# On the use of Synthetic Indexes based on Multi-Criteria Decision Making to Study the Efficiency of Teachers

Ana B. Ruiz[1] , Mariano Luque[1,*], Oscar D. Marcenaro-Gutierrez[2]

[1]*Department of Applied Economics (Mathematics)*

*Universidad de Málaga, C/ Ejido, 6, 29071, Málaga, Spain*

[2]*Department of Applied Economics (Statistics and Econometrics)*

*Universidad de Málaga, C/ Ejido, 6, 29071, Málaga, Spain*

abruiz@uma.es, mluque@uma.es, odmarcenaro@uma.es

## Abstract

Efficiency in education constitutes a dimension of the quality of education and gaining knowledge about the teachers' efficiency can contribute to improving the learning achievement of students. The main purpose of this paper is to evaluate the teachers' efficiency in order to determine which policies should be promoted in the future to reach a teaching environment enabling the teachers to be efficient. That is, we want to have a better insight about how to distribute the available resources to improve key factors of the teaching-learning context that are desirable to let them achieve efficiency. To this aim, we measure the teachers' efficiency from a wide multi-criteria decision making sense. We consider all the significant variables at the same importance level, assuming that efficient teachers are those achieving the best possible values for all of them simultaneously. Data from TIMSS and PIRLS 2011 for fourth-grade Spanish teachers in reading, mathematics and science are used. Firstly, three synthetic indexes (weak, strong and mixed) are built using a multi-criteria decision making approach that incorporates desirable ranges for the variable values. This allows us to know which teachers are performing better, within or worse than the desired limits. Afterward, a second stage analysis is performed to investigate the extent to which the controllable variables are

[*]Corresponding author. Tel.: +34 952131173; Fax: +34 952132061.

determinant in the teachers' rankings associated with these indexes. With all of this, key variables for being in the best/worst ranking positions are identified, which lets us better know which measures should be taken to promote efficiency in teaching.

***Keywords:*** efficiency; teachers; synthetic index; multi-criteria decision making.

**JEL classification**: A22 ● C44 ● C61.

## 1. Introduction

Education has a direct effect on the social capital, economic growth and well-being of a country. As underlined by e.g. Hanushek and Woessmann (2007), it plays an essential role in providing well-being to the population, placing the quality of education in a relevant position in the political debate. Indeed, the quality of education is so closely related to the development of a population's cognitive skills that authors such as Hanushek and Woessmann (2007) used them as its proxy.

Commonly, the quality of education has been analysed by means of comparative evaluations of the students' performances, such as the Programme for International Student Assessment (PISA). However, nowadays a higher importance is being given to the role of the efficiency and effectiveness of teachers in relation to their students' outcomes, as a way to improve the quality of the education systems (Dolton et al., 2018). Having effective and efficient teachers, and making an efficient use of the educational resources available, may be key issues to enhance the students' academic performance and their learning achievement, and to raise their implication in their own learning process. However, to promote the teachers' efficiency, firstly we need to identify the main features of the teachers that are considered as efficient, in order to gain knowledge about the policies that should be developed to let the other teachers achieve efficiency. This is the main focus of this paper, which –based on real data analyses– intends to provide a new methodological approach that allows to identify the most efficient teachers.

In pursuit of the efficiency of the education system through the efficiency of the teachers, there is a generalized belief that spending more time on teaching improves students' academic results (Locher and Pfost, 2020), the relationship between instruction time and academic achievement is not as obvious as it might seem, primarily because of the quality of time (Mullis and Martin, 2017). Lopez-Agudo and Marcenaro (2022) indicate that time for face-to-face teaching does not seem to be positively associated with students' academic performance for any of the countries under their analysis. Other authors (Egalite and Kisida, 2018; Lim and Meer, 2021; Calero and Escardíbul, 2019) also support that, when the instruction time is increased, other teacher's characteristics, such as gender, years of experience and level of education, has not a clear impact on the academic performance of the students.

In the education research field, an efficient teacher has been understood as the one achieving the maximum outputs (e.g. academic scores of the teachers' students) from a given amount of inputs of the teaching-learning process (e.g. resources devoted to the educational system, students' background variables, socio-economic characteristics of the teaching/learning context, etc.). In the other way around, teachers being capable to reach the same level of outputs with the minimum quantity of inputs are also considered as efficient teachers (Lockheed and Hanushek, 1994). This definition of teachers' efficiency corresponds to a conventional perspective, i.e. the output-oriented technical efficiency (a measure of how well decision making units maximize outputs when inputs and outputs mixes are predetermined), or the profit efficiency (a measure of how well decision making units maximize profits when inputs and outputs can be chosen freely). Several studies adopt this definition, such as Jürges and Schneider (2007), which proposes a regression-based method to construct "fair" rankings of teachers and schools (meaning by "fair" that important determinants of student achievement beyond the control of the individual teacher and that the imprecision inherent in measures of student achievement are accounted for).

However, teachers' performance evaluation is nothing but a multi-criteria decision making problem. As said, several quality attributes influence the efficiency of a potential teacher while guiding his /her students towards good academic outcomes. In this line, we adopt a wider multi-criteria decision making perspective to define teachers' efficiency in this paper. We propose quantifying their efficiency considering both inputs and outputs as goals or criteria to be improved as much as possible, and thus, we assume that efficient teachers are those achieving, at the same time, the best possible values for all the inputs and outputs. That is, efficient teachers would be these ones producing the maximum levels of outputs using the minimum inputs, which are in clear conflict. Therefore, the multi-criteria dimension of the teachers' efficiency is evident. Under this perspective, we would like to know the extent to which the best possible levels for all the inputs and outputs are attainable simultaneously or, by contrast, which type of trade-offs are needed among them.

In a broad sense, we seek to establish the characteristics of the teaching-learning context of the most efficient teachers. Trying to attain the best possible values for both inputs and outputs at the same time can lead us to interesting findings regarding the profile of the teachers with a balanced (optimal) situation for all of them, i.e. regarding the scenarios that would allow the teachers' efficiency. This justifies the consideration of the inputs not just as explanatory variables, but also as criteria to be improved as much as possible. Furthermore, although the outputs are factors that may not be directly controllable by the teacher, we are interested in evaluating them jointly to the inputs so that we can have a better knowledge about the profiles of the most/least efficient teachers from both points of views, that is, from the perspectives of the inputs and of the outputs. This will provide us some hints in relation to future measures that should be promoted by politicians or stakeholders in relation to these uncontrollable factors if they want to converge to a teaching system that enables the efficiency among the teachers (under the multi-criteria perspective considered here). That is, we want to have a better insight about how to manage the available resources to improve the key factors of the teaching-learning context that would let teachers achieve efficiency.

In our study, the output variables used as a proxy of the efficiency of teachers are the students' performance in subjects such as reading, mathematics and science, and an engagement index for these subjects indicating the capacity of teachers to increase his/her students' interest and implication in their studies. In addition, the inputs used to evaluate the efficiency are variables regarding the time devoted in class to teach each of the previous subjects. The PIRLS (Progress in International Reading Literacy Study) and TIMSS (Trends in Mathematics and Science Study) datasets have been combined in their 2011 wave for fourth-grade (9/10 years old) Spanish students.

To assess the teachers' efficiency using these inputs and outputs, an option is to build synthetic indicators or indexes. Such an index would combine into one composite number the different dimensions or aspects considered about the situation being studied, enabling to benchmark the performance of the teachers. Indeed, to rank the teachers under the aforementioned perspective of efficiency, we can calculate synthetic indexes based on multi-criteria decision making approaches (Nardo et al., 2005). To be more precise, in this study, we consider three synthetic indicators (weak, strong and mixed) based on the double reference point (DRP) multi-criteria decision making approach (Wierzbicki, 2000), which were originally proposed in Ruiz et al. (2011). Mainly, these indicators have two benefits. On the one hand, since they are based on the DRP approach, we can consider reference values that are desirable and acceptable for each of the inputs and outputs –

called aspiration and reservation levels, respectively– to build the synthetic indexes. With this, we introduce into the indexes' building process the use of a threshold with desirable values to be achieved by the inputs and outputs, enabling the efficiency of each teacher to be evaluated depending on her/his performance regarding these values (i.e. if (s)he performs better, within or worse than the desired limits).

On the other hand, the DRP-based synthetic indexes allow different aggregation schemes so that different grades of compensation among the inputs and outputs can be applied. This makes it possible to benchmark the teachers and to rank them according to their efficiency level under different compensation schemes, with the purpose of analysing the main features and the learning context of the most efficient ones from different points of views. In addition, as all inputs and outputs values are internally normalized using the aspiration and reservation levels, the three indexes are calculated in the same scale and, thus, the DRP-based synthetic indexes can be easily interpreted. Indeed, they provide information of the teachers' efficiency from a global point of view, and not only individually for each variable.

Once the most efficient teachers have been identified, it would be interesting to have a deeper knowledge about the influence of the variables which are controllable by teachers (i.e. the inputs considered) in the efficiency ranking obtained with the synthetic indexes. We seek to establish the contribution of each controllable input to achieve the efficiency according to these indexes. Let us recall that efficiency is understood here as the simultaneous achievement of the best possible values for both the inputs and outputs. To this aim, we perform a second stage analysis in which we study the existing correlation among the inputs used and the three synthetic indexes' values calculated. The coefficients obtained from estimating the regression model will reflect the dependencies observed, detecting the level of dependency of each synthetic index with regards to each input and, thus, providing us with more information about how to improve the teachers' efficiency rankings previously obtained.

The rest of the paper is structured as follows. The data used in our analysis to assess the teachers' efficiency is described in Section 2, providing descriptive statistics of the sample under scrutiny. Section 3 contains a literature review about efficiency in education. In Section 4, we present the methodological process followed to build the synthetic indexes and some further comments. The efficiency ranking of the teachers is analysed in Section 5, where we also include a robustness study and the second stage analysis. Finally, the main conclusions and future research lines are drawn in Section 6.

## 2. Data

We have considered information from TIMSS and PIRLS 2011, as this was the only year both studies coincided. The purpose of TIMSS is to measure the learning achievement of students in the areas of mathematics and science at the end of the fourth grade (9–10 years old), while PIRLS is focused on the reading achievement of fourth-grade students. We chose this grade as the students are more malleable at this age (Thompson-Schill et al., 2009) and, thus, they would better reflect their teachers' procedures. In addition, Spanish primary school students have the same teacher in each cycle; hence, a teacher in third and fourth grade teaches the same students, so their engagement could be attributed to the teachers under scrutiny. Last, but not least, both datasets contain a large number of teacher-level variables gathered in their teacher questionnaire –directly linkable to the students they taught– and a measurement of the students' engagement not included in other international assessments, such as e.g. PISA.

In our analysis, both TIMSS and PIRLS information were gathered by using a two-stage random sample design (Martin and Mullis, 2012). Firstly, schools were selected and, secondly, complete classes were sampled within the selected schools. Hence, each classroom (i.e. each student) is associated with the teacher who teaches the corresponding subject. The TIMSS and PIRLS studies consist of four models of context questionnaires: student, home (only in PIRLS), teacher and school. Other international assessment studies such as PISA did not contain information about teachers' characteristics and practices until the 2015 cycle, and this information is not linkable at a student level, only at a school level (OECD, 2017). Therefore, TIMSS/PIRLS 2011 are the most adequate and recent available data to study teachers' efficiency in Spain. The Spanish sample contains 4,043 students who participated in both tests and were taught by 174 teachers.

Both PIRLS 2011 (for reading) and TIMSS 2011 (for mathematics and science) include an index called "Students Engaged in Reading/Mathematics/Science Lessons index" to represent students' engagement. According to Mullis et al. (2012), this index was created by asking students the degree to which they agree a lot, agree a little, disagree a little or disagree a lot, with the following statements: "I know what my teacher expects me to do", "I think of things not related to the lesson", "My teacher is easy to understand", "I am interested in what my teacher says", and "My teacher gives me interesting things to do"[1]. Martin and Mullis (2012) provided a measure of the validity of this index using the Cronbach alpha reliability coefficient and concluded that it reached an acceptable level for Spain (0.68 for reading, 0.6 for mathematics and 0.64 for science). Based on these findings, the engagement index has been considered as a continuous variable in our study as a proxy of teachers' efficiency.

Next, we describe the procedure employed to select the variables of the teaching-learning process for our analysis. We have classified each available teacher-related variable according to its adscription to teacher efficiency characteristics. To determine whether a teacher variable represents efficiency, we have asked ourselves the following question: "Does this variable represent the available resources for teachers in order to develop their lessons, such as materials, classroom characteristics, available time or students' background, which are not easily alterable by the teacher?" The selected variables that answered this question and had a significant effect on explaining teacher efficiency were included in this study. Full descriptive statistics of the variables considered for the sample under scrutiny are provided in Table 1, reporting the mean, standard deviation and maximum and minimum values attained.

