	mean	abs. mean	st. dev.	min	max	range
EUQSA							
Temporal	65*	0.6	2.6	2.6	-2.8	4.9	7.7
Contemporaneous	30	-25.3	38.0	68.5	-304.2	57.0	361.2
MRTS							
Temporal	236	-0.1	0.5	0.6	-1.2	1.1	2.3
Contemporaneous	32	-0.1	0.7	0.9	-2.7	2.8	5.5

* Only 65 out of 175 of EUQSA variables present temporal discrepancies.

We extend the comparison in Di Fonzo and Marini (2011a) by including reconciliation procedures based on the GRP criterion. We apply two simultaneous procedures

- Sim GRP, through the application of the Hessian-based procedures presented in section 3 to minimize the global criterion (5);
- Sim PFD, through the direct solution of the linear system coming from the constrained minimization of the global criterion (6);

and four two-step procedures

- PFD-BB: the modified Denton PFD benchmarking is applied at the first step, and a least-squares adjustment based on (10) at the second step;
- PFD-ST: the modified Denton PFD benchmarking is applied at the first step, and a least-squares adjustment based on (11) at the second step;
- GRP-BB: the Newton's GRP benchmarking is applied at the first step, and a least-squares adjustment based on (12) at the second step;

⁹When the preliminary data presents a non-zero mean difference with the annual series, it is standardized to the overall level of the annual series according to the bias correction procedure described in Quenneville *et al.* (2009).

- GRP-ST: the Newton's GRP benchmarking is applied at the first step, and a least-squares adjustment based on (13) at the second step.

In order to assess the performance of the procedures, for each series we calculate the Mean Absolute Adjustment (*MAA*) to the percentage growth rates, that is

$$MAA_j = 100 \times \frac{1}{n-1} \sum_{t=2}^n |r_{j,t}^y - r_{j,t}^p|, \quad j = 1, \dots, m$$

where $r_{j,t}^y = (y_{j,t} - y_{j,t-1}) / y_{j,t-1}$ and $r_{j,t}^p = (p_{j,t} - p_{j,t-1}) / p_{j,t-1}$ are the growth rates of the reconciled and preliminary series, respectively. Overall indices for the whole system of time series are calculated accordingly. In our previous works, we used the Root Mean Squared Adjustment (*RMSA*) statistic to assess the adjustment. *RMSA* is, however, the square root of the GRP criterion, which is minimized by construction by GRP-based procedures. For this reason, in this paper we prefer the use of a more neutral statistic of movement preservation to evaluate and compare the performance of the procedures.

Table 2 shows summary statistics on the *MAA* values for the two systems. Sim GRP outperforms the other procedures in both systems, with the lowest mean, median and standard deviation. Sim GRP is also the procedure with the highest number of series with minimum *MAA* value (41.1% and 29.2%).

As expected, we note that Sim PFD is a good approximation of Sim GRP. The median *MAA* for EUQSA is 0.421% for Sim PFD and 0.378% for Sim GRP; on MRTS we find 0.487% for Sim PFD *vs.* 0.456% for Sim GRP. Nevertheless, using Sim PFD we notice a higher standard deviation of *MAA* (2.367% *vs.* 1.464%) and a maximum equal to 24.455%, much larger than 6.366% of Sim GRP (we come back on this difference below).

Concerning the two-step procedures, we note that using a GRP benchmarking at the first step slightly improves the results upon using the modified Denton PFD benchmarking procedure. However, the choice of the type of adjustment at the second step is much more important: whatever the first step is, using ST instead of BB guarantees an overall smaller adjustment in terms of *MAA*.

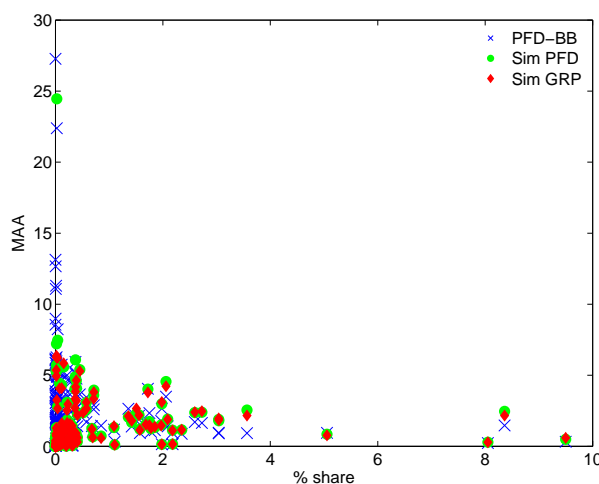
Looking at the two-step procedures using ST at the second step (PFD-ST and GRP-ST), we note a similar difference in the maximum *MAA* value identified when comparing Sim PFD and Sim GRP (24.396% for PFD-ST and 6.915% for GRP-ST). It appears that using GRP makes the adjustment process more balanced across the variables in the system, either when GRP is applied at the first step (provided ST is used) or when GRP is applied as a simultaneous procedure. Similarly to the already known relationship between PFD-ST and Sim PFD, the GRP-ST reconciliation procedure turns out to be a close approximation of Sim GRP.

