



Decision Support

Comparing school ownership performance using a pseudo-panel database: A Malmquist-type index approach

Juan Aparicio^a, Eva Crespo-Cebada^b, Francisco Pedraja-Chaparro^b, Daniel Santín^{c,*}^a Center of Operations Research, University Miguel Hernandez of Elche, Av.de la universidad s/n, 03202 Elche, Spain^b Department of Economics, University of Extremadura, Av. Elvas s/n, 06071 Badajoz, Spain^c Department of Applied Economics VI, Complutense University of Madrid, Campus de Somosaguas, 28223 Madrid, Spain

ARTICLE INFO

Article history:

Received 23 April 2014

Accepted 13 June 2016

Available online 16 June 2016

Keywords:

Data envelopment analysis

Education

Productivity and competitiveness

Malmquist index

ABSTRACT

In this paper, we propose a different way of using the Malmquist index that allows us to further analyze the relative performance divergences between two groups of decision-making units (DMUs) over time when only a pseudo-panel database is available. To do this, we extend the Camanho and Dyson (2006) one-period Malmquist-type index (CDMI) for a pseudo-panel database with a new pseudo-panel Malmquist index (PPMI). To illustrate the methodology, we apply it to examine how the performance gap between public and private government-dependent secondary schools in the Basque Country (Spain) performed across three PISA waves (2006, 2009 and 2012). The results suggest that performance is persistently and significantly higher for private government-dependent schools than for public schools.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Traditionally, the Malmquist index, proposed by Caves, Christensen, and Diewert (1982), is used to measure productivity changes between two or more periods. This implies that the same group of decision-making units (DMUs) is observed across all analyzed periods. In the educational or health sectors, however, it is common to extract random waves of representative samples of DMUs in order to analyze their resources, activities, results and performance.¹ The DMUs contained in each wave vary from 1 year to another and are mostly anonymous for researchers. So, we will refer to this information as a pseudo-panel database.

In many sectors related to services supply, it is likewise worth analyzing alternative approaches to carry out the production process. The groups of DMUs that employ one or other production scheme perform basically the same tasks using the same set of inputs to be transformed into the same set of outputs, but the internal organizational and managerial techniques used to deliver the service may differ. Camanho and Dyson (2006) originally proposed a Malmquist-type index (CDMI from now on) in order to achieve an average indicator of the relative performance of two or more

groups of DMUs within a period. The CDMI has been used in the literature (Ferreira & Marques, 2015; Thanassoulis, Shiraz, & Maniadakis, 2015; Vaz & Camanho, 2012) to provide straightforward performance comparisons of groups of DMUs in one period. However, it would also be interesting to follow up how the measured performance gap changes over time. This is the aim of this paper.

Regarding the education sector, a lot of the literature focuses on comparing performance by school ownership. For example, education is publicly funded in many countries, but education may be delivered by public schools (PS) or private government-dependent schools (PGDS). Previous research justifies the existence of PGDS arguing that privately run schools are likely to perform better than public schools because market competition should force private schools to achieve a more efficient use of resources while at the same time providing high standards of quality for their students (Alchian, 1950; Chubb & Moe, 1990; Friedman & Friedman, 1981; Hoxby 2003). Nevertheless, much of the research (Crespo-Cebada, Pedraja-Chaparro, & Santín, 2014; Kirjavainen & Loikkanen, 1998; Mancebón, Calero, Choi, & Ximenez, 2012; Perelman and Santín, 2011; Vandenberghe & Robin, 2004 among many others) provides mixed evidence about the superiority of either school type. Following Mancebón and Muñiz (2008) we think that the inconclusive results in the literature review can be due to country-specific heterogeneity in factors as the level of competition among schools or the admission policy among other factors. For this reason, analyzing the relative performance gap between both school groups is an important and not straightforward question to be answered empirically using the available educational data.

* Corresponding author. Fax: +34 91 394 2431.

E-mail addresses: j.aparicio@umh.es (J. Aparicio), ecreceb@unex.es (E. Crespo-Cebada), pedraja@unex.es (F. Pedraja-Chaparro), dsantin@ccee.ucm.es (D. Santín).¹ PISA, TIMSS and PIRLS are some international education databases that belong to this category. They will be described below. Some examples in the health sector would be NTDB (the National Trauma Data Bank) or NHCS (the National Hospital Care Survey), both based in the USA.

Note that many studies have used cross-sectional data in order to tackle the inefficiency measurement issue in education at school level with non-parametric techniques (De Witte & López-Torres, 2015). However, only a few have applied the standard Malmquist index to examine productivity changes in schools (some exceptions are Brennan, Haelermans, & Ruggiero, 2014; Essid, Oullette, & Vigeant, 2014; Maragos & Despotis, 2004; Portela, Camanho, & Keshvari, 2013). A possible explanation for the failure of empirical studies to apply the Malmquist index for benchmarking schools is the difficulty of obtaining administrative panel databases at school level from educational authorities.

The recent increase in national and international programs to evaluate educational achievement over the last few decades is indicative of a greater policy concern about educational performance (De la Fuente, 2011; Hanushek & Kimko, 2000; Hanushek & Woessman, 2008). Hence, there has been strong backing for international projects evaluating educational achievement for the vehicular disciplines: mathematics, reading and science. The most important international programs are TIMSS (*Third International Mathematics and Science Study*), PISA (*Program for International Student Assessment*) and PIRLS (*Progress in International Reading Literacy Study*), although many countries perform their own national evaluations. The main advantage of these international programs is that they provide an external evaluation of educational outcomes in an attempt to identify the causes of academic achievement levels and allow policy makers and school principals to thoroughly explore their management strengths and weaknesses. Nevertheless, it is not straightforward to compare school performance over time using these repeated cross sectional international studies because participant schools and students differ from one wave to another.

This paper proposes a new approach for comparing the performance of representative groups of DMUs using the Malmquist index when only a pseudo-panel database is available. This method extends Camanho and Dyson's Malmquist-type index for measuring and comparing the productivity of two groups of DMUs in one period.² With the aim of showing our proposal's potential, we include an empirical application in the educational context in order to test possible relative performance disparities between public (PS) and private government dependent schools (PGDS) over three time periods (from 2006 to 2012) on the publicly funded educational system in the Basque Country. This autonomous region of Spain participated in the *Program for International Student Assessment* (PISA) administered in 2006, 2009 and 2012 by the Organization for Economic Cooperation and Development (OECD) with an extended representative sample.

2. Methodology

The Malmquist index was proposed by Caves, Christensen and Diewert (1982) with the aim of measuring the total factor productivity changes between two data points within two time periods as the ratio of the distances of each data point relative to a common frontier. The index may be built and decomposed using different data envelopment analysis (DEA) programs to compute different distances between the evaluated production unit and the frontier for each period. According to Färe, Grosskopf, and Lovell (1994), the output-oriented Malmquist productivity index for two periods

of time t and $t + 1$ under a constant returns to scale technology³ can be written as:

$$M(x^{t+1}, y^{t+1}, x^t, y^t) = \left[\left(\frac{D^t(x^{t+1}, y^{t+1})}{D^t(x^t, y^t)} \right) \cdot \left(\frac{D^{t+1}(x^{t+1}, y^{t+1})}{D^{t+1}(x^t, y^t)} \right) \right]^{1/2}, \quad (1)$$

where the superscript indicates the time period, $x = (x_1, \dots, x_K) \in R_+^K$ and $y = (y_1, \dots, y_S) \in R_+^S$ are vectors of inputs and outputs, respectively, and $D^q(x^h, y^h) = \inf\{\theta : (x^h, y^h/\theta) \in T^q\}$ is the Shephard output distance function from the period h observation (x^h, y^h) to the frontier of the period q technology $T^q = \{(x^q, y^q) \in R_+^{K+S} : x^q \text{ can produce } y^q\}$.

Färe et al. (1994) also showed that this index may be decomposed into efficiency change and technical change as follows.

$$M(x^{t+1}, y^{t+1}, x^t, y^t) = \underbrace{\frac{D^{t+1}(x^{t+1}, y^{t+1})}{D^t(x^t, y^t)}}_{EC(x^{t+1}, y^{t+1}, x^t, y^t)} \cdot \underbrace{\left[\left(\frac{D^t(x^{t+1}, y^{t+1})}{D^{t+1}(x^{t+1}, y^{t+1})} \right) \cdot \left(\frac{D^t(x^t, y^t)}{D^{t+1}(x^t, y^t)} \right) \right]^{1/2}}_{TC(x^{t+1}, y^{t+1}, x^t, y^t)} \quad (2)$$

A Malmquist index higher (lower) than one implies productivity improvements (losses) from period t to period $t + 1$. Furthermore, Eq. (2) includes two components. The first ratio shows the technical efficiency change (EC), which captures the efficiency improvements (reductions) in period $t + 1$ with respect to period t if $EC > 1$ ($EC < 1$), whereas $EC = 1$ indicates that there are no changes in technical efficiency. The second measure (in square brackets) represents the technological change (TC) in period $t + 1$ with respect to period t , whose value may be analyzed in a similar way to EC ($TC > 1$ now means technological progress), although the two measures may point in different directions.