**Table 1** Selected variables and their descriptive statistics from TIMSS/PIRLS 2011 in fourth grade, Spain.

| **Variables** | **Type** | **Mean** | **Standard Deviation** | **Maximum value** | **Minimum value** |
|---|---|---|---|---|---|
| ***Outputs*** | | | | | |
| Mean scores in: | | | | | |
| Reading | Cont. | 519.18 | 33.54 | 592.29 | 392.96 |
| Mathematics | Cont. | 489.22 | 36.52 | 581.62 | 347.87 |
| Science | Cont. | 512.81 | 36.00 | 604.63 | 376.26 |
| Engagement in: | | | | | |
| Reading | Cont. | 9.94 | 0,87 | 12.16 | 6.48 |
| Mathematics | Cont. | 10.11 | 0.85 | 12.64 | 5.81 |
| Science | Cont. | 10.08 | 0.87 | 12.05 | 6.10 |
| ***Inputs*** | | | | | |
| Teaching minutes in: | | | | | |
| Reading | Cont. | 337.24 | 104.22 | 780.00 | 120.00 |

[1] PIRLS 2011 includes two additional questions in the engagement index: "I like what I read about in school" and "My teacher gives me interesting things to read".

| | | | | | |
|---|---|---|---|---|---|
| Mathematics | Cont. | 278.02 | 42.79 | 420.00 | 180.00 |
| Science | Cont. | 237.96 | 51.53 | 540.00 | 60.00 |

Source: Author’s own elaboration.

## 3. State of the Art

In the literature, we can find many studies on the efficiency of education, which mainly differ from this research work in the methodology used to quantify the efficiency. De Witte and Lopez-Torres (2017) provides a wide literature review of different methodological approaches applied to this research field.

Some studies use parametric methods to assess the efficiency, such as stochastic frontier analysis (e.g. see Jürges and Schneider (2007), De Witte et al. (2013), Scippacercola and D’Ambra (2014), and Lopez-Agudo and Marcenaro-Gutierrez (2017)). In many other works, non-parametric techniques are applied, such as data envelopment analysis (DEA) (Essid et al., 2010; Haelermans and Ruggiero, 2013; Blackburn et al., 2014; Wanke et al., 2016). Regarding the use of the students’ performance (in terms of average test scores in different subjects) to carry out educational efficiency analyses, as we do, some recent research works can be found in Crespo et al. (2014), Johnson and Ruggiero (2014), Podinovski et al. (2014), and Agasisti and Johnes (2015).

Note that the methodological differences existing in the literature do not necessarily influence the quality of the studies, but do affect the type of results obtained, their interpretation and the conclusions raised. However, there is a lack of works performing a study of the efficiency of teachers trying to optimize both inputs and outputs simultaneously, in order to know the characteristics of the most efficient teachers reaching a balanced (optimal) situation for all of them. To fill this gap, we apply a new technique (the DRP-based synthetic indicators) to gain new insights about the problem under study from a multi-criteria decision making perspective. Previously, the DRP-based synthetic indicators have been used in other research fields. It has been used to evaluate the weak and the strong sustainability of territories (Ruiz et al., 2011; Cabello et al., 2021; Ruiz and Cabello, 2021), to measure the human development (Luque et al., 2016), and to measure child and maternal health in developing countries (Perez-Moreno et al., 2016). In addition, the countries’ competitiveness was assessed with this methodology in Luque et al. (2017), and Ruiz et al. (2018) used them to build a ranking taking into account aspects related to the ease of doing business in terms of countries’ regulation.

To the best of our knowledge, these synthetic indexes have been just applied to assess teachers' efficiency in Luque et al. (2020). In that work, their efficiency was analysed considering different key variables describing the learning-teaching context from those considered here. Furthermore, no second stage post-analysis of the influence of the controllable variables was carried out there to identify potential correlations once the synthetic indexes were calculated. This post-analysis is now performed to have a deeper knowledge about the impact of these variables in the efficiency ranking of the teachers, i.e. to know the extent to which they could be used to improve the efficiency of a teacher. Therefore, the present work is intended to continue and complete the previous research in Luque et al. (2020). Besides, in the present paper, we also analyse several properties of the indexes used (see Section 4.3), which constitutes an additional contribution.

Finally, the consideration of the teaching time dedicated to the different subjects (i.e. teaching minutes in reading, mathematics and science) as inputs in our study deserves a discussion. Although the effect of instructional time is widely recognized as relevant to academic performance (e.g., Locher and Pfost, 2020), the effect on students' performance is complex and the relationship between instruction time and academic achievement is not as obvious as it might seem, primarily because of the quality of time (Mullis and Martin, 2017). The literature has not a clear consensus in this topic. On the one hand, some authors support that there is not a direct influence of the instruction time on the students' academic achievement (e.g. Woessmann, 2010; Marcenaro et al., 2016; Lopez-Agudo and Marcenaro-Gutierrez, 2018; Luque et al., 2020). Recently, Lopez-Agudo and Marcenaro-Gutierrez (2022) concluded that, although the quantity of instruction time varies by country, spending more time in teaching does not seem to be related to students' academic achievement, even when this increase in instruction time is accompanied by a higher engagement of students during lessons.

However, on the other hand, there exist works supporting a positive relationship between instruction time and students' academic achievement (e.g. Lavy, 2015; Rivkin and Schiman, 2015; Cattaneo et al., 2017; Meroni and Abbiati, 2016; Andersen et al., 2016). In fact, some of these studies identified differential influences of instruction time on the academic performance, such as e.g. Huebener et al. (2017), who found that, in

Germany, increasing instruction time by two hours had a more positive effect for high-achiever students than for lower-achievers. Besides, Sims (2008) determined that, at 4th grade, more instruction time was positive in mathematics in low-achieving districts of the United States. Moreover, other authors (e.g. OECD, 2011; Baker et al., 2004, Gromada and Shewbridge, 2016) also corroborated this positive relationship, but finally concluded that the quality of the instruction time is crucial and may condition the level of dependence between the time spent in the classroom and the academic achievement.

In our study, we have considered the teaching times devoted to reading, mathematics and science as inputs to reach the lowest values in order to see the impact of these variables on the efficiency of a teacher. Obviously, the purpose is not to find a teacher dedicating no instruction time to teach these subjects, which would represent an unrealistic –and unwanted– situation. Our main aim in doing so is to know to what extent the teaching times devoted to the three subjects contribute to and influence the efficiency of teachers (remember that our concept of efficiency is a wide one based on the minimization of the inputs and maximization of the outputs). Besides, we have decided to assume that the lower, the better for these teaching times (i.e., to minimize these variables), instead of the other way around, following the literature in the field supporting the idea that instruction time does not seem to influence students' academic achievement. Therefore, our study is performed without assuming that an inherent relationship exists between these two characteristics.

Finally, we are aware that teachers themselves do not completely control the instruction times of their subjects. In this regard, our study should be understood as follows: the results obtained can shed some light towards the profile of a potential efficient teacher (its teaching-learning context, regarding the inputs and outputs we have considered). In this way, the educational authorities can have a better knowledge to define policies directed at promoting such teaching-learning contexts that allow teachers to be efficient, as a way to improve the academic performance of the students. In particular, regarding the teaching times, our study can help to know if teaching for more or less time in reading, mathematics and science really conditions these efficiency in some way, based on the profile of the most and the least efficient teachers.

## 4. Methodology

The main purpose when studying a multi-criteria decision making problem is the analysis of the possible alternatives available considering the different conflicting criteria at the same time, in order to select the best one by classifying or ranking them according to some approach. In our case, the set of available alternatives is the sample of teachers under study, while the conflicting criteria are precisely the significant variables (inputs and outputs) shown in Table 1. Thus, we want to evaluate each teacher according to these criteria and to order them according to their efficiency level. Nonetheless, considering information about how far is each teacher from satisfying potential desirable values for the variables would greatly enrich the efficiency score that measures the goodness of each teacher. This would allow us to introduce external preferences into the ranking process, enabling the teachers to be analysed according to the desired expectations for the observed data.

To achieve this and to generate an efficiency ranking of the teachers that incorporates desirable variable values into the process, we build a set of synthetic indexes based on the DRP approach, which is described hereafter. In these synthetic indexes, the variable values achieved by the teachers are combined with some preferential information given in relation to the variables. To be more precise, the preferences are expressed by means of two reference levels for each variable: a reservation level (a value considered as acceptable), and an aspiration level (a value regarded as desirable). Furthermore, these synthetic indexes allow different compensation degrees among the values achieved in the variables. This means that bad performances in some variables can be compensated by good performances in others according to different scenarios, if desired, giving us the opportunity to obtain the efficiency ranking of teachers considering the desired balance among the variables. Additionally, a non-compensatory scheme can be also analysed using these synthetic indexes, driving the attention towards the worst variable value achieved by each teacher. This type of information can be useful if we want to develop improvement policies designed to enhance the achievement of the teaching-learning variable with the worst performance observed among the teachers in our sample.

### *4.1. Synthetic Indexes based on the DRP approach*

The methodology used to build the synthetic indexes is described in detail in Ruiz et al. (2011). As already mentioned, these indexes are based on the double reference point (DRP) preferential scheme introduced in Wierzbicki (2000) for multi-criteria decision analysis. If multiple conflicting criteria are used (denoted by $f_i$, $i = 1, \dots, m$ ($m \geq 2$)) to evaluate each possible known alternative, under the DRP scheme, the DM is asked to give, for each $i = 1, \dots, m$, a *reservation level* $q_i^r$ (a level below which a value for the criterion $f_i$ is considered unacceptable), and an *aspiration level* $q_i^a$ (a desirable value to be achieved by the criterion $f_i$). Then, a linear piecewise achievement scalarizing function is formulated to normalize the criteria values of each alternative considering the desirable ranges defined by the aspiration and reservation levels. Wierzbicki (2000) already suggested the use of this function as a way to rank different alternatives according to the achievement of the criteria values with respect to the aspiration and reservation levels, which is the basis of the synthetic indexes proposed in Ruiz et al. (2011).

Next, we describe the procedure followed to calculate these synthetic indexes for the study of the teachers' efficiency. As indicated in Section 2, 174 teachers constitute our sample (whose set is denoted by $S = \{t_1, \dots, t_{174}\}$), and we use the 9 input and output variables described in Table 1, which are related to their students' performance and engagement level in reading, mathematics and science, and the available time for each subject. Let us denote by $N_T$ the number of teachers $t$ in $S$ and by $N_V$ the number of variables used to evaluate them, i.e. $N_T$ = 174 and $N_V = 9$. For each variable considered, it is necessary to determine whether it is of the "more is better" type or if it belongs to the "less is better" type. In our case, the "more is better" variables are the mean scores and the engagement indexes for each subject, while the "less is better" variables are the teaching minutes dedicated to each subject. For simplicity, from now on, we refer to all the variables as $v_j$ ($j = 1, \dots, N_V$), considering that $v_1, \dots, v_6$ correspond to the outputs and $v_7, v_8, v_9$ denote the inputs.

Now, for each $i = 1, \dots, N_T$ and $j = 1, \dots, N_V$, let $x_{ij}$ be the sample value of the teacher $i$ for the variable $v_j$. Let us also denote by $q_j^M$ and $q_j^m$, respectively, the maximum and minimum values taken by the variable $v_j$ in the sample, that is:

$$q_j^m = x_{ij}\,, \quad \forall j = 1, \dots, N_V, \tag{3}$$

$$q_j^M = x_{ij}\,, \quad \forall j = 1, \dots, N_V. \tag{4}$$

To build the synthetic indexes based on the DRP approach, we follow the next steps:

1. *Preferential information*: let $q_j^a$ and $q_j^r$ be an aspiration and a reservation level, respectively, for each variable $v_j$ $(j = 1, \dots, N_V)$. In Section 4.2, we detail how these preferential levels have been set, since they play an important role for the indexes' building process and for the understanding of the results.

2. *Normalization*: for each $i = 1, \dots, N_T, j = 1, \dots, N_V$, the value $x_{ij}$ achieved by the teacher $i$ for the variable $v_j$ needs to be normalized to have all variable values in the same scale. Let us refer as $y_{ij}$ to the normalized value of $x_{ij}$. To construct the DRP synthetic indexes, $x_{ij}$ is normalized using the reservation and aspiration levels considered for $v_j$ as explained below. This normalization enables us to easily know the "position" of $x_{ij}$ regarding the corresponding reservation-aspiration range.

To normalize the variable values, we need to distinguish between the variables of the "more is better" type and those of the "less is better" type, since a different procedure is applied in each case. On the one hand, for the "more is better" type variables, we have $q_j^m \leq q_j^r \leq q_j^a \leq q_j^M$ and the normalized variable value $y_{ij}$ of $x_{ij}$ is calculated as follows:

- If $q_j^m \leq x_{ij} \leq q_j^r$, then $y_{ij} = \frac{x_{ij} - q_j^r}{q_j^r - q_j^m}$.
- If $q_j^r \leq x_{ij} \leq q_j^a$, then $y_{ij} = \frac{x_{ij} - q_j^r}{q_j^a - q_j^r}$. (5)
- If $q_j^a \leq x_{ij} \leq q_j^M$, then $y_{ij} = 1 + \frac{x_{ij} - q_j^a}{q_j^M - q_j^a}$.

