Performance measurement with multiple interrelated dimensions of performance and threshold target levels

Panagiotis D. Zervopoulos[1], Theodora S. Brisimi[2], Ali Emrouznejad[3], Gang Cheng[41]

[1] Bursa Orhangazi University, Department of International Trade, 440 Mimar Sinan Mh. Ankara Yolu Cd., Bursa, Turkey

[2] Boston University, Department of Electrical and Computer Engineering, 8 Saint Mary's Street, Boston, Massachusetts 02215, USA

[3] Aston University, Aston Business School, Birmingham, UK

[4] China National Health Development Research Center, 38 Xueyuan Rd, Haidian District, Beijing, China

Abstract

In this study we develop a DEA–based performance measurement methodology that is consistent with performance assessment frameworks such as the Balanced Scorecard. The methodology developed in this paper takes into account the direct or inverse relationships that may exist among the dimensions of performance to construct appropriate production frontiers. The production frontiers we obtain are deemed appropriate as they consist solely of units with desirable levels for all dimensions of performance. These levels should be at least equal to the critical values set by decision makers. The properties and advantages of our methodology against competing methodologies are presented through a numerical example and comparative analysis. This analysis explains the failure of existing studies to define appropriate production frontiers when directly or inversely related dimensions of performance are present.

Keywords: OR in service industries; Performance management; Data envelopment analysis; Balanced Scorecard

[1] Corresponding author: **Gang Cheng,** China National Health Development Research Center, 38 Xueyuan Rd, Haidian District, Beijing, China, maxdea@qq.com

1. Introduction

The performance of modern organizations that operate in competitive marketplaces is based on multiple interrelated dimensions, which are both endogenous (controllable by the organizations) and exogenous (either uncontrollable or partially controllable by the organizations). An important study in the area of performance management is the Balanced Scorecard (BSC) (Kaplan and Norton, 1992). The BSC goes beyond the traditional financial measures for assessing the performance of organizations as it also incorporates customers, internal processes, and learning and growth perspectives (Kaplan and Norton, 1996).

The fundamental drawback of the BSC is the ambiguity of putting the theory into practice by modeling the conceptual framework so as to yield specific and measurable results (Amado et al., 2012). In addition, the link among the dimensions of performance is vague, and the impact on performance of trade-offs that may exist among these dimensions is not explicit (Otley, 1998).

Several studies, most of which are performed within the area of Operational Research, have been published providing scientific underpinning to performance assessment frameworks. Many of these studies use Data Envelopment Analysis (DEA) for evaluating performance. DEA is a nonparametric methodology for assessing the production process of operational units. DEA provides a robust quantitative framework that enables the identification of the strengths and weaknesses of each unit under evaluation and yields measurable results that lead to the optimization of each unit's performance. Since the publication of the seminal paper by Charnes et al. (1978), a significant number of extensions of DEA have been developed (Emrouznejad et al., 2008). A selected list of DEA-based studies related to performance assessment in a multi-dimensional setting is provided in Table 1.

Table 1. DEA-based performance measurement studies

Study	Objectives	Applied method(s)	Outline of the methodology
Lim and Zhu (2013)	Performance measurement when targeted factors are incorporated in the analysis.	DEA	Modification of the radial, slacks-based, and Nerlove-Luenberger measures to treat unequally deviations of the factors from the targets that are selected by decision makers.

Amado et al. (2012)	Identification of the areas within the four perspectives of performance which need improvement.	BSC, DEA	Application of network DEA models to evaluate the four perspectives of performance according to the BSC. The relationships between the perspectives of performance are captured by network DEA.
Paradi et al. (2011)	Performance assessment and identification of firms' inefficiency when multiple dimensions are present.	DEA	A two-stage DEA-based methodology is developed to evaluate firms' performance. In the first stage, conventional DEA is applied to measure units' performance for every single dimension. In the second stage, a slacks-based measure is applied to develop a composite performance index for each unit.
Avkiran and Morita (2010)	Performance measurement based on the interpretations of stakeholders of the same performance metrics.	DEA	A modified range-adjusted super-efficiency metric of efficiency is applied to evaluate firms' performance when taking into account multiple stakeholders' (i.e. shareholders, customers, management, employees and regulators) perspectives.
García-Valderrama et al. (2009)	Development of a framework for the analysis of the relationships between the perspectives of performance.	BSC, DEA	DEA is applied five times to evaluate the relationships between the perspectives of performance in pairs: (a) financial perspective - innovation; (b) innovation - learning & growth; (c) learning & growth - internal processes; (d) internal processes - customers; (e) customers – financial perspective.
Ramanathan and Yunfeng (2009)	Development of a DEA-based framework to facilitate QFD calculations in a multi-dimensional setting.	Quality Function Deployment (QFD), DEA	DEA is applied to evaluate the relative importance of design requirements, which are regarded as decision making units, and the significance of several factors (e.g. customers' perspectives, cost, ease of development, environmental impact) to design requirements.
Eilat et al. (2008)	Evaluation of Research & Development (R&D) projects using multiple criteria.	BSC, DEA	A modified DEA program is applied to evaluate the performance of R&D projects, which provides the option to decision makers to set priorities and bounds to the perspectives of performance. The projects with the lowest performance are excluded from the analysis in order to facilitate the identification of the best-performing projects.
Eilat et al. (2006)	Evaluation of Research & Development (R&D) projects using multiple criteria.	BSC, DEA	A seven-step DEA-based methodology is applied to evaluate alternative portfolios when multiple objectives and possible interactions among the projects are present.
Sherman and Zhu (2006)	Performance measurement that incorporates quality metrics in addition to operational variables.	DEA	DEA and quality metrics are jointly used to measure performance. The units that are efficient but are assigned quality scores lower than a critical value are excluded from the evaluation process as these are not regarded as appropriate benchmarks for the remaining units.

Banker et al. (2004)	Evaluation of trade-offs between performance measures.	DEA	Application of modified DEA models, which do not include constraints for inputs, in conjunction with statistical analysis to define whether performance measures associated with the BSC are inversely or directly related.
Mukherjee et al. (2003)	Analysis of the linkage between resources, service quality and performance.	DEA	Application of DEA to evaluate the relationships between service quality and efficiency, and profitability and efficiency. In the end, an overall efficiency measure is obtained.

The studies presented in Table 1 either do not deal with trade-offs between dimensions of performance or omit a discussion of whether the targets set to a number of dimensions of performance by decision makers are satisfied. As a result, the benchmarking either does not express reality or is not flawless, as the production frontier consists of units for which acceptable scores for all dimensions of performance are not reported. Hence, such units are erroneously regarded as benchmarks for the remaining units.

In this paper, we address the issues raised in the existing studies. We modify and extend the methodology developed by Zervopoulos and Palaskas (2011) to measure performance considering the direct or inverse relationships among the dimensions of performance. The modified methodology relaxes the major assumption of the work of the said authors, which is that of fixed weights between the original and the adjusted variable levels. In addition, the modified methodology evaluates performance when multiple dimensions are present. This is not a straightforward extension of the work of Zervopoulos and Palaskas (2011), according to the analysis presented in Section 3 of this study. In the same Section, the significant differences between the modified and the original methodology are discussed. Beyond the improvements of the new methodology relative to the original, the former methodology ensures the identification of an appropriate production frontier, which consists solely of qualified units in all dimensions of performance.

This study unfolds as follows. Section 2 analyzes the performance assessment methodology. Section 3 justifies the selection of variables, clarifies the underlying relationships among them, and provides a numerical example to present the properties and advantages of our methodology relative to competing approaches. Conclusions, limitations, and future research are presented in the final section of this study.

2. Methodology

The methodology we develop in this paper extends the work of Zervopoulos and Palaskas (2011) to make it applicable to performance assessment frameworks such as the BSC. In particular, this methodology deals with multiple dimensions of performance, which are interrelated, to define a production frontier that consists of units that are efficient and qualified in the exogenous dimensions of performance. The distinction between qualified and disqualified exogenous dimensions of performance is based on a critical value (e.g. $\alpha^* = 0.800$) that is either derived from the measurement scale of the exogenous variables (Zervopoulos and Palaskas, 2011) or is a user-defined value.

Unlike most of the existing studies that do not consider the underlying relationships between the dimensions of performance (e.g. they deal with the dimensions of performance either in pairs or separately (García-Valderrama et al., 2009; Kamakura et al., 2002; Mukherjee et al., 2003), our methodology introduces a unified approach to performance measurement. The dimensions that are incorporated in the measurement of performance are controllable by the unit (e.g. efficiency) and non-controllable, or non-controllable in full, which are dimensions inversely related (e.g. customers' satisfaction) or directly related (e.g. profits) to the controllable dimension of performance. The methodology involves 4 steps (or, essentially, 3 steps, as the first one is applied merely for classification purposes).