Our next step is to relate the *MAA* values to the size of the variables in the system. We restrict our attention to Sim GRP, Sim PFD, and the two-step procedure PFD-BB¹⁰. In Di Fonzo and Marini (2011a) it is shown that PFD-BB preserves more the movements in larger variables than in smaller ones (in terms of *RMSA* statistic),

¹⁰For each two-step procedure using ST at the second step, the obtained results are very close to those of the relevant simultaneous procedure.

Table 2: Summary statistics of *MAA* for the EUQSA and MRTS systems

	Two-step				Simultaneous	
	PFD-BB	GRP-BB	PFD-ST	GRP-ST	Sim PFD	Sim GRP
EUQSA						
Mean	3.0949	2.9787	1.2717	1.1097	1.2799	1.0883
Median	2.1178	2.1635	0.3874	0.3884	0.4212	0.3780
St. dev.	3.2645	2.9854	2.3604	1.4987	2.3671	1.4641
Max	27.2699	29.1887	24.3963	6.9151	24.4550	6.3662
% Min	8.6	10.9	5.7	20.6	13.1	41.1
MRTS						
Mean	1.5312	1.5315	0.7064	0.7030	0.7054	0.6997
Median	1.4222	1.4214	0.4708	0.4601	0.4874	0.4555
St. dev.	0.7417	0.7484	0.6988	0.7021	0.6957	0.6928
Max	5.2620	5.3978	4.3494	4.3495	4.3538	4.3065
% Min	10.2	9.3	5.9	20.8	24.6	29.2

**Figure 1:** EUQSA system. *MAA* by share (%) of variables for Sim GRP, Sim PFD and PFD-BB

while Sim PFD distributes the adjustment more uniformly across the variables. We are interested to see if there is a similar picture using the *MAA* statistic and, more importantly, how is the distribution of adjustment resulting from Sim GRP. Figure 1 shows a scatter plot for EUQSA between the size of the variables (x -axis), each expressed in percentage of their total sum, and the *MAA* values (y -axis) for the three procedures. The presence of 'x' along the y -axis, and those associated with smaller *MAA* values along the x -axis, confirm that PFD-BB tends to over-adjust smaller variables in favour of the larger ones. The isolated green dot for Sim PFD on the y -axis signals that the maximum adjustment from this procedure (24.455%) is made to a very small-size variable. It can be noticed that the adjustments from Sim GRP are more evenly distributed than Sim PFD.

In previous GRP-PFD comparisons on univariate benchmarking (Di Fonzo and Marini, 2010), GRP was found to outperform PFD when large discrepancies and/or

high variability in the preliminary series were present. It is thus interesting to look at the scatter plots of MAA vs. (i) the average temporal discrepancy (in % and absolute values, figure 2), and (ii) the coefficient of variation (in absolute value, figure 3), taken as a standardized measure of variability of the series. Moving along the x -axis we find the variables with higher temporal discrepancy and larger variability, respectively. In figure 2 we notice that the maximum adjustment from Sim PFD moves to the right-end side of the plot, which identifies that variable as the one presenting the highest (percentage) temporal discrepancy in the system. Furthermore, from figure 3 we note that the highest values of MAA for PFD-BB (i.e. the highest 'x's in the scatter plot) are achieved for variables presenting medium-high variability in the system. These distributional aspects of the adjustment in the two systems confirm that, even in a reconciliation process of a system of time series, PFD deviates from GRP when discrepancies are large and the preliminary series is volatile.

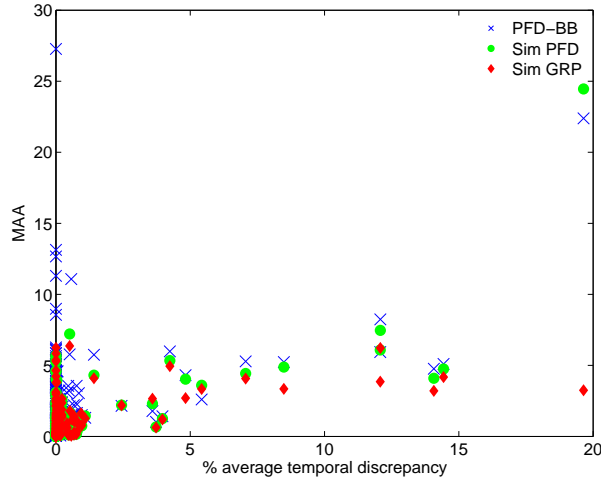


Figure 2: EUQSA system. MAA by average temporal discrepancy (in absolute value) for Sim GRP, Sim PFD and PFD-BB