On the other hand, for the "less is better" type variables, we have $q_j^m \leq q_j^a \leq q_j^r \leq q_j^M$ and the normalized variable value $y_{ij}$ is obtained by:

- If $q_j^m \leq x_{ij} \leq q_j^a$, then $y_{ij} = 1 + \frac{x_{ij} - q_j^a}{q_j^m - q_j^a}$.
- If $q_j^a \leq x_{ij} \leq q_j^r$, then $y_{ij} = \frac{x_{ij} - q_j^r}{q_j^a - q_j^r}$. (6)
- If $q_j^r \leq x_{ij} \leq q_j^M$, then $y_{ij} = \frac{x_{ij} - q_j^r}{q_j^r - q_j^M}$.

This piecewise normalization can be interpreted in the following way for any type of variable (“more is better” or “less is better”): if the normalized value $y_{ij}$ takes a negative value (between $-1$ and 0), then the original value $x_{ij}$ of the variable $v_j$ achieved by the teacher $i$ is worse than the reservation level $q_j^r$; a value of $y_{ij}$ between 0 and 1 means that $x_{ij}$ performs better than the reservation level $q_j^r$, but worse than its aspiration level $q_j^a$; and if $y_{ij}$ is between 1 and 2, then $x_{ij}$ reaches a better value than the aspiration level $q_j^a$. Note that, as a consequence, for all the normalized variable values $y_{ij}$, “more is better”, since reaching a (original) variable value better than the aspiration level would be much more preferred to achieving a value under the reservation level.

3. *Calculation of the synthetic indexes*: once the variable values are normalized, we build the next three indexes for each teacher $i$ ($i = 1, \dots, N_T$):

- A *weak index*, denoted by $WI_i$, which is the arithmetic mean of the normalized variable values:

$$WI_i = \frac{1}{N_V} \sum_{j=1}^{N_V} y_{ij}. \tag{7}$$

Note that $WI_i$ also takes values between −1 and 2, and therefore, it can be understood as the “position” of the teacher with respect to hypothetical global reservation and aspiration levels.

- A *strong index*, denoted by $SI_i$, measuring the worst performance of the teacher $i$, since it is defined as the minimum of all the normalized variable values:

$$SI_i = y_{ij} \, . \tag{8}$$

A negative value of $SI_i$ means that, at least, there is one variable value of this teacher $i$ that performs worse than its corresponding reservation level (i.e. at least, there exists one normalized variable value below 0). Besides, a value of $SI_i$ over 1 means that all variables of teacher $i$ have reached better values than their respective aspiration levels (i.e. all normalized variable values are above 1).

It is important to remark that, while the weak index allows full compensation among all the variables (substitutability, i.e. a poor performance in one variable can be compensated by a good performance in another one), the strong index does not enable any compensation among them since it represents the worst value achieved by the normalized variable values. Observe that $SI_i \leq WI_i, \forall i = 1, \dots, N_T$.

- Given that the weak and strong indexes represent extreme situations, enabling either full compensation or no compensation, a combination of them both can be used to enable a partial compensation among the values achieved by the variables. Thus, a *mixed index*, denote by $MI_i$, is calculated by linearly combining both of them:

$$MI_i = \mu\, WI_i \; + (1 - \mu)\, SI_i, \tag{9}$$

with $0 \leq \mu \leq 1$ being a real value. Note that $MI_i = SI_i$ if $\mu = 0$ (no compensation), while $MI_i = WI_i$ in case $\mu = 1$ (full compensation). Any value of $\mu$ between 0 and 1 represents an intermediate state between these two extremes, considering that the larger the value of $\mu$, the more compensation is allowed. Note that $SI_i \leq MI_i \leq WI_i, \forall i = 1, \dots, N_T$.

As described, the synthetic indexes have been calculated without using any weights in the process. Traditionally, weights have been used in the reference point schemes with a normalization role, but it is also possible to use weights with a preferential meaning (Luque et al., 2009). Based on this, $WI_i$, $SI_i$ and $MI_i$ can incorporate a weight factor for each variable, if desired. In case different weights want to be assigned to the variables, the building process would be adapted as follows. For each $j = 1, \dots, N_V$, let us denote by $\omega_j$ the weight value given to the variable $v_j$ ($\omega_j > 0$), and let us consider their normalized values as $\overline{\omega}_j = \frac{\omega_j}{\sum_{k=1}^{N_V} \omega_k}$ (observe that the normalized weights add up to 1). The *weighted weak index* $W - WI_i$ of teacher $i$ ($i = 1, \dots, N_T$) would be calculated as the weighted sum:

$$W - WI_i = \sum_{j=1}^{N_V} \overline{\omega}_j y_{ij}. \tag{10}$$

For building the strong index and avoiding unwanted effects due to the use of weights and to the fact that $y_{ij}$ can take positive and negative values, additional mathematical transformations are required. Let us consider the following modified weights:

$$\widehat{\omega}_j = \frac{\overline{\omega}_j}{\max\limits_{k=1,\dots,N_V} \overline{\omega}_j}, \quad \forall j = 1, \dots, N_V. \tag{11}$$

For each teacher $i$ and each variable $v_j$, we transform the values $y_{ij}$ in the following way:

$$\bar{y}_{ij} = y_{ij} - m_i, \quad \text{where} \quad m_i = \left[y_{ij}\right] + 1. \tag{12}$$

The $[\,\cdot\,]$ operator gives the integer part of any real number. Now, the *weighted strong index* $W - SI_i$ of teacher $i$ would be redefined as:

$$W - SI_i = m_i + \widehat{\omega}_j \bar{y}_{ij}\,. \tag{13}$$

Obviously, the mixed index is updated accordingly using the weighted weak and strong indexes. That is, the *weighted mixed index* $W - MI_i$ of teacher $i$ would be:

$$W - MI_i = \mu\, W - WI_i \; + (1-\mu)\, W - SI_i. \tag{14}$$

### *4.2. Aspiration and Reservation Levels Used*

For the normalization of the variable values carried out in (5) and (6) and for the building process, key parameters are the aspiration and reservation levels. Let us recall that they represent, respectively, a desirable level to be achieved by each variable and a level below which the variable values are regarded as unwanted. These preferential levels can be set in several ways. For example, they may be defined using absolute values, but such universally accepted values do not exist in the literature for the input and output variables under study in this paper. Another option may be to set appropriate values for them with the help of a panel of experts, but some inconsistencies might arise in practice, since these values may be set in a relative manner, i.e. considering the situation for each variable of one or several teachers with respect to the others.

Here, for each variable $v_j$, the aspiration and reservation levels $q_j^a$ and $q_j^r$ have been set based on the statistical descriptors given in Table 1, in order to set them using realistic values obtained according to the observational data considered. To be more precise, the values for $q_j^a$ and $q_j^r$ have been calculated as follows, based on the mean and the standard deviation attained by the variable $v_j$, denoted as $m_j$ and $sd_j$, respectively:

- For each of the “more is better” type variables, we have considered:

$$q_j^a = m_j + sd_j, \tag{15}$$
$$q_j^r = m_j - sd_j.$$

- For the “less is better” type variables, they have been set them as:

$$q_j^a = m_j - sd_j, \tag{16}$$
$$q_j^r = m_j + sd_j.$$

It should be mentioned that an outlier detection, using the interquartile range method, was carried out, and no outliers were detected. The aspiration and reservation levels calculated following equations (15) and (16) are provided in Table 2.

**Table 2** Aspiration and reservation levels used for the variables.

| **Variables** | **“More is better” or “Less is better” type** | **Aspiration level** | **Reservation level** | **Maximum value** | **Minimum value** |
|---|---|---|---|---|---|
| ***Outputs*** | | | | | |
| Mean scores in | | | | | |
| Reading | “More” | 553.42 | 486.33 | 592.29 | 392.96 |
| Mathematics | “More” | 525.74 | 452.70 | 581.62 | 347.87 |
| Science | “More” | 548.81 | 476.81 | 604.63 | 376.26 |
| Engagement in: | | | | | |
| Reading | “More” | 10.80 | 9.07 | 12.16 | 6.48 |
| Mathematics | “More” | 10.96 | 9.26 | 12.64 | 5.81 |
| Science | “More” | 10.95 | 9.21 | 12.05 | 6.10 |
| ***Inputs*** | | | | | |
| Teaching minu in: | | | | | |
| Reading | “Less” | 233.02 | 441.45 | 780.00 | 120.00 |
| Mathematics | “Less” | 235.22 | 320.81 | 420.00 | 180.00 |
| Science | “Less” | 186.43 | 289.49 | 540.00 | 60.00 |

Source: Author’s own elaboration.

### *4.3. Further Comments*

The weak, strong and mixed indexes satisfy the following properties:

a) *Existence and uniqueness.* The way of building three indexes assures that any of them can be always calculated for any teacher, and the index value is unique.

b) *Monotony*. The weak index is strictly monotonous since it is a linear combination of the normalized values using positive weights (the normalization is strictly increasing): if any normalized variable value is improved (resp., worsened) and the rest of them remain constant, the weak index reaches a higher value (resp., a lower value) than its former value. In addition, the strong index is also monotonous because, if any normalized value is improved (resp., worsened) and the rest remains constant, the strong index equals or reaches a higher value (resp., a lower value) than its former value. As consequence, the mixed index is also strictly monotonous if $\mu > 0$ in equation (9) since it is a linear combination of the other two. In case of using $\mu = 0$ in (9), the mixed index would coincide with the strong index and, thus, it would be only monotonous.

c) *Identity*. The three indexes hold this property both for the sum operation (if we add 0 -the neutral element of the sum- to all the normalized values, the three indexes remain invariant) and for the product operation (if we multiply all normalized value by 1-the neutral element of the product-, the three indexes remain invariant).

d) *Transitivity*. This property is verified by the three indexes: for any of them, if teacher A has a greater index value than teacher B, and in turn, teacher B achieves a greater index value than teacher C, then A has also a greater index value than C.

e) *Proportionality* (homogeneity of grade 1). It is trivial that, for the three indexes, if all the normalized values are multiplied by a positive constant $k \in R$, the new index value is its former value multiplied by $k$. To be more precise, this property is held by the weak index because it is defined as a linear combination of the normalized values. Furthermore, for the strong index, which is calculated as the maximum of the normalized values, it is evident that the maximum value would be the previous maximum value multiplied by $k$ if all the normalized values are multiplied by $k$ (note that we assume that $k$ has a positive value). And consequently, this property is also held for the mixed index, given that it is a linear combination of the weak and strong indexes.

In the literature, there are many methodologies to build synthetic indexes (see Nardo et al. (2005)). As stated there, the quality of a composite indicator depends not only on the building methodology used but also on the quality of the data used. We do not question the benefits of the different procedures available in the literature, but we have selected the DRP approach for building our synthetic indexes because it enables us to assess the teachers' efficiency using desirable scenarios for the variables considered, i.e. we can introduce preferential information into the efficiency scores. Thus, the resulting indexes can provide us information about how to distribute the available resources to improve key factors of the teaching-learning context that are desirable to let teachers achieve efficiency, according to the specified preferences.

Obviously, other methods for building the indexes would offer other advantages and would lead to different conclusions. For example, so-called "benefit of the doubt" (BOD) approaches (Cherchye et al., 2007) are DEA-based methodologies estimating weights for the indexes' calculation which do consider the relative contribution of each variable or dimension under study to the index calculation. Specifically, based on the observational data, weights are assigned to each unit in a way to evaluate its composite measure, subject to a set of specified constraints. It has the advantage that higher weights are assigned to the dimensions on which the unit performs well, and lower weights to those on which it performs bad. However, the weights obtained are unit specific (i.e. weights are not uniquely determined for all the units) and it is likely that many units are classified as efficient (Nardo et al., 2005).

In Despotis (2005), a "fair" assessment of the Human Development Index (HDI), a composite index of socio-economic indicators, was proposed also based on the DEA methodology. Its main purpose is the calculation of common weights for the socio-economic sub-indicators composing the HDI according to the observational data, in a manner that the resulting efficiency (global) scores are as close as possible to the ideal ones. For the calculation of such common weights, an optimization problem (named as problem (2) in Despotis (2005)) is solved based on a linear combination of the $L_1$ and the $L_\infty$ metrics. Somehow, this linear combination may be seen similar to the one used to calculate our mixed synthetic index. Nevertheless, we combine the maximum (strong index) and the mean (weak index) of the normalized variable values, while in (Despotis, 2005), problem (2) considers a linearly combination of the maximum and the mean deviations between DEA scores obtained in model (1) of (Despotis, 2005) and the adjusted global efficiency scores.

