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

Tommaso Di Fonzo and Marco Marini

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.

Key words: Time series analysis, Combining data from different sources, National economic and social accounts, Benchmarking, Movement preservation, Modified Denton PFD, Causey and Trager GRP

1 Introduction

Benchmarking monthly and quarterly series to annual series is a common practice in many National Statistical Institutes. The benchmarking problem arises when

---

Tommaso Di Fonzo
University of Padua, Department of Statistical Sciences, Via C. Battisti 241, Padova
e-mail: difonzo@stat.unipd.it

Marco Marini
ISTAT, National Accounts Department, Via A. Depretis 74b, Roma e-mail: marco.marini@istat.it

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 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 (e.g. Bloem *et al.*, 2001) 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 obervational 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. $\sum\limits_{t\in T} p_t \neq Y_T$, $T = 1,\dots,N$, we look for benchmarked estimates y_t^b such that $\sum\limits_{t\in T} y_t^b = Y_T$. 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.

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

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

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 of the linear system (Di Fonzo and Marini, 2009)

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

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 $\mathbf{P}^{-1}\mathbf{p} = \mathbf{1}_n$ and $\Delta_n \mathbf{1}_n = \mathbf{0}$, so that it is $\mathbf{Qp} = \mathbf{0}$. In addition, the coefficient matrix of the system 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

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

The solution of the above linear system 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

$$\begin{bmatrix} \mathbf{y}^{PFD} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \end{bmatrix} \begin{bmatrix} \Delta_n' \Delta_n & \mathbf{PC}' \\ \mathbf{CP} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{Y} \end{bmatrix}. \tag{2}$$

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:

$$\min_{y_t} \sum_{t=2}^{n} \left(\frac{y_t}{y_{t-1}} - \frac{p_t}{p_{t-1}} \right)^2 \qquad \text{subject to} \quad \sum_{t \in T} y_t = Y_T, \quad T = 1, \dots, N. \tag{3}$$

Looking at the criterion to be minimized in (3), 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 (2), the minimization problem in (3) is inherently non–linear. Trager (1982; see Bozik and Otto, 1988) suggests to use a technique based on the steepest descent method, using $\mathbf{y}^{PFD}$ as starting values, in order to calculate the benchmarked estimates y_t^{GRP}, $t = 1, \ldots, n$, solution to problem (3). We employ the Sequential Quadratic Programming (SQP) method available in the Optimization Toolbox™ of MATLAB® (version 2009b). It consists of an iterative procedure that splits the problem into easier subproblems to be solved by standard (quadratic) nonlinear programming methods. In each iteration an approximation of the Hessian of the Lagrangian function is calculated using a quasi-Newton updating method, which enables to make informed decisions regarding directions of search and step length.

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):

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

Expression (4) 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 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' $\dfrac{Y_T}{\sum_{t \in T} p_t}$, $T = 1, \ldots, N$ (Bloem *et al.*, 2001), 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. 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 time 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. The results will be evaluated in terms of the ratio of the 'Growth Rate Preservation' criteria computed from the two solutions, that is

$$r = \frac{\sum_{t=2}^{n} \left(\frac{y_t^{GRP}}{y_{t-1}^{GRP}} - \frac{p_t}{p_{t-1}} \right)^2}{\sum_{t=2}^{n} \left(\frac{y_t^{PFD}}{y_{t-1}^{PFD}} - \frac{p_t}{p_{t-1}} \right)^2}. \tag{5}$$

Obviously, we expect the GRP technique always reaches a lower value of the chosen criterion than PFD, and thus the ratio (5) is never larger than 1.

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 2: as expected, the minimum C_{GRP} is achieved by the GRP procedure (441.2 against 1442.8 of PFD, with r=0.306). 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.

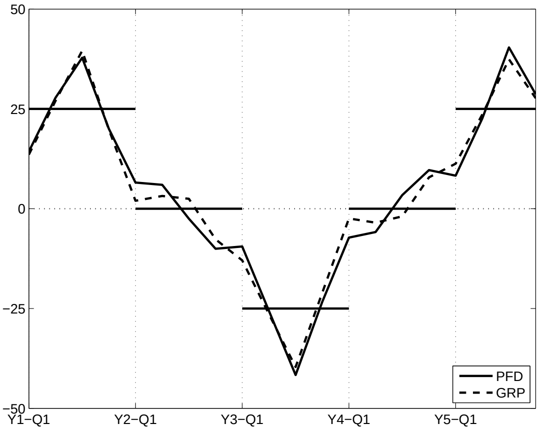

Fig. 1 Adjustments to the artificial series produced by the PFD and GRP procedures

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

The second example is a real-life economic series coming from the European 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. Figure 2 shows 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 ($r = 0.355$). The quarterly adjustments in the two cases are also 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 ($n = 28$).