2.1. Camanho–Dyson Malmquist index (CDMI) approach

In the Malmquist index methodology the same group of DMUs has to be observed in different periods, for which purpose a panel database needs to be implemented. However, this approach was not originally defined to explore the potential performance disparities among different groups of units whose organizational structure and background circumstances differ over time. To overcome part of this problem, Camanho and Dyson (2006) originally proposed a Malmquist-type index in order to achieve an average indicator of the relative performance of two or more groups of DMUs within the same period when the organizational and managerial guidelines differ as applies in the school ownership case.

The CDMI is an adaptation of the Malmquist index to provide a cross-sectional comparison of the performance of DMUs operating under different conditions rather than a measurement of the productivity change between two periods. Camanho and Dyson (2006) defined an overall measure for comparing the performance between two (or more) groups of DMUs by replacing the superindices t and $t + 1$ related to the period by G (which we assume to denote PGDS) and P (which we assume in the empirical

² A similar strategy was developed previously by Berg et al. (1993) in order to compare banking efficiencies in three Nordic countries or Balk and Althin (1996) to compare the evolution of Swedish pharmacy productivity over the 1980–89 period. Both papers propose new strategies for calculating the Malmquist index, taking a particular unit as the comparison reference or taking a fixed period as the baseline in order to calculate multi-period Malmquist indices, respectively. Both alternatives satisfy the transitivity property.

³ There are other studies, such as Balk and Althin (1996), Ray and Desli (1997), Grifell-Tatjé and Lovell (1999), Forsund (2002) and Brennan et al. (2014), that also consider variable returns to scale. However, it is well known (Afsharian & Ahn, 2015; Pastor & Lovell, 2005) that infeasibilities can arise when DEA is used to compute the distance functions constituting the Malmquist decomposition when the scale component is taken into account in the productivity growth. For the sake of simplicity, we follow Camanho and Dyson (2006) in this paper and assume a constant returns to scale technology avoiding the infeasibility issue (Xue & Harker, 2002). In the case of our empirical application to education, scale is not a major problem and variables are usually normalized to avoid school size issues.

exercise to be PS) related to the groups to be compared. Consider N DMUs ($j = 1, \dots, N$) in group G , using an input vector $x^G \in R_+^K$ to produce outputs $y^G \in R_+^S$, where the input-output vector for DMU $_j$ in group G is denoted by (x_j^G, y_j^G) , and $D^q(x_j^G, y_j^G)$ represents the Shephard output distance function for DMU j with respect to the frontier of group q ($q = G, P$). Similarly, consider M DMUs ($i = 1, \dots, M$) in group P using an input vector $x^P \in R_+^K$ to produce outputs $y^P \in R_+^S$, where the input-output vector for DMU $_i$ in group P is denoted by (x_i^P, y_i^P) , and $D^q(x_i^P, y_i^P)$ represents the Shephard output distance function for DMU i with respect to the frontier of group q ($q = G, P$). The CDMI for comparing the performance of two groups of DMUs G and P associated with different programs, ownerships or practices in one time period t can be defined in Eq. (3) as follows:

$$CDMI_t^{GP} = \left[\frac{\left(\prod_{j=1}^N D^G(x_j^G, y_j^G) \right)^{1/N} \cdot \left(\prod_{i=1}^M D^P(x_i^G, y_i^G) \right)^{1/M}}{\left(\prod_{i=1}^M D^G(x_i^P, y_i^P) \right)^{1/M} \cdot \left(\prod_{i=1}^M D^P(x_i^P, y_i^P) \right)^{1/M}} \right]^{1/2} \quad (3)$$

The ratios inside square brackets evaluate the distance of the DMUs to a single reference technology. The first ratio assesses the geometric average distance of units belonging to group G divided by the geometric average distance of DMUs from group P , all measured relative to group G 's frontier. The second ratio is a similar quotient but using group P 's frontier as benchmark. Following Camanho and Dyson (2006), $CDMI_t^{GP}$ may be interpreted as an overall measure for the comparison of the relative performance between the units of the two groups, G and P . Moreover, this relative performance gap may be decomposed into the following terms:

$$CDMI_t^{GP} = \underbrace{\frac{\left(\prod_{j=1}^N D^G(x_j^G, y_j^G) \right)^{1/N}}{\left(\prod_{i=1}^M D^P(x_i^P, y_i^P) \right)^{1/M}}}_{EG_t^{GP}} \times \underbrace{\left[\frac{\left(\prod_{j=1}^N D^P(x_j^G, y_j^G) \right)^{1/N} \cdot \left(\prod_{i=1}^M D^P(x_i^P, y_i^P) \right)^{1/M}}{\left(\prod_{j=1}^N D^G(x_j^G, y_j^G) \right)^{1/N} \cdot \left(\prod_{i=1}^M D^G(x_i^P, y_i^P) \right)^{1/M}} \right]^{1/2}}_{TG_t^{GP}} \quad (4)$$

The ratio EG_t^{GP} outside square brackets in (4) compares within-group efficiency spreads, measuring the technical efficiency gap between both groups, while the ratio TG_t^{GP} , inside square brackets, evaluates the productivity gap between the frontiers of the two analyzed groups, measuring the distance between the best-practice frontiers of G and P . In terms of the interpretation of $CDMI_t^{GP}$, and both components, a value greater than one indicates better performance in group G than in group P . The decomposition in (4) illustrates that better overall performance may be related to two possible factors: less dispersion in the technical efficiency levels of the units in one group compared to the other, and/or the dominance of the best-practice frontier.

Fig. 1 illustrates these concepts in a simple one input-one output setting in order to compare two groups of DMUs in one period. Let us assume that the G Frontier and P Frontier represent the constant returns to scale technologies for both $PGDS$ and PS , respectively.

The circles in Fig. 1 represent the N DMUs in group G , whereas crosses represent the M DMUs in group P . This production activity information is used to estimate both production frontiers using DEA. Consequently, we estimate the distance of each school to its

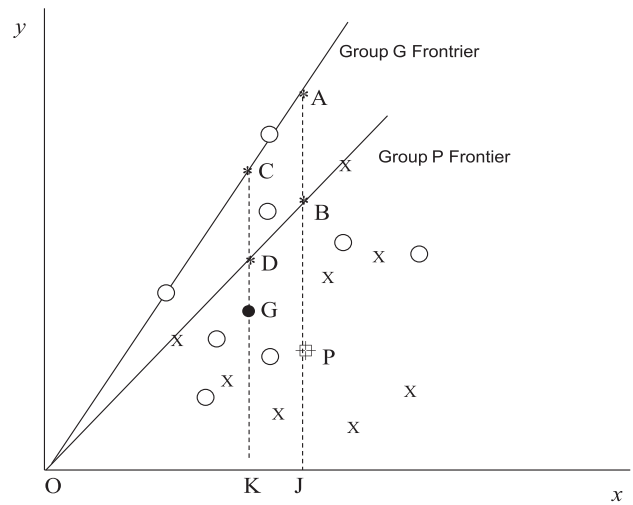


Fig. 1. Productivity divergences between PS and GDPS.

own frontier. Then, we estimate the geometric average distance of each group with respect to its own frontier and the comparison group frontier. The dots G and P in Fig. 1 denote this theoretical average production activity for both groups, considering the performance of DMUs belonging to the same group only. Therefore, according to Eq. (3), we calculate the four required distance functions, measured as the geometric average distance of units belonging to each of both groups relative to each of both group frontiers in order to estimate the CDMI in Fig. 1 as follows:

$$\frac{KG}{KC} = \left(\prod_{j=1}^N D^G(x_j^G, y_j^G) \right)^{1/N} \quad (5)$$

$$\frac{JP}{JB} = \left(\prod_{i=1}^M D^P(x_i^P, y_i^P) \right)^{1/M} \quad (6)$$

$$\frac{KG}{KD} = \left(\prod_{j=1}^N D^P(x_j^G, y_j^G) \right)^{1/N} \quad (7)$$

$$\frac{JP}{JA} = \left(\prod_{i=1}^M D^G(x_i^P, y_i^P) \right)^{1/M} \quad (8)$$

The CDMI approach is capable of evaluating the average discrepancies between both school groups⁴, but it was not defined for analyzing this behavior over time. One advantage of this approach over Charnes, Cooper, and Rhodes (1981) well-known methodology is that the CDMI does not need to pool the DMUs together to build a common meta-frontier⁵. In summary, note that the CDMI is essentially a standard Malmquist index where the two time periods are replaced by group superindices and where only the average result is meaningful.

2.2. CDMI extended: the pseudo-panel Malmquist index

The CDMI provides an opportunity to monitor a sector in which panel databases do not exist, for example, an educational system. As mentioned before, the number of national and international pseudo-panel databases for schools available for research

⁴ Camanho and Dyson (2006) show that their index can also be used to compare more than two groups in a more general setting. The CDMI for more than two groups satisfies Frisch's circularity condition (Frisch, 1936), so if we assume three groups A , B and C , then $CDMI^{AB} \times CDMI^{BC} = CDMI^{AC}$.

