

The version of record of this article has been published and is available in
School Effectiveness and School Improvement: An International Journal of Research, Policy and
Practice (date of publication 1 August 2017)

<https://doi.org/10.1080/09243453.2017.1364272>

Engaging children in lessons: the role of effective teachers

L. A. Lopez-Agudo^a and O. D. Marcenaro-Gutierrez^{b*}

^a*Departamento de Economía Aplicada (Estadística y Econometría, 15), Universidad de Malaga, C/ Ejido, 6, 29013, Malaga, Spain. Programa de Doctorado en Economía y Empresa de la Universidad de Malaga. Tel.: +34 95 213 7003; Fax: +34 95 213 7262. Email: lopezagudo@uma.es.*; ^b*Departamento de Economía Aplicada (Estadística y Econometría, 15), Universidad de Malaga, C/ Ejido, 6, 29013, Malaga, Spain. Tel.: +34 95 213 7003; Fax: +34 95 213 7262.*

Acknowledgements and Funding

This work has been partly supported by the *Consejería de Innovación, Ciencia y Empresa de la Junta de Andalucía* (PAI group SEJ-532 and Excellence research group SEJ-2727); by the *Ministerio de Economía y Competitividad* (Research Project ECO2014-56397-P) and scholarship FPU2014 04518 of the *Ministerio de Educación, Cultura y Deporte*. We also acknowledge the training received from the University of Malaga PhD Programme in Economy and Business [*Programa de Doctorado en Economía y Empresa de la Universidad de Malaga*].

* Corresponding author. Email: odmarcenaro@uma.es.

Abstract

The training of effective teachers has been one of the main aims of educational systems, in so far as it could be an important tool to improve the education performance of children and, consequently, contribute to their career progression and foster social mobility towards a more meritocratic society. The present study intends to identify these teachers by their ability to engage young students in lessons, which may reflect the capacity of their teaching practices and efficient use of available resources to increase students' learning.

We focus on fourth grade reading and mathematics teachers in Spain –using TIMSS and PIRLS 2011 data– and we propose a two-step procedure for this analysis: the first step obtains teachers' efficiency scores, which are later employed in a second step –together with teachers' practices– to explain children's engagement in lessons.

Keywords: children engagement; teachers' effectiveness; teachers' efficiency; time-invariant stochastic frontier analysis; multilevel regression analysis.

Introduction

The training of effective teachers has been one of the main aims of educational systems, in so far as it could be an important tool to improve the education performance of children and, consequently, contribute to their career progression. In this sense, it is necessary to first indicate how an effective teacher can be defined. In the effectiveness literature it is common to rely on the achievement of students to measure the effectiveness of a particular teacher. Nevertheless, we use the capacity of teachers to engage students in their lessons as an alternative to student achievement, focusing on the definition of effective teachers proposed by Creemers and Reezigt (1996) and Kyriakides (2008) “effective teachers are expected (...) to maximize engagement rates”. The idea is that so-defined effective teachers are expected to increase their students’ interest and implication in their studies. This way, we use students’ engagement as the outcome of an effective teacher in order to provide useful information to orientate educational policy decisions.

Following the recent literature, effective teachers could also be defined as “(...) those who provide their students with knowledge that is useful in future learning, presumably require their students to exert effort by paying attention and being concentrated in class and by doing demanding homework” (Braga, Paccagnella, & Pellizzari, 2014, p. 84). Thus, an effective teacher has the ability to reach out to the students and make them learn.

With regard to the concept of efficiency, it basically means getting the maximum output from the amount of inputs teachers use in the teaching-learning process, or achieving the same output with a minimum quantity of inputs¹. The definition of Creemers and Reezigt (1996) and Kyriakides (2008) also highlights that effective teachers are expected “(...) to organize and manage the classroom environment as an efficient learning environment”, so effective teachers may be also efficient. Hence, teachers’ efficiency might be defined as their ability to perform fruitful work in an organized way, using available resources and time without wasting any of them.