To be more precise, Step 1 draws on a modified expression of the directional distance function developed by Cheng and Zervopoulos (2014). This step is applied to the units under assessment only for classification reasons, taking into account the scores of all dimensions of performance in the evaluation.

$$\zeta^* = \min \frac{1-\frac{1}{m}\sum_{i=1}^{m}\beta g_i / x_{io}}{1+\frac{1}{s}\sum_{r=1}^{s}\beta g_r / y_{ro}}$$

$$\begin{aligned} s.t. \ & \sum_{j=1}^{n}\lambda_j x_{ij} + \beta g_x \le x_{io} \\ & \sum_{j=1}^{n}\lambda_j y_{rj} - \beta g_y \ge y_{ro} \\ & \sum_{j=1}^{n}\lambda_j = 1 \\ & \lambda_j \ge 0 \\ & g_x = 1, \ g_y = 0 \end{aligned} \quad (1)$$

where g_x and g_y denote the direction vectors associated with the inputs (x_i) and outputs (y_r), respectively, and the ratios $\beta g_i / x_{io}$ and $\beta g_r / y_{ro}$ express the proportion of the inputs' decrease and outputs' increase, respectively.

If at least one of the exogenous variables that are inversely related to efficiency is disqualified (e.g. when a score is reported that is lower than a selected critical value) then Step 2 should be applied. Step 2 uses a modified formula relative to the one originally presented in Zervopoulos and Palaskas (2011). This stage is regarded as preprocessing, aiming to estimate new (increased) inputs or new (decreased) outputs, depending on the orientation of the analysis, in order to bring disqualified exogenous variables to desirable levels.

In case more than one variables of the same unit that are inversely related to efficiency are disqualified, expressions (2) and (3) should be applied to the variable with the lowest score (i.e. $\min(z_{kh}^d)$, where $h \subset j$).

The efficiency score of the units with disqualified exogenous variables is estimated as follows:

$$\left(\zeta_h\right)^{ad} = \zeta_o + \left(\frac{(\min(z_{kh}^d) - z^*)^2 \cdot (\zeta_o - \zeta_h)^2 \cdot (\min(z_{kh}^d)^* - z^*)^2}{\left((\min(z_{kh}^d) - z^*)^2 + (\zeta_o - \zeta_h)^2\right) \cdot (\min(z_{kh}^d)^* - z^*)^2 - (\min(z_{kh}^d) - z^*)^2 \cdot (\zeta_o - \zeta_h)^2} \right)^{1/2}$$

(2)

where $\left(\zeta_h\right)^{ad} < 1$

In addition, ζ_o and z^* are cut-off levels for the efficiency scores and the scores of the non-controllable variables, which can be omitted from formula (2) if they are considered unnecessary. Moreover, $\min(z_{kh}^d)$ indicates the score of the k th disqualified exogenous variable that is inversely related to efficiency, which mostly deviates from the critical value (e.g. α^* = 0.800), and $\min(z_{kh}^d)^*$ stands for an acceptable score for $\min(z_{kh}^d)$, which lies within a given interval (e.g. $0.8 \le \min(z_{kh}^d)^* \le 1.0$) and is user defined.

Formula (2) satisfies the inverse relationship between some of the exogenous variables (e.g. $\min(z_{kh}^d)$) and efficiency. The managerial interpretation of this relationship is that the utilization of additional resources is a prerequisite for the improvement of some dimensions of performance (e.g. customers' satisfaction) when the outputs and technology are fixed.

In this context, the new (increased) inputs are measured by applying program (3):

$$\max \sum_{i=1}^{m} v_i^{ad} x_{ih}^{ad}$$

$$s.t. \quad \sum_{i=1}^{m} v_i^{ad} x_{ih}^{ad} \ge 1$$

$$\left(\zeta_h\right)^{ad} \sum_{i=1}^{m} v_i^{ad} x_{ih}^{ad} \ge \sum_{r=1}^{s} u_r y_{rh} + u^*$$

$$x_{ih}^{ad} \ge x_{ih}$$

$$x_{ih}^{ad} \leq \left(2-\left(\zeta_h\right)^{ad}\right)x_{ih} \tag{3}$$

$u,\ v^{ad} \geq 0,\ u^*$ is free in sign.

where v_i^{ad} and u_r are input and output multipliers, respectively, and x_{ih}^{ad} denotes the adjusted *i*th input of the *h*th unit.

Program (3) relaxes the major assumption made in the study by Zervopoulos and Palaskas (2011) i.e. that of fixed weights between the original and the adjusted inputs.

The increased inputs (i.e. x_{ih}^{ad}) that are defined from program (3) affect the scores of the remaining exogenous variables, both of the disqualified (i.e. $z_{(k-1)h}^{d}$, where $\min(z_{kh}^{d}) \notin z_{(k-1)h}^{d}$) and the qualified variables (i.e. z_{lh}^{q}), of the adjusted *h* units. The new scores of the remaining disqualified exogenous variables are measured as follows:

$$\left(z_{(k-1)h}^{d}\right)^{ad} = z^* + \left(\frac{(\left(\zeta_h\right)^{ad} - \zeta_o)^2 \cdot (z_{(k-1)h}^{d} - z^*)^2 \cdot (\zeta_o - \zeta_h)^2}{(\left(\zeta_h\right)^{ad} - \zeta_o)^2 \cdot \left((z_{(k-1)h}^{d} - z^*)^2 + (\zeta_o - \zeta_h)^2\right) - (z_{(k-1)h}^{d} - z^*)^2 \cdot (\zeta_o - \zeta_h)^2}\right)^{1/2}$$

(4)

Similarly, the new scores for z_{lh}^{q} are defined as follows:

$$\left(z_{lh}^{q}\right)^{ad} = z^* + \left(\frac{(\left(\zeta_h\right)^{ad} - \zeta_o)^2 \cdot (z_{lh}^{q} - z^*)^2 \cdot (\zeta_o - \zeta_h)^2}{(\left(\zeta_h\right)^{ad} - \zeta_o)^2 \cdot \left((z_{lh}^{q} - z^*)^2 + (\zeta_o - \zeta_h)^2\right) - (z_{lh}^{q} - z^*)^2 \cdot (\zeta_o - \zeta_h)^2}\right)^{1/2} \tag{5}$$

The new scores (i.e. $\left(z_{(k-1)h}^{d}\right)^{ad}$, $\left(z_{lh}^{q}\right)^{ad}$) of the remaining exogenous variables are never lower than a user-defined critical value (α).

In Step 3, the efficiency scores of all the sample units are measured. The dataset that is used for measuring the efficiency scores (η) consists of the original inputs (i.e. $x_{ij,j\neq h}^{or}$) and the adjusted inputs of the *h* units (i.e. x_{ih}^{ad}), which were defined in program (3):

$$
\eta^{*} = \min \frac{1 - \frac{1}{m}\sum_{i=1}^{m} \beta g_i / x_{io}^{T}}{1 + \frac{1}{s}\sum_{r=1}^{s} \beta g_r / y_{ro}}
$$

$$
\begin{aligned}
s.t.\ & \sum_{j=1}^{n} \lambda_j x_{ij}^{T} + \beta g_x \le x_{io}^{T}, \quad x_{ij}^{T} = x_{ij,\, j\neq h}^{or} + x_{ih}^{ad},\ h \subset j \\
& \sum_{j=1}^{n} \lambda_j y_{rj} - \beta g_y \ge y_{ro} \\
& \sum_{j=1}^{n} \lambda_j = 1 \\
& \lambda_j \ge 0 \\
& g_x = 1,\ g_y = 0
\end{aligned} \tag{6}
$$

Formula (2) and program (6) measure two perspectives of efficiency. The formula measures stand-alone efficiency, which is unit specific, while the program measures relative efficiency. The two perspectives of efficiency are used to adjust the exogenous variables that are directly related to efficiency (e.g. profits). For instance, the profits of a unit are negatively affected by the utilization of additional resources, even if the unit's relative efficiency score remains unchanged. In addition, particularly in mature and declining markets, the profits of a unit are negatively affected by relative efficiency deterioration (Oral and Yolalan, 1990). However, this may not always be the case (Taylor et al., 1997).

Formula (7) includes the twofold impact of efficiency changes on the exogenous variable(s) (i.e. b_{uh}) that are directly related to it:

$$
b_{uh}^{ad} = \left(w_{\zeta} \cdot \left(\zeta_h \right)^{ad} + w_{\eta} \cdot \eta_h^{*} \right) \cdot b_{uh}, \quad w_{\zeta} + w_{\eta} = 1; \quad u = 1, \ldots, \mathrm{c} \tag{7}
$$

where w_{ζ} and w_{η} denote user-defined weights that are adjusted to the particular characteristics of the market to which the performance assessment methodology is applied.

