



Department of Statistical Sciences
University of Padua
Italy

UNIVERSITÀ
DEGLI STUDI
DI PADOVA
DIPARTIMENTO
DI SCIENZE
STATISTICHE

Benchmarking and movement preservation Evidences from real-life and simulated series

Tommaso Di Fonzo

Department of Statistical Sciences
University of Padua
Italy

Marco Marini

Statistics Department
International Monetary Fund
USA

Abstract: 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 annual benchmarks and corresponding sums of the sub-annual values. Two widely used benchmarking procedures are the modified Denton *Proportional First Differences (PFD)* and the Causey and Trager *Growth Rates Preservation (GRP)* techniques. The *PFD* procedure, which looks for benchmarked estimates aimed at minimizing the sum of squared proportional differences between the target and the unbenchmarked values, involves simple matrix operations. The *GRP* technique is a non-linear procedure based on a ‘true’ movement preservation principle, according to which the sum of squared differences between the growth rates of the target and of the unbenchmarked series is minimized. In the literature it is often claimed that the *PFD* procedure produces results very close to those obtained through the *GRP* procedure. In this paper we study the conditions under which this result holds, by looking at an artificial and a real-life economic series, and by means of a simulation exercise.

Keywords: Combining data from different sources, Benchmarking, Movement Preservation, Modified Denton *PFD*, Causey and Trager *GRP*.

JEL classification codes: C22, C61, C82.

Contents

1	Introduction	1
2	Two benchmarking procedures	2
2.1	Modified Denton <i>PFD</i>	3
2.2	Causey and Trager <i>GRP</i>	4
2.3	Relationship between <i>PFD</i> and <i>GRP</i> criteria	4
3	Evidences from artificial and real life series	5
4	The simulation exercise	7
	Appendix. Non-singularity of the coefficient matrix of system (2)	12
	References	13

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>

Benchmarking and movement preservation

Evidences from real-life and simulated series

Tommaso Di Fonzo

Department of Statistical Sciences
University of Padua
Italy

Marco Marini

Statistics Department
International Monetary Fund
USA

Abstract: 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 annual benchmarks and corresponding sums of the sub-annual values. Two widely used benchmarking procedures are the modified Denton *Proportional First Differences* (*PFD*) and the Causey and Trager *Growth Rates Preservation* (*GRP*) techniques. The *PFD* procedure, which looks for benchmarked estimates aimed at minimizing the sum of squared proportional differences between the target and the unbenchmarked values, involves simple matrix operations. The *GRP* technique is a non-linear procedure based on a ‘true’ movement preservation principle, according to which the sum of squared differences between the growth rates of the target and of the unbenchmarked series is minimized. In the literature it is often claimed that the *PFD* procedure produces results very close to those obtained through the *GRP* procedure. In this paper we study the conditions under which this result holds, by looking at an artificial and a real-life economic series, and by means of a simulation exercise.

Keywords: Combining data from different sources, Benchmarking, Movement Preservation, Modified Denton *PFD*, Causey and Trager *GRP*.

JEL classification codes: C22, C61, C82.

1 Introduction

Benchmarking monthly and quarterly series to annual series is a common practice in many National Statistical Institutes. 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 annual benchmarks and corresponding sums of the sub-annual values. The most widely used benchmarking procedures are the modified Denton *Proportional First Differences* (*PFD*) technique (Denton, 1971; Cholette, 1984), and the Causey and Trager (1981) *Growth Rates Preservation* (*GRP*) procedure (see also Trager, 1982,

and Bozik and Otto, 1988). The *PFD* procedure looks for benchmarked estimates aimed at minimizing the sum of squared proportional differences between the target and the unbenchmarked values, and is characterized by an explicit benchmarking formula involving simple matrix operations. The *GRP* technique is a non-linear procedure based on a ‘true’ movement preservation principle, according to which the sum of squared differences between the growth rates of the target and of the unbenchmarked series is minimized. As 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 those obtained through the *GRP* procedure, in this paper we study the conditions under which this result holds. We do that by showing how the two procedures work in practice, by looking at an artificial and a real-life economic series. Then a simulation exercise is performed in order to appreciate the impact on the benchmarked series of the variance of the observational error and of possible ‘steps’ in the annual benchmarks.

The paper is organized as follows. In section 2 the two benchmarking procedures are described, and the way they take into account a ‘movement preservation principle’ is discussed. In section 3 the artificial time series of Denton (1971) and a quarterly preliminary series of the EU Quarterly Sector Accounts (Di Fonzo and Marini, 2009) are benchmarked to their annual counterparts, using both modified Denton *PFD* and Causey and Trager *GRP* benchmarking procedures, and the results are discussed. In section 4 we design a simulation exercise in order to analyze the distinctive features of the two procedures.