⁵ The CDMI does not assume convex combinations of group-specific frontiers to be feasible.

has grown in recent years. A pseudo-panel database consists of different waves of representative school samples, where participant schools and students vary from one wave to another.

With the aim of analyzing the evolution of the school ownership divergences⁶ across two time periods, we propose making the ratio of two CDMIs calculated for two time periods t and $t+1$. Applying this strategy we can check which organizational pattern behaves better over time in order to infer relevant implications for policy makers. The relative performance gap change between *PGDS* and *PS* within t and $t+1$ is defined as the pseudo-panel Malmquist index (PPMI) as follows:

$$PPMI_{t,t+1}^{GP} = \frac{CDMI_{t+1}^{GP}}{CDMI_t^{GP}} \quad (9)$$

Eq. (9) measures how CDMI changes over time. In order to interpret PPMI in a suitable way, it is necessary to analyze both values of its components, $CDMI_t^{GP}$ and $CDMI_{t+1}^{GP}$, and the value of PPMI itself. Next, we show all the possible settings.

Setting 1: $CDMI_t^{GP}, CDMI_{t+1}^{GP} < 1$. This means that, on average, group P had a better relative performance than group G both in t and $t+1$ periods. Additionally and regarding the value of PPMI, there are two possibilities.

- (1a) $PPMI_{t,t+1}^{GP} < 1$, which means that the relative performance gap was opened up by P over G. In fact, $100 \cdot (1 - PPMI_{t,t+1}^{GP})$ indicates the percentage in which the relative performance of G compared to P has worsened.
- (1b) $PPMI_{t,t+1}^{GP} > 1$, which means that the group G is catching up on the group P. In this case, $100 \cdot (PPMI_{t,t+1}^{GP} - 1)$ indicates the percentage in which the relative performance of G compared to P has improved.

Setting 2: $CDMI_t^{GP}, CDMI_{t+1}^{GP} > 1$. This means that group G had a better relative performance than group P both in t and $t+1$. Moreover, regarding PPMI, we have again two scenarios.

- (2a) $PPMI_{t,t+1}^{GP} < 1$, which means that the group P is catching up on the group G. In fact, $100 \cdot (1 - PPMI_{t,t+1}^{GP})$ indicates the percentage in which the relative performance of G compared to P has decreased.
- (2b) $PPMI_{t,t+1}^{GP} > 1$, which means that the relative performance gap was opened up by G over P. In this case, $100 \cdot (PPMI_{t,t+1}^{GP} - 1)$ indicates the percentage in which the relative performance of G compared to P has increased.

Setting 3: $CDMI_t^{GP} > 1$ and $CDMI_{t+1}^{GP} < 1$. Under this scenario, group G had a better relative performance than group P in period t but P had a better relative performance than G in the second period, $t+1$. As for the value of PPMI, we have only one possibility: $PPMI_{t,t+1}^{GP} < 1$. In this case, the status of group G worsened drastically from period t to period $t+1$, and the value of the PPMI measures how many the relative performance gap between the two analyzed groups of units reduced over time through $100 \cdot (1 - PPMI_{t,t+1}^{GP})$ percent.

Setting 4: $CDMI_t^{GP} < 1$ and $CDMI_{t+1}^{GP} > 1$. In this case, P had a better relative performance than G in period t but the relation between both groups changed drastically in period $t+1$, where G had a better relative performance than P. Regarding the PPMI, we have again only one possibility: $PPMI_{t,t+1}^{GP} > 1$. In this case, the status of group G improved from period t to period $t+1$, and the value of

the PPMI signals how many the relative performance gap between the two analyzed groups of units increased over time through $100 \cdot (PPMI_{t,t+1}^{GP} - 1)$ percent.

Note that above, seeking simplicity, we did not consider scenarios where a sign of equality is possible. However, its interpretation is straightforward. For the CDMI index, it would mean that no group stands out over the other and for the PPMI measure, and regarding the relative performance gap, it would imply that no changes over time occurred.

As was suggested by Camanho and Dyson (2006), the results of the comparison of the relative performance of group G and P through the CDMI index should not be interpreted isolated but alongside the disaggregated measures corresponding to the technical efficiency spread dimension and the frontier productivity dimension. Accordingly, the PPMI can also be decomposed into efficiency gap change (*EGC*) and technological gap change (*TGC*), as Eq. (10) shows:

$$PPMI_{t,t+1}^{GP} = \frac{CDMI_{t+1}^{GP}}{CDMI_t^{GP}} = \frac{EGC_{t+1}^{GP}}{EGC_t^{GP}} \cdot \frac{TGC_{t+1}^{GP}}{TGC_t^{GP}} = EGC_{t,t+1}^{GP} \cdot TGC_{t,t+1}^{GP} \quad (10)$$

For the correct interpretation of the values of the two sub-components, $EGC_{t,t+1}^{GP}$ and $TGC_{t,t+1}^{GP}$, we need to follow the same causality than that introduced above for the PPMI measure. For the sake of simplicity in the exposition, we omit this part in the text.

Additionally, it is worth mentioning that the PPMI for two groups satisfies the circular relationship, so $PPMI_{t,t+2}^{GP} = PPMI_{t,t+1}^{GP} \cdot PPMI_{t+1,t+2}^{GP}$. This property is demonstrated by Eq. (11) as follows:

$$\frac{CDMI_{t+2}^{GP}}{CDMI_t^{GP}} = \frac{CDMI_{t+1}^{GP}}{CDMI_t^{GP}} \cdot \frac{CDMI_{t+2}^{GP}}{CDMI_{t+1}^{GP}} \quad (11)$$

The CDMI, despite being a Malmquist-type index, does not measure productivity change but relative performance between two groups, or more, of decision making units. As a consequence, the PPMI does not have a direct relationship with the productivity change over time of units in groups G and P. Nevertheless, next we will identify when the PPMI is related to the ratio between an aggregated measure of productivity changes, from t to $t+1$, of the units corresponding to group G and an aggregated measure of productivity changes, from t to $t+1$, of the units belonging to P, i.e.

$$PPMI_{t,t+1}^{GP} = \frac{\left[\frac{\left(\prod_{j=1}^{N^{t+1}} D^{G^{t+1}}(x_j^{G^{t+1}}, y_j^{G^{t+1}}) \right)^{1/N^{t+1}}}{\left(\prod_{j=1}^{N^t} D^{G^t}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t}} \cdot \frac{\left(\prod_{j=1}^{N^{t+1}} D^G(x_j^{G^{t+1}}, y_j^{G^{t+1}}) \right)^{1/N^{t+1}}}{\left(\prod_{j=1}^{N^t} D^G(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t}} \right]^{1/2}}{\left[\frac{\left(\prod_{i=1}^{M^{t+1}} D^{P^{t+1}}(x_i^{P^{t+1}}, y_i^{P^{t+1}}) \right)^{1/M^{t+1}}}{\left(\prod_{i=1}^{M^t} D^{P^t}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t}} \cdot \frac{\left(\prod_{i=1}^{M^{t+1}} D^{P^t}(x_i^{P^{t+1}}, y_i^{P^{t+1}}) \right)^{1/M^{t+1}}}{\left(\prod_{i=1}^{M^t} D^{P^t}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t}} \right]^{1/2}}, \quad (12)$$

where the superscripts t and $t+1$ denotes that the technology of reference and/or the corresponding vector of inputs and outputs was observed in t and $t+1$, respectively.

Note that both the numerator and the denominator in (12) are consistent with the expression of the Malmquist productivity index developed by Färe et al. (1994) as reported in (1), but adapted following the philosophy of Camanho and Dyson (2006) for aggregating distances. In this sense, the numerator could be interpreted as a relative measure of performance change from t to $t+1$ experienced by the units belonging to group G. The denominator has the same interpretation than the numerator changing G by P. Additionally, we want to highlight that in the case of working with an usual panel data, where the same set of DMUs has been observed across all analyzed periods and, therefore, $N^t = N^{t+1} = N$

⁶ The empirical application of this paper compares two types of school ownership. However, this methodology could be applied to any other comparison of two or more groups of DMUs whose practices differ significantly. This methodology can also be used in impact evaluation programs to compare results for treated and non-treated DMUs.

and $M^t = M^{t+1} = M$, expression (12) would be equivalent to the ratio of the geometric average of the Malmquist productivity indexes for all the units in G and the geometric average of the Malmquist productivity indexes for all the units in P:

$$PPMI_{t,t+1}^{GP} = \frac{\left(\prod_{j=1}^N M(x_j^{G^{t+1}}, y_j^{G^{t+1}}, x_j^{G^t}, y_j^{G^t}) \right)^{1/N}}{\left(\prod_{i=1}^M M(x_i^{P^{t+1}}, y_i^{P^{t+1}}, x_i^{P^t}, y_i^{P^t}) \right)^{1/M}}, \tag{13}$$

In general, PPMI does not coincide with expression (12). We next show under which conditions both expressions are equivalent.