The development of a measure to identify effective teachers is a relevant issue, as the traditional “league tables”, which are based only on the average score of teachers’ students, do not take into account sampling variability and other sources of error (Leckie & Goldstein, 2009). A proper classification of teachers in terms of effectiveness is particularly important to reward effective teachers based on their performance, fostering higher levels of student engagement and learning.

Regarding the methodology used for this analysis, a two-step procedure has been followed, which –to the best of our knowledge– also entails a novelty in the study of teacher effectiveness. Firstly, teachers’ efficiency has been estimated by using stochastic frontier analysis (SFA), using as output students’ academic achievement in reading and mathematics and, as inputs, available school resources for teachers. Then, the efficiency scores obtained from the first procedure have been employed in a second step, together with teachers’ learning practices in the classroom, to explain students’ engagement in reading and mathematics lessons, by making use of multilevel regression models.

To carry out this research, we made use of PIRLS (Progress in International Reading Literacy Study) and TIMSS (Trends in Mathematics and Science Study) combined databases in their 2011 wave for fourth grade –9/10 years old– Spanish students. Fourth grade students were chosen as they are more malleable at this age (Thompson-Schill, Ramscar, & Chrysikou, 2009) and they would, thus, show a better reflection of teachers’ procedures; besides, Spanish primary school students have the same teacher in each cycle and for all subjects, so students’ engagement could be attributed to the teacher under study. In addition, both statistical sources contain a rich set of teacher level variables gathered in their teacher questionnaire.

¹ Own adaptation of Lockheed and Hanushek (1994) –and references therein– general definition of efficient system.

The rest of the paper is organized as follows: the second section provides a literature review. The third section deals with the data, while the fourth section shows the methodology employed to perform this analysis. Section five contains the results obtained from estimating the proposed models. In section six we finalize with a summary of the most relevant conclusions and educational policy implications of this analysis.

Literature review

In general, the concepts of effectiveness and efficiency in teaching have been approached separately by previous literature (Lockheed & Hanushek, 1994). There are many definitions and elements which characterize effective teachers. The latter are said to create an environment focused on academic activities rather than others as social relationships (Creemers, 1994; Reynolds, Sammons, Stoll, Barber, & Hillman, 1996). They are also able of organizing the development of activities with short transitions (Brophy & Good, 1986), dealing at the same time with potential bad behavior or other problems which could affect the flow of the lesson (Borich, 1996).

The contribution of effective teachers to students' learning is so important that Borich (1996) and Brophy and Good (1986) stated that students learn more during their lessons rather than doing homework or other school tasks at home, as teachers provide the knowledge directly to the student, which does not happen with other methods of self-learning. In addition, individual or group tasks are essential in effective teaching, it being necessary to explain and monitor the tasks, and also to provide feedback to students.

The ability of getting students engaged in lessons has been highlighted in the literature as a distinctive characteristic of effective teachers. Willms (2003) concluded –after analyzing PISA 2000– that students who are not engaged gradually lose interest in school and become disruptive, which can affect the rest of the class negatively. Timothy, Rupley, and Nichols (2007) stated that “Effective teachers not only use appropriate materials but also attend to actively engaging students in learning from the materials”. In addition, Creemers and Reezigt (1996) and Muijs et al. (2014) denoted that effective teachers “are expected to organize and manage the classroom environment as an efficient learning environment and thereby to maximize engagement rates”. These two definitions contain the main characteristics of an effective teacher we are considering in our research: a teacher who can use his/her learning materials and resources in an efficient way and, together with effective teaching methodologies, is able to engage students in the lesson and make them learn.

The relevance of engaging students has been remarked by some other authors, such as Taylor, Pearson, Clark, and Walpole (1999) for kindergarten to third grade United States students. They found that effective teachers engaged their students 96% of the time, while less effective teachers engaged them only 63% of the time. These effective teachers are able to link the new information to student's prior knowledge (McLaughlin et al., 2005) and also employ teaching methods in order to keep their students engaged, activating them cognitively, which can be achieved by a properly managed and supportive classroom environment (Lipowsky et al., 2009). Carini, Kuh, and Klein (2006) found that student engagement is linked in a positive way to high learning outcomes such as critical thinking.