2 Two benchmarking procedures

Let Y_T , $T = 1, \dots, N$, and p_t , $t = 1, \dots, n$, be, respectively, the (say annual) totals and the (say quarterly) preliminary values of an unknown quarterly target variable y_t . The preliminary values being not in line with the annual benchmarks, i.e. $\sum_{t \in T} p_t \neq Y_T$, $T = 1, \dots, N$, we look for benchmarked estimates y_t^b such that

$$\sum_{t \in T} y_t^b = Y_T.$$

As Bozik and Otto (1988, p. 2) stress, “Just forcing a series to sum to its benchmark totals does not make a unique benchmark series”. Some characteristic of the original series p_t should be considered in addition, in order to get benchmarked estimates ‘as close as possible’ to the preliminary values. In an economic time series framework, the preservation of the temporal dynamics (however defined) of the preliminary series is often a major interest of the practitioner. Thus in what follows we consider two procedures designed to preserve at the best the movement of the series p_t : modified Denton *PFD* and Causey and Trager *GRP*¹.

¹Empirical comparisons between the Cholette-Dagum regression based benchmarking approach, which can be seen (Dagum and Cholette, 2006) as a generalization of the seminal contribution by Denton (1971), and the Causey and Trager approach, are shown in Harvill Hood (2005) and Titova *et al.* (2010).

2.1 Modified Denton *PFD*

Denton (1971) proposed a benchmarking procedure grounded on the *Proportionate First Differences* between the target and the original series. Cholette (1984) slightly modified the result of Denton, in order to correctly deal with the starting conditions of the problem. The *PFD* benchmarked estimates are thus obtained as the solution to the constrained quadratic minimization problem

$$\min_{y_t} \sum_{t=2}^n \left(\frac{y_t}{p_t} - \frac{y_{t-1}}{p_{t-1}} \right)^2 \quad \text{subject to} \quad \sum_{t \in T} y_t = Y_T, \quad T = 1, \dots, N. \quad (1)$$

In matrix notation, denoting \mathbf{p} and \mathbf{Y} the $(n \times 1)$ and $(N \times 1)$, respectively, vectors of preliminary and benchmark values, the *PFD* benchmarked series is contained in the $(n \times 1)$ vector \mathbf{y}^{PFD} solution to the linear system (Cholette, 1984, p. 40)

$$\begin{bmatrix} \mathbf{Q} & \mathbf{C}' \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{y}^{PFD} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{Qp} \\ \mathbf{Y} \end{bmatrix}, \quad (2)$$

where λ is a $(N \times 1)$ vector of Lagrange multipliers, $\mathbf{Q} = \mathbf{P}^{-1} \Delta'_n \Delta_n \mathbf{P}^{-1}$, $\mathbf{P} = \text{diag}(\mathbf{p})$, \mathbf{C} is a $(N \times n)$ temporal aggregation matrix converting quarterly values in their annual sums, and Δ_n is the $((n-1) \times n)$ first differences matrix.

Notice that $\Delta'_n \Delta_n$ has rank $n-1$ (Cohen *et al.*, 1971, p. 122), so \mathbf{Q} is singular. However, provided no preliminary value is equal to zero², the coefficient matrix of system (2) has full rank (see the Appendix). In addition, this matrix can be factorized as

$$\begin{bmatrix} \mathbf{Q} & \mathbf{C}' \\ \mathbf{C} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{P}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \end{bmatrix} \begin{bmatrix} \Delta'_n \Delta_n & \mathbf{PC}' \\ \mathbf{CP} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \end{bmatrix},$$

whose inverse can be expressed as

$$\begin{bmatrix} \mathbf{Q} & \mathbf{C}' \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \end{bmatrix} \begin{bmatrix} \Delta'_n \Delta_n & \mathbf{PC}' \\ \mathbf{CP} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \end{bmatrix}.$$

Given that $\mathbf{P}^{-1}\mathbf{p} = \mathbf{1}_n$ and $\Delta_n \mathbf{1}_n = \mathbf{0}$, it is $\mathbf{Qp} = \mathbf{0}$. The solution of the linear system (2) can thus be written as

$$\begin{bmatrix} \mathbf{y}^{PFD} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \end{bmatrix} \begin{bmatrix} \Delta'_n \Delta_n & \mathbf{PC}' \\ \mathbf{CP} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{Y} \end{bmatrix},$$

that is

$$\begin{bmatrix} \mathbf{y}^{PFD} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \end{bmatrix} \begin{bmatrix} \Delta'_n \Delta_n & \mathbf{PC}' \\ \mathbf{CP} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{Y} \end{bmatrix}. \quad (3)$$

²When some p_t is null, a standard practice in the benchmarking literature (see, for example, Cholette and Chhab, 1991, p. 413) consists in setting the originally null preliminary data at a very small value, e.g. $p_t = 0.001$.