Finally, table 3 presents an indication on the computational time needed to complete the reconciliation process for each procedure. Times are expressed in seconds and are derived as average from five sequential executions of the same reconciliation process. The simultaneous GRP solution (with the interior-point algorithm) takes less than one second to reconcile EUQSA and about 18 seconds for MRTS. The Newton's method for the unconstrained problem solves the EUQSA problem in about 50 seconds (reaching the same optimal point of the interior-point method), but it fails to converge for MRTS due to an out of memory error generated by the lack of sparsity of the reduced Hessian matrix (see section 3). The interior-point method has proved to be much faster and reliable than the Newton's method for reconciliation problems. Nevertheless, noting the difference in the computational times between EUQSA and MRTS we may expect that even the interior-point method could become an unfeasible procedure when the system dimension is very large. In

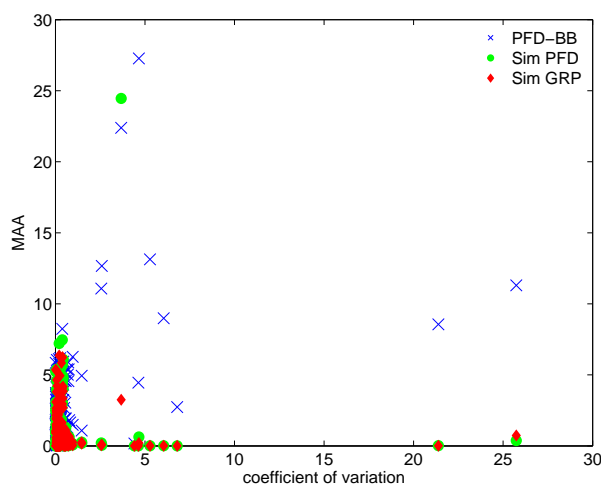


Figure 3: EUQSA system. *MAA* by coefficient of variation (in absolute value) of variables for Sim GRP, Sim PFD and PFD-BB

that case, the two-step procedure GRP-ST would represent a valuable alternative to Sim GRP.

Table 3: Average computational times (in seconds) to perform the reconciliation process*

Method	EUQSA	MRTS
Sim GRP**	0.8	18.0
Sim PFD	0.1	2.0
GRP-BB/ST	1.3	3.5
PFD-BB/ST	0.1	1.5

* Using processor Intel Core i5-2430M @2.40 GHz, Ram 6.0GB, Windows 7 OS, Matlab R2008b. ** Using IP algorithm. Sim GRP with the Newton's method with Hessian modification takes about 50 seconds to solve EUQSA, while it does not achieve a solution for MRTS.

6 Conclusions

In this paper we have proposed simultaneous and two-step reconciliation procedures based on the GRP movement principle, which is widely recognized as the most natural choice for preserving the movements in an economic series. To solve the nonlinear GRP problem we have made recourse to Newton's optimization methods that exploit the full derivative information of the problem, namely the gradient and the Hessian. Using two systems of economic series with both temporal and cross-sectional constraints, we have shown that these procedures are accurate, feasible, and time-efficient in finding an optimal solution of the GRP problem.

This work has largely benefited from our previous findings on benchmarking and reconciliation. First, we had already solved efficiently the nonlinear GRP problem

for benchmarking (i.e. one variable at a time) using the same Newton's optimization methods proposed in this paper. Second, we knew that only by preserving and exploiting the sparsity structure of the matrices involved would allow us to solve a reconciliation problem with many variables and many observations simultaneously, namely with all constraints in the system (temporal and contemporaneous) considered at the same time. Finally, we were aware that a satisfactory approximation of the optimal simultaneous solution could be obtained through a more convenient and simplified two-step procedure.

The purpose of this paper was twofold. First and foremost, we wanted to prove that the nonlinear GRP reconciliation problem could be solved accurately and efficiently using Newton's optimization algorithms. Then, we wanted to compare the GRP solution with the Denton PFD solution to verify, as often claimed, that PFD is a close approximation of GRP. We believe that both objectives have been successfully achieved.