Nevertheless, few studies have used data from TIMSS to measure teacher effectiveness, e.g. Zuzovsky (2009) for the case of Israel by means of multi-level regression analysis. She found that having an advanced degree and majoring in the field of teaching in mathematics have a differential effect on students' achievement. However, she did not focus on students' engagement to measure teacher effectiveness.

Focusing on teachers' efficiency, Ryan and Cooper (2007) claimed that the most efficient teachers could engage their students about thirty minutes per day more than average efficient teachers, and an hour more when compared to inefficient teachers. According to them,

the higher number of hours taught in a year by efficient teachers explained why their students learned more than those taught by inefficient teachers.

To the best of our knowledge, there is a virtual absence of literature which studies teacher effectiveness by also considering teacher efficiency as one of its elements at the Spanish primary education level. However, there are examples in which efficiency and effectiveness are confronted for other countries. This is the case, e.g., of McEwan and Carnoy (2000), who analyzed Chile's public and private schools and found that public schools were more effective, but that their efficiency was similar.

Data

Our analysis focuses on the effectiveness of the Spanish teachers who participated in TIMSS 2011 and PIRLS 2011. The purpose of TIMSS is to measure learning achievement in the areas of mathematics and science of students at the end of fourth (9–10 years) and eighth grades (13–14 years), while PIRLS is focused on reading achievement of fourth grade students.

PIRLS and TIMSS coincided in 2011 for the first time, so a total of 4,183 students who participated in both tests, and were taught by a total of 200 teachers², were chosen for our study. This sample was reduced in the efficiency analysis to 3,438 students taught by 184 teachers –for the case of reading– and 3,432 students taught by 184 teachers –for mathematics–, due to the missing information for some variables related to efficiency.

Students' engagement –proxy of teachers' effectiveness– is represented by an index provided by PIRLS 2011 –for reading– and TIMSS 2011 –for mathematics–, called “Students Engaged in Reading/Mathematics Lessons index”. Following Mullis, Martin, Foy, and Arora (2012), this index was created by asking students the degree to which they agreed (“Agree a lot”; “Agree a little”; “Disagree a little”; “Disagree a lot”) with five 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”³. This index has been used in this research as a continuous variable.

In what follows, we describe the procedure employed to select the variables of the teaching-learning process dealt with in our analysis. We have classified each of the available variables according to their adscription to teacher effectiveness or efficiency characteristics –supporting this classification in the previous literature–. When determining whether a teacher variable represents his/her effectiveness or efficiency, we have asked ourselves the following two questions: “Does this variable represent the available resources for teachers in order to develop their lessons, as materials, classroom characteristics, available time or students' background, which are not easily alterable by the teacher –efficiency–?”; “Does this variable describe a method or procedure defined by the teacher when teaching his/her students to make them exert effort in learning –effectiveness–?” The selected variables which respond to one of these questions and had a significant effect on explaining teacher efficiency or effectiveness were included in the estimations of Table 1 –efficiency– or Table 2 –effectiveness–⁴.

In addition to the variables presented in the effectiveness estimation, many of them related to teachers' methodologies were analyzed –as they fulfilled the criteria of the second question–, but they did not result significant in explaining teacher effectiveness.

Methodology

² The mixed TIMSS and PIRLS database has been employed by other researchers, like Foy and O'Dwyer (2013) and Grilli, Pennoni, Rampichini, and Romero (2014), although with different objectives.

³ In PIRLS 2011 there are two additional questions included in this index: “I like what I read about in school” and “My teacher gives me interesting things to read”.

⁴ The main descriptive statistics of the selected variables are not shown to conserve space; they are available upon request to the authors.