Overall, observe that the main contribution of (Despotis, 2005) is the calculation of common weights for the socio-economic sub-indicators composing the HDI, while our main contribution is not relative to the weighting scheme, but to the incorporation of preferential information into the efficiency evaluation of teachers. To this aim, as explained, we perform a normalization of the input and output variable values based on aspiration and reservation levels used for the variables, which constitutes the main difference with the index building scheme followed in (Despotis, 2005). As aforementioned, these aspiration and reservation levels are set according to the observational data in our sample. This allows us to know if the teachers are performing better, within or worse than the desired limits, and gives us the opportunity to assess the teachers' efficiency using desirable scenarios.

## 5. Results and Discussion

In this section, the results obtained for the Spanish teachers according to the synthetic indexes described in Section 4 are discussed in Section 5.1, and the dependence of the efficiency ranking with respect to the inputs considered is analysed in Section 5.2.

### *5.1. Calculation of the Weak, Strong and Mixed Indexes*

In our analysis, all the inputs and outputs have received the same importance, meaning that the same weights have been assigned to all the variables in the index's calculation. Therefore, we have obtained the weak ($WI_i$), strong ($SI_i$) and mixed ($MI_i$) indexes following equations (7), (8) and (9), respectively. To study the teachers' efficiency using a balanced compensation degree between the full compensation represented by the weak index and the zero-compensation implied by the strong index, we have considered $\mu = 0.5$ to calculate the mixed index.

The three indexes values attained by the Spanish teachers according to our sample are given in Tables 3 and 4, which show the teachers in the first (Q1) and the fourth quartile (Q4), respectively, when they are ranked in descending order according to the mixed synthetic index. Due to the space limit, only teachers in these two quartiles are provided, although the values of the synthetic indexes for the rest of teachers are available upon request. Note that we have decided to study the teachers ranked according to the mixed index since it is a combination of the weak and strong indexes.

The information provided in Tables 3 and 4 allows us to have a broad understanding of the most efficient teachers and the least ones (according to the mixed index) and to identify key variables. The ranking position of each teacher regarding each index is indicated in the three first columns (the mixed index's position is in boldface to highlight the order considered to rank the teachers). In addition to the values reached by the three indexes and the normalized variable values –calculated following expressions (5) and (6)–, their means and their standard deviations are also reported in the first two rows. Actually, normalized variable values above 1 are colored in green, while those between 0 and 1 are in yellow, and the ones lower than 0 are in red. In this way, for each variable, we can easily visualise at a glance teachers performing better than the aspiration level (in green), between the reservation and aspiration levels (in yellow), or worse than the reservation level (in red). Finally, the worst variable value of each teacher (which corresponds to the strong index) is highlighted with a bold red face boxed.

Next, let us revise the results in detail. The green (values above the aspiration levels) and the yellow (values among the aspiration and the reservation levels) colours of the teachers at the highest positions (see Table 3) inform us that the most efficient teachers attain balanced levels for all the variables, mostly above or close to the most desirable situation. This is a reasonable performance of the teachers in the higher rank positions. On average, these teachers achieve values for the mixed index around $0.603$, around $0.383$ for the strong index, and around $0.823$ for the weak index (note that $0.383 \leq 0.603 \leq 0.823$). The positive value attained by the mean of the strong index indicates that the worst variable performance of the most efficient teachers is still within the desirable limits, on average.

Observe that, in general terms, the teachers in the highest positions got values above the aspiration levels (in green) mainly for the scores and the engagement indexes, except for some cases (such as the teacher in the first position, which is further described hereafter). This fact states the key role of these variables for the efficiency of the teachers, as expected. However, to reach the most efficient positions, the minutes spent on teaching in class are also above or within the desirable limits, thus these variables result also determinant in the ranking. Actually, the means of all the variables are above $0.5$ (the mid value of the desirable interval), most of them even above $0.8$. This implies that, on average, teachers with acceptable values, near the aspiration levels, occupy the highest positions of the ranking for the means scores, the engagement and the teaching minutes for the three subjects. However, note that the best-averaged values correspond to the scores in mathematics and science, and the worst ones to the teaching minutes in science.

Regarding the profile of the most efficient teacher in Table 3, it is noteworthy that the teaching minutes invested in reading and maths have been crucial to be ranked in the first position, given that these two variables get the highest possible values (which is $2$). Besides, the outstanding performance of this teacher has been also possible because the other variables are within their desirable intervals, but very close to 1, meaning that they are very close to their corresponding aspiration levels. Thus, the combination of all these factors becomes crucial for reaching the most efficient position. Particularly interesting is the fact that this teacher also gets the first position for the strong index, and it is the second most efficient one regarding the weak index. This implies that, in addition, (s)he is the most efficient if we seek for the worst variable performance (no compensation), and the second most efficient when a full compensation among the variables is used. Finally, note that the second most efficient teacher in Table 3 obtains this position due to the fact that, (s)he gets values above the aspiration level for the variables for which the most efficient teacher performs within the limits but not above the aspiration levels. This reveals once more that having a balanced performance among all the inputs and outputs is decisive, and therefore actions aimed at promoting all of them as a whole, would be beneficial when searching for efficiency in teachers.

It is very important to remind that a synthetic index combines individual indicators, in order to obtain a global aggregated measure that cannot be quantified by each of the indicators considered. However, the aggregation into a single synthetic index always implies a certain loss of information along the way, and the type of compensation used is crucial to understand the results. A total compensation of the individual indicators (as in the weak index) implies that poor results of some indicators can be compensated with good ones of some others. A zero-compensation scheme can be applied to reflect the worst performance (as in the strong index). However, we consider the mixed index to order the teachers and a partial compensation between the total and zero schemes is used. That is why being the best in the three inputs (teaching times) makes teacher 1 in Table 3 to be the most efficient (actually, this teacher gets the best values in all the sample for the three inputs), although teacher 2 is the best one in all the outputs (scores and engagements). However, the output values of teacher 2 are close to their aspiration levels, but not as close as the input values of teacher 1 are from their corresponding aspiration values. This explains the ranking obtained, based on the partial compensation used.

Regarding the teachers in Table 4 (these teachers in Q4 according to the mixed index), we can see that the number of cells coloured in red (corresponding to negative normalized variable values) is quite significant, while the number of these in green colour decreases, as it may be expected. This indicates a poor performance, under their reservation levels, in a higher number of variables. Indeed, the averaged values achieved by the strong and the mixed indexes lead us to the same conclusion, since they have negative mean values ($-0.514$ and $-0.139$, respectively). The full compensation internally performed in the mixed index is the reason explaining the positive mean value reached by this indicator (0.235). In relation to the variables, despite reaching averaged values above 0 (i.e. between the desirable ranges defined by the aspiration and reservation levels), their mean values tend to be below 0.5 –the mid value of the desirable interval. However, the teaching minutes in science attains the highest mean value among the least efficient teachers, while the scores in maths and science are the variables with the lowest average performances. Remember that exactly the same variables reported, respectively, the worst and the best mean values among the most efficient teachers. With this, a conflicting relation among the variables is shown, concluding that teaching for more (resp. less) time in science does not necessarily imply an increase (resp. a decrease) of the scores in maths and science, making even the teacher to be less (resp. more) efficient than in the other way around, as shown by the results obtained.

The least efficient teacher, ranked in the last position in Table 4, attains the worst values for the scores in the three subjects (since their normalized values are –1), thus making its strong index to reach the lowest possible value (–1). The same situation is observed for the teacher in the previous position regarding the three engagement levels, but the performance of this teacher is a bit better for the rest of variables.

To have a more precise view of the results, we indicate the descriptive statistics of the variables at the top and bottom deciles and quartiles of the teacher's distribution ranked according to their mixed index in Tables 5 and 6, using their original values and their normalized values, respectively. The mean scores and engagement indexes reached by the students of those teachers ranked within the top decile/quartile –in the three subjects–, are clearly above the figures found for those teachers at the bottom of the efficiency distribution, which is logical since these variables are outputs (of the "more is better" type). A similar situation is detected regarding the teaching times for the three subjects: the teachers at the bottom of the ranking spend significantly higher times in the lessons than those at the most efficient positions (as these variables are inputs, "less is better" if we revise their original values shown in Table 5). Besides, it is interesting to see that the minutes' difference is significantly high for reading but, at the same time, quite low for the science competence. All of this is in line with the findings observed when analysing Tables 3 and 4. This conclusion must have important implications in terms of educational policies, given that revising and improving the reading skills for teaching can lead to significant advances regarding the learning achievement of students.

**Table 3** Weak ($WI_i$), strong ($SI_i$) and mixed ($WI_i$) indexes and normalized variable values for the Spanish teachers in the first quartile (Q1).

| Ranking | | | Index | | | Mean scores | | | Engagement | | | Teaching Minutes | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $WI_i$ | $SI_i$ | $MI_i$ | $WI_i$ | $SI_i$ | $MI_i$ | Reading | Math | Science | Reading | Math | Science | Reading | Math | Science |
| **Mean** | | | **0,823** | **0,383** | **0,603** | **0,850** | **0,896** | **0,882** | **0,856** | **0,831** | **0,866** | **0,775** | **0,816** | **0,638** |
| **Stand. Desv.** | | | **0,149** | **0,137** | **0,112** | **0,346** | **0,402** | **0,391** | **0,276** | **0,291** | **0,332** | **0,336** | **0,429** | **0,273** |
| 2 | 1 | **1** | 1,193 | 0,804 | 0,998 | 0,919 | 0,978 | 0,858 | 0,804 | 0,915 | 1,210 | 2,000 | 2,000 | 1,051 |
| 1 | 14 | **2** | 1,239 | 0,477 | 0,858 | 1,633 | 1,813 | 1,729 | 1,361 | 1,201 | 0,969 | 0,679 | 0,477 | 1,288 |
| 15 | 3 | **3** | 0,903 | 0,594 | 0,749 | 0,886 | 0,995 | 1,239 | 1,059 | 1,014 | 1,110 | 0,607 | 0,594 | 0,626 |
| 4 | 5 | **4** | 1,004 | 0,480 | 0,742 | 0,881 | 1,140 | 0,978 | 0,953 | 1,313 | 1,380 | 0,966 | 0,944 | 0,480 |
| 24 | 2 | **5** | 0,842 | 0,626 | 0,734 | 0,939 | 0,982 | 0,718 | 0,689 | 0,869 | 0,816 | 0,751 | 1,185 | 0,626 |
| 13 | 4 | **6** | 0,908 | 0,542 | 0,725 | 0,713 | 0,550 | 0,542 | 0,979 | 0,882 | 0,711 | 1,292 | 1,638 | 0,868 |
| 7 | 5 | **7** | 0,942 | 0,480 | 0,711 | 0,898 | 1,091 | 1,417 | 0,833 | 1,169 | 1,110 | 0,535 | 0,944 | 0,480 |
| 9 | 5 | **8** | 0,932 | 0,480 | 0,706 | 0,954 | 1,105 | 0,725 | 1,214 | 1,120 | 0,883 | 0,966 | 0,944 | 0,480 |
| 11 | 5 | **9** | 0,917 | 0,480 | 0,698 | 0,750 | 1,186 | 1,247 | 1,334 | 0,865 | 0,910 | 0,535 | 0,944 | 0,480 |
| 19 | 5 | **10** | 0,885 | 0,480 | 0,682 | 0,711 | 0,563 | 0,608 | 0,883 | 1,323 | 1,482 | 0,966 | 0,944 | 0,480 |
| 20 | 5 | **11** | 0,879 | 0,480 | 0,679 | 1,100 | 0,876 | 1,076 | 0,786 | 0,896 | 1,071 | 0,679 | 0,944 | 0,480 |
| 18 | 17 | **12** | 0,890 | 0,440 | 0,665 | 1,066 | 1,126 | 1,073 | 1,207 | 1,139 | 0,440 | 0,535 | 0,944 | 0,480 |
| 23 | 5 | **13** | 0,843 | 0,480 | 0,661 | 1,022 | 0,755 | 0,946 | 1,203 | 0,664 | 0,724 | 0,607 | 1,185 | 0,480 |
| 12 | 21 | **14** | 0,909 | 0,413 | 0,661 | 0,611 | 0,429 | 0,413 | 1,139 | 0,906 | 0,883 | 1,292 | 1,638 | 0,868 |
| 3 | 33 | **15** | 1,046 | 0,243 | 0,645 | 1,614 | 1,455 | 1,460 | 1,546 | 1,028 | 0,913 | 0,679 | 0,243 | 0,480 |
| 27 | 5 | **16** | 0,795 | 0,480 | 0,638 | 0,884 | 0,964 | 1,048 | 0,810 | 0,598 | 0,750 | 0,679 | 0,944 | 0,480 |
| 25 | 16 | **17** | 0,805 | 0,469 | 0,637 | 0,760 | 1,188 | 0,974 | 0,882 | 0,869 | 0,469 | 0,679 | 0,944 | 0,480 |
| 30 | 5 | **18** | 0,776 | 0,480 | 0,628 | 0,710 | 0,549 | 0,671 | 0,844 | 0,962 | 0,855 | 0,966 | 0,944 | 0,480 |
| 32 | 15 | **19** | 0,764 | 0,475 | 0,619 | 0,684 | 0,662 | 0,611 | 0,720 | 0,657 | 0,475 | 1,115 | 1,276 | 0,674 |
| 21 | 28 | **20** | 0,875 | 0,344 | 0,610 | 0,574 | 0,344 | 0,453 | 0,918 | 1,132 | 1,136 | 1,204 | 1,638 | 0,480 |
| 5 | 61 | **21** | 0,966 | 0,183 | 0,575 | 0,793 | 0,964 | 0,183 | 1,115 | 1,526 | 2,000 | 0,823 | 0,243 | 1,051 |
| 14 | 33 | **22** | 0,905 | 0,243 | 0,574 | 1,067 | 1,294 | 1,120 | 0,908 | 0,739 | 1,047 | 0,679 | 0,243 | 1,051 |

Source: Author's own elaboration.