Proposition 1.

$$PPMI_{t,t+1}^{GP} = \left[\frac{\left(\prod_{j=1}^{N^{t+1}} D^{G^{t+1}}(x_j^{G^{t+1}}, y_j^{G^{t+1}}) \right)^{1/N^{t+1}} \cdot \left(\prod_{j=1}^{N^t} D^{G^t}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t}}{\left(\prod_{j=1}^{N^t} D^{G^{t+1}}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t} \cdot \left(\prod_{j=1}^{N^{t+1}} D^{G^t}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^{t+1}}} \right]^{1/2}$$

$$= \left[\frac{\left(\prod_{i=1}^{M^{t+1}} D^{P^{t+1}}(x_i^{P^{t+1}}, y_i^{P^{t+1}}) \right)^{1/M^{t+1}} \cdot \left(\prod_{i=1}^{M^t} D^{P^t}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t}}{\left(\prod_{i=1}^{M^t} D^{P^{t+1}}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t} \cdot \left(\prod_{i=1}^{M^{t+1}} D^{P^t}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^{t+1}}} \right]^{1/2}$$

if and only if

$$DC_{t,t+1}^{GP} = \left[\frac{\left(\prod_{j=1}^{N^{t+1}} D^{P^{t+1}}(x_j^{G^{t+1}}, y_j^{G^{t+1}}) \right)^{1/N^{t+1}} \cdot \left(\prod_{i=1}^{M^{t+1}} D^{P^t}(x_i^{P^{t+1}}, y_i^{P^t}) \right)^{1/M^{t+1}}}{\left(\prod_{i=1}^{M^t} D^{P^{t+1}}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t} \cdot \left(\prod_{j=1}^{N^t} D^{P^t}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t}} \right]^{1/2} = 1.$$

Proof. Combining (3) and (9) and operating, we get

$$PPMI_{t,t+1}^{GP} = \left[\frac{\left(\prod_{j=1}^{N^{t+1}} D^{G^{t+1}}(x_j^{G^{t+1}}, y_j^{G^{t+1}}) \right)^{1/N^{t+1}} \cdot \left(\prod_{j=1}^{N^t} D^{G^t}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t}}{\left(\prod_{j=1}^{N^t} D^{G^{t+1}}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t} \cdot \left(\prod_{j=1}^{N^{t+1}} D^{G^t}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^{t+1}}} \right]^{1/2}$$

$$= \left[\frac{\left(\prod_{i=1}^{M^{t+1}} D^{P^{t+1}}(x_i^{P^{t+1}}, y_i^{P^{t+1}}) \right)^{1/M^{t+1}} \cdot \left(\prod_{i=1}^{M^t} D^{P^t}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t}}{\left(\prod_{i=1}^{M^t} D^{P^{t+1}}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t} \cdot \left(\prod_{i=1}^{M^{t+1}} D^{P^t}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^{t+1}}} \right]^{1/2}$$

$$\times \left[\frac{\left(\prod_{j=1}^{N^{t+1}} D^{P^{t+1}}(x_j^{G^{t+1}}, y_j^{G^{t+1}}) \right)^{1/N^{t+1}} \cdot \left(\prod_{i=1}^{M^{t+1}} D^{P^t}(x_i^{P^{t+1}}, y_i^{P^t}) \right)^{1/M^{t+1}}}{\left(\prod_{i=1}^{M^t} D^{P^{t+1}}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t} \cdot \left(\prod_{j=1}^{N^t} D^{P^t}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t}} \right]^{1/2}$$

$$\times \left[\frac{\left(\prod_{i=1}^{M^{t+1}} D^{G^{t+1}}(x_i^{P^{t+1}}, y_i^{P^{t+1}}) \right)^{1/M^{t+1}} \cdot \left(\prod_{j=1}^{N^{t+1}} D^{G^t}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^{t+1}}}{\left(\prod_{j=1}^{N^t} D^{G^{t+1}}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t} \cdot \left(\prod_{i=1}^{M^t} D^{G^t}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t}} \right]^{1/2}.$$

In this way, the result of the proposition is true if and only if the second ratio equals one.

In the proposition, we introduced a new term, denoted as $DC_{t,t+1}^{GP}$, and baptized as the ‘Divergence Component’ of the PPMI. It indicates how far is the PPMI from the expression that appears in (12) with a meaning of ratio of aggregated productivity changes in G and P. Note that mixed group distances appear in the formulation of $DC_{t,t+1}^{GP}$ as, for example,

$D^{P^{t+1}}(x_j^{G^{t+1}}, y_j^{G^{t+1}})$, which resorts to the frontier of the technology associated with the group P in period $t + 1$ for calculating the distance from a unit of the group G observed in the same period of time.

Regarding the sub-components of the PPMI, the case of the efficiency gap change is the easiest to be studied since $EGC_{t,t+1}^{GP}$ always coincides with the ratio between an aggregated measure of technical efficiency changes, from t to $t + 1$, of the units in G and an aggregated measure of technical efficiency changes, from t to $t + 1$, of the units belonging to P, i.e.

$$EGC_{t,t+1}^{GP} = \frac{\left(\prod_{j=1}^{N^{t+1}} D^{G^{t+1}}(x_j^{G^{t+1}}, y_j^{G^{t+1}}) \right)^{1/N^{t+1}}}{\left(\prod_{j=1}^{N^t} D^{G^t}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t}} \cdot \frac{\left(\prod_{i=1}^{M^{t+1}} D^{P^{t+1}}(x_i^{P^{t+1}}, y_i^{P^{t+1}}) \right)^{1/M^{t+1}}}{\left(\prod_{i=1}^{M^t} D^{P^t}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t}}. \tag{14}$$

Something similar to the PPMI happens with the technological gap change component, $TGC_{t,t+1}^{GP}$, when the aim is to relate this last measure to the technological change term, TC, of the traditional Malmquist productivity index in (2) for G and P. $TGC_{t,t+1}^{GP}$ does not coincide, in general, with the ratio of aggregated measures of technological changes for the units in the group G and P. Indeed, this sub-component of PPMI can always be rewritten as follows.

$$TGC_{t,t+1}^{GP} = \left[\frac{\left(\prod_{j=1}^{N^t} D^{G^t}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t} \cdot \left(\prod_{j=1}^{N^{t+1}} D^{G^t}(x_j^{G^{t+1}}, y_j^{G^{t+1}}) \right)^{1/N^{t+1}}}{\left(\prod_{j=1}^{N^t} D^{G^{t+1}}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t} \cdot \left(\prod_{j=1}^{N^{t+1}} D^{G^{t+1}}(x_j^{G^{t+1}}, y_j^{G^{t+1}}) \right)^{1/N^{t+1}}} \right]^{1/2}$$

$$= \left[\frac{\left(\prod_{i=1}^{M^t} D^{P^t}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t} \cdot \left(\prod_{i=1}^{M^{t+1}} D^{P^t}(x_i^{P^{t+1}}, y_i^{P^t}) \right)^{1/M^{t+1}}}{\left(\prod_{i=1}^{M^t} D^{P^{t+1}}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t} \cdot \left(\prod_{i=1}^{M^{t+1}} D^{P^{t+1}}(x_i^{P^{t+1}}, y_i^{P^{t+1}}) \right)^{1/M^{t+1}}} \right]^{1/2}$$

$$\times \left[\frac{\left(\prod_{j=1}^{N^{t+1}} D^{P^{t+1}}(x_j^{G^{t+1}}, y_j^{G^{t+1}}) \right)^{1/N^{t+1}} \cdot \left(\prod_{i=1}^{M^{t+1}} D^{P^t}(x_i^{P^{t+1}}, y_i^{P^t}) \right)^{1/M^{t+1}}}{\left(\prod_{i=1}^{M^t} D^{P^{t+1}}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t} \cdot \left(\prod_{j=1}^{N^t} D^{P^t}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t}} \right]^{1/2}$$

$$\times \underbrace{\left[\frac{\left(\prod_{i=1}^{M^{t+1}} D^{G^{t+1}}(x_i^{P^{t+1}}, y_i^{P^{t+1}}) \right)^{1/M^{t+1}} \cdot \left(\prod_{j=1}^{N^{t+1}} D^{G^t}(x_j^{G^{t+1}}, y_j^{G^{t+1}}) \right)^{1/N^{t+1}}}{\left(\prod_{j=1}^{N^t} D^{G^{t+1}}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t} \cdot \left(\prod_{i=1}^{M^t} D^{G^t}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t}} \right]^{1/2}}_{DC_{t,t+1}^{GP}}. \tag{15}$$

By analogy with the TC component in (2), the first ratio in (15) is, in some sense, a quotient that compares an aggregated measure of technological changes, from t to $t + 1$, of the units in G to an aggregated measure of technological changes, from t to $t + 1$, of the units in P, while the second ratio coincides, as expected, with the divergence component of Proposition 1. Therefore, we have the following direct consequence.