As indicated in the introduction section, the first step of this analysis –efficiency– begins with the prediction of teachers’ efficiency by using stochastic frontier analysis, considering as output students’ achievement and as inputs teachers’ available learning resources. This efficiency measure is employed in the second step of the analysis – effectiveness– as a regressor, together with some teacher characteristics, learning practices and procedures, which will explain students’ engagement.

First step: Stochastic frontier analysis

In the first step of our model, we make use of stochastic frontier analysis to obtain a measure of teacher efficiency. Concretely, a time-invariant stochastic frontier model (Battese & Coelli, 1988) has been proposed. It helps us define a parametric function in which there are two main levels: the first level is students, whose units –students– are grouped in the second level – teachers–. Although the definition of a proper functional form for SFA could give rise to problems, measurement errors are a major concern when dealing with students’ score data, which is a field where SFA outperforms non-parametric approaches. This, together with the use of a moderate set of inputs, made SFA more suitable for our intended study.

This procedure has the particularity that it presents a composed error term, which is the sum of a measurement or specification error and a one-sided disturbance, which represents inefficiency. We have employed a log-linear Cobb-Douglas education function, in line with Chakraborty, Biswas, and Lewis (2001), who remark that its use would be more appropriate than other forms as translog when there are fewer degrees of freedom –as in our case–. This frontier model can show the best combination of inputs that maximize average students’ performance.

Formally:

<Insert equation 1>

where y_{ji} denotes the output (scores in reading or mathematics) of student “ i ” taught by teacher “ j ”; $x_{kji}^{(1)}$ is the k^{th} input of the socio-economic background characteristics of student “ i ” – taught by teacher “ j ”– and β_k its technology parameter, while $z_{sj}^{(1)}$ represents the s^{th} input of teacher learning resources –time, materials, etc.– of teacher “ j ” and γ_s its technology parameter. The composite error consists of the sum or the difference of the disturbance u_j (representing inefficiency) –as it is a time-invariant model $u_{ji} = u_j$ – and v_{ji} , which is a normally distributed disturbance, these two terms being independent of each other and *i. i. d.* between observations. After performing our analysis we obtained a better fit for distribution (F) of the inefficiency term u_j with the truncated-normal one $-u_j \sim N^+(\mu, \sigma_u^2)$ –. Due to the assumptions to identify the inefficiency term, the estimation has to be performed by maximum likelihood. Once this model has been estimated, we predict a measure of efficiency – u_j – for each teacher.

Second step: Multilevel regression analysis

Standard regression models do not take into account the hierarchical structure which educational data usually presents. Because of that, and following works such as Lenkeit (2013) or Afshartous and Wolf (2007), multilevel regression analysis (Hierarchical Linear Models, HLM) constitutes a proper approach in order to study effectiveness when cross-section data are available and is thus used in the second step of our analysis.

Specifically, we estimate the contribution to the variance (total heterogeneity) of students’ characteristics in the first level and that of teachers in the second level. Then, a regression is defined for each teacher in the second level. The estimated model can be specified in a general way for $h = 1, \dots, q$ variables by:

<Insert equation 2>

where E_{ij} denotes the engagement index in the subjects of reading or mathematics of student “ i ” who is taught by teacher “ j ”; δ_{0j} is a level parameter that reflects the differential effect on E_{ij} of

each teacher; $x_{hij}^{(2)}$ reflects the h^{th} characteristic of student “ i ” –taught by teacher “ j ”–; δ_h is the slope for the relationship between E_{ij} and $x_{hij}^{(2)}$ and ε_{ij} is the idiosyncratic error term.