2.2 Causey and Trager *GRP*

Causey and Trager (1981; see also Monsour and Trager, 1979, and Trager, 1982) consider a different quadratic minimization problem, 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:

$$\min_{y_t} \sum_{t=2}^n \left(\frac{y_t}{y_{t-1}} - \frac{p_t}{p_{t-1}} \right)^2 \quad \text{subject to} \quad \sum_{t \in T} y_t = Y_T, \quad T = 1, \dots, N. \quad (4)$$

Looking at the criterion to be minimized in (4), it clearly appears that, differently from (1), 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).

It should be noted that, while problem (1) has linear first-order conditions for a minimum, and thus gives rise to an explicit solution as shown in (3), the minimization problem in (4) is inherently non-linear. Trager (1982; see Bozik and Otto, 1988) suggests to use a technique based on the steepest descent method³, using \mathbf{y}^{PFD} as starting values, in order to calculate the benchmarked estimates y_t^{GRP} , $t = 1, \dots, n$, solution to problem (4).

We employ the Interior Point method (Byrd *et al.*, 1999) made available by the Optimization ToolboxTM of MATLAB[®] (version 2009b, The Mathworks, 2009). It consists of an iterative procedure that solves a sequence of approximate unconstrained minimization problems by standard (quadratic) nonlinear programming methods. In each iteration the procedure exploits the exact gradient vector and hessian matrix of the Lagrangian function, which enables to make informed decisions regarding directions of search and step length. This fact makes the procedure feasible and robust, in terms of reduced numbers of iterations required for the convergence, as far as of quality of the found minimum.

2.3 Relationship between *PFD* and *GRP* criteria

It is interesting to go deep into the relationship between the criteria optimized by the two alternative procedures. Let

$$C_{PFD} = \sum_{t=2}^n \left(\frac{y_t}{p_t} - \frac{y_{t-1}}{p_{t-1}} \right)^2 \quad \text{and} \quad C_{GRP} = \sum_{t=2}^n \left(\frac{y_t}{y_{t-1}} - \frac{p_t}{p_{t-1}} \right)^2$$

be the objective functions of the *PFD* and *GRP* benchmarking procedures, respectively. After a bit of algebra, we can write (U.S. Bureau of the Census, 2009, p. 96):

$$C_{PFD} = \sum_{t=2}^n \left[\frac{y_{t-1}}{p_t} \left(\frac{y_t}{y_{t-1}} - \frac{p_t}{p_{t-1}} \right) \right]^2. \quad (5)$$

Expression (5) makes clear the relationship between C_{PFD} and C_{GRP} . The term in parentheses is the difference between the growth rates of the target and the preliminary series, namely the addendum of C_{GRP} . In C_{PFD} these terms are weighted by

³For a recent survey on this issue, see Brown (2010).

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’ (Bloem *et al.*, 2001)

$$\frac{Y_T}{\sum_{t \in T} p_t}, \quad T = 1, \dots, N$$

is a smooth series, C_{PFD} and C_{GRP} are very close to each other. On the contrary, when the ratios (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.

The situation is rather similar to the one described by Dagum and Cholette (2006, p. 121), when the risk of producing negative benchmarked values is discussed: ‘These situations occur when the benchmarks dramatically change from one year to the next, while the sub-annual series changed very little in comparison; or, when the benchmarks change little, while the annual sums of the sub-annual series change dramatically’.

Keeping in mind this relationship, we move to investigate on the differences between the *PFD* and the *GRP* benchmarking solutions in simulated and real-life cases.

3 Evidences from artificial and real life series

In this section we apply both the *PFD* and *GRP* benchmarking procedures to two illustrative examples, in order to show to what extent the former solution can be used effectively to approximate the “ideal” movement preservation criterion based on growth rates. We consider also a distance measure between the growth rates of the preliminary and target series given by the absolute, rather than the squared, value of their difference. The results are thus evaluated looking at the two ratios

$$r_\alpha = \left(\frac{\sum_{t=2}^n \left| \frac{y_t^{GRP}}{y_{t-1}^{GRP}} - \frac{p_t}{p_{t-1}} \right|^\alpha}{\sum_{t=2}^n \left| \frac{y_t^{PFD}}{y_{t-1}^{PFD}} - \frac{p_t}{p_{t-1}} \right|^\alpha} \right)^{\frac{1}{\alpha}} \quad \alpha = 1, 2. \quad (6)$$