**Table 3** (Continued).

| Ranking | | | Index | | | Mean scores | | | Engagement | | | Teaching Minutes | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $WI_i$ | $SI_i$ | $MI_i$ | $WI_i$ | $SI_i$ | $MI_i$ | Reading | Math | Science | Reading | Math | Science | Reading | Math | Science |
| 16 | 33 | **23** | 0,902 | 0,243 | 0,573 | 1,356 | 1,208 | 1,295 | 0,920 | 0,792 | 1,144 | 0,679 | 0,243 | 0,480 |
| 17 | 33 | **24** | 0,891 | 0,243 | 0,567 | 1,385 | 1,715 | 1,708 | 0,481 | 0,822 | 0,438 | 0,751 | 0,243 | 0,480 |
| 8 | 57 | **25** | 0,937 | 0,189 | 0,563 | 1,093 | 1,430 | 1,522 | 0,723 | 0,937 | 0,914 | 0,679 | 0,944 | 0,189 |
| 10 | 57 | **26** | 0,924 | 0,189 | 0,557 | 1,224 | 1,404 | 1,454 | 1,117 | 1,193 | 1,100 | 0,391 | 0,243 | 0,189 |
| 37 | 23 | **27** | 0,719 | 0,391 | 0,555 | 0,812 | 0,764 | 0,796 | 0,766 | 0,424 | 0,527 | 0,391 | 0,944 | 1,051 |
| 54 | 18 | **28** | 0,660 | 0,440 | 0,550 | 0,750 | 0,482 | 0,597 | 0,440 | 0,528 | 0,751 | 0,966 | 0,944 | 0,480 |
| 38 | 27 | **29** | 0,717 | 0,381 | 0,549 | 0,381 | 0,456 | 0,501 | 0,779 | 0,524 | 0,484 | 1,336 | 0,944 | 1,051 |
| 6 | 64 | **30** | 0,952 | 0,141 | 0,546 | 1,377 | 1,658 | 1,404 | 1,282 | 0,975 | 0,975 | 0,223 | 0,535 | 0,141 |
| 57 | 19 | **31** | 0,649 | 0,439 | 0,544 | 0,466 | 0,473 | 0,468 | 0,751 | 0,989 | 1,198 | 0,439 | 0,477 | 0,577 |
| 46 | 26 | **32** | 0,686 | 0,385 | 0,535 | 0,564 | 0,714 | 0,851 | 1,018 | 0,535 | 0,385 | 0,679 | 0,944 | 0,480 |
| 55 | 25 | **33** | 0,651 | 0,387 | 0,519 | 0,387 | 0,487 | 0,507 | 0,723 | 0,748 | 0,619 | 0,966 | 0,944 | 0,480 |
| 66 | 22 | **34** | 0,617 | 0,396 | 0,506 | 0,662 | 0,543 | 0,458 | 0,396 | 0,584 | 0,514 | 0,679 | 0,944 | 0,771 |
| 65 | 23 | **35** | 0,618 | 0,391 | 0,504 | 0,609 | 0,707 | 0,606 | 0,514 | 0,590 | 0,722 | 0,391 | 0,944 | 0,480 |
| 86 | 20 | **36** | 0,555 | 0,435 | 0,495 | 0,482 | 0,501 | 0,562 | 0,764 | 0,639 | 0,435 | 0,535 | 0,594 | 0,480 |
| 53 | 30 | **37** | 0,661 | 0,316 | 0,489 | 0,463 | 0,316 | 0,411 | 0,582 | 0,496 | 0,613 | 1,071 | 0,944 | 1,051 |
| 36 | 33 | **38** | 0,728 | 0,243 | 0,485 | 0,866 | 0,867 | 0,920 | 0,485 | 0,654 | 0,785 | 0,679 | 0,243 | 1,051 |
| 58 | 29 | **39** | 0,644 | 0,317 | 0,481 | 0,427 | 0,354 | 0,383 | 0,794 | 0,317 | 0,506 | 1,204 | 0,944 | 0,868 |
| 43 | 33 | **40** | 0,693 | 0,243 | 0,468 | 1,469 | 1,172 | 1,293 | 0,291 | 0,377 | 0,517 | 0,391 | 0,243 | 0,480 |
| 34 | 60 | **41** | 0,752 | 0,184 | 0,468 | 0,479 | 1,463 | 0,768 | 0,865 | 0,184 | 0,921 | 0,751 | 0,418 | 0,917 |
| 45 | 33 | **42** | 0,686 | 0,243 | 0,464 | 0,402 | 0,720 | 0,842 | 0,609 | 0,986 | 0,856 | 0,463 | 0,243 | 1,051 |
| 44 | 53 | **43** | 0,688 | 0,239 | 0,464 | 0,239 | 0,271 | 0,367 | 0,484 | 1,015 | 1,426 | 0,966 | 0,944 | 0,480 |
| 47 | 33 | **44** | 0,681 | 0,243 | 0,462 | 1,358 | 0,720 | 0,994 | 0,691 | 0,466 | 0,929 | 0,247 | 0,243 | 0,480 |

Source: Author's own elaboration.

**Table 4** Weak ($WI_i$), strong ($SI_i$) and mixed ($WI_i$) indexes and normalized variable values for the Spanish teachers in the fourth quartile (Q4).

| **Ranking** | | | **Index** | | | **Mean scores** | | | **Engagement** | | | **Teaching Minutes** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\boldsymbol{WI_i}$ | $\boldsymbol{SI_i}$ | $\boldsymbol{MI_i}$ | $\boldsymbol{WI_i}$ | $\boldsymbol{SI_i}$ | $\boldsymbol{MI_i}$ | **Reading** | **Math** | **Science** | **Reading** | **Math** | **Science** | **Reading** | **Math** | **Science** |
| **Mean** | | | **0,235** | **-0,514** | **-0,139** | **0,207** | **0,181** | **0,181** | **0,207** | **0,219** | **0,226** | **0,264** | **0,232** | **0,395** |
| **Stand. Desv.** | | | **0,168** | **0,280** | **0,168** | **0,578** | **0,490** | **0,499** | **0,485** | **0,426** | **0,441** | **0,540** | **0,611** | **0,453** |
| 162 | 105 | **131** | 0,144 | -0,093 | 0,026 | 0,024 | 0,086 | 0,086 | 0,380 | 0,120 | 0,398 | 0,103 | -0,093 | 0,189 |
| 144 | 123 | **132** | 0,253 | -0,205 | 0,024 | -0,180 | -0,205 | -0,009 | 0,130 | 0,227 | 0,358 | 0,535 | 0,944 | 0,480 |
| 145 | 122 | **133** | 0,243 | -0,196 | 0,024 | -0,183 | -0,173 | -0,196 | 0,201 | 0,158 | -0,003 | 0,391 | 0,944 | 1,051 |
| 135 | 135 | **134** | 0,324 | -0,286 | 0,019 | -0,286 | 0,518 | 0,255 | 0,204 | 0,378 | 0,161 | 0,391 | 0,243 | 1,051 |
| 161 | 115 | **135** | 0,168 | -0,139 | 0,015 | -0,139 | -0,104 | -0,023 | -0,074 | 0,223 | 0,516 | 0,391 | 0,243 | 0,480 |
| 129 | 138 | **136** | 0,367 | -0,342 | 0,013 | -0,312 | -0,144 | -0,342 | 0,657 | 0,902 | 0,727 | 0,391 | 0,944 | 0,480 |
| 116 | 141 | **137** | 0,417 | -0,395 | 0,011 | 0,916 | 0,655 | 0,830 | 0,176 | 0,475 | 0,639 | -0,025 | -0,395 | 0,480 |
| 121 | 141 | **138** | 0,403 | -0,395 | 0,004 | 0,734 | -0,028 | 0,191 | 0,519 | 0,836 | 0,829 | -0,114 | -0,395 | 1,051 |
| 125 | 141 | **139** | 0,399 | -0,395 | 0,002 | 0,091 | 0,309 | 0,327 | 0,679 | 0,797 | 0,626 | 0,679 | -0,395 | 0,480 |
| 146 | 128 | **140** | 0,239 | -0,242 | -0,001 | -0,242 | 0,029 | 0,132 | 0,275 | -0,016 | 0,507 | 0,391 | 0,594 | 0,480 |
| 151 | 127 | **141** | 0,225 | -0,229 | -0,002 | 0,306 | -0,229 | 0,061 | 0,292 | 0,392 | 0,611 | 0,391 | 0,243 | -0,042 |
| 149 | 129 | **142** | 0,229 | -0,243 | -0,007 | 0,458 | -0,243 | -0,023 | 0,048 | -0,176 | 0,599 | 0,679 | 0,243 | 0,480 |
| 156 | 126 | **143** | 0,209 | -0,222 | -0,007 | -0,072 | -0,222 | -0,024 | 0,223 | 0,107 | -0,042 | 0,487 | 0,944 | 0,480 |
| 112 | 155 | **144** | 0,429 | -0,459 | -0,015 | 1,312 | 1,133 | 1,134 | -0,227 | -0,313 | -0,459 | 0,751 | -0,244 | 0,771 |
| 85 | 161 | **145** | 0,555 | -0,587 | -0,016 | -0,348 | -0,555 | -0,587 | 1,264 | 1,143 | 1,082 | 1,336 | 1,185 | 0,480 |
| 131 | 141 | **146** | 0,360 | -0,395 | -0,018 | 0,443 | 0,552 | 0,527 | 0,402 | 0,364 | 0,411 | -0,114 | -0,395 | 1,051 |
| 152 | 132 | **147** | 0,224 | -0,278 | -0,027 | 0,074 | 0,284 | -0,278 | 0,349 | -0,141 | 0,225 | 0,823 | -0,093 | 0,771 |
| 73 | 162 | **148** | 0,584 | -0,646 | -0,031 | 1,543 | 1,258 | 1,276 | 0,137 | 0,188 | 0,077 | -0,646 | 0,944 | 0,480 |
| 155 | 134 | **149** | 0,210 | -0,282 | -0,036 | -0,224 | -0,125 | -0,084 | 0,320 | -0,103 | -0,282 | 0,966 | 0,944 | 0,480 |
| 136 | 141 | **150** | 0,320 | -0,395 | -0,038 | -0,244 | 0,215 | 0,051 | 0,762 | 0,904 | 0,648 | -0,114 | -0,395 | 1,051 |
| 101 | 160 | **151** | 0,483 | -0,561 | -0,039 | 1,147 | 0,676 | 0,928 | -0,561 | -0,273 | 0,037 | 0,966 | 0,944 | 0,480 |
| 154 | 137 | **152** | 0,213 | -0,310 | -0,048 | 0,817 | 0,302 | 0,498 | -0,310 | -0,076 | -0,071 | -0,025 | 0,594 | 0,189 |

Source: Author's own elaboration.