Equation (2) assumes that the intercept of the regression is the same for all transversal units. However, we need to control each teacher’s “individual” characteristics, so we use a random effects model, concretely, a random intercept model. This methodology suggests that each teacher has a different level parameter (interception term), which is not fixed and behaves as a random variable with an average value (λ_0) affected by deviations represented by the random variable w_j , so δ_{0j} could be defined as:

<Insert equation 3>

When substituting equation (3) in (2) we obtain the expression of the random intercept model:

<Insert equation 4>

The random effects estimation is preferable in the case that the disturbance term associated with the teacher (w_j) and the exogenous variables ($x_{hij}^{(2)}$) are not correlated. The ε_{ij} is assumed independent among students and also independent of w_j . Finally, we add the effect on the engagement (E_{ij}) of the $r = 1, \dots, g$ variables considered at the teacher level ($z_{1j}^{(2)}, \dots, z_{gj}^{(2)}$) to equation (4), reaching the following model:

<Insert equation 5>

which represents the multilevel model using random effects, where ϑ_r is the slope for the relationship between E_{ij} and $z_{rj}^{(2)}$.

In order to establish a quantitative measure of teachers’ effectiveness, teacher-level residuals are usually used, because they can be defined as predictions of the random effects representing teacher effectiveness. Following Raudenbush and Willms (1995), these residuals can be obtained by using the Empirical Bayes (EB) method, which produces the so-called “shrunk residuals” –also known as best linear unbiased predictors (BLUPs)–. They are shrunk but efficient estimates of teacher effectiveness and avoid the accidental assignment of a teacher from the bottom to the top of the ranking (Arpino & Varriale, 2010), although they are affected by the sampling variability and other errors. Because of this, Goldstein and Healy (1995) suggested creating pairwise confidence intervals in order to overcome this uncertainty –this methodology has been used in the caterpillar plots in Figure 1 and Figure 2 in the Results section–. Hence, as our analysis focuses on determining which teachers have top or bottom effectiveness, we will make use of the Empirical Bayes method.

Results

The main results obtained from both teacher efficiency and effectiveness analysis for the subjects of reading and mathematics, performed for the selected variables in each step, are discussed below.

First step: Efficiency analysis

Table 1 presents the results for the estimates of teacher efficiency in reading and mathematics. These efficiency estimations have been made by considering students’ achievement in reading and mathematics as dependent variables.

<Insert Table 1>

Focusing on human resources –i.e. students– teachers work with, we make use of the Home Resources for Learning index, as this variable is a good “proxy” of students’ background characteristics –which resulted significant in the analysis carried out by Foy and O’Dwyer

(2013) and Grilli et al. (2014)–. It was found that those teachers who have students from high socio-economic backgrounds are able to obtain better results from them in reading and mathematics. This positive relationship between families’ socio-economic background and students’ achievement was also found for Spain by authors such as Calero, Choy, and Waisgrais (2010). The fact that students present diglossia is detrimental for students’ achievement in both subjects –students who present it perform 4% and 3.3% lower in reading and mathematics, respectively, than those who do not–. Carabaña (2013) also found –for PISA 2012 Spanish students– that diglossia decreases students’ achievement and, thus, the probability of grade repetition increases due to its correlation with students’ immigrant status. Additionally, this variable could be reflecting side effects of co-official languages in the teaching-learning process.

In the case of material resources, when the lack of an adequate provision of materials is not a serious problem, it may entail higher achievement in reading for students taught by the teacher –approximately 6%–. Better school facilities and conditions are highlighted in the literature as essential in order to make teachers’ tasks easier, helping them not only improve their efficiency when performing their activities, but also their retention in the school (Leung, Chan, & Wang, 2006; Schneider, 2002). However, this lack of materials does not seem to affect scores in mathematics.

Our results show that when overcrowded classrooms are not a serious problem they may have a positive effect –4% on average– on scores in mathematics⁵. Blatchford and Martin (1998) found that primary school students in the United Kingdom received better support in smaller classes than in overcrowded ones, and that teacher-pupil interactions were higher in smaller ones. Finally, for time resources, students’ achievement in reading is increased 2.3% when teachers teach between 4 and 5 hours in a week –compared to 4 hours–, although a higher number of hours does not seem to have a significant effect.

To sum up, the behavior of these variables shows that a favorable environment for teaching may help teachers obtain better results from their students.