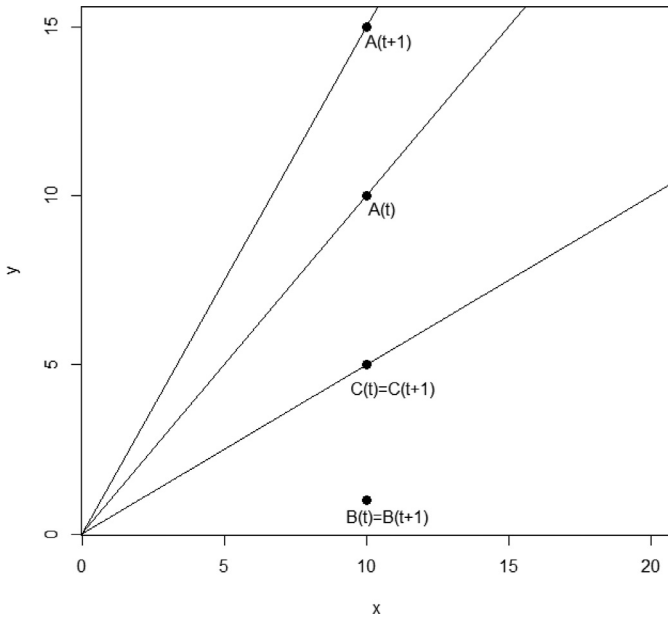


Fig. 2. Graphical illustration of example 1.

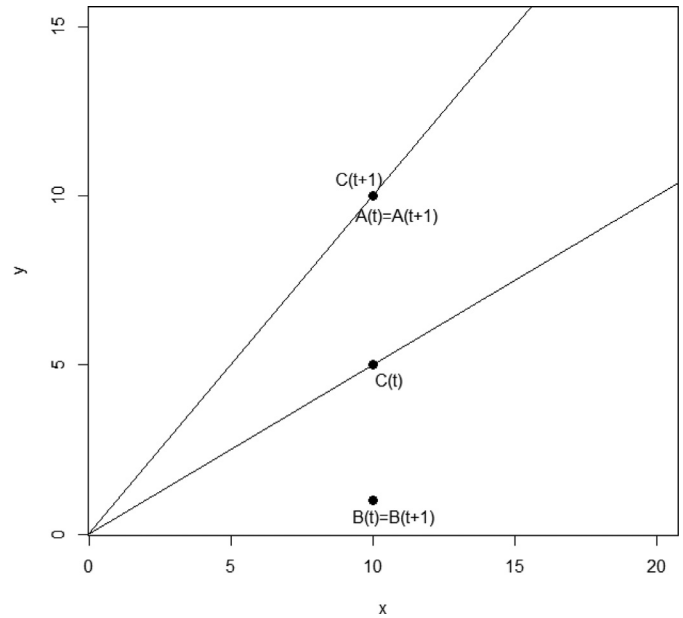


Fig. 3. Graphical illustration of example 2.

Proposition 2.

$$TGC_{t,t+1}^{GP} = \frac{\left[\frac{\left(\prod_{j=1}^{N^t} D^{G^t}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^{t+1}} \cdot \left(\prod_{j=1}^{N^{t+1}} D^{G^t}(x_j^{G^{t+1}}, y_j^{G^{t+1}}) \right)^{1/N^{t+1}}}{\left(\prod_{j=1}^{N^t} D^{G^{t+1}}(x_j^{G^t}, y_j^{G^t}) \right)^{1/N^t} \cdot \left(\prod_{j=1}^{N^{t+1}} D^{G^{t+1}}(x_j^{G^{t+1}}, y_j^{G^{t+1}}) \right)^{1/N^{t+1}}} \right]^{1/2}}{\left[\frac{\left(\prod_{i=1}^{M^t} D^{P^t}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t} \cdot \left(\prod_{i=1}^{M^{t+1}} D^{P^t}(x_i^{P^{t+1}}, y_i^{P^{t+1}}) \right)^{1/M^{t+1}}}{\left(\prod_{i=1}^{M^t} D^{P^{t+1}}(x_i^{P^t}, y_i^{P^t}) \right)^{1/M^t} \cdot \left(\prod_{i=1}^{M^{t+1}} D^{P^{t+1}}(x_i^{P^{t+1}}, y_i^{P^{t+1}}) \right)^{1/M^{t+1}}} \right]^{1/2}}$$

if and only if $DC_{t,t+1}^{GP} = 1$.

As happens with the PPMI, if we work with a standard panel data instead of a pseudo-panel database, the efficiency gap change coincides with the ratio of the geometric mean of the traditional efficiency changes calculated using exclusively group G in t and $t+1$ and the geometric mean of the traditional efficiency changes calculated using exclusively group P in t and $t+1$, i.e. $EGC_{t,t+1}^{GP} =$

$$\frac{\left(\prod_{j=1}^N EC(x_j^{G^{t+1}}, y_j^{G^{t+1}}, x_j^{G^t}, y_j^{G^t}) \right)^{1/N}}{\left(\prod_{i=1}^M EC(x_i^{P^{t+1}}, y_i^{P^{t+1}}, x_i^{P^t}, y_i^{P^t}) \right)^{1/M}}$$

As for the technological gap change and under the satisfaction of Proposition 2, this term coincides with the ratio of the geometric mean of the traditional technological changes calculated using exclusively group G in t and $t+1$ and the geometric mean of the traditional technological changes calculated using exclusively group P in t and $t+1$, i.e. $TGC_{t,t+1}^{GP} =$

$$\frac{\left(\prod_{j=1}^N TC(x_j^{G^{t+1}}, y_j^{G^{t+1}}, x_j^{G^t}, y_j^{G^t}) \right)^{1/N}}{\left(\prod_{i=1}^M TC(x_i^{P^{t+1}}, y_i^{P^{t+1}}, x_i^{P^t}, y_i^{P^t}) \right)^{1/M}}$$

Next, we illustrate the introduced methodology through three simple numerical examples, before showing how this performs in a real situation.

Example 1. Let us assume that we have observed three different units in two periods: A, B and C. Units A and B belong to the group G, whereas unit C belongs to the group P. Let us also suppose that the technology produces a single output with a single input. The data are as follows: $A(t)=(10,10)$, $B(t)=(10,1)$, $C(t)=(10,5)$, $A(t+1)=(10,15)$, $B(t+1)=(10,1)$ and $C(t+1)=(10,5)$. The example is illustrated by Fig. 2. Applying the definitions

of the CDMI, PPMI and their sub-components, we obtain that $CDMI_t^{GP} = 0.632 = \underbrace{0.316}_{EG_t^{GP}} \cdot \underbrace{2}_{TG_t^{GP}}$, $CDMI_{t+1}^{GP} = 0.775 = \underbrace{0.258}_{EG_{t+1}^{GP}} \cdot \underbrace{3}_{TG_{t+1}^{GP}}$ and

$PPMI_{t,t+1}^{GP} = 1.225 = \underbrace{0.816}_{EG_{t,t+1}^{GP}} \cdot \underbrace{1.5}_{TGC_{t,t+1}^{GP}}$. This means that group P (unit

C) had a better performance than group G (units A and B) both in t and $t+1$ periods (setting 1) and it was due exclusively to the dimension associated with the technical efficiency spread ($EG_t^{GP} < 1$ and $EG_{t+1}^{GP} < 1$). However, the group G is catching up on the group P since $PPMI_{t,t+1}^{GP} = 1.225 > 1$ (setting 1(b)). In particular, the relative performance of G compared to P improved by 22.5 percent from t to $t+1$, despite going behind the group P in both periods. Finally, and regarding Proposition 1, $DC_{t,t+1}^{GP} = 1.107 \neq 1$. Therefore, (12) does not hold in this example and the PPMI cannot be interpreted as the ratio of aggregated productivity changes from t to $t+1$ in the two groups of units.

Example 2. Let us assume now that we have observed again three different units in two periods: A, B and C. Units A and B belong to the group G, whereas unit C belongs to the group P. Let us also suppose that the technology produces a single output with a single input. In this case, the data are as follows: $A(t)=(10,10)$, $B(t)=(10,1)$, $C(t)=(10,5)$, $A(t+1)=(10,10)$, $B(t+1)=(10,1)$ and $C(t+1)=(10,10)$. This example is illustrated by Fig. 3. Applying the definitions of the CDMI, PPMI and their sub-components, we get $CDMI_t^{GP} = 0.632 = \underbrace{0.316}_{EG_t^{GP}} \cdot \underbrace{2}_{TG_t^{GP}}$, $CDMI_{t+1}^{GP} = 0.316 = \underbrace{0.316}_{EG_{t+1}^{GP}} \cdot \underbrace{1}_{TG_{t+1}^{GP}}$

and $PPMI_{t,t+1}^{GP} = 0.5 = \underbrace{1}_{EG_{t,t+1}^{GP}} \cdot \underbrace{0.5}_{TGC_{t,t+1}^{GP}}$. This means that group P

(unit C) had a better performance than group G (units A and B) both in t and $t+1$ periods (setting 1) and, as in the first example, the dimension associated with the technical efficiency spread was to blame ($EG_t^{GP} < 1$ and $EG_{t+1}^{GP} < 1$). Additionally, the relative performance gap was opened up by P over G ($PPMI_{t,t+1}^{GP} = 0.5 < 1$). In fact, PPMI indicates the percentage in which the relative performance of G compared to P has worsened (setting 1(a)). Specifically, the relative performance of G compared to P worsened significantly by 50 percent from t to $t+1$. Taking into account Proposition 1, $DC_{t,t+1}^{GP} = 1$. Therefore, (12) holds.