**Table 4** (Continued).

| Ranking | | | Index | | | Mean scores | | | Engagement | | | Teaching Minutes | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $WI_i$ | $SI_i$ | $MI_i$ | $WI_i$ | $SI_i$ | $MI_i$ | Reading | Math | Science | Reading | Math | Science | Reading | Math | Science |
| 140 | 152 | **153** | 0,285 | -0,416 | -0,065 | 1,293 | 0,949 | 0,746 | -0,258 | -0,220 | -0,416 | 0,391 | -0,395 | 0,480 |
| 147 | 141 | **154** | 0,237 | -0,395 | -0,079 | 0,969 | 0,729 | 0,592 | -0,213 | -0,136 | -0,286 | 0,391 | -0,395 | 0,480 |
| 159 | 140 | **155** | 0,179 | -0,355 | -0,088 | -0,189 | 0,168 | -0,355 | -0,065 | 0,332 | 0,035 | 0,966 | 0,243 | 0,480 |
| 163 | 154 | **156** | 0,142 | -0,424 | -0,141 | -0,424 | 0,427 | 0,167 | -0,350 | -0,058 | -0,068 | 0,679 | 0,944 | -0,042 |
| 164 | 153 | **157** | 0,132 | -0,420 | -0,144 | -0,420 | -0,302 | -0,216 | 0,645 | -0,051 | 0,301 | 0,679 | 0,068 | 0,480 |
| 160 | 158 | **158** | 0,173 | -0,485 | -0,156 | 0,436 | 0,516 | 0,337 | -0,321 | -0,143 | -0,485 | 0,535 | -0,093 | 0,771 |
| 168 | 141 | **159** | 0,049 | -0,395 | -0,173 | -0,345 | -0,320 | -0,355 | 0,856 | 0,850 | 0,480 | -0,291 | -0,395 | -0,042 |
| 169 | 141 | **160** | 0,043 | -0,395 | -0,176 | 0,046 | -0,051 | -0,098 | -0,142 | 0,089 | -0,002 | -0,114 | -0,395 | 1,051 |
| 167 | 159 | **161** | 0,092 | -0,507 | -0,207 | -0,312 | -0,507 | -0,460 | 0,576 | 0,638 | 0,667 | -0,202 | 0,243 | 0,189 |
| 171 | 141 | **162** | -0,025 | -0,395 | -0,210 | -0,186 | -0,205 | -0,216 | 0,419 | 0,067 | -0,061 | 0,391 | -0,395 | -0,042 |
| 78 | 167 | **163** | 0,574 | -1,000 | -0,213 | 0,862 | 0,677 | 0,860 | 1,180 | 0,948 | 0,720 | 0,679 | 0,243 | -1,000 |
| 138 | 165 | **164** | 0,315 | -0,816 | -0,250 | -0,077 | 0,623 | 0,422 | -0,816 | 0,276 | 0,122 | 0,823 | 1,276 | 0,189 |
| 142 | 166 | **165** | 0,266 | -0,823 | -0,278 | 0,216 | 0,545 | 0,503 | 0,180 | 0,212 | 0,135 | -0,823 | 0,944 | 0,480 |
| 114 | 167 | **166** | 0,426 | -1,000 | -0,287 | 0,657 | 0,710 | 0,714 | 1,009 | 0,740 | 1,155 | 0,247 | -1,000 | -0,401 |
| 139 | 167 | **167** | 0,305 | -1,000 | -0,348 | 0,671 | 0,793 | 0,641 | -0,008 | 0,363 | 0,361 | 0,679 | 0,243 | -1,000 |
| 173 | 163 | **168** | -0,075 | -0,682 | -0,378 | -0,682 | -0,647 | -0,666 | 0,134 | 0,013 | -0,152 | 0,607 | 0,243 | 0,480 |
| 170 | 164 | **169** | 0,006 | -0,793 | -0,394 | -0,276 | -0,135 | -0,219 | -0,301 | -0,208 | -0,232 | -0,793 | 1,638 | 0,577 |
| 158 | 167 | **170** | 0,181 | -1,000 | -0,409 | 0,418 | -0,004 | 0,149 | 0,868 | 0,542 | 0,459 | -1,000 | 0,243 | -0,042 |
| 165 | 167 | **171** | 0,113 | -1,000 | -0,444 | 0,350 | -0,101 | 0,044 | 0,389 | 0,406 | 0,725 | -1,000 | 0,243 | -0,042 |
| 166 | 167 | **172** | 0,110 | -1,000 | -0,445 | 0,698 | 0,670 | 0,711 | -0,151 | -0,372 | -0,152 | 0,103 | -1,000 | 0,480 |
| 172 | 167 | **173** | -0,028 | -1,000 | -0,514 | 0,761 | 0,431 | 0,923 | -1,000 | -1,000 | -1,000 | 0,247 | -0,093 | 0,480 |
| 174 | 167 | **174** | -0,174 | -1,000 | -0,587 | -1,000 | -1,000 | -1,000 | 0,615 | 0,243 | 0,058 | -0,202 | 0,243 | 0,480 |

Source: Author's own elaboration.

**Table 5** Descriptive statistics of teachers in the first decile (D1), first quartile (Q1), tenth decile (D10), and the fourth quartile (Q4), ordered by the mixed index, using the original variable values.

| | | **Mean scores** | | | **Engagement** | | | **Teaching Minutes** | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **Reading** | **Math** | **Science** | **Reading** | **Math** | **Science** | **Reading** | **Math** | **Science** |
| **D1** | **Mean** | 548,43 | 524,46 | 546,84 | 10,82 | 10,93 | 10,78 | 273,24 | 237,65 | 224,71 |
| | **Stand. Dev.** | 13,35 | 20,76 | 21,42 | 0,35 | 0,34 | 0,40 | 56,57 | 27,92 | 25,87 |
| **Q1** | **Mean** | 540,36 | 516,29 | 538,19 | 10,54 | 10,68 | 10,66 | 282,91 | 252,33 | 223,02 |
| | **Stand. Dev.** | 19,12 | 26,55 | 25,66 | 0,45 | 0,49 | 0,49 | 56,62 | 31,76 | 29,02 |
| **Q4** | **Mean** | 495,71 | 461,93 | 486,33 | 9,33 | 9,50 | 9,48 | 403,30 | 303,75 | 257,50 |
| | **Stand. Dev.** | 41,53 | 40,85 | 40,49 | 0,98 | 0,96 | 0,97 | 142,28 | 53,57 | 73,64 |
| **D10** | **Mean** | 488,28 | 453,79 | 479,98 | 9,33 | 9,43 | 9,34 | 466,94 | 314,17 | 292,78 |
| | **Stand. Dev.** | 41,95 | 46,11 | 45,92 | 1,23 | 1,14 | 1,13 | 169,65 | 56,77 | 98,08 |

Source: Author's own elaboration.

**Table 6** Descriptive statistics of teachers in the first decile (D1), first quartile (Q1), tenth decile (D10), and the fourth quartile (Q4), ordered by the mixed index, using the normalized variable values.

| | | **Mean scores** | | | **Engagement** | | | **Teaching Minutes** | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **Reading** | **Math** | **Science** | **Reading** | **Math** | **Science** | **Reading** | **Math** | **Science** |
| **D1** | **Mean** | 0,96 | 1,01 | 1,00 | 1,04 | 0,99 | 0,93 | 0,85 | 1,03 | 0,62 |
| | **Stand. Dev.** | 0,27 | 0,32 | 0,34 | 0,24 | 0,20 | 0,27 | 0,37 | 0,41 | 0,24 |
| **Q1** | **Mean** | 0,84 | 0,90 | 0,88 | 0,86 | 0,84 | 0,86 | 0,79 | 0,83 | 0,64 |
| | **Stand. Dev.** | 0,34 | 0,41 | 0,40 | 0,28 | 0,29 | 0,34 | 0,33 | 0,43 | 0,27 |
| **Q4** | **Mean** | 0,21 | 0,18 | 0,18 | 0,21 | 0,22 | 0,23 | 0,26 | 0,23 | 0,40 |
| | **Stand. Dev.** | 0,58 | 0,49 | 0,50 | 0,49 | 0,43 | 0,44 | 0,54 | 0,61 | 0,45 |
| **D10** | **Mean** | 0,10 | 0,09 | 0,12 | 0,23 | 0,20 | 0,17 | 0,03 | 0,13 | 0,17 |
| | **Stand. Dev.** | 0,53 | 0,52 | 0,54 | 0,59 | 0,46 | 0,49 | 0,60 | 0,66 | 0,53 |

Source: Author's own elaboration.

### *5.2. Robustness analysis*

A robustness check of the results obtained has been carried out, in order to know the impact on the teachers' efficiency ranking when any of the parameters involved in the synthetic indexes' calculation is varied.

#### *5.2.1. Changing the Weights*

In Section 5.1, we analysed the ranking when all the inputs and outputs were equally weighted. The weights assigned to the variables play a crucial role in the calculation of the synthetic indexes, and the efficiency of a teacher may be affected if we give more importance to some of the variables than to the rest. In this case, the normalized values of the highest weighted variables count more for the calculation of the (weighted) weak, strong and mixed indexes, as shown in equations (10), (13) and (14), respectively. However, a change in the weights does not alter the normalized values obtained by the input and output variables.

Let us study what happens if we vary the weights assigned to the inputs and outputs (in Table 1) to study the following scenarios:

- Giving more importance to the scores in the three subjects than to the other variables, and equal importance to the engagements and to the teaching times. That is:
- Giving more importance to the engagements in the three subjects than to the other variables, and equal importance to the scores and to the teaching times.
- Giving more importance to the teaching times in the three subjects than to the other variables, and equal importance to the scores and to the engagements.

The weights $\omega_j$ given to the variables $v_j$ (with $j = 1, \ldots, N_V$) to represent these three scenarios can be depicted in Table 7. As previously explained in Section 4.1, when different weights are given to the variables, the (weighted) weak, strong and mixed indexes are obtained using equations (10), (13) and (14).

**Table 7** Different weights used for the variables.

| **Variables** | | **More importance to the Scores** | **More importance to the Engagements** | **More importance to the Teaching Times** |
|---|---|---|---|---|
| ***Outputs*** | | | | |
| Mean scores in: | | | | |
| Reading | $\omega_1$ | 2 | 1 | 1 |
| Mathematics | $\omega_2$ | 2 | 1 | 1 |
| Science | $\omega_3$ | 2 | 1 | 1 |
| Engagement in: | | | | |
| Reading | $\omega_4$ | 1 | 2 | 1 |
| Mathematics | $\omega_5$ | 1 | 2 | 1 |
| Science | $\omega_6$ | 1 | 2 | 1 |
| ***Inputs*** | | | | |
| Teaching minutes in: | | | | |
| Reading | $\omega_7$ | 1 | 1 | 2 |
| Mathematics | $\omega_8$ | 1 | 1 | 2 |
| Science | $\omega_9$ | 1 | 1 | 2 |

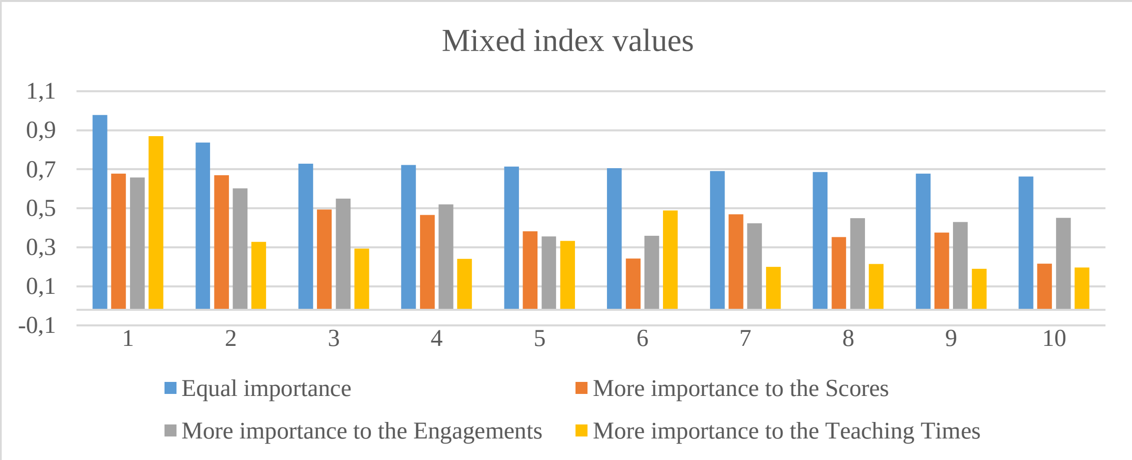
Source: Author's own elaboration.

To save space, we just show the mixed index values achieved in the three scenarios. The rest of the results are available upon request to the authors. To compare the new results against the ones in Section 5.1, in Figure 1, we have represented the mixed index values achieved at the four scenarios by the 10 most efficient teachers identified in Section 5.1, and we show their positions in the ranking produced at each scenario in Table 8. By "*Equal importance*" we refer to the results previously described in Section 5.1, where all variables were equally weighted. Logically, there are some changes in the values obtained and in the rankings generated. However, the teachers at the first three positions in the ranking "*Equal importance*" (teachers 1, 2 and 3) preserve the same efficiency position when giving more weights to the scores and to the engagements, which prove the robustness of the efficiency of these teachers at least in regards to these scenarios. However, except for teacher 1, the ranking of all teachers are altered in the last scenario ("*More importance to the Teaching Times*"), which corroborates the role of the instruction times for a teacher to be efficient. This happens since there are other teachers achieving better values for these variables, making them to be ranked higher.

In the same way, we have also included the same information for the 10 teachers at the bottom of the ranking in Section 5.1, which can be seen in Figure 2 and Table 9. According to Figure 2, if different weights are used, the values of the mixed index become much lower than when all variables are equally weighted. Indeed, the least efficient teacher at Section 5.1 (teacher 174) is again at the last position when more importance is assigned to the Scores, but (s)he is at position 170 in the scenario giving more emphasis to the engagements and at position 166 when the teaching times are more weighted.

It can thus be concluded that, in the light of the results obtained, the use of different weights for the inputs and outputs can offer different information, so the weights must be set according to the purpose of the study to be carried out to compare the teachers according to their efficiency.