On the implementation aspects, we derived the analytical expressions of the gradient and Hessian of the global GRP function. These expressions are necessary to feed any Newton's optimization methods, and to our knowledge they have never been derived before. Next, we used two different (but related) algorithms to solve the GRP reconciliation problem: an interior-point algorithm applied to the original constrained problem, and a Newton's method with Hessian modification applied to a suitably reduced unconstrained problem. The interior-point algorithm turned out to be very fast, accurate and robust, solving both the problems faced in this paper. The only drawback of the interior-point algorithm used is that it is available through a commercial software, i.e. Matlab, which is not affordable for many potential "customers" of reconciliation procedures (in general public data-producing agencies). Being the interior-point an algorithm very difficult to replicate, we decided to develop an alternative Newton's method for the unconstrained system, whose steps are fully described in our paper. With the Newton's method with Hessian modification, however, we were able to solve accurately only one of the two systems considered. For the largest system, the procedure stopped due to an out of memory error caused by multiplication of dense matrices generated by the transformation of the system into the unconstrained form (a problem on which we are actively working).

To assess the GRP and PFD results we compared the two simultaneous solutions along with four two-step procedures. The latter are derived from using alternatively the PFD and GRP benchmarking techniques at the first step, and from using the level or the squared level of the temporal benchmarked series as normalizing factor at the second step. The comparison, based on a metric different from both the GRP and PFD criteria, showed that the simultaneous GRP solution is always the best method at preserving the movements in the preliminary series. We found, in particular, that using the simultaneous GRP guarantees a more balanced adjustment process across the variables. In general, the simultaneous PFD solution was shown to be very close to the GRP; but for a few series in the system, the most volatile ones with large temporal discrepancies to be distributed, PFD resulted in markedly higher adjustments than GRP.

As regards the two-step procedures, the same issue of robustness noted above was noted from using the GRP or the PFD benchmarking techniques at the first step. In

general, choosing GRP at the first step guarantees a more balanced adjustment than PFD. However, the choice made at the second step counts much more in terms of quality of the adjustment, as already highlighted by Di Fonzo and Marini (2011a) for the PFD case. The simultaneous GRP solution is best approximated when the GRP is used at the first step, and the squared temporal benchmarked series is considered as a normalizing factor at the second step (i.e. the GRP-ST procedure). The use of the (absolute) level, instead, was found to penalize too much the smaller series in the system in favour of the larger ones.

Practitioners may want to know which are the implications of our findings on the GRP on the routine work they conduct in benchmarking and reconciliation. As we have shown in this paper, and in the companion paper Di Fonzo and Marini (2011a) on benchmarking, the nonlinear GRP problem can be solved accurately and efficiently through optimization algorithms exploiting the full derivative information of the problem. If one aims at preserving the growth rates of the variables in the system under adjustment, our simultaneous GRP solution is undoubtedly the most accurate and reliable approach. Nonetheless, we have also shown that very similar performance can be reached by using a two-step procedure where the GRP is more conveniently applied at the first step on each individual series of the system.

We note, however, that in the great majority of the cases the PFD solution is very close to the GRP one. Large differences in the adjustment arise only when the series are volatile and present large discrepancies with respect to their low-frequency benchmark values. When the series are smooth and discrepancies are consistently small across the system – two desirable conditions in any reconciliation problem –, simultaneous or two-step reconciliation procedures based on the Denton PFD principle can approximate very well the GRP results.

Other aspects of benchmarking and reconciliation merit further investigation and are left for future research. First, so far we have only dealt with systems with binding and linear constraints. Our solutions may be easily extended to include in the system soft (e.g., only positive variables) and nonlinear (e.g., ratios between variables) constraints (Danilov and Magnus, 2008), as well as users' information about the reliability of the preliminary figures (Bikker and Bujithenek, 2006, Quenneville and Fortier, 2012). Second, we are already looking at different criteria to preserve the yearly growth rates (period compared to the same period of the previous year) rather than the period-to-period change preserved by both the GRP and PFD criteria, which may be interesting especially for seasonal time series. Finally, benchmarking and reconciliation in extrapolation (i.e. for quarters with no annual/quarterly benchmarks yet available) are yet to be fully explored¹¹.

Appendix A. Matrix representation of the constraints

Let Y_T be the low-frequency measurement of a given phenomenon for $T = 1, \dots, N$. Let p_t be some high-frequency preliminary information related to the same phe-

¹¹Di Fonzo and Marini (2012) illustrates an enhancement for the Denton PFD benchmarking method in extrapolation, based on the projection of the unavailable annual benchmark-to-indicator ratio.