Second step: Effectiveness analysis

Once an efficiency score prediction for reading and mathematics teachers is obtained from the efficiency estimates in Table 1, this efficiency measure is included as a regressor in the estimations presented in Table 2, where teachers’ effectiveness, proxied by students’ engagement in reading and mathematics lessons, is analyzed. The results of the ICCs and the rejection of the null hypothesis of the LR tests –namely that a one-level regression model would be better than a multilevel model– presented in Table 2 denote that students taught by the same teacher are more alike –so there exist differences between teachers–, which supports the use of multilevel analysis. In addition, the Snijders/Bosker R^2 presented by both models showed their high explicative capacity at student and teacher levels.

<Insert Table 2>

Firstly, teachers’ general characteristics and those not related to their teaching methodologies are analyzed. In the case of reading, the fact that students are girls may suppose an increase in their engagement compared to boys, which could also be reflected in that female teachers are able to make their students more engaged in their reading lessons than male teachers, maybe due to the transmission of their enjoyment of this subject to their students. However, there are authors who have obtained results showing that sex was not a factor which could condition teachers’ effectiveness (Slater, Davies, & Burgess, 2012, for the case of secondary schools in the United Kingdom). Furthermore, teachers’ years of experience have

⁵ A test of the equality of coefficients for the categories “Not a problem”, “Minor problem” and “Moderate problem” for each of the variables “Classrooms are overcrowded” and “Teachers do not have adequate instructional materials and supplies” has been performed, respectively. In both the null hypothesis of the equality of coefficients has been accepted, so the three categories of each variable show the same effect on academic achievement, compared to the reference category “Serious problem”.

appeared to be positive to engage students in lessons. Wolters and Daugherty (2007) indicated that more experienced teachers are more effective instructors than those with less experience and Bandura (1997) remarked that this could be due to that these teachers have faced more challenging situations than others with less experience. In contrast, Klassen and Chiu (2010) found that teachers' years of experience increased their effectiveness from early career until it starts falling afterwards, denoting a nonlinear behavior. In our research this non-linear relationship of students' engagement index and years of experience was checked, finding that it is not accomplished. In the case of efficiency, an increase in teachers' efficiency score⁶ of one normalized point implies a higher index of teacher engagement of approximately 0.8 in both subjects. These variables are often accompanied by others which account for teacher's highest level of education. Nevertheless, in the case of our data, 153 out of 156 reading teachers (154 of 158 for mathematics teachers) had ISCED level 5a (first), and only 3 (4 for mathematics) had ISCED level 5a (second), making it irrelevant to control for this characteristic.

Dealing with the effectiveness of teachers' methodologies, the discussion with other teachers on how to teach a concept in reading –only for daily or almost daily frequencies– and mathematics is positive for students' engagement. The higher effect in mathematics may be explained by the higher complexity of this subject, which could require further discussion and greater teacher consensus on how to approach a certain topic. Grossman, Wineburg, and Woolworth (2001) stated that this kind of teacher collaboration foments the intellectual renewal of the school, which in turn translates into learning benefits for students. In the case of daily or almost daily visits to other teachers' classrooms in mathematics, it has a negative effect on students' engagement, which could be due to the higher propensity of students in these classes to become distracted. However, this practice of visiting other teachers' classrooms daily has been highlighted in the literature as being positive for student learning by authors such as Fullan (2002) –who named this practice “intervisitation”–.

The frequent discussion in class of students' homework has a positive effect on engagement in reading, as reading activities may have different interpretations depending on the reader's point of view, while mathematics, as a deterministic subject, does not need a deep discussion of its results. Related to this issue, Muijs and Reynolds (2000) found –for students of the United Kingdom– that time spent teaching interactively instead of using seat work or group work contributes to increasing teacher effectiveness. This interactive approach also includes engaging with students and asking them open questions, allowing multiple answers, etc. (Gagne, Yekovich, & Yekovich, 1993). Furthermore, the control of homework by teachers in reading is also positive for student engagement –increasing engagement by 0.659–. In relation to this practice, Bonesrønning (2004) found that, in Norway, monitoring students' activities has a positive influence on their learning, while summarizing the content and repeating explanations has a positive and negative influence, respectively.