The scope of Step 4 is the measurement of performance (θ) of all sample units incorporating the original and adjusted exogenous variables that are directly related to efficiency (i.e. $b_{uj}^{T}=b_{uj,j\neq h}^{or}+b_{uh}^{ad}$) and the original and adjusted exogenous variables that are inversely related to efficiency (i.e. $z_{tj}^{T}=z_{tj,j\neq h}^{or}+\left(z_{(k-1)h}^{d}\right)^{ad}+\min\left(z_{kh}^{d}\right)^{*}+\left(z_{lh}^{q}\right)^{ad}$, where $t=k+l$). Since the exogenous variables that are inversely related to efficiency cannot be regulated, they are regarded as freely disposable in the same way as inputs. In this context, program (8) captures the relationships among all variables, as minimization of inputs aiming at performance optimization is expected to decrease exogenous variables z_{tj}^{T} and increase exogenous variables b_{uj}^{T}.

$$\theta^{*}=\min\ \frac{1-\frac{1}{m}\sum_{i=1}^{m}\beta g_{i}\,/\,x_{io}^{T}}{1+\frac{1}{s+w+p}\left(\sum_{r=1}^{s}\beta g_{r}\,/\,y_{ro}+\sum_{u=1}^{c}\beta g_{u}\,/\,b_{uo}^{T}+\sum_{t=1}^{p}\beta g_{t}\,/\,z_{to}^{T}\right)}$$

$$\begin{aligned}
s.t.\ &\sum_{j=1}^{n}\lambda_{j}x_{ij}^{T}+\beta g_{x}\leq x_{io}^{T}\\
&\sum_{j=1}^{n}\lambda_{j}y_{rj}-\beta g_{y}\geq y_{ro}\\
&\sum_{j=1}^{n}\lambda_{j}b_{uj}^{T}-\beta g_{b}\geq b_{uo}^{T}\\
&\sum_{j=1}^{n}\lambda_{j}z_{tj}^{T}-\beta g_{z}\leq z_{to}^{T}\\
&\sum_{j=1}^{n}\lambda_{j}=1\\
&\lambda_{j}\geq 0\\
&g_{x}=1,\ g_{y}=0,\ g_{b}=1,\ g_{z}=-1
\end{aligned}\tag{8}$$

The best-practice units that are obtained from program (8) always attain qualified scores for their exogenous variables.

Program (8) satisfies the four properties that Tone (2001) regarded as important for designing measures of efficiency, which are: (a) units invariance; (b) monotonicity; (c) translation invariance; (d) reference-set dependence. Proofs for properties (a), (b) and (d) are provided in Tone (2001). Proof for property (c) is provided in Färe and Grosskopf (2010).

3. Numerical example

3.1 Links of the selected variables with performance assessment frameworks

The dataset we use for applying the new performance assessment methodology comes from Greek privatized citizen service centers (Appendix A – Table A1). The dataset consists of 50 units that employ six inputs to produce three outputs. The exogenous variables selected are as follows: (a) citizens' satisfaction (CS), (b) employees' satisfaction (ES), and (c) profits per employee (P/E). The first two exogenous variables were originally measured with a five-point Likert scale and then translated into percentages so that their scale can match that of efficiency (Zervopoulos and Palaskas, 2011). The same scale (i.e. [0.0, 1.0]) was applied to the third exogenous variable using formula (9):

$$\left\{\frac{\text{Profits}_j}{\text{Employees}_j}\right\}_{j=1}^{n} \Bigg/ \left\{\frac{\max\left(\text{Profits}_o\right)}{\text{Employees}_o}\right\} \tag{9}$$

The selected dimensions of performance are in line with the concepts of the BSC. In particular, in the BSC context, profits are an indicator of the financial perspectives (Amado et al., 2012); customers' (or citizens', for the purpose of this study) satisfaction expresses the customers' perspective (Dyson, 2000; Kaplan and Norton, 1996); employees' satisfaction, which encompasses employees' morale and perception of the working environment, indicates the learning and growth perspective (Kaplan and Norton, 1996); finally, efficiency is a measure of the internal perspective (Dyson, 2000). The scope of the BSC is to determine the actions that can lead the firm to long-term success. Similarly, the proposed methodology identifies and measures the appropriate adjustments in the resources in order to accomplish high standards for every dimension of performance.

There is a significant number of studies that evaluate the underlying relationships among the dimensions of performance. In particular, a direct link is present between customers' satisfaction and employees' satisfaction (Kaplan and Norton, 2001; Soteriou and Zenios, 1999). Both measures are perceptions of the quality of services/goods and the quality of the working environment. Higher quality commonly requires additional investments in personnel, in training for personnel, in operating systems, in the reward system, and in the tangibles of an organization. It is inevitable that, in the pursuit of improving customers' and employees' satisfaction, the profits of a firm will decline, especially in the short run, because of the increased cost (Banker et al., 2004; Gustafsson and Johnson, 2002; Kamakura et al., 2002). Profits are directly related to efficiency, regardless of the chosen orientation (downsizing or upsizing). Unlike the relationship between profits and efficiency, customers' and employees' satisfaction are inversely related to efficiency (Anderson et al., 1997).

In case additional dimensions beyond those selected in this paper are incorporated in the assessment of performance, without any information about the underlying relationships among the dimensions, the method developed by Banker et al. (2004) can be applied prior to our methodology to identify the unknown relationships.

In this paper, the dimension 'profits per employee' is used instead of 'profits' in order to remove the size effects from the assessment of performance. Large-sized organizations are more likely to report higher profits than small- or even medium-sized ones. As a result, large-sized firms are likely to be qualified in this particular dimension, unlike small- and medium-sized firms. We remove the size effects from profits by dividing profits by the number of employees, which serves as a proxy for the organization's size (Schilke and Goerzen, 2010).

3.2 Application of the proposed methodology

By applying Step 1 of the present methodology, we classify the units according to their performance in each dimension. High-efficiency units are considered to be only those attaining efficiency scores equal to unity. The user-defined critical value (α^*) for labeling exogenous variables as high-performing (H), which is identical to "qualified", is 0.800. This value is obtained from the transformation of the five-point Likert scale into percentages and the translation of satisfied customers' and employees' scores into a qualified status for the variables (Zervopoulos and Palaskas, 2011). The same critical value is applied to profits per employee. In this context, in Table 2, variables with scores equal to or greater than 0.800 are classified as high-performing (H), and those with scores lower than 0.800 are classified as low-performing (L). Table 2 illustrates the results of Step 1. According to these results, 14 units out of 50 are found to be top performers in all dimensions. The units listed in bold (i.e. 19 units out of 50) need adjustment because at least one of their exogenous variables that is inversely related to efficiency is disqualified (i.e. CS < 0.800 or/and ES < 0.800).

Table 2. Classification of the units

Units	Eff. (ζ)	CS	ES	P/E	Status	Units	Eff. (ζ)	CS	ES	P/E	Status
1	1.000	0.923	0.945	0.970	HE-HCS-HES-HPE	26	1.000	0.816	0.863	0.958	HE-HCS-HES-HPE
2	0.965	0.930	0.901	0.970	LE-HCS-HES-HPE	**27**	1.000	0.836	**0.790**	0.894	HE-HCS-LES-HPE
3	1.000	0.943	0.898	0.983	HE-HCS-HES-HPE	**28**	1.000	0.801	**0.697**	0.968	HE-HCS-LES-HPE
4	0.858	0.821	0.891	0.897	LE-HCS-HES-HPE	29	1.000	0.914	0.850	0.894	HE-HCS-HES-HPE
5	0.800	0.860	0.830	0.654	LE-HCS-HES-LPE	30	1.000	0.933	0.966	0.890	HE-HCS-HES-HPE
6	0.772	0.874	0.890	0.835	LE-HCS-HES-HPE	**31**	1.000	**0.779**	**0.733**	0.986	HE-LCS-LES-HPE
7	0.778	0.819	**0.752**	0.855	LE-HCS-LES-HPE	**32**	1.000	**0.776**	0.812	0.979	HE-LCS-HES-HPE
8	0.679	0.870	0.829	0.699	LE-HCS-HES-LPE	**33**	0.722	**0.790**	**0.770**	0.868	LE-LCS-LES-HPE
9	0.730	0.873	0.900	0.796	LE-HCS-HES-LPE	34	0.754	0.934	0.978	0.769	LE-HCS-HES-LPE
10	1.000	0.811	**0.760**	0.931	HE-HCS-LES-HPE	35	1.000	0.906	0.825	0.933	HE-HCS-HES-HPE
11	0.831	**0.781**	**0.770**	0.895	LE-LCS-LES-HPE	36	0.870	0.841	0.812	0.814	LE-HCS-HES-HPE
12	0.832	0.864	0.890	0.868	LE-HCS-HES-HPE	**37**	1.000	0.823	**0.790**	0.967	HE-HCS-LES-HPE
13	0.693	**0.793**	0.810	0.599	LE-LCS-HES-LPE	**38**	1.000	0.811	**0.733**	0.979	HE-HCS-LES-HPE
14	1.000	0.969	0.911	0.933	HE-HCS-HES-HPE	39	0.996	0.817	0.867	0.837	LE-HCS-HES-HPE
15	1.000	0.950	0.897	1.000	HE-HCS-HES-HPE	40	1.000	0.961	0.922	0.936	HE-HCS-HES-HPE
16	0.739	0.943	0.981	0.710	LE-HCS-HES-LPE	**41**	1.000	**0.790**	0.890	0.974	HE-LCS-HES-HPE
17	1.000	0.904	0.849	0.864	HE-HCS-HES-HPE	**42**	0.823	**0.769**	0.815	0.902	LE-LCS-HES-HPE
18	0.776	0.927	0.908	0.677	LE-HCS-HES-LPE	43	0.867	0.846	0.899	0.889	LE-HCS-HES-HPE
19	1.000	0.947	0.956	0.947	HE-HCS-HES-HPE	**44**	0.731	0.823	**0.728**	0.809	LE-HCS-LES-HPE
20	1.000	0.945	0.920	0.989	HE-HCS-HES-HPE	45	0.859	0.885	0.817	0.933	LE-HCS-HES-HPE
21	1.000	0.969	0.860	0.991	HE-HCS-HES-HPE	46	0.850	0.947	0.882	0.929	LE-HCS-HES-HPE
22	0.833	0.808	**0.745**	0.935	LE-HCS-LES-HPE	47	0.755	0.920	0.981	0.525	LE-HCS-HES-LPE
23	1.000	0.808	**0.720**	0.999	HE-HCS-LES-HPE	48	0.745	0.956	0.982	0.617	LE-HCS-HES-LPE
24	1.000	0.810	**0.763**	0.979	HE-HCS-LES-HPE	**49**	1.000	**0.666**	**0.592**	1.000	HE-LCS-LES-HPE
25	1.000	0.872	0.831	0.916	HE-HCS-HES-HPE	**50**	1.000	**0.694**	**0.733**	0.990	HE-LCS-LES-HPE