nomenon, observed for $t = 1, \dots, n$. The benchmarking problem looks for the high-frequency values of an unknown target variable y_t , for $t = 1, \dots, n$, which should be consistent with the benchmarks Y_T and show movements as close as possible to the movements in p_t ¹². Let s be the aggregation order (e.g., $s = 4$ for quarterly-to-annual aggregation, $s = 12$ for monthly-to-annual aggregation, $s = 3$ for monthly-to-quarterly aggregation), and let \mathbf{C} be a $(N \times n)$ temporal aggregation matrix, converting n high-frequency values in N low-frequency ones (we assume $n = s \cdot N$). If we denote with \mathbf{x} the $(n \times 1)$ vector of desired high-frequency values, and with \mathbf{Y} the $(N \times 1)$ vector of low-frequency values, the aggregation constraints of a benchmarking problem can be expressed as

$$\mathbf{C}\mathbf{x} = \mathbf{Y}. \quad (14)$$

Depending on the nature of the involved variables (e.g., flows, averages, stocks), the temporal aggregation matrix \mathbf{C} usually can be written as

$$\mathbf{C} = \mathbf{I}_N \otimes \mathbf{c}^T, \quad (15)$$

where the $(s \times 1)$ vector \mathbf{c} may assume one of the following forms:

1. flows: $\mathbf{c} = \mathbf{1}_s = (1 \ 1 \ \dots \ 1)^T$,
2. averages: $\mathbf{c} = \frac{1}{s}\mathbf{1}_s$,
3. stocks (end-of-the-period): $\mathbf{c} = (0 \ 0 \ \dots \ 1)^T$,
4. stocks (beginning-of-the-period): $\mathbf{c} = (1 \ 0 \ \dots \ 0)^T$.

In a reconciliation problem, the target high-frequency series are also required to be in line with known totals, or satisfy certain given equalities, in any observed periods at the same high-frequency level. Let us suppose there are m series to adjust. We introduce an additional sub-index j in our notation to identify each series in this system, so that the input data become the m n -dimensional vectors \mathbf{p}_j and the m N -dimensional vectors \mathbf{Y}_j , for $j = 1, \dots, m$. The target variables of a reconciliation process are denoted as \mathbf{y}_j .

The temporal constraints linking the high-frequency component series to their temporally aggregated counterparts in (14) can be extended to cover all the variables in the system. Denoting now $\mathbf{Y} = [\mathbf{Y}_1^T \mathbf{Y}_2^T \dots \mathbf{Y}_m^T]^T$, we have

$$(\mathbf{I}_m \otimes \mathbf{C})\mathbf{y} = \mathbf{Y}, \quad (16)$$

where \otimes is the Kronecker product and $\mathbf{y} = [\mathbf{y}_1^T \mathbf{y}_2^T \dots \mathbf{y}_m^T]^T$ denotes the $(mn \times 1)$ vector of all the unknown component series.

The cross-sectional constraints are less obvious to define and vary according to the nature of the relationships between the variables in the system. Let \mathbf{G} be a $(k \times m)$ matrix of known constants (often 0, 1 and -1) defining the (contemporaneous) accounting relationships between \mathbf{y}_j , k being the number of linear relationships to

¹²In general, we use capital letters to denote low-frequency series (and matrices as well) and small letters to indicate high-frequency series.

be fulfilled. Let \mathbf{z}_h , $h = 1, \dots, k$, be the $(n \times 1)$ vectors of high-frequency known quantities associated to the k accounting constraints in \mathbf{G} . The contemporaneous aggregation constraints can be written in compact form as

$$(\mathbf{G} \otimes \mathbf{I}_n) \mathbf{y} = \mathbf{z}, \quad (17)$$

where $\mathbf{z} = [\mathbf{z}_1^T \mathbf{z}_2^T \dots \mathbf{z}_k^T]^T$ has dimension $(kn \times 1)$. When \mathbf{G} is a $(1 \times m)$ row vector of constant values we have a single contemporaneous constraint. In addition, $\mathbf{G} = \mathbf{1}'_m$ when a simple summation constraint links the variables, $\mathbf{1}_m$ being a $(m \times 1)$ vector of one. Di Fonzo and Marini (2011a) show how the matrix \mathbf{G} can be derived using some examples.

Let \mathbf{A} be the $((kn + Nm) \times nm)$ aggregation matrix

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

and $\mathbf{b} = [\mathbf{z}' \mathbf{Y}']'$ the $((kn + Nm) \times 1)$ vector containing both high-frequency and low-frequency known values. The complete set of constraints between the unknown high-frequency component series and the available constraints can be expressed in compact form as

$$\mathbf{A} \mathbf{y} = \mathbf{b}. \quad (18)$$

Notice that the contemporaneous aggregation of temporally aggregated series implies

$$(\mathbf{G} \otimes \mathbf{I}_N) \mathbf{Y} = (\mathbf{I}_k \otimes \mathbf{C}) \bar{\mathbf{Z}}, \quad (19)$$

where¹³

$$(\mathbf{I}_k \otimes \mathbf{C}) \mathbf{z} = \bar{\mathbf{Z}} = [\mathbf{z}_1^T \mathbf{z}_2^T \dots \mathbf{z}_k^T]^T. \quad (20)$$