When $\alpha = 2$, this index is simply the square root of the ratio between the Causey and Trager ‘Growth Rate Preservation’ criteria computed from the two solutions. Obviously, we expect the *GRP* technique always reaches a lower value of the chosen criterion than *PFD*, and thus the ratio r_2 should be never larger than 1. Put in other words, r_2 is the ratio between the Root Mean Squared Adjustments to the preliminary growth rates produced by the Causey and Trager *GRP* and the Denton *PFD* benchmarking procedures. On the other hand, r_1 can be seen as the ratio between the Mean Absolute Adjustments: sometimes this index can be larger than 1, thus indicating a better performance of Denton *PFD* when the size of the

Table 1: Adjustments to the artificial series produced by the *PFD* and *GRP* procedures

Year	Qtr	p_t	y_t^{PFD}	y_t^{GRP}	$y_t^{PFD} - p_t$	$y_t^{GRP} - p_t$
1	1	50	64.3	63.6	14.3	13.6
	2	100	127.8	127.0	27.8	27.0
	3	150	187.8	189.6	37.8	39.6
	4	100	120.0	119.8	20.0	19.8
2	1	50	56.6	52.0	6.6	2.0
	2	100	106.0	103.2	6.0	3.2
	3	150	147.5	152.5	-2.5	2.5
	4	100	90.0	92.3	-10.0	-7.7
3	1	50	40.5	37.1	-9.5	-12.9
	2	100	74.4	73.6	-25.6	-26.4
	3	150	108.3	110.3	-41.7	-39.7
	4	100	76.7	79.0	-23.3	-21.0
4	1	50	42.8	47.6	-7.2	-2.4
	2	100	94.1	96.5	-5.9	-3.5
	3	150	153.4	148.1	3.4	-1.9
	4	100	109.7	107.9	9.7	7.9
5	1	50	58.3	61.3	8.3	11.3
	2	100	122.6	123.6	22.6	23.6
	3	150	190.4	187.4	40.4	37.4
	4	100	128.7	127.7	28.7	27.7

corrections to the preliminary growth rates is measured according to an absolute rather than a squared form.

The first example we consider is the artificial preliminary series used in the seminal paper of Denton (1971). It consists of a five-year artificial quarterly series, with a fixed seasonal pattern invariant from year to year. The values are 50, 100, 150 and 100 in the four quarters, for a total yearly amount of 400. The annual benchmarks are assumed to be 500, 400, 300, 400 and 500 in the five successive years. The corresponding discrepancies (i.e., the differences between the known benchmarks and the annual sums of the preliminary series) are therefore 100, 0, -100, 0 and 100, respectively. The benchmarked values are shown in table 1: as expected, the minimum C_{GRP} is achieved by the *GRP* procedure (0.04412 against 0.14428 of *PFD*, with $r_2=0.553$). The *GRP* procedure shows better results as regards the movement preservation, also when the distance between the preliminary and the target growth rates is measured by the absolute differences ($r_1 = 0.539$). Figure 1 shows the adjustments to the levels of the original series in the two cases. The horizontal lines in each year denote the (average) annual discrepancy to be distributed.

The second example is a real-life economic series coming from the European

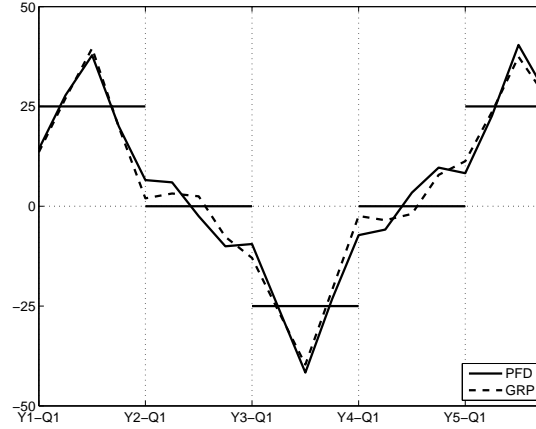


Figure 1: Adjustments to the artificial series produced by the *PFD* and *GRP* procedures

Quarterly Sector Accounts (EU-QSA). The EU-QSA system has been dealt with by Di Fonzo and Marini (2009) in a *reconciliation* exercise, where several time series have to be adjusted in order to be in line with both temporal and contemporaneous known aggregates (Dagum and Cholette, 2006). In this paper we consider the series 'Other Property Income' of the Financial Corporation sector, showing a considerable amount of temporal discrepancies.

Table 2 presents the original quarterly series (p_t) and the annual benchmarks (Y_T) in the third and fourth columns, respectively. The annual discrepancies

$$Y_T - \sum_{t \in T} p_t, \quad T = 1, \dots, N,$$

are shown in the fifth one: it can be noticed the large discrepancy in 2002, when the original series accounts for just 65% of the annual target. From 2003 onwards the discrepancies are much more contained. This is a typical practical situation where the preservation of the original growth rates can be better guaranteed by the *GRP* procedure ($r_2 = 0.579$ and $r_1 = 0.615$). The quarterly adjustments in the two cases are displayed in figure 2. The differences are large in the years with large discrepancies (2001-2002), but they are also remarkable in 2003, when the discrepancy is limited. In this case the smoother distribution produced by the *GRP* procedure is clearly visible.