The 19 units disqualified in at least one exogenous variable that is inversely related to efficiency are introduced in Step 2. The adjustment of the efficiency score (ζ), and of the scores of both disqualified and qualified exogenous variables of the same unit is based on models (2) - (5). The critical value (i.e. α^{*} = 0.800) was arbitrarily selected as the target for the improvement of the originally disqualified variable that deviated mostly from the critical value. However, in this study, any target score that lies within the interval [0.8, 1.0] could be set. The results obtained from the adjustment process of efficiency, CS and ES scores are displayed in Table 3. After the adjustment, there is no disqualified exogenous variable, while the efficiency scores ($(\zeta_h)^{ad}$) fall short of the original efficiency scores (ζ_h).

The adjusted scores that are assigned to efficiency and the exogenous variables, which are inversely related to efficiency, measured according to the methodology of Zervopoulos and Palaskas (2011), are illustrated in Table 3 under the heading adjusted scores (2). Based on the formula in their work, which corresponds to our formula (2), the adjustments were made only to the variables of the units with: (a) efficiency score (ζ) equal to unity and (b) at least one exogenous variable among those that are inversely related to efficiency, which does not satisfy the critical value (i.e. α^{*}). In the above study, there is no explicit reference to the exogenous variable that should be adjusted first in case more than one exogenous variables score lower than the critical value. We tackled this problem by using $\min(z_{kh}^{d})$ as the adjustment starting point, according to formula (2) discussed in Section 2 of this paper.

Table 3. Adjustment of efficiency and exogenous variables that are inversely related to efficiency

Units	Original Scores			Adjusted Scores (1)			Adjusted Scores (2)		
	Eff. (ζ)	CS	ES	Eff. $(\zeta)^{ad}$	CS^{ad}	ES^{ad}	Eff. $(\zeta)^{ad}$	CS^{ad}	ES^{ad}
7	0.778	0.819	0.752	0.735	0.888	0.800	-	-	-
10	1.000	0.811	0.760	0.912	0.864	0.800	0.912	0.864	0.800
11	0.831	0.781	0.770	0.796	0.814	0.800	-	-	-
13	0.693	0.793	0.810	0.689	0.800	0.818	-	-	-
22	0.833	0.808	0.745	0.769	0.888	0.800	-	-	-
23	1.000	0.808	0.720	0.834	0.948	0.800	0.834	0.948	0.800

24	1.000	0.810	0.763	0.917	0.859	0.800	0.917	0.859	0.800
27	1.000	0.836	0.790	0.978	0.848	0.800	0.978	0.848	0.800
28	1.000	0.801	0.697	0.794	1.000	0.800	0.794	1.000	0.800
31	1.000	0.779	0.733	0.858	0.869	0.800	0.858	0.869	0.800
32	1.000	0.776	0.812	0.946	0.800	0.841	0.946	0.800	0.841
33	0.722	0.790	0.770	0.701	0.823	0.800	-	-	-
37	1.000	0.823	0.790	0.977	0.835	0.800	0.977	0.835	0.800
38	1.000	0.811	0.733	0.859	0.919	0.800	0.859	0.919	0.800
41	1.000	0.790	0.890	0.978	0.800	0.906	0.978	0.800	0.906
42	0.823	0.769	0.815	0.788	0.800	0.854	-	-	-
44	0.731	0.823	0.728	0.679	0.952	0.800	-	-	-
49	1.000	0.666	0.592	0.635	1.000	0.800	0.635	1.000	0.800
50	1.000	0.694	0.733	0.789	0.800	0.874	0.789	0.800	0.874

(1): Adjusted scores are obtained from the methodology presented in Section 2

(2): Adjusted scores are obtained from the methodology presented in Zervopoulos and Palaskas (2011)

Taking into account the inverse relationship between CS and ES, and efficiency and also the input orientation of the analysis, we note that the increase in the CS and ES scores requires the employment of additional resources. The new inputs (x_{ih}^{ad}), obtained from a modified version of program (3), are presented in Table 4 under heading (1). The modification incorporates an additional constraint that is associated with weekly working hours (x_{2h}^{ad}) which, according to a directive of the Ministry of Administrative Reform, cannot be more than 66. The modified program (3) is as follows:

$$\begin{aligned}
&\max \sum_{i=1}^{m} v_i^{ad} x_{ih}^{ad} \\
&s.t. \quad \sum_{i=1}^{m} v_i^{ad} x_{ih}^{ad} \geq 1 \\
&\quad \left(\zeta_h\right)^{ad} \sum_{i=1}^{m} v_i^{ad} x_{ih}^{ad} \geq \sum_{r=1}^{s} u_r y_{rh} + u^* \\
&\quad x_{ih}^{ad} \geq x_{ih} \\
&\quad x_{ih,i\neq 2}^{ad} \leq \left(2 - \left(\zeta_h\right)^{ad}\right) x_{ih,i\neq 2} \\
&\quad x_{2h}^{ad} \leq 66
\end{aligned} \tag{10}$$

The values in brackets denote the change (increase in 70 out of 114 cases) in the new input levels compared to the original input levels. In Table 4, the adjusted inputs obtained from the methodology of Zervopoulos and Palaskas (2011) are displayed under heading (2). Drawing on the formula that these two authors used to adjust the input levels (i.e. $x_{ih}^{ad} = 1/\left(\zeta_h\right)^{ad} \cdot x_{ih}$), the adjusted inputs cannot be controlled in order to satisfy the constraint for the weekly working hours.

Table 4. Adjusted inputs

Units	Employees (1)		Employees (2)	Hours (1)		Hours (2)	PC (1)		PC (2)	FAX (1)		FAX (2)	Printer (1)		Printer (2)	Surface (1)		Surface (2)
7	[3]	16	-	[0]	66	-	[3]	15	-	[1]	4	-	[2]	10	-	[19]	119	-
10	[0]	5	5	[25]	55	33	[0]	5	5	[0]	1	1	[0]	2	2	[6]	86	88
11	[1]	6	-	[5]	65	-	[1]	6	-	[0]	1	-	[1]	5	-	[11]	81	-
13	[3]	14	-	[0]	66	-	[5]	24	-	[1]	4	-	[3]	12	-	[20]	110	-
22	[1]	7	-	[25.3]	55	-	[2]	11	-	[0]	2	-	[1]	4	-	[16]	106	-
23	[1]	7	7	[0]	66	79	[1]	9	10	[0]	0	0	[0]	1	1	[29]	279	300
24	[0]	5	5	[7.5]	65	63	[1]	8	8	[0]	1	1	[0]	2	2	[7]	107	109
27	[0]	5	5	[0]	66	68	[0]	7	7	[0]	2	2	[0]	3	3	[2]	92	92
28	[1]	6	6	[0]	66	83	[2]	14	15	[0]	0	0	[1]	5	5	[16]	116	126
31	[1]	6	6	[23]	56	38	[1]	5	5	[0]	0	0	[0]	2	2	[6]	56	58
32	[1]	19	19	[3]	66	67	[1]	15	15	[0]	2	2	[0]	4	4	[4]	84	85
33	[2]	8	-	[0]	66	-	[2]	9	-	[0]	1	-	[1]	4	-	[18]	98	-
37	[0]	5	5	[24.75]	56	32	[0]	6	6	[0]	0	0	[0]	1	1	[1]	46	46
38	[5]	44	45	[3]	66	73	[3]	25	26	[1]	5	5	[1]	8	8	[16]	166	175
41	[0]	5	5	[20.5]	58	38	[0]	9	9	[0]	1	1	[0]	3	3	[2]	82	82
42	[1]	5	-	[0]	66	-	[1]	5	-	[0]	1	-	[0]	2	-	[19]	139	-
44	[2]	7	-	[3]	66	-	[2]	8	-	[0]	1	-	[1]	4	-	[23]	123	-
49	[1]	4	5	[21.5]	58	57	[1]	3	3	[0]	0	0	[0]	1	2	[36]	186	236
50	[1]	5	5	[23.5]	56	41	[0]	2	3	[0]	1	1	[0]	2	3	[26]	206	228

(1): Adjusted inputs are obtained from the methodology presented in Section 2

(2): Adjusted scores are obtained from the methodology presented in Zervopoulos and Palaskas (2011)

The new adjusted inputs replace the original inputs in the dataset, and we proceed to the application of Step 3 in order to determine the new efficiency scores (η). The results of Step 3 are illustrated in Table 5. The units listed in bold indicate those for which adjustment was made.