Moreover, (13) also holds since we are working with a panel data in this numerical example. If the traditional Malmquist productivity indexes are calculated for group G and P separately, then we obtain that $M(x_A^{G,t+1}, y_A^{G,t+1}, x_A^G, y_A^G) = 1$, $M(x_B^{G,t+1}, y_B^{G,t+1}, x_B^G, y_B^G) = 1$ and $M(x_C^{P,t+1}, y_C^{P,t+1}, x_C^P, y_C^P) = 2$. Therefore, $(\prod_{j=1}^N M(x_j^{G,t+1}, y_j^{G,t+1}, x_j^G, y_j^G))^{1/N} / (\prod_{i=1}^M M(x_i^{P,t+1}, y_i^{P,t+1}, x_i^P, y_i^P))^{1/M} = 0.5$, which coincides with the value calculated for the $PPMI_{t,t+1}^{GP}$ in this example.

In the first two examples, we resorted to standard panel databases. In the next simple numerical example, we will use a pseudo-panel database. Analyzing the results obtained in the Example 2, group B has a worse relative performance than group P both in periods t and $t+1$, and the reason is its behavior in the technical efficiency spread with respect to group P ($EG_t^{GP} < 1$ and $EG_{t+1}^{GP} < 1$). In particular, unit B was the culprit, since it was very inefficient regarding the frontier in G. In this way, we will delete unit B from period t of the database for generating a new example (Example 3), improving the technical efficiency spread for group G in the first period.

Example 3. Let us assume that we have observed two units in period t (A and C) and three units in period $t+1$ (A, B and C). Units A and B belong to the group G, whereas unit C belongs to the group P. Let us also suppose that the technology produces a single output with a single input. The data are: $A(t) = (10, 10)$, $C(t) = (10, 5)$, $A(t+1) = (10, 10)$, $B(t+1) = (10, 1)$ and $C(t+1) = (10, 10)$. In this example, $CDMI_t^{GP} = 2 = \underbrace{1}_{EG_t^{GP}} \cdot \underbrace{2}_{TC_t^{GP}}$, $CDMI_{t+1}^{GP} = 0.316 = \underbrace{0.316}_{EG_{t+1}^{GP}} \cdot \underbrace{1}_{TC_{t+1}^{GP}}$ and $PPMI_{t,t+1}^{GP} = 0.158 = \underbrace{0.316}_{EG_{t,t+1}^{GP}} \cdot \underbrace{0.5}_{TC_{t,t+1}^{GP}}$. Consequently, this situation

is related to setting 3. Therefore, group G had a better relative performance than group P in period t but P had a better behavior than G in the second period, $t+1$. Moreover, the status of group G worsened drastically over time since $PPMI_{t,t+1}^{GP} = 0.158 < 1$. Specifically, the relative performance gap is reduced by 84.2 percent ($0.842 = 1 - 0.158$). Going into detail, $EG_{t,t+1}^{GP} = 0.316 < 1$ and comes from the ratio of $EG_t^{GP} = 1$ and $EG_{t+1}^{GP} = 0.316 < 1$. So, setting 1(a) applies for this sub-component. Groups G and P had the same behavior, regarding efficiency spread, in the first period. However, P had a better performance than group G in $t+1$. Moreover, the efficiency gap was opened up by P over G (the efficiency spread of G compared to P has worsened by 68.4 percent). Regarding the frontier productivity gap, we have that $TC_t^{GP} = 2$ and $TC_{t+1}^{GP} = 1$ with $TC_{t,t+1}^{GP} = 0.5$ (setting 2(a)). Therefore, the group G had a frontier with greater productivity than the group P in t and there were no differences between the two groups in period $t+1$. Additionally, we observe that the group P is catching up on the group G with respect to the technological dimension. The frontier productivity gap decreased by 50 percent from t to $t+1$. Finally, regarding Proposition 1, $DC_{t,t+1}^{GP} = 1$. Therefore, (12) holds and the PPMI may be interpreted as the ratio of aggregated measures of productivity changes in the two groups separately. However, (13) cannot be applied because we are working with a pseudo-panel database instead of a standard panel database.

3. Educational production function, dataset and variables

In our empirical analysis, we use data from three PISA waves: 2006, 2009 and 2012. PISA is an initiative that the OECD started up in the late 1990s to assess 15-year-old students. The assessment focuses on measuring the extent to which students are able to apply their knowledge and skills to achieve future real-life challenges rather than evaluating how well they have mastered a specific school curriculum. Because home, school, and regional en-

vironments can play an important role in how students learn, PISA also collects general information about such background factors.

In order to illustrate the potential of the approach proposed here, we provide an application to the Basque Country. The Basque Country is an Autonomous Region of Spain that has been entirely responsible for deciding how many educational resources to allocate to the educational sector and its management since 1980.⁷ This region has opted to participate with an extended representative sample of its population for PISA assessment since 2003. Publicly funded schools in the Basque Country and Spain, i.e., schools receiving their core funding from government agencies, may be classified as either *PS* or *PGDS*. The difference lies in whether a public entity or a private agency, respectively, has decision-making authority concerning their management. *PS* are monitored and managed by a public education authority or agency. *PGDS* are governed by a non-public organization,⁸ which implies that their governing board is not elected by a government agency. Private schools are classified as *PGDS* if they receive more than 50 percent of their core funding from government agencies.⁹ It is important to note here that the academic program; the time allocated to each subject, as well as the academic degree required to qualify as a teacher is regulated by law and is exactly the same for both school types.

3.1. Educational production function

In many earlier papers on the economics of education, the common conceptual framework for estimating the educational production function at school level has taken the following form (Hanushek, 1979; Hanushek, Link, & Woessmann, 2013; Levin, 1974):

$$A^t = f(B^t, S^t) \tag{8a}$$

Where A^t is the single output that now equals the average achievement of students attending the evaluated school at time t ; B^t is the school's average socio-economic background and S^t are average school resources. Both B^t and S^t are the input components of x^t in Eq. (1), $x^t = (B^t, S^t)$. We can generalize the production function for the case where a school produces more than one output $A^t = y^t = (y_1^t, \dots, y_5^t)$ using the transformation function $H^t(x^t, y^t)$ assuming that there are inefficient behaviors in schools (Grosskopf, Hayes, & Taylor, 2014; Johnes, 2014; Nechyba, 2000; Perelman and Santín, 2011; Woessman, 2001). In this way, we can calculate all the Malmquist index components.

3.2. Outputs and inputs

One of the main advantages of the PISA study is that each student receives a score in each test on a continuous scale. Likewise, PISA also collects a large set of data about the educational context from two questionnaires: one completed by the students themselves and another filled out by the principals. From these data, it is possible to extract information referred to the main determining factors of educational performance represented by variables associated with the family and the educational environments, as well as with the school management and the educational supply.

The school output includes student knowledge and skills in different dimensions measured by the results that they achieve in the

⁷ The Basque Country was the first Autonomous Region (or Community) inside Spain upon which educational policy powers were devolved by the central government after the Spanish transition to democracy in 1977.

⁸ Most of these organizations are Catholic schools, teachers' cooperatives, non-profit organizations or simply private enterprises.

⁹ There are also government-independent private schools, controlled by private organizations which receive most of their core funding from fees paid by students. In this paper, we focus only on publicly funded schools.

three vehicular competences evaluated in PISA (mathematics, reading and science). Fair comparisons of the two groups of schools over time inside a country or region with a representative sample are possible because PISA uses a common scale for the purpose of trends. However, results in different PISA waves are only directly comparable for some years and subjects (OECD, 2014a, p. 52–53). It is important to remark here under what circumstances PISA results in different waves are comparable. As OECD (2014b, p. 159) states “for PISA 2012 the decision was made to report the reading, mathematics and science scores on these previously developed scales. That is the reading scales used for PISA 2000, PISA 2003, PISA 2006, PISA 2009 and PISA 2012 are directly comparable. PISA 2012 mathematics reporting scale is directly comparable to PISA 2003, PISA 2006 and PISA 2009 and the science reporting scale is directly comparable to PISA 2006 and PISA 2009 scale”. As a consequence, in this work we have not included PISA 2003 data because science results are only comparable from 2006 on, and this only includes PISA 2006, 2009 and 2012.

In order to calculate the output-oriented CDMI and the PPMI we use four inputs directly involved in the learning process. We consider that schools have to operate with and try to achieve the best outputs from the resources that they have been allocated (including staff, school resources and also student background¹⁰). Therefore, the output-oriented CDMI and PPMI are the correct specifications for characterizing this production process and are selected to interpret performance gaps between both school types.