4 The simulation exercise

By means of this experiment we wish to shed light on the conditions under which the *PFD* benchmarking procedure produces results 'close' to the *GRP* technique in terms of differences between the growth rates of the benchmarked and preliminary series. We consider quarterly series covering a period of 7 years ($n = 28$).

Let $\theta_t = \theta_{t-1} + \varepsilon_t$ be a random walk process, where ε_t is a Gaussian white noise with unit variance ($\sigma_\varepsilon^2 = 1$) and $\theta_0 = \varepsilon_0$. The target series of the exercise, y_t , is

Table 2: Adjustments to the EU-QSA series produced by the *PFD* and *GRP* procedures

Year	Qtr	p_t	Y_T	$Y_T - \sum_{t \in T} p_t$	y_t^{PFD}	y_t^{GRP}	$y_t^{PFD} - p_t$	$y_t^{GRP} - p_t$
1999	1	27,311	127,087.5	1,134.5	27,471.8	27,414.8	160.8	103.8
	2	40,983			41,269.3	41,160.7	286.3	177.7
	3	23,210			23,435.8	23,527.6	225.8	317.6
	4	34,449			34,910.6	34,984.4	461.6	535.4
2000	1	29,026	170,140.5	7,378.5	29,561.7	29,707.1	535.7	681.1
	2	50,778			52,259.3	52,132.6	1,481.3	1,354.6
	3	39,534			41,502.7	41,681.4	1,968.7	2,147.4
	4	43,424			46,816.8	46,619.4	3,392.8	3,195.4
2001	1	32,004	190,586.4	37,557.4	35,682.5	36,094.2	3,678.5	4,090.2
	2	49,598			58,312.9	56,747.5	8,714.9	7,149.5
	3	22,432			28,569.9	30,476.0	6,137.9	8,044.0
	4	48,995			68,021.1	67,268.6	19,026.1	18,273.6
2002	1	24,082	161,395.1	55,977.1	37,079.6	40,173.0	12,997.6	16,091.0
	2	36,436			58,420.2	61,288.1	21,984.2	24,852.1
	3	27,061			41,518.1	41,164.9	14,457.1	14,103.9
	4	17,839			24,377.2	18,769.1	6,538.2	930.1
2003	1	34,201	159,755.6	-5,906.4	38,776.5	34,839.9	4,575.5	638.9
	2	53,574			51,637.6	52,756.0	-1,936.4	-818.0
	3	36,873			32,902.0	34,074.8	-3,971.0	-2,798.2
	4	41,014			36,439.5	38,085.0	-4,574.5	-2,929.7
2004	1	32,304	161,472.1	2,154.1	31,009.9	31,963.1	-1,294.1	-340.9
	2	51,855			52,512.9	51,934.1	657.9	79.1
	3	33,668			34,859.0	34,741.6	1,191.0	1,073.6
	4	41,491			43,090.3	42,833.2	1,599.3	1,342.2
2005	1	34,198	177,814.7	-362.3	34,802.0	34,522.0	604.0	324.0
	2	59,280			59,326.8	59,604.5	46.8	324.5
	3	39,534			39,172.7	39,121.3	-361.3	-412.7
	4	45,165			44,513.2	44,566.9	-651.8	-598.1

derived as

$$y_t = \theta_t^* + \mu_t, \quad t = 1, \dots, n$$

with $\theta_t^* = 100 + \theta_t$, where the constant term 100 is large enough to prevent negative values, and μ_t is given by

$$\mu_t = \begin{cases} \mu & t = 9, \dots, 16 \\ -\mu & t = 17, \dots, 24 \\ 0 & \text{elsewhere} \end{cases}.$$

The preliminary series p_t is related to y_t as follows:

$$p_t = \theta_t^* + e_t, \quad t = 1, \dots, n$$

where e_t is a Gaussian white noise with variance σ_e^2 . It is clear that preliminary and target series are different for the effects of μ_t and e_t . The former is introduced in the model for y_t in order to simulate yearly biases of the preliminary series. The first control parameter of the experiment is thus μ . When $\mu > 0$, the target series