Unlike the efficiency score ($(\zeta)^{ad}$), which was calculated through a unit-specific procedure, the efficiency score (η) is obtained from a relative evaluation of the production process of the sample units, which is based on program (6).

Step 3 does not exist in the methodology of Zervopoulos and Palaskas (2011).

Table 5. New efficiency scores

Units	Eff. (ζ)	Eff. (η)	Units	Eff. (ζ)	Eff. (η)	Units	Eff. (ζ)	Eff. (η)	Units	Eff. (ζ)	Eff. (η)
1	1.000	1.000	14	1.000	1.000	**27**	1.000	1.000	40	1.000	1.000
2	0.965	1.000	15	0.965	1.000	**28**	1.000	1.000	**41**	1.000	1.000
3	1.000	1.000	16	1.000	1.000	29	1.000	1.000	**42**	0.823	0.782
4	0.858	0.934	17	0.858	0.934	30	1.000	1.000	43	0.867	0.914
5	0.800	0.833	18	0.800	0.833	**31**	1.000	1.000	**44**	0.731	0.711
6	0.772	0.821	19	0.772	0.821	**32**	1.000	1.000	45	0.859	0.916
7	0.778	0.756	20	0.778	0.756	**33**	0.722	0.720	46	0.850	1.000
8	0.679	0.732	21	0.679	0.732	34	0.754	0.834	47	0.755	0.782
9	0.730	0.769	**22**	0.730	0.769	35	1.000	1.000	48	0.745	0.778
10	1.000	0.784	**23**	1.000	0.784	36	0.870	0.873	**49**	1.000	1.000
11	0.831	0.751	**24**	0.831	0.751	**37**	1.000	1.000	**50**	1.000	1.000
12	0.832	0.966	25	0.832	0.966	**38**	1.000	1.000			
13	0.693	0.685	26	0.693	0.685	39	0.996	1.000			

The adjustment process led to a decrease in the efficiency score (η) for 9 out of the 19 units that were originally disqualified. The originally disqualified units are displayed in bold in Table 5. Essentially, the adjustment process changed the production frontier. To be more precise, prior to the adjustment, the frontier consisted of 19 units, while after the adjustment, the number of units located on the frontier increased to 21.

The preprocessing stage ends with the adjustment of profits per employee, which are directly related to efficiency. The adjusted profits-per-employee ratio is less than the original ratio when the input-oriented approach is followed. The proof of this conclusion is straightforward and is provided in Appendix B. We used formula (7), with $w_{\zeta} = w_{\eta} = 0.5$, to measure the adjusted profits per employee. In Table 6, we present the original (P/E) variable and the adjusted $(P/E)^{ad(1)}$ variable. The measurement of the $(P/E)^{ad(2)}$ variable draws on the methodology of Zervopoulos and Palaskas (2011). In particular, the $(\text{P/E})^{\text{ad}(2)}$ scores are defined as follows: $(\text{P/E})^{\text{ad}(2)} = (\zeta)^{\text{ad}} \cdot (\text{P/E})$.

Table 6. Profits per employee (P/E)

Units	(P/E)	$(\text{P/E})^{\text{ad}(1)}$	$(\text{P/E})^{\text{ad}(2)}$	Units	(P/E)	$(\text{P/E})^{\text{ad}(1)}$	$(\text{P/E})^{\text{ad}(2)}$	Units	(P/E)	$(\text{P/E})^{\text{ad}(1)}$	$(\text{P/E})^{\text{ad}(2)}$
7	0.855	0.637	-	27	0.894	0.884	0.874	41	0.974	0.963	0.952
10	0.931	0.790	0.849	28	0.968	0.868	0.768	42	0.902	0.708	-
11	0.895	0.692	-	31	0.986	0.916	0.846	44	0.809	0.562	-
13	0.599	0.412	-	32	0.979	0.953	0.927	49	1.000	0.818	0.635
22	0.935	0.669	-	33	0.868	0.617	-	50	0.990	0.886	0.781
23	0.999	0.916	0.833	37	0.967	0.956	0.945				
24	0.979	0.939	0.898	38	0.979	0.910	0.840				

(1): Adjusted P/E obtained from the methodology presented in Section 2

(2): Adjusted P/E obtained from the methodology presented in Zervopoulos and Palaskas (2011)

The original and adjusted inputs and exogenous variables obtained from the methodology presented in Section 2 and from that of Zervopoulos and Palaskas (2011) are introduced in program (8) to assess performance. Column 2 in Table 7 displays the performance (θ) of all units measured by the two methodologies (headings (1) and (2), respectively). In the following columns, the adjusted and target scores of the three exogenous variables are presented. The target scores are projections of the adjusted scores, which will be attained when the units become top performing (i.e. $\theta = 1.000$).

Drawing on the results of Table 7, there is no benchmark unit identified from the methodology discussed in Section 2 which is disqualified in any of the exogenous variables. Therefore, the methodology we applied managed to construct an appropriate frontier for the assessment of performance of the units that lie below this frontier. In addition, it enables some units to become top performing (e.g. 23, 24, 27), reporting scores for their exogenous variables that, at a minimum, meet the critical value ($\alpha^* = 0.800$). Taking into consideration the underlying relationships among the numerous dimensions of performance, it is unlikely to construct a frontier that consists of units that are qualified in every dimension of performance without the application of the preprocessing stage (formulas (2) - (5)). The methodology of Zervopoulos and

Palaskas (2011) failed to identify qualified benchmark units. For instance, units 7, 28, and 49 are assigned performance scores equal to unity while the target ES score for unit 7 and the target P/E scores for units 28 and 49 lie below the critical value (i.e. $\alpha^* = 0.800$).

In Table 7, focusing on the results of the methodology presented in Section 2, we see that some units (e.g. 4, 43), which are not regarded as benchmarks but are qualified in CS, ES, and P/E, remain qualified when they are projected to the frontier. However, most of the non-benchmark units will not be able to attain acceptable scores for their exogenous variables when they are projected to the frontier. The units for which target scores are reported that are lower than the critical value for CS and ES but greater than the critical value for P/E can invest in the improvement of the working environment (e.g. provision of in-house training, development of a reward system) in order to increase employees' satisfaction. To increase customers' satisfaction, given that the resources employed will be decreased in order to attain top performance, the decision makers should place emphasis on intangibles (e.g. politeness, readiness to provide the services asked for, additional training of the personnel) and tangibles (e.g. newly restored working area, appropriate room temperature). In case decision makers are willing to increase resources of the units that lie below the frontier and their exogenous variables do not meet the targeted critical value, then for these particular units, formulas (2) - (7) can be applied, by setting a new critical value higher than the initial targeted critical value ($\alpha^{**} > \alpha^*$).

Table 7. Performance measurement

Units	Performance (θ)		CS			ES			P/E		
	(1)	(2)	Adjusted$^{(1)}$	Target$^{(1)}$	Target$^{(2)}$	Adjusted$^{(1)}$	Target$^{(1)}$	Target$^{(2)}$	Adjusted$^{(1)}$	Target$^{(1)}$	Target$^{(2)}$
1	1.000	1.000	0.923	0.923	0.923	0.945	0.945	0.945	0.970	0.970	0.970
2	1.000	1.000	0.930	0.930	0.930	0.901	0.901	0.901	0.970	0.970	0.970
3	1.000	1.000	0.943	0.943	0.943	0.898	0.898	0.898	0.983	0.983	0.983
4	0.993	0.972	0.821	0.814	0.793	0.891	0.817	0.839	0.897	0.905	0.925
5	0.872	0.879	0.860	0.705	0.701	0.830	0.689	0.697	0.654	0.795	0.787
6	0.943	0.956	0.874	0.796	0.814	0.890	0.804	0.823	0.835	0.912	0.894
7	0.897	1.000	0.888	0.677	0.819	0.800	0.671	0.752	0.637	0.766	0.855
8	0.892	0.891	0.870	0.726	0.732	0.829	0.703	0.702	0.699	0.825	0.826
9	0.913	0.914	0.873	0.769	0.770	0.900	0.796	0.796	0.796	0.901	0.900
10	0.927	0.962	0.864	0.751	0.817	0.800	0.727	0.764	0.790	0.862	0.886
11	0.891	1.000	0.814	0.699	0.781	0.800	0.685	0.770	0.692	0.807	0.895
12	0.981	0.956	0.864	0.844	0.818	0.890	0.797	0.798	0.868	0.888	0.913