Relationship (19) reflects the fact that the temporally aggregated information has to be consistent with the system constraints. Thus, the low-frequency component series, when ‘longitudinally’ aggregated through matrix \mathbf{G} , must be equal to the series obtained by temporal aggregation of the high-frequency series in \mathbf{z} . In other words, we are assuming that \mathbf{Y} and $\bar{\mathbf{Z}}$ fulfill, respectively, all contemporaneous and temporal aggregation constraints. This point must be stressed, because it is a strong pre-requisite in order the reconciliation procedure may work.

It must be noted that the target series of the reconciliation problem \mathbf{y}_j , for $j = 1, \dots, m$, are different from the benchmarked series \mathbf{x}_j , which are only required to be consistent with the annual benchmarks \mathbf{Y}_j . Most likely the following inequality holds true after the application of independent univariate benchmarking processes:

$$(\mathbf{G} \otimes \mathbf{I}_n) \mathbf{x} \neq \mathbf{z}. \quad (21)$$

In summary, provided that $\mathbf{A} \mathbf{p} \neq \mathbf{b}$, a reconciliation problem looks for estimates of the high-frequency component series \mathbf{y} for which, while the temporal profile of the original preliminary series is preserved “at the best” (*movement preservation principle*), the aggregation constraints (18) must hold.

¹³We use the symbol $\bar{\mathbf{Z}}$ to avoid confusion with the null-space matrix \mathbf{Z} described in section 3.2.

Appendix B. Gradient and Hessian of the global GRP criterion

In this appendix we present the analytical expressions of the gradient vector and of the Hessian matrix for the global GRP criterion (5), which can be exploited by Newton-type nonlinear programming optimization procedures. The derivation is a straightforward extension of the expressions shown by Di Fonzo and Marini (2011b) for the univariate GRP criterion.

Let us denote with $y_{j,t}$ and $p_{j,t}$, respectively, the j^{th} target and preliminary series of the system observed in the period t , with $j = 1, \dots, m$, $t = 1, \dots, n$, where m is the number of variables and n the number of the high-frequency periods. The value $\{y_{j,t}\}$ is the $[i + (j - 1)n] - \text{th}$ element in the stacked nm -dimensional vector \mathbf{y} . The gradient vector of function (5) is the $(nm \times 1)$ vector

$$\nabla f(\mathbf{y}) = \mathbf{g}(\mathbf{y}) = \{g_i\}_{i=1}^{nm},$$

where

$$\begin{aligned} g_{1+(j-1)n} &= -2 \frac{y_{j,2}}{y_{j,1}^2} \left(\frac{y_{j,2}}{y_{j,1}} - \frac{p_{j,2}}{p_{j,1}} \right) \\ g_{t+(j-1)n} &= \frac{2}{y_{j,t-1}} \left(\frac{y_{j,t}}{y_{j,t-1}} - \frac{p_{j,t}}{p_{j,t-1}} \right) - 2 \frac{y_{j,t+1}}{y_{j,t}^2} \left(\frac{y_{j,t+1}}{y_{j,t}} - \frac{p_{j,t+1}}{p_{j,t}} \right) \quad t = 2, \dots, n-1 \\ g_{n+(j-1)n} &= \frac{2}{y_{j,n-1}} \left(\frac{y_{j,n}}{y_{j,n-1}} - \frac{p_{j,n}}{p_{j,n-1}} \right). \end{aligned}$$

for $j = 1, \dots, m$.

Let us denote the elements of the Hessian matrix, $\nabla^2 f(\mathbf{y}) = \mathbf{H}(\mathbf{y})$, as

$$h_{r,s} = \frac{\partial^2 f(\mathbf{y})}{\partial y_r \partial y_s} = \frac{\partial g_r}{\partial y_s}, \quad r, s = 1, \dots, nm.$$