**Figure 1** Weak index values at the different scenarios of the 10 most efficient Spanish teachers in Section 5.1.



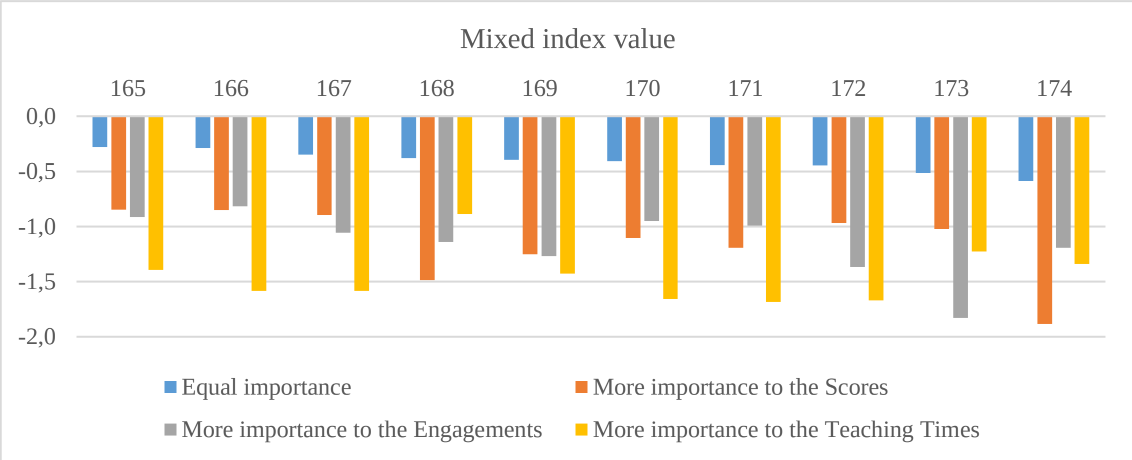

Source: Author's own elaboration.

**Table 8** Ranking positions according to the weak index at the different scenarios of the 10 most efficient Spanish teachers in Section 5.1.

| Equal importance | More importance to the Scores | More importance to the Engagements | More importance to the Teaching Times |
|---|---|---|---|
| 1 | 1 | 1 | 1 |
| 2 | 2 | 2 | 5 |
| 3 | 3 | 3 | 7 |
| 4 | 5 | 4 | 10 |
| 5 | 10 | 12 | 4 |
| 6 | 21 | 11 | 2 |
| 7 | 4 | 8 | 13 |
| 8 | 12 | 6 | 12 |
| 9 | 11 | 7 | 15 |
| 10 | 24 | 5 | 14 |

Source: Author's own elaboration.

**Figure 2** Weak index values at the different scenarios of the 10 least efficient Spanish teachers in Section 5.1.




Source: Author's own elaboration.

**Table 9** Ranking positions according to the weak index at the different scenarios of 10 ten least efficient Spanish teachers in Section 5.1.

| Equal importance | More importance to the Scores | More importance to the Engagements | More importance to the Teaching Times |
|---|---|---|---|
| 165 | 150 | 157 | 167 |
| 166 | 152 | 151 | 170 |
| 167 | 157 | 166 | 171 |
| 168 | 173 | 168 | 149 |
| 169 | 172 | 171 | 168 |
| 170 | 168 | 160 | 172 |
| 171 | 170 | 162 | 174 |
| 172 | 160 | 172 | 173 |
| 173 | 163 | 174 | 165 |
| 174 | 174 | 170 | 166 |

Source: Author's own elaboration.

### *5.2.2. Changing the Aspiration and Reservation Levels*

Logically, the results obtained highly depend on the reference levels chosen (i.e. aspiration and reservation levels for the inputs and outputs), which in turn must be set according to the final aim of the analysis. As described in Section 4.2, we have used relative values to set them, which were fixed as the mean plus/minus the standard deviation attained by each variable. The use of these statistical values implies that the final result for each teacher indicate its situation (regarding efficiency), as compared to the rest of the teachers in our sample. That is, the measure of efficiency is a relative one.

A change in these reference levels affects, on the one hand, to the normalization of the variables and, on the second hand, to the calculation of the three synthetic indexes.

Let us see how the use of other reference levels would influence the results. Concretely, let us now define the aspiration and reservation levels of each output variable as percentiles 75 and 25, respectively (as these variables are of the type “more is better”). For the input variables (of the “less is better” type), let us use percentile 25 as the aspiration level and percentile 75 as the reservation level. Table 10 shows the new reference levels considered.

**Table 10** New aspiration and reservation levels used for the variables.

| **Variables** | **“More is better” or “Less is better” type** | **Aspiration level** | **Reservation level** | **Maximum value** | **Minimum value** |
|---|---|---|---|---|---|
| ***Outputs*** | | | | | |
| Mean scores in: | | | | | |
| Reading | “More” | 545.44 | 501.56 | 592.29 | 392.96 |
| Mathematics | “More” | 511.30 | 467.04 | 581.62 | 347.87 |
| Science | “More” | 538.21 | 490.43 | 604.63 | 376.26 |
| Engagement in: | | | | | |
| Reading | “More” | 10.53 | 9.42 | 12.16 | 6.48 |
| Mathematics | “More” | 10.71 | 9.53 | 12.64 | 5.81 |
| Science | “More” | 10.69 | 9.48 | 12.05 | 6.10 |
| ***Inputs*** | | | | | |
| Teaching minutes in: | | | | | |
| Reading | “Less” | 285.00 | 360.00 | 780.00 | 120.00 |
| Mathematics | “Less” | 240.00 | 300.00 | 420.00 | 180.00 |
| Science | “Less” | 210.00 | 240.00 | 540.00 | 60.00 |

Source: Author’s own elaboration.

Table 11 shows the positions of the 10 most (left) and of the 10 least (right) efficient teachers of Section 5.1 at the new ranking generated by the mixed index obtained according to the new aspiration and reservation levels. As said, this change affects the normalized values of the variables, which differ from the normalized values shown in Tables 3 and 4. To save space, this information has not been included in the article, but it is available upon request to the authors. It can be observed that, for the old and the new reference levels, the same teachers occupy the first and second positions at the top and at the bottom of the ranking. The rest of teachers slightly change their positions. This means that the results described in Section 5.1 are quite stable, at least in comparison to the new aspiration and reservation levels used.

**Table 11** Ranking positions with the initial and new reference levels according to the weak index of the 10 most and 10 least efficient Spanish teachers in Section 5.1.

| *The 10 most efficient teachers* | | *The 10 least efficient teachers* | |
|---|---|---|---|
| **Initial Aspiration and Reservation Levels** | **New Aspiration and Reservation Levels** | **Initial Aspiration and Reservation Levels** | **New Aspiration and Reservation Levels** |
| 1 | 1 | 165 | 167 |
| 2 | 2 | 166 | 162 |
| 3 | 4 | 167 | 166 |
| 4 | 8 | 168 | 170 |
| 5 | 5 | 169 | 169 |
| 6 | 3 | 170 | 168 |
| 7 | 14 | 171 | 172 |
| 8 | 11 | 172 | 171 |
| 9 | 17 | 173 | 173 |
| 10 | 19 | 174 | 174 |

Source: Author's own elaboration.

Note that there is no unique way to set these reference values. Indeed, they may be set in an absolute manner, according to the expectations expressed by a group of experts in the problem domain. Nevertheless, these experts should be available and willing to be actively involved in the process, which may be very cognitively demanding if the number of variables considered is elevated.

### 5.3. *Second Stage Analysis*

Besides the analysis of the efficiency ranking obtained with the synthetic indexes, it would be interesting to have a deeper knowledge about the influence of the variables which are controllable (i.e. the inputs, which are the time spent to teach each of the three subjects according to Table 1) in this ranking. In this second stage analysis, we seek to establish the contribution of each controllable input to achieve the highest efficiency positions, i.e. we want to know the extent to which these variables have been decisive for a teacher to be ranked in the best positions. This will shed some light about how to distribute the available resources in order to promote efficiency among teachers.

To this aim, simple regression models have been estimated in which the weak, strong and mixed indexes' values are regressed on the three inputs. The regression coefficients are estimated using ordinary least square (OLS), whose idea is to minimize the so-called statistical noise as much as possible. Let us remind that, respectively, $v_6, v_7$ and $v_8$ denote the input variables for the teaching minutes dedicated to reading, mathematics and science, $x_{i6}$, $x_{i7}$ and $x_{i8}$ are the original values of these variables for the teacher $i$, and $y_{i6}$, $y_{i7}$ and $y_{i8}$ are the normalized values of these variables for the teacher $i$. Then, the three synthetic indexes for each teacher $i$ can be predicted with respect to these normalized variable values by the following three equations:

$$
\begin{aligned}
WI_i &= \hat{\alpha}^W + \hat{\beta}_6^W y_{i6} + \hat{\beta}_7^W y_{i7} + \hat{\beta}_8^W y_{i8} + \epsilon_{Wi}, \\
SI_i &= \hat{\alpha}^S + \hat{\beta}_6^S y_{i6} + \hat{\beta}_7^S y_{i7} + \hat{\beta}_8^S y_{i8} + \epsilon_{Si}, \\
MI_i &= \hat{\alpha}^M + \hat{\beta}_6^M y_{i6} + \hat{\beta}_7^M y_{i7} + \hat{\beta}_8^M y_{i8} + \epsilon_{Mi},
\end{aligned}
\tag{17}
$$

where, for each index, $\hat{\beta}^W = \left(\hat{\beta}_6^W, \hat{\beta}_7^W, \hat{\beta}_8^W\right)^T$, $\hat{\beta}^S = \left(\hat{\beta}_6^S, \hat{\beta}_7^S, \hat{\beta}_8^S\right)^T$ and $\hat{\beta}^M = \left(\hat{\beta}_6^M, \hat{\beta}_7^M, \hat{\beta}_8^M\right)^T$ are the vectors of regression coefficients (*slopes*), $\hat{\alpha}^W$, $\hat{\alpha}^S$ and $\hat{\alpha}^M$ are the estimated population intercept and $\epsilon_{Wi}$, $\epsilon_{Si}$ and $\epsilon_{Mi}$ are the error terms. In this way, we are assuming that each index is affected by random factors, which are inherently unobservable and distributed normally. Table 7 shows the estimated coefficients, where the standard deviations and the significance levels for each coefficient are also reported.

To interpret the estimation results obtained in the regression model, it is important to have in mind that we are using the normalized values of the three inputs, for which "more is better" due to the normalization, although the original values of these three inputs are of the "less is better" type.

**Table 7** OLS estimates of the three index values.

| **Variables** | **Weak index** ($WI_i$) | **Strong index** ($SI_i$) | **Mixed index** ($WI_i$) |
|---|---|---|---|
| Teaching minutes in: | | | |
| Reading | 0.175*** | 0.323*** | 0.249*** |
| | (0.044) | (0.059) | (0.046) |
| Mathematics | 0.118*** | 0.168*** | 0.143*** |
| | (0.038) | (0.051) | (0.040) |
| Science | 0.068 | 0.207*** | 0.136*** |
| | (0.046) | (0.061) | (0.048) |
| Constant | 0.338*** | -0.406*** | -0.033 |
| | (0.035) | (0.047) | (0.037) |
| Observations | 174 | 174 | 174 |
| R-squared | 0.231 | 0.352 | 0.343 |

Source: Author's own elaboration. OLS estimates. Standard errors in parentheses. Dependent variable: Values of the weak, strong and mixed indexes. Labels: *** significant at 1%, ** significant at 5%, * significant at 10%.

We can see that the normalized values of the three inputs (of the "more is better" type) are positively correlated with the three synthetic indexes, although it seems that the teaching time for science was not so relevant for the weak index. In addition, it can be observed that spending more time in reading is decisive in the calculation of the three indexes, given that its correlation coefficient is the highest one from three subjects, either for the weak, strong and mixed indexes. Overall, the strong index would be the furthest improved by reaching better inputs (normalized) values, given that, in the hypothetical absence of teaching times, this index would be estimated with a negative value (because its intercept is negative, as shown in Table 7). In this hypothetical case, the mixed index would experience the next highest improvement, and lately, the weak one.

If we interpret these findings in terms of the original values of the inputs (of the "less is better" type), our results indicate that the three indexes are negatively correlated with the three inputs, being the role of the time devoted to science the least relevant, and having the reading teaching time the most important impact. Therefore, we can conclude that, on average, when more time is spent in teaching any of the subjects, the efficiency of the teachers is not improved. This conclusion corroborates the findings previously deduced from Tables 3, 4, 5 and 6, where we found that the teaching times for the teachers at the bottom of the efficiency ranking are significantly higher than the times employed for teaching by the most efficient ones. This is in line with previous works such as e.g. Gromada and Shewbridge (2016), Lopez-Agudo and Marcenaro (2019), Luque et al., (2020) and Lopez-Agudo and Marcenaro (2022), which stated that, in general, weekly instruction time has no effect on children's academic achievement for Spain[2].