13	0.800	0.875	0.800	0.555	0.642	0.818	0.574	0.659	0.412	0.656	0.750
14	1.000	1.000	0.969	0.969	0.969	0.911	0.911	0.911	0.933	0.933	0.933
15	1.000	1.000	0.950	0.950	0.950	0.897	0.897	0.897	1.000	1.000	1.000
16	0.844	0.847	0.943	0.759	0.763	0.981	0.797	0.800	0.710	0.894	0.891
17	1.000	1.000	0.904	0.904	0.904	0.849	0.849	0.849	0.864	0.864	0.864
18	0.847	0.852	0.927	0.754	0.740	0.908	0.735	0.743	0.677	0.850	0.843
19	1.000	1.000	0.947	0.947	0.947	0.956	0.956	0.956	0.947	0.947	0.947
20	1.000	1.000	0.945	0.945	0.945	0.920	0.920	0.920	0.989	0.989	0.989
21	1.000	1.000	0.969	0.969	0.969	0.860	0.860	0.860	0.991	0.991	0.991
22	0.891	1.000	0.888	0.711	0.808	0.800	0.672	0.745	0.669	0.797	0.935
23	1.000	1.000	0.948	0.948	0.948	0.800	0.800	0.800	0.916	0.916	0.833
24	1.000	1.000	0.859	0.859	0.859	0.800	0.800	0.800	0.939	0.939	0.898
25	1.000	1.000	0.872	0.872	0.872	0.831	0.831	0.831	0.916	0.916	0.916
26	1.000	1.000	0.816	0.816	0.816	0.863	0.863	0.863	0.958	0.958	0.958
27	1.000	1.000	0.848	0.848	0.848	0.800	0.800	0.800	0.884	0.884	0.874
28	1.000	1.000	1.000	1.000	1.000	0.800	0.800	0.800	0.868	0.868	0.768
29	1.000	1.000	0.914	0.914	0.914	0.850	0.850	0.850	0.894	0.894	0.894
30	1.000	1.000	0.933	0.933	0.933	0.966	0.966	0.966	0.890	0.890	0.890
31	1.000	1.000	0.869	0.869	0.869	0.800	0.800	0.800	0.916	0.916	0.846
32	1.000	1.000	0.800	0.800	0.800	0.841	0.841	0.841	0.953	0.953	0.927
33	0.859	0.980	0.823	0.670	0.770	0.800	0.647	0.750	0.617	0.770	0.888
34	0.873	0.869	0.934	0.791	0.785	0.978	0.810	0.830	0.769	0.913	0.918
35	1.000	1.000	0.906	0.906	0.906	0.825	0.825	0.825	0.933	0.933	0.933
36	0.941	0.945	0.841	0.758	0.753	0.812	0.748	0.753	0.814	0.878	0.873
37	1.000	1.000	0.835	0.835	0.835	0.800	0.800	0.800	0.956	0.956	0.945
38	1.000	1.000	0.919	0.919	0.919	0.800	0.800	0.800	0.910	0.910	0.840
39	0.962	0.964	0.817	0.781	0.782	0.867	0.780	0.774	0.837	0.873	0.872
40	1.000	1.000	0.961	0.961	0.961	0.922	0.922	0.922	0.936	0.936	0.936
41	1.000	1.000	0.800	0.800	0.800	0.906	0.906	0.906	0.963	0.963	0.952
42	0.897	1.000	0.800	0.694	0.769	0.854	0.726	0.815	0.708	0.813	0.902
43	0.957	0.958	0.846	0.802	0.803	0.899	0.838	0.843	0.889	0.934	0.932
44	0.840	0.971	0.952	0.649	0.750	0.800	0.621	0.701	0.562	0.741	0.836
45	1.000	0.998	0.885	0.885	0.841	0.817	0.817	0.816	0.933	0.933	0.934
46	1.000	0.997	0.947	0.947	0.908	0.882	0.882	0.879	0.929	0.929	0.932
47	0.782	0.784	0.920	0.669	0.671	0.981	0.710	0.732	0.525	0.776	0.773
48	0.812	0.815	0.956	0.731	0.735	0.982	0.758	0.762	0.617	0.842	0.837
49	1.000	1.000	1.000	1.000	1.000	0.800	0.800	0.800	0.818	0.818	0.635
50	1.000	0.961	0.800	0.800	0.763	0.874	0.874	0.728	0.886	0.886	0.819

(1): Scores obtained from the methodology presented in Section 2

(2): Scores obtained from the methodology presented in Zervopoulos and Palaskas (2011)

The appropriateness of our methodology, which was discussed in Section 2 of this paper, for identifying qualified units in all dimensions of performance as benchmarks, while interrelated dimensions of performance exist, is also illustrated in Table 8. The results of the new methodology are tested against those obtained from the methodology developed by Lim and Zhu (2013). These authors developed a performance measurement methodology which is applicable when there are target levels for some variables, which the units should strive to achieve. Similarly, in this study, target levels were set for CS, ES and P/E (i.e. $\alpha^* = 0.800$), with the aim of not having any of the units under evaluation report scores for the exogenous variables that are lower than the target level.

Further, we developed three scenarios to cross-check the results of our methodology against those based on the methodology of Lim and Zhu (2013). In particular, in scenario 1, the program used was the following:

$$
\begin{aligned}
&\min \theta / \varphi \\
&s.t. \sum_{j=1}^{n} \lambda_j x_{ij}^T \leq \theta x_{io}^T \\
&\quad \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro} \\
&\quad \sum_{j=1}^{n} \lambda_j b_{uj}^T \geq \varphi b_{uo}^T \\
&\quad \sum_{j=1}^{n} \lambda_j \left| z_{tj}^T - \omega_t \right| \leq \theta \left| z_{to}^T - \omega_t \right| \\
&\quad \sum_{j=1}^{n} \lambda_j = 1 \\
&\quad \lambda_j \geq 0
\end{aligned}
\tag{11}
$$

where $\omega_t = 0.800$. The ratio θ/φ expresses the performance of the units under evaluation.

In scenarios 2 and 3, the program used was as follows:

$$\min \theta / \varphi$$

$$s.t. \sum_{j=1}^{n} \lambda_j x_{ij}^T \le \theta x_{io}^T$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{ro}$$

$$\sum_{j=1}^{n} \lambda_j \left| b_{uj}^T - \omega_u \right| \ge \varphi \left| b_{uo}^T - \omega_u \right|$$

$$\sum_{j=1}^{n} \lambda_j z_{tj}^T \le \theta z_{to}^T$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \ge 0 \tag{12}$$

where, in scenario 2: $\omega_u = 0.800$, and in scenario 3: $\omega_u = 0.956$. In scenario 3, the value of ω_u is the average P/E score of all the efficient units according to program (1) (Table 2). Scenario 3 was included in the analysis to satisfy the managerial and microeconomic rationale which underlies profit maximization. From such a perspective, it is reasonable that the target score for the P/E ratio is equal to the average P/E scores of the best-practice units rather than being an approximation of the average P/E of all the units under evaluation.

Table 8. Comparative analysis

Units	Performance				Targets								
	Program	Scenarios			Scenario 1			Scenario 2			Scenario 3		
	(8)	1	2	3	CS	ES	P/E	CS	ES	P/E	CS	ES	P/E
1	1.000	1.000	1.000	1.000	0.677	0.655	0.970	0.923	0.945	0.630	0.923	0.945	0.946
2	1.000	1.000	1.000	0.177	0.670	0.699	0.970	0.930	0.901	0.630	0.834	0.813	0.899
3	1.000	1.000	1.000	1.000	0.657	0.702	0.983	0.943	0.898	0.617	0.943	0.898	0.933
4	0.993	0.895	0.497	0.770	0.780	0.727	0.960	0.788	0.757	0.612	0.821	0.840	0.874
5	0.872	0.515	0.652	1.000	0.778	0.777	0.974	0.817	0.830	0.576	0.860	0.830	0.650
6	0.943	0.752	0.168	1.000	0.765	0.735	0.965	0.823	0.795	0.603	0.874	0.890	0.830
7	0.897	0.779	0.296	1.000	0.784	0.758	0.975	0.756	0.752	0.615	0.819	0.752	0.851
8	0.892	0.611	0.463	1.000	0.758	0.777	0.936	0.826	0.817	0.584	0.870	0.829	0.695