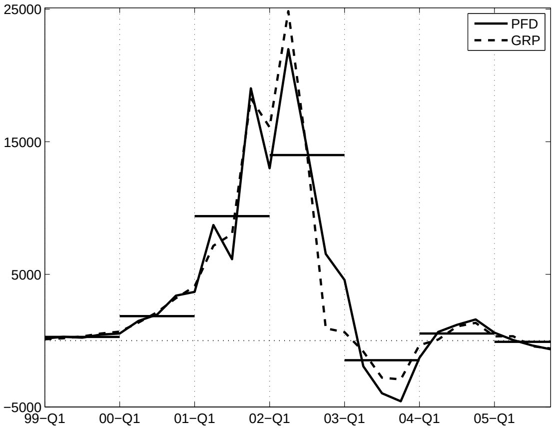

Fig. 2 Adjustments to the real-life series produced by the PFD and GRP procedures

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

$$y_t = \theta^*_t + \mu_t, \qquad t = 1,\dots,n$$

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:

$$p_t = \theta^*_t + e_t, \qquad t = 1,\dots,n$$

where e_t is a Gaussian white noise with variance σ^2_e. 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 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 obtained over the 1,000 experiments.

Table 2 shows the median[1] value of r under different values of σ_e (rows) and μ (columns). As expected, the median ratio is always smaller than 1. However, the PFD procedure provides very similar results to GRP when discrepancies are small and unsystematic (median ratio ≥ 0.9 when $\sigma_e \leq 10$ and $\mu \leq 15$). The reduction is stronger as both σ_e and μ increase. This implies that, when the uncertainty of the preliminary series and/or its bias with respect to the target variable are high, the GRP procedure is likely to provide better results (in terms of growth rates) than PFD.

Table 2 Median value of r for different values of σ_e and μ (across 1,000 experiments)

			μ		
σ_e	0	15	30	45	60
5	0.984	0.969	0.919	0.816	0.583
10	0.936	0.922	0.872	0.770	0.542
15	0.859	0.846	0.799	0.700	0.489
20	0.759	0.744	0.705	0.610	0.429
25	0.641	0.624	0.592	0.507	0.364

References

1. Bloem, A., Dippelsman, R., Mæhle, N.: Quarterly National Accounts Manual. Concepts, Data Sources, and Compilation, International Monetary Fund, Washington DC. (2001)
2. Bozik, J.E., Otto, M.C.: Benchmarking: Evaluating methods that preserve month-to-month changes. Bureau of the Census - Statistical Research Division, RR-88/07 (1988). http://www.census.gov/srd/papers/pdf/rr88-07.pdf
3. Causey, B., Trager, M.L.: Derivation of Solution to the Benchmarking Problem: Trend Revision. Unpublished research notes, U.S. Census Bureau, Washington D.C. (1981)
4. Cholette, P.A.: Adjusting sub-annual series to yearly benchmarks. Survey Methodology **10**, 35–49 (1984)
5. Dagum, E.B., Cholette, P.A.: Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series. Springer, New York (2006)
6. Denton, F.T.: Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization. JASA **333**, 99–102 (1971)
7. Di Fonzo, T., Marini, M.: Simultaneous and Two-step Reconciliation of Systems of Time Series. Joint Statistical Meeting 2009, Proceedings of the Business and Economic Statistics Section, 803–817 (2009)
8. Monsour, N.J., Trager, M.L.: Revision and benchmarking of business time series. American Statistical Association, Proceedings of the Business and Economic Statistics Section, pp. 333-337 (1979)
9. Trager, M.L.: Derivation of Solution to the Benchmarking Problem: Relative Revision. Unpublished research notes, U.S. Census Bureau, Washington D.C. (1982)
10. U.S. Census Bureau, X-12-ARIMA Reference manual, Version 0.3. U.S. Census Bureau, U.S. Department of Commerce, Washington D.C. (2009). http://www.census.gov/srd/www/x12a/

---

[1] The median is more representative than the mean in the case of atypical values. We also calculated mean, standard deviation, minimum, maximum and range of the ratios, available on request from the authors.