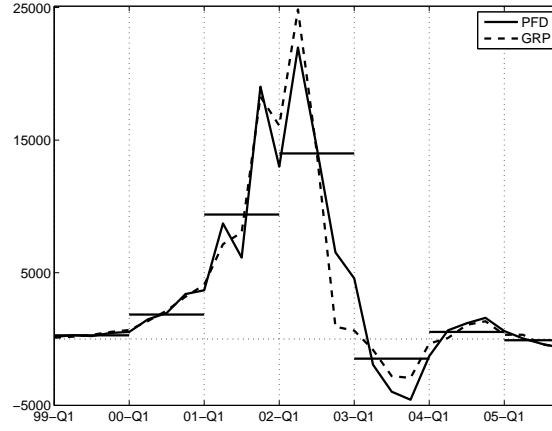


Figure 2: Adjustments to the real-life series produced by the *PFD* and *GRP* procedures

Table 3: Median of r_2 for different values of σ_e and μ (across 1,000 experiments)

σ_e	μ				
	0	15	30	45	60
5	0.992	0.984	0.959	0.903	0.763
10	0.967	0.961	0.935	0.879	0.739
15	0.928	0.922	0.895	0.839	0.702
20	0.870	0.866	0.841	0.785	0.655
25	0.800	0.795	0.771	0.718	0.598

contains a positive drift from p_t in years 3 and 4, followed by a negative step (of the same amount) in years 5 and 6. We set $\mu = 0, 15, 30, 45, 60$. The second control parameter is σ_e , the standard deviation of the innovation process e_t . The larger this parameter is, the larger the observational error in the preliminary series will be. We set $\sigma_e = 5, 10, 15, 20, 25$.

We drew two sets of 1,000 n -dimensional vectors as $N(0, 1)$. One set is used to simulate ε_t ; the other is used to derive the innovation e_t according to the 5 levels of σ_e . By using the 5 values of μ , we achieved 1,000 experiments for each of the 25 combinations. For each combination, we computed summary statistics on the ratios r_1 and r_2 obtained over the 1,000 experiments.

Tables 3 and 4 show, respectively, the median and the maximum values⁴ of r_2 under different values of σ_e (rows) and μ (columns).

According to Bozik and Otto (1988), we used the series benchmarked *via* modified Denton *PFD* as starting values of the *GRP* procedure, and this turned out to be

⁴The median is more representative than the mean in the case of atypical values. We also calculated mean, standard deviation, minimum and range of r_1 and r_2 , available on request from the authors.

Table 4: Maximum of r_2 for different values of σ_e and μ (across 1,000 experiments)

σ_e	μ				
	0	15	30	45	60
5	0.999	0.995	0.982	0.946	0.870
10	0.996	0.993	0.978	0.950	0.888
15	0.992	0.987	0.977	0.946	0.903
20	0.987	0.982	0.973	0.941	0.907
25	0.980	0.977	0.965	0.943	0.885

a good choice⁵: as one would expect, from table 4 it appears that in all cases the *GRP* procedure improves on the modified *PFD* starting values and reaches a lower value of the criterion.

However, from table 3 we observe that the *PFD* procedure provides very similar results to *GRP* when discrepancies are small and unsystematic (median $r_2 \geq 0.9$ when $\sigma_e \leq 10$ and $\mu \leq 15$). The reduction is stronger as both σ_e and μ increase.

These results are confirmed by r_1 , whose median and maximum values are shown in tables 5 and 6, respectively, with an important remark: from table 6 we observe that if the absolute differences between preliminary and target growth rates are considered, and when the bias is either absent or small ($\mu \leq 5$), there are cases where Denton *PFD* gives benchmarked estimates whose dynamics is ‘closer’ to the preliminary series than Causey and Trager *GRP* does⁶.

Table 5: Median of r_1 for different values of σ_e and μ (across 1,000 experiments)

σ_e	μ				
	0	15	30	45	60
5	0.993	0.977	0.941	0.880	0.748
10	0.969	0.956	0.922	0.860	0.725
15	0.928	0.917	0.884	0.825	0.693
20	0.873	0.864	0.832	0.773	0.650
25	0.805	0.795	0.765	0.710	0.597

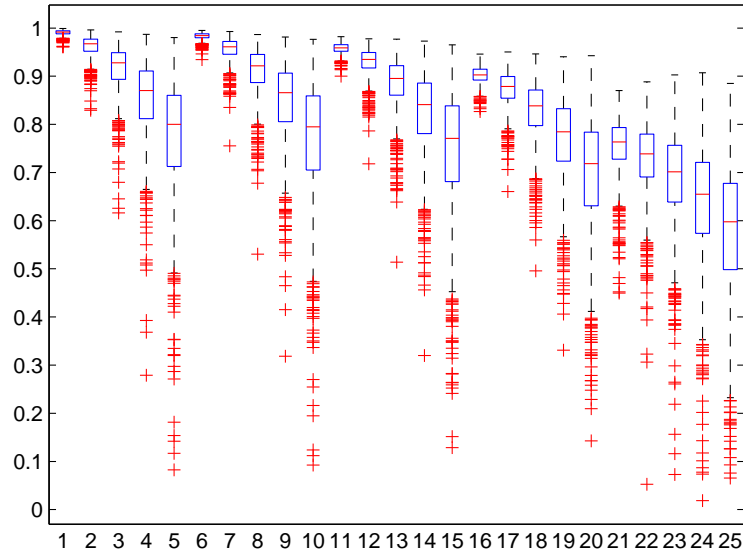
A visual, more comprehensive overview of the results of the simulation experiment is given by the boxplots of r_2 (figure 3), and r_1 (figure 4). The 25 boxplots in each figure are ordered (from left to right) so that the first group of five boxplots corresponds to $\sigma_e = 5$ and $\mu = 0, 15, 30, 45, 60$, respectively, the second group to $\sigma_e = 10$ and $\mu = 0, 15, 30, 45, 60$, and so on.