9	0.913	0.661	0.018	0.702	0.741	0.719	0.976	0.873	0.867	0.579	0.873	0.871	0.722
10	0.927	1.000	1.000	1.000	0.789	0.760	0.931	0.811	0.760	0.669	0.811	0.760	0.927
11	0.891	1.000	0.480	0.863	0.781	0.770	0.895	0.781	0.743	0.603	0.781	0.752	0.880
12	0.981	0.883	0.370	1.000	0.740	0.716	0.922	0.864	0.840	0.616	0.864	0.890	0.864
13	0.800	1.000	0.950	1.000	0.793	0.790	0.599	0.793	0.791	0.589	0.793	0.810	0.595
14	1.000	0.915	0.709	1.000	0.786	0.720	0.980	0.787	0.720	0.620	0.969	0.911	0.928
15	1.000	1.000	1.000	1.000	0.650	0.703	1.000	0.950	0.897	0.600	0.950	0.897	0.916
16	0.844	0.509	0.350	0.895	0.709	0.742	0.966	0.773	0.739	0.590	0.869	0.924	0.677
17	1.000	1.000	1.000	1.000	0.696	0.751	0.864	0.904	0.849	0.736	0.904	0.849	0.860
18	0.847	0.480	0.488	0.872	0.780	0.742	0.978	0.869	0.908	0.549	0.869	0.908	0.632
19	1.000	1.000	1.000	1.000	0.653	0.644	0.947	0.947	0.956	0.653	0.947	0.956	0.943
20	1.000	1.000	1.000	1.000	0.655	0.680	0.989	0.945	0.920	0.611	0.945	0.920	0.927
21	1.000	1.000	1.000	1.000	0.631	0.740	0.991	0.969	0.860	0.609	0.969	0.860	0.925
22	0.891	1.000	1.000	1.000	0.792	0.745	0.935	0.808	0.745	0.665	0.808	0.745	0.931
23	1.000	1.000	1.000	1.000	0.792	0.720	0.999	0.808	0.720	0.601	0.808	0.720	0.917
24	1.000	1.000	1.000	1.000	0.790	0.763	0.979	0.810	0.763	0.621	0.810	0.763	0.937
25	1.000	1.000	1.000	1.000	0.728	0.769	0.916	0.872	0.831	0.684	0.872	0.831	0.912
26	1.000	1.000	1.000	1.000	0.784	0.737	0.958	0.816	0.863	0.642	0.816	0.863	0.954
27	1.000	1.000	1.000	1.000	0.764	0.790	0.894	0.836	0.790	0.706	0.836	0.790	0.889
28	1.000	1.000	1.000	1.000	0.799	0.697	0.968	0.801	0.697	0.632	0.801	0.697	0.948
29	1.000	1.000	1.000	1.000	0.686	0.750	0.894	0.914	0.850	0.706	0.914	0.850	0.890
30	1.000	1.000	1.000	1.000	0.667	0.634	0.890	0.933	0.966	0.710	0.933	0.966	0.886
31	1.000	1.000	1.000	1.000	0.779	0.733	0.986	0.779	0.733	0.614	0.779	0.733	0.930
32	1.000	1.000	1.000	1.000	0.776	0.788	0.979	0.776	0.812	0.621	0.776	0.812	0.936
33	0.859	1.000	0.325	0.657	0.790	0.770	0.868	0.790	0.765	0.590	0.790	0.767	0.816
34	0.873	0.673	0.140	0.995	0.712	0.740	0.970	0.819	0.763	0.605	0.869	0.896	0.764
35	1.000	1.000	1.000	1.000	0.694	0.775	0.933	0.906	0.825	0.667	0.906	0.825	0.929
36	0.941	0.788	0.067	1.000	0.777	0.788	0.971	0.830	0.812	0.598	0.841	0.812	0.809
37	1.000	1.000	1.000	1.000	0.777	0.790	0.967	0.823	0.790	0.633	0.823	0.790	0.948
38	1.000	1.000	1.000	1.000	0.789	0.733	0.979	0.811	0.733	0.621	0.811	0.733	0.937
39	0.962	1.000	1.000	1.000	0.783	0.733	0.837	0.817	0.867	0.763	0.817	0.867	0.833
40	1.000	1.000	1.000	1.000	0.639	0.678	0.936	0.961	0.922	0.664	0.961	0.922	0.932
41	1.000	1.000	1.000	1.000	0.790	0.710	0.974	0.790	0.890	0.626	0.790	0.890	0.942
42	0.897	1.000	0.491	0.560	0.769	0.786	0.902	0.728	0.666	0.604	0.754	0.730	0.852
43	0.957	0.855	0.455	0.568	0.757	0.733	0.976	0.802	0.749	0.614	0.846	0.847	0.831
44	0.840	0.780	0.041	1.000	0.780	0.752	0.911	0.757	0.720	0.592	0.823	0.728	0.804
45	1.000	0.984	0.667	0.255	0.756	0.783	0.935	0.816	0.766	0.613	0.823	0.817	0.850
46	1.000	0.926	0.678	0.340	0.716	0.720	0.974	0.832	0.795	0.609	0.868	0.882	0.866
47	0.782	0.357	1.000	1.000	0.778	0.733	0.980	0.920	0.981	0.525	0.920	0.981	0.520
48	0.812	0.472	0.761	1.000	0.729	0.689	0.945	0.872	0.926	0.564	0.956	0.982	0.613
49	1.000	1.000	1.000	1.000	0.666	0.592	1.000	0.666	0.592	0.600	0.666	0.592	0.916
50	1.000	1.000	1.000	1.000	0.694	0.733	0.990	0.694	0.733	0.610	0.694	0.733	0.925

Taking into account the critical value ($\alpha^* = 0.800$) we set in our analysis to distinguish the qualified units (which can only be appropriate benchmarks) from the disqualified, it is clear from Table 8 that none of the three scenarios can identify appropriate benchmarks. Most of the top-performing units are assigned scores lower than the critical value in at least one of their exogenous variables. Furthermore, programs (11) and (12) fail to capture the inverse relationship between CS, ES, and P/E. For instance, the target scores for CS and ES for unit 17 in scenario 1 are lower than their original scores, while the target score for P/E is unchanged. Similarly, the target score for P/E for unit 1 in scenario 3 is lower than the original score, while the target scores for CS and ES are unchanged.

4. Conclusions and future research

The performance measurement methodology presented in this paper is applicable to a setting with multiple interrelated dimensions of performance when direct or inverse relationships are present among these dimensions. This methodology constructs an appropriate production frontier, which solely consists of units that are simultaneously qualified in all dimensions of performance. Our methodology is a modified and extended version of the one put forth by Zervopoulos and Palaskas (2011). A major novelty of the new methodology is the relaxation of the assumption of fixed-weights between actual and adjusted input or output variables. In addition, the new methodology can be regarded as a scientific underpinning of performance assessment frameworks such as the Balanced Scorecard.

The properties of the new methodology were presented through a numerical example. Based on the same dataset, a comparative analysis between our methodology and those of Zervopoulos and Palaskas (2011) and of Lim and Zhu (2013) was performed. This analysis explicitly presented the advantages of the new methodology compared to the methodologies developed in the other studies.

The managerial implications of our methodology were discussed in both Sections 2 and 3. Its fundamental advantage for decision makers is the possibility of developing an appropriate production frontier that facilitates realistic benchmarking based on all of the incorporated dimensions of performance. However, a limitation of this study is that it does not identify target levels for the exogenous variables of many of the non-benchmark units which at least satisfy a user-defined critical value, when the units are projected to the frontier. This limitation is mainly due to the underlying inverse relationship between a number of variables. Further research is needed to define the minimal distance of the non-benchmark units from the frontier, giving priority to the satisfaction of the targets set by the decision makers for the dimensions of performance.

Appendix A

Table A1. Dataset

Units	Inputs						Outputs			Exogenous variables		
	Employees[1]	Hours[2]	PC[1]	FAX[1]	Printer[1]	Surface[3]	Online services[1]	Services[1]	Served-citizens[1]	Citizens' satisfaction[4]	Employees' satisfaction[4]	Profits/ Employee[4]
1	8	66	7	1	3	90	29311	27384	34570	0.9230	0.9453	0.9699
2	8	66	10	1	5	50	18723	15241	22054	0.9304	0.9014	0.9701
3	5	66	9	1	5	50	15102	56607	18434	0.9431	0.8984	0.9832
4	5	40	7	1	3	70	6516	20082	9203	0.8208	0.8905	0.8973
5	13	60	13	1	6	100	20730	38324	32269	0.8600	0.8302	0.6539
6	44	66	36	6	8	130	42426	337310	154994	0.8736	0.8896	0.8346
7	13	66	12	3	8	100	30470	102836	65346	0.8185	0.7524	0.8548
8	7	63	9	2	4	65	13717	11004	13810	0.8704	0.8286	0.6989
9	7	66	9	2	4	80	18128	12775	16011	0.8733	0.9003	0.7961
10	5	30	5	1	2	80	5610	2890	5902	0.8111	0.7603	0.9308
11	5	60	5	1	4	70	3166	9962	5402	0.7815	0.7697	0.8947
12	6	32.25	7	1	3	95	8523	21680	15730	0.8637	0.8900	0.8681
13	11	66	19	3	9	90	18608	3879	11187	0.7926	0.8103	0.5991
14	6	63	8	0	4	70	16275	7325	14658	0.9689	0.9105	0.9326
15	5	30	7	1	4	50	8406	2032	8154	0.9496	0.8974	1.0000
16	6	63	8	1	4	115	6960	3385	6989	0.9430	0.9805	0.7103
17	6	63.5	6	1	3	100	17549	76644	26759	0.9037	0.8489	0.8639
18	9	62	7	1	3	110	16625	27373	14764	0.9274	0.9084	0.6773
19	5	63	6	1	3	55	7403	55763	9181	0.9467	0.9558	0.9471
20	5	63	8	2	1	75	14305	30013	22270	0.9452	0.9201	0.9891
21	8	66	7	0	1	60	1746	3254	3448	0.9689	0.8598	0.9908