- *PARED* is the index of the highest level of parental education, measured by the number of years of schooling according to the International Standard Classification of Education (ISCED; OECD, 1999).
- *HISEI* is the index of the highest parental occupational status according to International Socio-Economic Index of Occupational Status (ISEI; Ganzeboom, De Graaf, Treiman, & De Leeuw, 1992).
- *SCHRESOURCES* is an index of the quality of the school resources derived from school principal responses. All questionnaires contain several items related to school deficiencies regarding such issues, but some items are different across the four waves. So ten coincident items were selected for each sample, and the school receives one point for each item for which the principal's response is *not deficient*.¹¹ The maximum (minimum) score for each school is 10 (zero) points, which indicates an excellent (poor) educational input.¹²
- *STRATIO* is a ratio between the total number of teachers weighted by their dedication (part-time teachers contribute 0.5 and full-time teachers 1) and the total number of pupils and the total number of students.

Table 1 shows the mean values for the three outputs—student results in mathematics, reading and science—and the four inputs named above. The figures below indicate that student results were generally higher for students attending *PGDS* in all disciplines from 2006 to 2009. However, we observe a change in this trend in the last PISA 2012 wave, where it is slightly higher in *PS*. A similar trend applies to the average socioeconomic background, measured by the variables *PARED* and *HISEI*. These variables used to be lower

in *PS*, but socioeconomic background is very similar in both school types in the last wave.

With respect to school resources, we observe that there are no significant differences in the *SCHRESOURCES* variable over the years. Nevertheless, the teacher-student ratio (*STRATIO*) is favorable to students attending *PS*, where the ratio is higher. This implies that each teacher is in charge of a smaller group of students. This advantage of *PS* can be explained by the fact that *PGDS* tend to minimize the number of teachers that the school hires and conform strictly to the requirements of educational law as a way to cut costs. Also, the number of teachers in *PGDS* is relatively lower because they are assigned to teach general knowledge areas (sciences, arts, sports, etc.), whereas most teachers in *PS* are public servants and their teaching profile is more specialized.

4. Results

This section reports the main results of our analysis for the Basque Country. Using the methodology described in Section 2, we can compare *PGDS* and *PS* relative performances over time from 2006 to 2012.

Table 2 reports the results after applying the CDMI approach on school data for 2006, 2009 and 2012, considering the set of three outputs and four inputs. Column 1 in Table 2 shows the $CDMI_t^{GP}$ for each year t and is calculated according to Eq. (3). Results show relevant relative performance gaps, ranging from 36.95 percent (2006) to 25.66 percent (2012), every year, according to which *PGDS* have an edge over *PS*. The technological gap component TC_t^{GP} , calculated through Eq. (4), clearly accounts for this gap in every wave, so the circumstances shown by the empirical analysis reveal that it is very similar to the situation illustrated in Fig. 1. *PGDS* presents the most productive frontiers, leading to a persistent productivity difference with respect to their public counterparts ($TC_t^{GP} > 1$). Finally, the technical efficiency spread gap EG_t^{GP} appears to play a minor role in explaining the relative performance gap since, in Table 2, EG_t^{GP} is close to 1 for each year except for 2006 where $EG_t^{GP} = 0.9677$.

Apart from the $CDMI_t^{GP}$ and its components, Table 3 reports the time evolution of the relative performance gap ($PPMI_{t,t+1}^{GP}$) and its components, the efficiency spread gap change ($EGC_{t,t+1}^{GP}$) and the productivity frontier gap change ($TGC_{t,t+1}^{GP}$).

Although *PGDS* had a better relative performance than *PS*, Table 3 shows that, on average, *PS* was catching up on the *PGDS* schools from 2006 to 2009 (by 10.44 percent), while relative performance gap widened by an additional 2.46 percent in favor of *PGDS* from 2009 to 2012. In both cases, the relative gap change can be explained by the technological component, which presents the same behavior over time. The productivity frontier gap improved in favor of *PS* by 12.8 percent with respect to *PGDS* from 2006 to 2009, whereas it worsened by 1.04 percent from 2009 to 2012. In the case of the efficiency dimension, *PS* were able to narrow the gap with respect to *PGDS* by around 2.37 percent from 2006 to 2009, while the status of *PGDS* improved from period 2009 to 2012 by 1.74 percent.

Regarding the divergence component, it equals 0.9138 for the first biannual period and 0.9304 for the second one. It means that the *PPMI* does not coincide with the ratio of productivity changes in group *PGDS* and *PS*. Indeed, the *PPMI* is smaller than the ratio of productivity changes by 8.62 percent from 2006 to 2009 and by 6.96 percent from 2009 to 2012.

To summarize, the results show a persistent relative performance difference in favor of *PGDS* across all three waves. This means, according to an output orientation interpretation, that, with the same resources, student backgrounds and school inputs, the

¹⁰ As Bradley et al (2001, p. 554) state, socioeconomic level reflects the quality of the raw material.

¹¹ The selected items are: 'Qualified science teachers', 'Qualified mathematics teachers', 'Qualified reading teachers', 'Any other personal support', 'Science laboratory equipment', 'Educational material', 'Computers', 'Software', 'Library resources', 'Audiovisual resources'. This variable is subjective and for this reason is only a *proxy* of a more realistic 'physical school resources' variable.

¹² This variable has been rescaled so that the minimum value is one in order to avoid zero values in the empirical analysis.

Table 1
Descriptive statistics of outputs and inputs in the Basque Country over three PISA waves.

Public Schools (PS)						
Variables	2006		2009		2012	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
OUTPUTS						
Math	477.99	53.47	497.61	40.53	507.13	34.92
Read	464.82	54.95	480.02	34.07	502.60	39.77
Science	474.26	44.51	481.58	28.73	506.78	32.65
INPUTS						
Pared	11.86	1.71	12.96	1.30	52.25	9.62
Hisei	44.47	6.78	45.55	6.06	13.76	1.16
Schresources	8.48	2.51	9.01	1.55	9.22	1.87
Stratio	15.54	2.89	15.02	2.99	14.17	3.37
Obsv.	56		68		72	
Private Government-Dependent Schools (PGDS)						
Variables	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
OUTPUTS						
Math	512.45	38.62	520.73	34.77	499.91	48.75
Read	500.62	39.03	509.52	36.64	491.41	50.41
Science	505.40	36.91	507.56	35.15	502.73	47.63
INPUTS						
Pared	12.89	1.57	13.80	1.37	13.48	10.57
Hisei	49.33	8.42	51.53	8.07	49.13	1.54
Schresources	8.06	2.73	9.08	1.85	9.72	1.63
Stratio	6.89	1.25	7.41	1.72	7.17	1.68
Obsv.	81		90		83	

Source: Own compilation from PISA 2006, 2009 and 2012.

Table 2
The CDMI for GDPS and PS in the Basque Country over time.

t	$CDMI_{t,t+1}^{GP}$	EG_t^{GP}	TG_t^{GP}
2006	1.3695	0.9677	1.4152
2009	1.2265	0.9906	1.2341
2012	1.2566	1.0078	1.2469

A value greater than one indicates better performance in PGDS than in the PS group.

Table 3
Pseudo-panel Malmquist index (PPMI). Productivity gains in PS and PGDS from 2006 to 2012.

t; t + 1	$PPMI_{t,t+1}^{GP}$	$EG_{t,t+1}^{GP}$	$TGC_{t,t+1}^{GP}$
2006; 2009	0.8956	1.0237	0.8720
2009; 2012	1.0246	1.0174	1.0104

A value greater (less) than one indicates that the relative performance, efficiency spread or technological gaps open for PGDS (PS) with respect to PS (PGDS).

best PGDS are producing more than PS. We suggest two possible explanations for these results.

First, according to the descriptive statistics in Table 1, having more pupils per full-time equivalent teachers in PGDS does not generally appear to deteriorate the learning process. This result is consistent with Hanushek's (1986, 1997, 2003) previous findings that pouring more educational resources into schools may not automatically translate into better academic results.

Second, it is unclear which other factors account for these relative performance differences, but their existence warrants more research to monitor and correct this gap in the near future. Remember, however, that the theoretical literature referenced in Section 1 points out that private institution and teachers employed in the private sector have more incentives to make efficient use of resources and tend to be more productive. Some people prefer the PS option because they believe that teacher quality is higher as PS teachers need to pass a competitive state exam to enter the public school system. This is considered to possibly lead to better overall academic achievement. But teachers in PS are automatically

granted tenure once they pass the entrance exam. This leads some people to argue that PS teachers do not have definite incentives to improve their teaching methods and activities once they have joined the system as civil servants (basically, because they cannot be dismissed). This could affect PS performance, rendering a more flexible private management preferable. Although more research is needed, it looks as if this second hypothesis is, at least in our empirical case, more plausible.

5. Conclusions

The Malmquist index methodology is widely used in the literature in order to measure the productivity growth within two time periods as the distance between each DMU and the frontier for each period. However, the traditional Malmquist index cannot be implemented without a panel database, so it focuses on analyzing the evolution of the same units over time.

In this paper, we extend the Malmquist-type index proposed by Camanho and Dyson (2006) that is used to measure the average relative performance divergences between different groups of DMUs within the same year. Thus, a pseudo-panel Malmquist index is built for comparing the evolution of average performance discrepancies between publicly funded schools, including both public and private government-dependent schools, when only a pseudo-panel database is available.