⁵When the preliminary series were used as starting values, in 50 out of 25,000 cases (0.2%) the *GRP* procedure produced benchmarked series with $r_2 > 1$.

⁶The index r_1 is greater than one for 488 out of 25,000 series (1.95%). The highest number of cases with $r_1 > 1$ (270) is observed for $(\sigma_e, \mu) = (5, 0)$, followed by 101 cases for $(\sigma_e, \mu) = (10, 0)$. The remaining cases are: 50 for $(\sigma_e, \mu) = (15, 0)$, 15 for $(\sigma_e, \mu) = (20, 0)$, 4 for $(\sigma_e, \mu) = (25, 0)$, 20 for $(\sigma_e, \mu) = (5, 5)$, 15 for $(\sigma_e, \mu) = (10, 5)$, 8 for $(\sigma_e, \mu) = (15, 5)$, 2 for $(\sigma_e, \mu) = (20, 5)$, and 3 for $(\sigma_e, \mu) = (25, 5)$.

Table 6: Maximum of r_1 for different values of σ_e and μ (across 1,000 experiments)

σ_e	μ				
	0	15	30	45	60
5	1.035	1.015	0.997	0.936	0.841
10	1.053	1.013	0.997	0.964	0.871
15	1.047	1.017	0.982	0.952	0.896
20	1.028	1.022	0.982	0.941	0.903
25	1.013	1.057	0.970	0.929	0.874

**Figure 3:** Boxplots of r_2

From these evidences, and for this dataset, we conclude that the modified Denton *PFD* benchmarking procedure can be viewed as a sensible approximation of the Causey and Trager benchmarking procedure 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 (high variability and/or large bias), the quality of the approximation clearly worsens. In addition, as regards the ‘movement preservation’, we have found that generally the Causey and Trager *GRP* benchmarking procedure gives better performances as compared to Denton *PFD*. This turned out to be always the case (as one would expect) when the comparison criterion is the one optimized by the Causey and Trager procedure, and in the very largest amount (more than 98%) of the 25,000 series of our simulation experiment, when the distance between the growth rates of the preliminary and the benchmarked series is measured by the absolute difference.

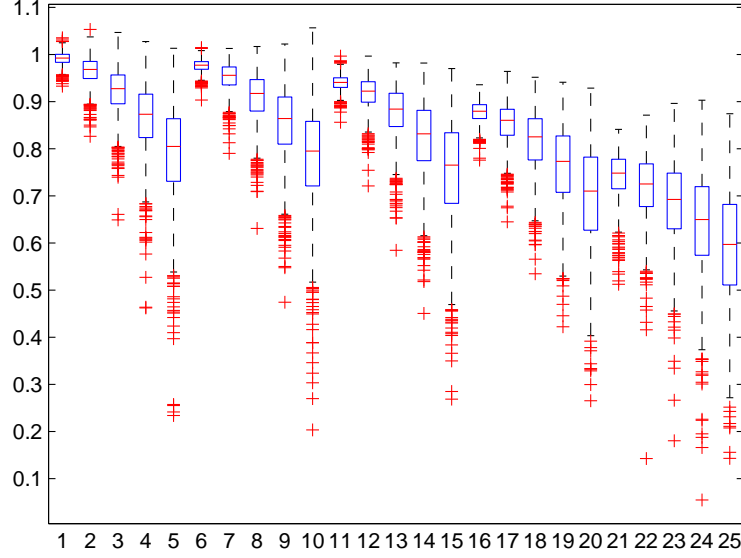


Figure 4: Boxplots of r_1

Appendix. Non-singularity of the coefficient matrix of system (2)

Luenberger (1984, p. 424) shows that a unique solution of the constrained quadratic minimization problem

$$\min_{\mathbf{y}} \mathbf{y}^T \mathbf{Q} \mathbf{x} \quad \text{subject to} \quad \mathbf{C} \mathbf{y} = \mathbf{Y}$$

exists if the matrix \mathbf{C} is of full rank, and the matrix \mathbf{Q} is positive definite on the subspace $\mathcal{N}(\mathbf{C}) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{C}\mathbf{x} = \mathbf{0}\}$, i.e. the null space of matrix \mathbf{C} (Harville, 1997).