## 6. Conclusion

In this paper, a set of synthetic indexes based on multi-criteria decision making techniques have been analysed to study the teachers' efficiency. The synthetic indexes calculated enable us to use either a full compensation among all the variables (weak index), no compensation informing about the worst variable value (strong index), or a specific degree of compensation of the two formers (mixed indicator). The novelty yields in the fact that these indexes are built using desirable ranges of values to be achieved by each of the variables used, enabling us to include preferences (defined according to the observational data considered in this study) to evaluate the teachers' efficiency. These

[2] Nevertheless, the data available do not allow measuring the quality of the weekly instruction time received by children in a precise manner for our purposes, so our results should be taken with caution in this area.

desirable ranges, defined by the aspiration and reservation levels, are calculated by the statistical descriptors of the indicators according to the observational data considered.

According to the data used (corresponding to TIMSS and PIRLS 2011 for fourth-grade Spanish teachers for reading, mathematics and science), the most efficient teachers are those with students reaching higher learning performances in the three subjects. Conversely, the times dedicated to teach are significantly lower for these teachers than the teaching minutes spent by the least efficient teachers, which supports similar conclusions stated by other studies (Gromada and Shewbridge, 2016; López-Agudo and Marcenaro, 2018; Luque et al., 2020). This leads us to the idea that a better distribution of the instruction time in class may directly imply an improvement of the academic performance of Spanish students, thus making a more efficient use of the resources, not only in terms of time but also in what concerns the monetary costs of education.

Increasing the instruction times (our inputs) may produce better outputs (i.e. scores and/or engagements), although there is not a direct relationship of the teaching times with the academic performance according to the literature. However, from the efficiency point of view, our findings suggest that teaching for more time, all alone, does not improve the efficiency of a teacher. It would be interesting to research in the future if an increase in the teaching time would imply the efficiency to be higher if, at the same time, other aspects of the teaching-learning context, such as e.g. the quality of the teaching or resources in the classroom, are also improved. We plan to carry out a more in-depth study of these aspects in a future research. Finally, it is convenient to point out that our findings do not categorically conclude that an increase of the instruction times negatively influences the efficiency of a teacher. In our study, our concept of efficiency must be understood from a global perspective, since it has been defined based on the optimization of both inputs and outputs at the same time. Therefore, our results cannot be interpreted from just the inputs' point of view, nor from the outputs' one. The compensation and balance among the inputs and outputs is what counts to decide the efficiency score of a teacher in our study.

In addition, we have performed a second stage analysis whose main purpose is to know the impact of the controllable inputs considered (i.e. times spent in reading, mathematics and science) in the efficiency ranking given by the three indexes. In line with our previous findings, this analysis has indicated that a strategy to enhance the performance of the least efficient teachers is to reorganise the instruction time in class. Furthermore, at this redistribution of the lessons, our analysis has corroborated the key role of the time dedicated to reading at class. Therefore, educational policy makers should have this in mind to dictate policies and measures that aim to attain efficiency among teachers, and thus, an efficient educational system.

To our knowledge, to date, no one has investigated the efficiency of teachers using a discrete multi-criteria approach as the one used. The dimensional character of what has been analysed has been reflected with the double benchmark methodology and with the definition of synthetic indicators. This, together with the conclusions just mentioned, makes this work a novel contribution to the field of economics of education. A possible future research line may consist of replicating our study using absolute, and not relative, values to set the aspiration and reservation levels, as mentioned before. Their role is crucial when normalizing the variables and, thus, when obtaining the indexes and the efficiency ranking. One or a group of policy decision makers -experts in the economics of education domain- can establish them according to different purposes (i.e. they can be set using benchmarks from other countries). In this case, the synthetic indexes would give us an absolute measure of performance of each teacher, with respect to the absolute values the experts want to use.

## Acknowledgements

The authors would like to thank for the financial support received from Spanish government (project PID2020-115429GB-I00), from the Andalusian Regional Ministry of Economy and Knowledge (PAIDI group SEJ-532) and from the University of Málaga (grant B1-2020_18).

## References

Agasisti, T., and Johnes, G. (2015). Efficiency, costs, rankings and heterogeneity: The case of US higher education. *Studies in Higher Education*, 40(1): 60–82.

Andersen, S.C., Humlum, M.K., and Nandrup, A.B. (2016). Increasing instruction time in school does increase learning. *Proceedings of the National Academy of Sciences of the United States of America*, 113 (27): 7481–7484.

Baker, D.P., Fabrega, R., Galindo, C., and Mishook, J. (2004). Instructional time and national achievement: cross-national evidence. *Prospects*, 34 (3): 311–334.

Blackburn, V., Brennan, S., and Ruggiero, J. (2014). Measuring efficiency in Australian Schools: A preliminary analysis. *Socio-Economic Planning Sciences*, 48(1): 4–9.

Cabello, J.M., Navarro, E., Thiel, D., Rodriguez, B., and Ruiz, F. (2021). Assessing environmental sustainability by the double reference point methodology: the case of the provinces of Andalusia (Spain). *International Journal of Sustainable Development & World Ecology*, 28: 4–17.

Calero, J., and Escardíbul, J.O. (2020). Teacher quality and student skill acquisition. An analysis based on PIRLS-2011 outcomes. *Educational Studies*, 46(6): 676-692.

Cattaneo, M.A., Oggenfuss, C., and Wolter, S.C. (2017). The more, the better? The impact of instructional time on student performance. *Education Economics*, 25 (5): 433-445.

Cherchye, L., Moesen, W., Rogge, N., and Van Puyenbroeck, T. (2007). An introduction to ‘benefit of the doubt’ composite indicators. *Social Indicators Research*, 82: 111–145.

Crespo-Cebada, E., Pedraja-Chaparro, F., and Santin, D. (2014). Does school ownership matter? An unbiased efficiency comparison for regions of Spain. *Journal of Productivity Analysis*, 41(1): 153–172.

Despotis, D.K. (2005). A reassessment of the human development index via data envelopment analysis. *Journal of the Operational Research Society*, 56(8): 969–980.

De Witte, K., Groot, W., and Maassen van den Brink, H. (2013). The efficiency of education in generating literacy: a stochastic frontier approach. *Review of Economics & Finance*, 2(2): 25–37.

De Witte, K., and Lopez-Torres, L. (2017). Efficiency in education: a review of literature and a way forward. *Journal of the Operational Research Society*, 68(4): 339–363.

Dolton, P. Marcernaro-Gutierrez, O., De Vries, R., and She P.-W. (2018). *Global Teacher Efficiency Index 2018*. Varkey Foundation, London.

Egalite, A.J., and Kisida, B. (2018). The effects of teacher match on students' academic perceptions and attitudes. *Educational Evaluation and Policy Analys*is, 40 (1): 59-81.

Essid, H., Ouellette, P., and Vigeant, S. (2010). Measuring efficiency of Tunisian schools in the presence of quasi-fixed inputs: A bootstrap data envelopment analysis approach. *Economics of Education Review*, 29(4): 589–596.

Gromada, A., and Shewbridge, C. (2016). Student learning time: a literature review. *OECD Education Working Papers*, No. 127. OECD Publishing, Paris.

Haelermans, C., and Ruggiero, J. (2013). Estimating technical and allocative efficiency in the public sector: A nonparametric analysis of Dutch schools. *European Journal of Operational Research*, 227(1): 174–181.

Hanushek, E.A., and Woessmann, L. (2007). *Education Quality and Economic Growth*. World Bank, Washington, DC.

Huebener, M., Kuger, S., and Marcus, J. (2017). Increased instruction hours and the widening gap in student performance. *Labour Economics*, 47: 15–34.

Johnson, A.L., and Ruggiero, J. (2014). Nonparametric measurement of productivity and efficiency in education. *Annals of Operations Research*, 221(1): 197–210.

Jürges, H., and Schneider, K. (2007). Fair ranking of teachers. *Empirical Economics,* 32: 411–431.

Lavy, V. (2015). Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries. *The Economic Journal*, 125: F397–F424.

Lim, J., and Meer, J. (2021). Persistent effects of teacher-student gender matches. *Journal of Human Resources*, 0320-10796R2

Locher, F., and Pfost, M. (2020). The relation between time spent reading and reading comprehension throughout the life course. *Journal of Research in Reading*, 43 (1): 57-77.

Lockheed, M., and Hanushek, E. (1994). Concepts of educational efficiency and effectiveness. In Husén, T., and Postlethwaite, T.N. (eds.), *International encyclopedia of education*, pp. 1779–1784. Pergamon, Oxford.

Lopez-Agudo, L.A., and Marcenaro-Gutierrez, O.D. (2017). Engaging children in lessons: the role of efficient and effective teachers. *School Effectiveness and School Improvement*, 28(4): 650–669.

Lopez-Agudo, L., and Marcenaro-Gutierrez, O.D. (2019). Are Spanish children taking advantage of their weekly classroom time? *Child Indicators Research*, 12: 187–211.

Lopez-Agudo, L.A., and Marcenaro-Gutierrez, O.D. (2022). Instruction time and students’ academic achievement: a cross-country comparison. *Compare: A Journal of Comparative and International Education*, 52(1): 75-91.

Luque, M., Perez-Moreno, S., and Rodriguez, B. (2016). Measuring human development: a multi-criteria approach. *Social Indicators Research*, 125(3): 713–733.

Luque, M., Perez-Moreno, S., Robles, J.A., and Rodriguez, B. (2017). Measuring child and maternal health in developing countries: a proposal of new hybrid MDG composite indices. *Applied Research in Quality of Life*, 12(3): 737–758.

Luque, M., Marcenaro-Gutierrez, O.D., and Ruiz, A.B. (2020). Evaluating the global efficiency of teachers through a multi-criteria approach. *Socio-Economic Planning Sciences*, 70:100676.

Marcenaro-Gutierrez, O.D., Luque, M., and Lopez-Agudo, L.A. (2016). Balancing teachers' math satisfaction and other indicators of the education system's performance. *Social Indicators Research*, 129 (3): 1319-1348.

Martin, M.O., and Mullis, I.V.S. (2012). *Methods and Procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA, TIMSS & PIRLS International Study Center, Boston College.

Martin, M.O., Mullis, I.V.S., Foy, P., and Stanco, G.M. (2012). *TIMSS 2011 International Results in Science*. Chestnut Hill, MA, TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., and Martin, M.O. (2017). *PIRLS 2016 international results in reading*. Chestnut Hill, MA, TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., Foy, P., and Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*. Chestnut Hill, MA, TIMSS & PIRLS International Study Center, Boston College.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., and Giovannini, E. (2005). Handbook on constructing composite indicators: methodology and user guide, *OECD Statistics Working Papers*, No. 2005/03, D.O.I. https://doi.org/10.1787/533411815016. OECD Publishing, Paris.

OECD (2017). *PISA 2015 Technical Report*. Retrieved from http://www.oecd.org/pisa/data/2015-technical-report/, last accessed on 25 March 2021.

Perez-Moreno, S., Luque, M., and Rodriguez, B. (2016). Assessing global competitiveness under multi-criteria approach. *Economic Modelling*, 53(1): 398–408.

Podinovski, V.V., Ismail, I., Bouzdine-Chameeva, T., and Zhang, W. (2014). Combining the assumptions of variable and constant returns to scale in the efficiency evaluation of secondary schools. *European Journal of Operational Research*, 239(2): 504–513.

Rivkin, S.G. and Schiman, J.C. (2015). Instruction time, classroom quality, and academic achievement. *The Economic Journal*, 125: F425–F448.

Ruiz, F., Cabello, J.M., and Luque, M. (2011). An application of reference point techniques to the calculation of synthetic sustainability indicators. *Journal of the Operational Research Society,* 62: 189–197.

Ruiz, F., Cabello, J.M., and Perez-Gladish, B. (2018). Building Ease-of-Doing-Business synthetic indicators using a double reference point approach. *Technological Forecasting and Social Change*, 131: 130–140.

Ruiz, F. and Cabello, J.M. (2021). MRP-PCI: A multiple reference point based partially compensatory composite indicator for sustainability assessment. *Sustainability*, 13(3): 1261.

Sims, D.P. (2008). Strategic responses to school accountability measures: it's all in the timing. *Economics of Education Review,* 27 (1): 58–68.

Scippacercola, S., and D'Ambra, L. (2014). Estimating the relative efficiency of secondary schools by stochastic frontier analysis. *Procedia Economics and Finance*, 17: 79–88.

Thompson-Schill, S.L., Ramscar, M., and Chrysikou, E.G. (2009). Cognition without control: when a little frontal lobe goes a long way. *Current Directions in Psychological Science*, 18: 259–263.

Wanke, P., Blackburn, V., and Pestana-Barros, C. (2016). Cost and learning efficiency drivers in Australian schools: a two-stage network DEA approach. *Applied Economics*, 48(38): 3577–3604.

Wierzbicki, A.P., Makowski, M., and Wessels, J. (2000). *Model-Based Decision Support Methodology with Environmental Applications*. Kluwer Academic Publishers, Dordrecht.

Woessmann, L. (2010). Institutional determinants of school efficiency and equity: German states as a microcosm for OECD countries. *Journal of Economics and Statistics*, 230 (2): 234–270.

## Ethics declarations

### *Conflict of interest*

The authors declare that they have no conflict of interest.