22	6	30	9	2	3	90	8354	572	10501	0.8081	0.8049	0.9351
23	6	66	8	0	1	250	15699	26231	33544	0.8076	0.7198	0.9989
24	5	57.5	7	1	2	100	16062	38678	25699	0.8103	0.7626	0.9791
25	5	63	6	1	3	30	11382	16070	8687	0.8719	0.8309	0.9159
26	6	66	4	2	2	90	25072	31586	16281	0.8156	0.8627	0.9583
27	5	66	7	2	3	90	22669	24675	11491	0.8356	0.7904	0.8937
28	5	66	12	0	4	100	24781	61382	27353	0.8007	0.6967	0.9681
29	5	63.5	3	1	1	50	4274	6087	5286	0.9141	0.8498	0.8943
30	5	27.5	7	1	3	60	15823	18166	17166	0.9333	0.9661	0.8901
31	5	33	4	0	2	50	11764	9721	8769	0.7793	0.7328	0.9861
32	18	63	14	2	4	80	42216	322231	177779	0.7763	0.8123	0.9794
33	6	66	7	1	3	80	5492	14034	5334	0.7896	0.7697	0.8681
34	5	55	6	1	3	90	7841	4291	6286	0.9342	0.9782	0.7691
35	3	37.5	3	1	2	120	3905	4012	3368	0.9059	0.8254	0.9331
36	14	66	9	2	2	80	15199	20200	17788	0.8415	0.8123	0.8135
37	5	31.25	6	0	1	45	3719	5802	4690	0.8234	0.7901	0.9673
38	39	63	22	4	7	150	50696	621331	202417	0.8111	0.7329	0.9789
39	4	60	6	1	2	80	5910	3902	5303	0.8170	0.8673	0.8371
40	5	66	7	1	3	70	20704	36535	10095	0.9607	0.9219	0.9362
41	5	37.5	9	1	3	80	16901	62846	16208	0.7904	0.8903	0.9736
42	4	66	4	1	2	120	2639	9303	2407	0.7689	0.8145	0.9017
43	5	63.5	7	1	3	70	12123	14813	9038	0.8459	0.8991	0.8893
44	5	63	6	1	3	100	2921	11521	3531	0.8230	0.7281	0.8085
45	5	47.5	6	1	3	70	11468	7521	10209	0.8849	0.8173	0.9329
46	7	33.5	7	1	3	140	11373	40662	10084	0.9467	0.8824	0.9294
47	8	63	7	1	3	90	15126	7701	11528	0.9200	0.9805	0.5246
48	7	57.5	9	1	4	90	16720	19573	14176	0.9556	0.9824	0.6172
49	3	36.5	2	0	1	150	1699	1015	1483	0.6659	0.5923	1.0000
50	4	32.5	2	1	2	180	3786	1348	2112	0.6941	0.7328	0.9903

[1] number, [2] per week, [3] square meters, [4] scale [0.0, 1.0] (see Section 3.1)

Appendix B

It is known from microeconomic theory that:

$$\text{Profits } (\mathrm{P}_j) = \text{Total Revenue } (\mathrm{TR}_j) - \text{Total Cost } (\mathrm{TC}_j)$$

or, $$\mathrm{P}_h = \mathrm{TR}_h - \mathrm{TC}_h, \text{ where } h \subset j$$

Equivalently, $$\frac{\mathrm{P}_h}{\mathrm{E}_h} = \frac{\mathrm{TR}_h - \mathrm{TC}_h}{\mathrm{E}_h}$$

where $\mathrm{TC}_h = \sum_{h=1}^{g} c_{ih} x_{ih}$ and E stands for employees, where $\mathrm{E}_h \in x_{ih}$.

Therefore, $$\frac{\mathrm{P}_h^{ad}}{\mathrm{E}_h^{ad}} = \frac{\mathrm{TR}_h - \mathrm{TC}_h^{ad}}{\mathrm{E}_h^{ad}}$$

If $$x_{ih}^{ad} > x_{ih} \text{ and } \mathrm{E}_h^{ad} > \mathrm{E}_h$$

then $$\mathrm{TC}_h^{ad} > \mathrm{TC}_h \text{ and } \frac{\mathrm{P}_h^{ad}}{\mathrm{E}_h^{ad}} < \frac{\mathrm{P}_h}{\mathrm{E}_h} \text{ holding } \mathrm{TR}_h \text{ fixed}$$

The above assumption for TR is based on the input orientation of the analysis which considers the outputs as fixed. To be more precise, $\mathrm{TR}_h = \sum_{h=1}^{g} p_{rh} y_{rh}$ where p denotes the unit price of the disposable services, and $\mathrm{TR}_h^* = \sum_{h=1}^{g} p_{rh}^* y_{rh}^*$ assuming that the unit price is fixed.

References

Amado, C. A. F., Santos, S. P., & Marques, P. M. (2012). Integrating the Data Envelopment Analysis and the Balanced Scorecard approaches for enhanced performance assessment. *Omega, 40,* 390-403.

Anderson, E. W., Fornell, C., & Rust, R. T. (1997). Customer satisfaction, productivity, and profitability: Differences between goods and services. *Marketing Science, 16*, 129-145.

Avkiran, N. K., & Morita, H. (2010). Benchmarking firm performance from a multiple-stakeholder perspective with an application to Chinese banking. *Omega, 38,* 501-508.

Banker, R. D., Chang, H., Janakiraman, N., & Konstans, C. (2004). A balanced scorecard analysis of performance metrics. *European Journal of Operational Research, 154,* 423-436.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*, 429-444.

Cheng, G., & Zervopoulos, P. D. (2014). Estimating the technical efficiency of health care systems: A cross-country comparison using the directional distance function. *European Journal of Operational Research, 238*, 899-910.

Dyson, R. G. (2000). Performance and operational research. *Journal of the Operational Research Society, 51*, 5-11.

Eilat, H., Golany, B., & Shtub, A. (2006). Constructing and evaluating balanced portfolios of R&D projects with interactions: A DEA based methodology. *European Journal of Operational Research, 172*, 1018-1039.

Eilat, H., Golany, B., & Shtub, A. (2008). R&D project evaluation: An integrated DEA and balanced scorecard approach. *Omega, 36*, 895-912.

Emrouznejad, A., Parker, B. R., & Tavares, G. (2008). Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. *Socio-Economic Planning Sciences, 42*, 151-157.

Färe, R., & Grosskopf, S. (2010) Directional distance functions and slacks-based measures of efficiency. *European Journal of Operational Research, 200*, 320-322.

García-Valderrama, T., Mulero-Mendigorri, E., & Revuelta-Bordoy, D. (2009). Relating the perspectives of the balanced scorecard for R&D by means of DEA. *European Journal of Operational Research, 196*, 1177-1189.

Gustafsson, A., & Johnson, M. D. (2002). Measuring and managing the satisfaction-loyalty-performance links at Volvo. *Journal of Targeting, Measurement and Analysis for Marketing, 10*, 249-258.

Kamakura, W. A., Mittal, V., de Rosa, F., & Mazzon, J. A. (2002). Assessing the Service-Profit Chain. *Marketing Science, 21*, 294-317.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard - measures that drive performance. *Harvard Business Review, 70*, 71-79.

Kaplan, R. S., & Norton, D. P. (1996). Using the Balanced Scorecard as a strategic management system. *Harvard Business Review, January-February*, 75-85.

Kaplan, R. S., & Norton, D. P. (2001). Transforming the balanced scorecard from performance measurement to strategic management: Part I. *Accounting Horizons, 15*, 87-104.

Lim, S., & Zhu, J. (2013). Incorporating performance measures with target levels in data envelopment analysis. *European Journal of Operational Research, 230*, 634-642.

Mukherjee, A., Nath, P., & Pal, M. (2003). Resource, service quality and performance triad: A framework for measuring efficiency of banking services. *Journal of the Operational Research Society, 54*, 723-735.

Oral, M., & Yolalan, R. (1990). An empirical study on measuring operating efficiency and profitability of bank branches. *European Journal of Operational Research, 46*, 282-294.

Otley, D. (1998). Performance management: A framework for management control systems research. *Management Accounting Research, 10*, 363-382.

Paradi, J. C., Rouatt, S., & Zhu, H. (2011). Two-stage evaluation of bank branch efficiency using data envelopment analysis. *Omega, 39*, 99-109.

Ramanathan, R., & Yunfeng, J. (2009). Incorporating cost and environmental factors in quality function deployment using data envelopment analysis. *Omega, 37*, 711-723.

Schilke, O., & Goerzen, A. (2010). Alliance management capability: An investigation of the construct and its measurement. *Journal of Management, 36*, 1192-1219.

Sherman, D., & Zhu, J. (2006). Benchmarking with quality-adjusted DEA (Q-DEA) to seek lower-cost high-quality service: Evidence from a US bank application. *Annals of Operations Research, 145*, 301-319.

Soteriou, A., & Zenios, S. (1999). Operations, quality, and profitability in the provision of banking services. *Management Science, 45*, 1221-1238.

Taylor,W. M., Thompson, R. G., Thrall, R. M., & Dharmapala, P. S. (1997). DEA/AR efficiency and profitability of Mexican banks: A total income model. *European Journal of Operational Research, 98*, 346-363.

Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research, 130*, 498-509.

Zervopoulos, P., & Palaskas, T. (2011). Applying quality-driven, efficiency-adjusted DEA (QE-DEA) in the pursuit of high-efficiency - high-quality service units: An input-oriented approach. *IMA Journal of Management Mathematics, 22,* 401-417.