Let us consider the matrix $\mathbf{Q} = \mathbf{P}^{-1} \Delta^T \Delta \mathbf{P}^{-1}$, and let the vector \mathbf{y} belong to $\mathcal{N}(\mathbf{C})$. We assume $p_t \neq 0$, $t = 1, \dots, n$ (otherwise the objective function is not defined), and $\mathbf{C}\mathbf{p} \neq \mathbf{Y}$, which corresponds to exclude the trivial solution $\mathbf{y}^* = \mathbf{p}$, valid when there is no benchmarking problem. Given that

$$\mathbf{y}^T \mathbf{Q} \mathbf{y} = \sum_{t=2}^n \left(\frac{y_t}{p_t} - \frac{y_{t-1}}{p_{t-1}} \right)^2,$$

it is immediately recognized that the expression above is strictly positive, which means that matrix $\mathbf{Q} = \mathbf{P}^{-1} \Delta^T \Delta \mathbf{P}^{-1}$ is positive definite on the null space spanned by the columns of matrix \mathbf{C} , and thus the coefficient matrix of system (2)

$$\begin{bmatrix} \mathbf{P}^{-1} \Delta^T \Delta \mathbf{P}^{-1} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix}$$

is nonsingular.

Notice that here we have considered the first differences proportional variant of the Denton's objective criterion. For purely additive variants of the objective function, the nonsingularity of the system matrix is discussed by Boot *et al.* (1967), which consider first and second differences, and by Cohen *et al.* (1971), which give results for a generic order of differences $d \geq 1$.

References

- [1] Bloem, A., Dippelsman, R., and Mæhle, N. (2001), *Quarterly National Accounts Manual. Concepts, Data Sources, and Compilation*, International Monetary Fund, Washington DC.
- [2] Boot, J.C.G., Feibes, W., Lisman, J.H.C. (1967), "Adjusting sub-annual series to yearly benchmarks," *Applied Statistics*, **16**, 65–75.
- [3] 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>
- [4] Brown, I. (2010), An empirical comparison of constrained optimization methods for benchmarking economic time series. In *JSM Proceedings*, Business and Economic Statistics Section. Alexandria, VA: American Statistical Association, 2131–2143.
- [5] 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.
- [6] 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).
- [7] Cholette, P. (1984), "Adjusting sub-annual series to yearly benchmarks," *Survey Methodology*, **10**, 35–49.
- [8] Cholette, P.A., Chhab, N.B. (1991), "Converting aggregates of weekly data into monthly values", *Applied Statistics*, **40**, 411–422.
- [9] Dagum, E.B., Cholette, P.A. (2006), *Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series*. Springer, New York.
- [10] 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.
- [11] Di Fonzo, T., Marini, M. (2009), Simultaneous and Two-step Reconciliation of Systems of Time Series. In *JSM Proceedings*, Business and Economic Statistics Section. Alexandria, VA: American Statistical Association, pp. 803–817.
- [12] Harvill Hood, C.C. (2005), An empirical comparison of methods for benchmarking seasonally adjusted series to annual totals. Paper presented at the

- Workshop on Frontiers in Benchmarking Techniques and their Applications to Official Statistics, Luxembourg, 7-8 April 2005.
- [13] Harville, D.A. (1997), *Matrix Algebra from a Statistician's Perspective*. Springer, New York.
 - [14] Luenberger, D.G. (1984), *Linear and Nonlinear Programming*. Addison-Wesley, Reading.
 - [15] 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.
 - [16] The MathWorks (2009), *Optimization ToolboxTM 4 User's Guide*. Natick, MA.
 - [17] 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 Proceedings*, Business and Economic Statistics Section. Alexandria, VA: American Statistical Association, 3007–3021.
 - [18] 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).
 - [19] U.S. Census Bureau (2009), *X-12-ARIMA Reference manual, Version 0.3*. U.S. Census Bureau, U.S. Department of Commerce, Washington D.C..
<http://www.census.gov/srd/www/x12a/>

Acknowledgements

An early version of this paper was presented at the 45th Scientific Meeting of the Italian Statistical Society, Padua, 16-18 June 2010. A reduced version of this paper has been submitted for publication on a book of the Springer Series 'Studies in Theoretical and Applied Statistics'. Thanks to David Findley and Claudio Lupi for their valuable comments. We alone are responsible for the remaining errors. The views expressed herein are those of the authors and should not be attributed to the IMF, its Executive Board, or its management. Financing from MIUR is also gratefully acknowledged by T. Di Fonzo.

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>