Notice that the Hessian matrix is both symmetric and tri-diagonal, that is its non-zero items are $h_{s,s}$, $s = 1, \dots, nm$, $h_{s-1,s}$, $s = 2, \dots, nm$, and $h_{s+1,s}$, $s = 1, \dots, nm - 1$. After some calculations, we find:

$$\begin{aligned} h_{1+(j-1)n, 1+(j-1)n} &= 2 \frac{y_{j,2}}{y_{j,1}^3} \left(3 \frac{y_{j,2}}{y_{j,1}} - 2 \frac{p_{j,2}}{p_{j,1}} \right) \\ h_{i+(j-1)n, i+(j-1)n} &= \frac{2}{y_{j,t-1}^2} + 2 \frac{y_{j,t+1}}{y_{j,t}^3} \left(3 \frac{y_{j,t+1}}{y_{j,t}} - 2 \frac{p_{j,t+1}}{p_{j,t}} \right) \quad t = 2, \dots, n-1 \\ h_{n+(j-1)n, n+(j-1)n} &= \frac{2}{y_{j,n-1}^2} \\ h_{i+(j-1)n, k+(j-1)n} &= -\frac{2}{y_{j,i}^2} \left(2 \frac{y_{j,k}}{y_{j,i}} - \frac{p_{j,k}}{p_{j,i}} \right) \\ &\quad i = k + 1, k = 1, \dots, n-1 \\ &\quad \vee i = k - 1, k = 2, \dots, n, \end{aligned}$$

for $j = 1, \dots, m$.

References

- [1] Beaulieu, J.J., Bartelsman, E.J. (2004), “Integrating Expenditure and Income Data: What to Do with the Statistical Discrepancy?”, Tinbergen Institute Discussion Paper TI 2004-078/3. <http://www.tinbergen.nl/discussionpapers/04078.pdf>
- [2] Bikker, R., Buijtenhek, S. (2006), “Alignment of Quarterly Sector Accounts to annual data”, CBS Voorburg. http://www.cbs.nl/NR/rdonlyres/D918B487-45C7-4C3C-ACD0-0E1C86E6CAFA/0/Benchmarking_QSA.pdf.
- [3] Bloem, A., Dippelsman, R., Mæhle, N. (2001), *Quarterly National Accounts Manual. Concepts, Data Sources, and Compilation*, International Monetary Fund, Washington DC.
- [4] Boyd, S., Vandenberghe, L. (2004), *Convex Optimization*, Springer, New York.
- [5] Bozik, J.E., Otto, M.C. (1988), Benchmarking: Evaluating methods that preserve month-to-month changes. Bureau of the Census - Statistical Research Division, RR-88/07. <http://www.census.gov/srd/papers/pdf/rr88-07.pdf>
- [6] Brown, I. (2010), An empirical comparison of constrained optimization methods for benchmarking economic time series. In *JSM 2009 Proceedings, Business and Economic Statistics Section*, 2131–2143.
- [7] Brown, I. (2012), “Modernization of benchmarking economic time series at the U.S. Census Bureau”, Federal Committee on Statistical Methodology 2012 Research Conference, Washington, D.C. <http://www.fcsm.gov/12papers/Brown2012FCSMII-D.pdf>.
- [8] Byrd, R.H., Gilbert, J.C., Nocedal J. (2000), “A Trust Region Method Based on Interior Point Techniques for Nonlinear Programming”, *Mathematical Programming*, **89**, 149–185.
- [9] Byrd, R.H., Hribar, M.E., Nocedal, J. (1999), “An Interior Point Algorithm for Large-Scale Nonlinear Programming,” *SIAM Journal of Optimization*, **9**, 877–900.
- [10] Causey, B., Trager, M.L. (1981), *Derivation of Solution to the Benchmarking Problem: Trend Revision*. Unpublished research notes, U.S. Census Bureau, Washington D.C.. Available as an appendix in Bozik and Otto (1988).
- [11] Chen, B. (2006), “A Balanced System of Industry Accounts for the U.S. and Structural Distribution of Statistical Discrepancy”, BEA Working Papers WP2006-8. http://www.bea.gov/papers/pdf/reconciliation_wp.pdf.
- [12] Cholette, P. (1984), “Adjusting sub-annual series to yearly benchmarks”, *Survey Methodology*, **10**, 35–49.
- [13] Cholette, P., Dagum, E.B. (1994), “Benchmarking time series with autocorrelated survey errors”, *International Statistical Review*, **62**, 365–377.
- [14] Dagum, E.B., Cholette, P.A. (2006), *Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series*. Springer, New York.