To illustrate this approach, we use data on publicly funded schools (PGDS and PS) in the educational context of the Spanish Basque Country region from the PISA 2006, 2009 and 2012 databases that provide us with a wide range of information. Results indicate that PGDS persistently obtain better relative performance than PS over time. PGDS outperform PS due to PGDS's technological superiority. This means that the within-group technical efficiency variance is on average similar at both school types, and there will be good and bad schools within each group. However, the best schools that define the most productive frontier belong to the PGDS group. Also, the evolution of the technological gap from 2006 to 2012 is flat, so the initial differences in favor of PGDS are still there at the end of the analyzed 6-year period. From a policy-making viewpoint, this result suggests that more

attention and resources should be spent on improving teaching practices inside classrooms and on monitoring the educational system in order to close the performance gap between all schools funded with public resources.

Finally, in their empirical application Camanho and Dyson (2006) ran the CDMI with different, but similar, group sizes. As they recognize, there may be an unaccounted-for group size effect when group sizes are severely unbalanced, although it is expected to be insignificant when both groups have similar sizes. For this reason we suggest, as a future line of research, to develop a bootstrap procedure to build confidence intervals for the real $CDMI_{t,t+1}^{GP}$, $PPMI_{t,t+1}^{GP}$ and their components to face this potential problem when comparing unbalanced groups of DMUs.

Acknowledgments

We thank three anonymous referees for helpful discussions and suggestions. Eva Crespo-Cebada, Francisco Pedraja-Chaparro and Daniel Santín acknowledges the financial support of the Spanish Ministry of Economy and Competitiveness (ECO2014-53072-P). Additionally, Juan Aparicio is grateful to the same institution for supporting this research through grant MTM2013-43903-P.

References

- Afsharian, M., & Ahn, H. (2015). The overall Malmquist index: A new approach for measuring productivity changes over time. *Annals of Operations Research*, 226(1), 1–27.
- Alchian, A. A. (1950). Uncertainty, evolution and economic theory. *Journal of Political Economy*, 58, 211–221.
- Balk, B. M., & Althin, R. (1996). A new, transitive productivity index. *The Journal of Productivity Analysis*, 7, 19–27.
- Berg, S. A., Forsund, F. R., Hjalmarsson, L., & Suominen, M. (1993). Banking efficiency in the Nordic countries. *Journal of Banking and Finance*, 17, 371–388.
- Bradley, S., Johnes, G., & Millington, J. (2001). The effect of competition on the efficiency of secondary schools in England. *European Journal of Operational Research*, 135(3), 545–568.
- Brennan, S., Haelermans, C., & Ruggiero, J. (2014). Nonparametric estimation of education productivity incorporating nondiscretionary inputs with an application to Dutch schools. *European Journal of Operational Research*, 234(3), 809–818.
- Camanho, A. S., & Dyson, R. G. (2006). Data envelopment analysis and Malmquist indices for measuring group performance. *Journal of Productivity Analysis*, 26, 35–49.
- Caves, D. W., Christensen, L. R., & Diewert, W. E. (1982). The economics theory of index numbers and the measurement of input, output and productivity. *Econometrica*, 50, 1393–1414.
- Charnes, A., Cooper, W. W., & Rhodes, E. (1981). Evaluating program and managerial efficiency: An application of data envelopment analysis to program follow through. *Management Science*, 27, 668–697.
- Chubb, J. E., & Moe, T. M. (1990). *Politics, markets and America's schools*. Washington: The Brookings Institution.
- Crespo-Cebada, E., Pedraja-Chaparro, & Santín, D. (2014). Does school ownership matter? An unbiased efficiency comparison for regions of Spain. *Journal of Productivity Analysis*, 41(1), 153–172.
- De La Fuente, A. (2011). Human capital and productivity. *Nordic Economic Policy Review*, 2, 103–131.
- De Witte, K., & López-Torres, L. (2015). Efficiency in education: A review of literature and a way forward. *Journal of the Operational Research Society*. doi:10.1057/jors.2015.92.
- Essid, H., Oullette, P., & Vigeant, S. (2014). Productivity, efficiency and technical change of Tunisian schools: A bootstrapped Malmquist approach with quasi-fixed inputs. *OMEGA*, 42(1), 88–97.
- Färe, R., Grosskopf, S., & Lovell, C. A. K. (1994). *Production frontiers*. Cambridge University Press.
- Ferreira, D., & Marques, R. C. (2015). Did the corporatization of Portuguese hospitals significantly change their productivity? *European Journal of Health Economics*, 16(3), 289–303.
- Forsund, F.R. (2002). On the circularity of the Malmquist productivity index, ICR working papers, pp. 29.
- Friedman, M., & Friedman, R. (1981). *Free to choose*. New York: Avon.
- Frisch, R. (1936). Annual survey of general economic theory: The problem of index numbers. *Econometrica*, 4, 1–38.
- Ganzeboom, H., De Graaf, P., Treiman, J., & De Leeuw, J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56.
- Grifell-Tatjé, E., & Lovell, C. K. (1999). A generalized Malmquist productivity index. *Top*, 7(1), 81–101.
- Grosskopf, S., Hayes, K. J., & Taylor, L. L. (2014). Efficiency in education: Research and implications. *Applied Economic Perspectives and Policy*, 36(2), 175–210.
- Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources*, 14, 351–388.
- Hanushek, E. A. (1986). The economics of schooling. *Journal of Economic Literature*, 24(3), 1141–1171.
- Hanushek, E. A. (1997). Assessing the effects of school resources on student performance, an update. *Educational Evaluation and Policy Analysis*, 19, 141–164.
- Hanushek, E. A. (2003). The failure of input based schooling policies. *The Economic Journal*, 113, 64–98.
- Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review*, 90(5), 1184–1208.
- Hanushek, E. A., Link, S., & Woessmann, L. (2013). Does school autonomy make sense everywhere? Panel estimates from PISA. *Journal of Development Economics*, 104, 212–232.
- Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3), 607–668.
- Hoxby, C. M. (Ed.). (2003). *The economics of school choice*. Chicago: University of Chicago Press.
- Johnes, J. (2014). Operational research in education. *European Journal of Operational Research*. doi:10.1016/j.ejor.2014.10.043.
- Kirjavainen, T., & Loikkanen, H. A. (1998). Efficiency differences of Finnish senior secondary schools: An application of DEA and Tobit analysis. *Economics of Education Review*, 17(4), 377–394.
- Levin, H. M. (1974). Measuring efficiency in educational production. *Public Finance Quarterly*, 2, 3–24.
- Mancebón, M. J., Calero, J., Choi, A., & Ximenez, D. (2012). Efficiency of public and publicly-subsidized high schools in Spain. Evidence from PISA 2006. *Journal of Operational Research Society*, 63, 1516–1533.
- Mancebón, M. J., & Muñiz, M. A. (2008). Private versus public high schools in Spain: Disentangling managerial and programme efficiencies. *Journal of the Operational Research Society*, 59(7), 892–901.
- Maragos, E. K., & Despotis, D. K. (2004). Evaluating school performance over time in the frame of regional socio-economic specificities. *WSEAS Transactions on Mathematics*, 3(3), 664–670.
- Nechyba, T. J. (2000). Mobility targeting and private-school vouchers. *American Economic Review*, 90(1), 130–146.
- OECD (1999). *Classifying educational programmes, manual for ISCED-97 implementation in OECD countries*. Paris: Organisation for Economic Co-operation and Development.
- OECD (2014a). *PISA 2012 results: What students know and can do—Student performance in mathematics, reading and science: I*. OECD Publishing revised edition, PISA.
- OECD (2014b). *PISA 2012 Technical Report*. OECD Publishing.
- Pastor, J. T., & Lovell, C. A. (2005). A global Malmquist productivity index. *Economics Letters*, 88(2), 266–271.
- Perelman, S., & Santín, D. (2011). Measuring educational efficiency at student level with parametric stochastic distance functions: An application to Spanish PISA results. *Education Economics*, 19(1), 29–49.
- Portela, M. C., Camanho, A. S., & Keshvari, A. (2013). Assessing the evolution of school performance and value-added: Trends over four years. *Journal of Productivity Analysis*, 39(1), 1–14.
- Ray, S., & Desli, E. (1997). Productivity growth, technical progress, and efficiency change in industrialized countries: Comment. *American Economic Review*, 87(5), 1033–1039.
- Thanassoulis, E., Shiraz, R. K., & Maniadakakis, N. (2015). A cost Malmquist productivity index capturing group performance. *European Journal of Operational Research*, 241(3), 796–805.
- Vandenbergh, V., & Robin, S. (2004). Evaluating the effectiveness of private education across countries: A comparison of methods. *Labour Economics*, 11(4), 487–506.
- Vaz, C. B., & Camanho, A. S. (2012). Performance comparison of retailing stores using a Malmquist-type index. *Journal of the Operational Research Society*, 63(5), 631–645.
- Woessman, L. (2001). Why students in some countries do better. *Education Matters*, 1(2), 67–74.
- Xue, M., & Harker, P. T. (2002). Note: Ranking DMUs with infeasible super-efficiency DEA models. *Management Science*, 48(5), 705–710.