-
- [15] Danilov, D., Magnus, J.R. (2008), “On the estimation of a large sparse Bayesian system: The Snaer program”, *Computational Statistics and Data Analysis*, **52**, 4203–4224.
- [16] Davis, T. (2011), “Algorithm 915, SuiteSparseQR: Multifrontal multithreaded rank-revealing sparse QR factorization”, *ACM Transactions on Mathematical Software*, **38**.
- [17] 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.
- [18] Di Fonzo, T., Marini, M. (2010), *Benchmarking and movement preservation. Evidences from real-life and simulated series*. Department of Statistical Sciences, University of Padua, Working Paper Series, n. 14. <http://www.stat.unipd.it/ricerca/fulltext?wp=422>
- [19] Di Fonzo, T., Marini, M. (2011a), “Simultaneous and Two-step Reconciliation of Systems of Time Series: Methodological and Practical Issues”, *Journal of the Royal Statistical Society C*, **60**, 143–164.
- [20] Di Fonzo, T., and Marini, M. (2011b), *A Newton’s Method for Benchmarking Time Series according to a Growth Rates Preservation Principle*, IMF Working Paper WP/11/179. <http://www.imf.org/external/pubs/ft/wp/2011/wp11179.pdf>.
- [21] Di Fonzo, T., Marini, M. (2012), *On the Extrapolation with the Denton Proportional Benchmarking Method*, IMF Working Paper WP/12/169. <http://www.imf.org/external/pubs/ft/wp/2012/wp12169.pdf>.
- [22] Helfand, S., Monsour, N., Trager, M. (1977), “Historical revision of current business survey estimates”, *American Statistical Association, Proceedings of the Business and Economic Statistics Section*, 246–250.
- [23] Karmarkar, N. (1984), “A polynomial-type algorithm for linear programming”, *Combinatorica*, **4**, 373–395.
- [24] Lesaja, G. (2009), “Introducing Interior-Point Methods for Introductory Operations Research Courses and/or Linear Programming Courses”, *The Open Operational Research Journal*, 2009, **3**, 1–12.
- [25] Monsour, N.J., Trager, M.L. (1979), Revision and benchmarking of business time series. American Statistical Association, *Proceedings of the Business and Economic Statistics Section*, 333-337.
- [26] Nocedal, J., Wright, S. (2006), *Numerical Optimization*. 2nd edition. Springer, New York.
- [27] Quenneville, B., and Fortier, S. (2012), “Restoring accounting constraints in time series - Methods and software for a statistical agency”, in Bell, W.R., Holan, S.H., and McElroy, T.S. (eds.), *Economic Time Series Modeling and Seasonality*, Taylor & Francis, New York, 231–253.

-
- [28] Quenneville, B., Fortier, S., and Gagnè, C. (2009), “A non-parametric iterative smoothing method for benchmarking and temporal distribution”, *Computational Statistics & Data Analysis*, **53**, 3386–3396.
- [29] Quenneville, B., and Rancourt, E. (2005), “Simple methods to restore the additivity of a system of time series”, paper presented at the workshop on ‘Frontiers in benchmarking techniques and their application to official statistics’, Luxembourg, 7-8 April 2005.
- [30] Round, J.I. (2003), “Constructing SAMs for Development Policy Analysis: Lessons Learned and Challenges Ahead”, *Economic Systems Research*, **15**, 161–183.
- [31] Schmidt, M. (2006), The minFunc Toolbox for Matlab. <http://www.cs.ubc.ca/~schmidtm>
- [32] Stuckey, A., Zhang, X., McLaren, C.H. (2004), “Aggregation of Seasonally Adjusted Estimates by a Post-Adjustment”. www.uow.edu.au/~craigmc/abs_agg_2004.pdf.
- [33] The MathWorks (2009), *Optimization ToolboxTM 4 User’s Guide*. Natick, MA.
- [34] Titova, N, Findley, D., Monsell, B.C. (2010), Comparing the Causey-Trager method to the multiplicative Cholette-Dagum regression-based method of benchmarking sub-annual data to annual benchmarks. In *JSM 2010 Proceedings, Business and Economic Statistics Section*, 3007–3021.
- [35] Trager, M.L. (1982), *Derivation of Solution to the Benchmarking Problem: Relative Revision*. Unpublished research notes, U.S. Census Bureau, Washington D.C.. Available as an appendix in Bozik and Otto (1988).
- [36] U.S. Census Bureau (2009), *X-12-ARIMA Reference manual, Version 0.3*. U.S. Department of Commerce, Washington D.C.. <http://www.census.gov/srd/www/x12a/>
- [37] Waltz, R. A., J. L. Morales, J. Nocedal, Orban, D. (2006), “An interior algorithm for nonlinear optimization that combines line search and trust region steps”, *Mathematical Programming*, **107**, 391–408.

Acknowledgements

A preliminary version of this paper has been presented at the JSM 2012 of the American Statistical Association, San Diego, July 28th - August 2nd, 2012. We are grateful to Gary Brown, Irene Brown, Baoline Chen, Estela Bee Dagum, Vittorio Nicolardi and Neil Parkin for their helpful comments. We alone are responsible for the errors of the paper. The views expressed herein are those of the authors and should not be attributed to the IMF, its Executive Board, or its management.

Working Paper Series
Department of Statistical Sciences, University of Padua

You may order paper copies of the working papers by emailing wp@stat.unipd.it
Most of the working papers can also be found at the following url: <http://wp.stat.unipd.it>