Nowadays, teacher effectiveness is usually measured by using the standardized scores obtained by students in regional or national achievement tests. The fact that teachers' salaries or other working conditions may be conditioned by these scores can make teachers who are not really effective –and thus, do not have proper methods to increase students' effort and learning– to be incorrectly classified as such (Rothstein et al., 2010). Because of that, we have also controlled for the approach of “teaching to the test”, which has been extended currently due to the increasing importance of international assessment tests in the so-called “teacher rankings” (Jürges & Schneider, 2007). However, our results have shown that some emphasis of teachers in national achievement tests increases their students' engagement in mathematics –by 0.286 points–, as well as the frequent control of students with a written test or quiz –by 0.388 for every or almost every lesson and 0.448 for about half the lessons–, which could increase the

⁶ These efficiency scores have been re-scaled following the formula for a “j” efficiency score –for teacher “j”– in subject c ($c = 1$ for reading and $c = 2$ for mathematics): $EfficiencyScore_{cj} - MinimumEfficiencyScore_c / (MaximumEfficiencyScore_c - MinimumEfficiencyScore_c)$. This procedure was employed in order to make the coefficients easily interpretable, as the range of the efficiency scores in reading is 83.62 – 99.24, while for mathematics it is 82.21 – 98.85.

interest of students in paying attention in class, as mathematics may require a greater effort by students during lessons in order to understand certain concepts. The latter practice also enhances academic achievement in mathematics, as indicated by authors such as Shirvani (2009).

In the case of the methodologies which are more related to a particular subject –which is why we only have information about them for one subject– it appears that when students are not used to working on problems while their teacher is doing other tasks, they may lose interest in the subject –which only reduces engagement in mathematics by 0.281 when done in some lessons–. This has been highlighted by Muijs and Reynolds (2000), who stated that teachers should actively monitor students’ individual and group work, being approachable and going around the classroom, instead of sitting at his/her desk. In the reading subject the practice of making students read aloud to the class has shown a negative effect on engagement, which could be due to a lack of interest of the other students who are not following the reading. According to Tindal, Heath, Hollenbeck, Almond, and Harniss (1998), reading aloud does not permit self-pacing for students, which undermines students’ attention and, then, engagement. Furthermore, Elbaum, Arguelles, Campbell, and Saleh (2004) found –for the United States– that reading out loud increases retention in working memory, but that it slows down reading speed.

Figure 1 and Figure 2 show caterpillar plots of teachers’ effectiveness scores in reading and mathematics –sorted in an ascending trend from the left, representing a “ranking” of effectiveness–, whose confidence intervals have been calculated at a level of 95%. As can be seen, most confidence intervals of effectiveness scores in the middle of the distribution overlap with the horizontal line of the average of BLUPs⁷ –showing the existence of little difference between teachers who obtained an “average” score–. There are 14 (9%) reading teachers and 10 (6%) mathematics teachers who perform significantly lower than the average, so we denote them as “bottom teachers”. However, there are 13 (8%) reading teachers and 6 (4%) mathematics teachers who perform over the average, so they are called “top teachers”. The procedure employed for this analysis assures that the effectiveness values obtained for these two groups do not place a “bottom teacher” as a “top teacher” –as Arpino and Varriable (2010) remarked–. In addition, the predicted engagement that teachers achieve shows a clear higher mean for “top teachers” –10.54 for reading and 10.35 for mathematics– than for “bottom teachers” –9.06 for reading teachers and 8.95 for mathematics teachers–. Hence, the structure of these figures may indicate that effective and non-effective teachers present differences in the way they engage their students in lessons and, thus, the work of those teachers who are more effective should be recognized.

Figure 1. Caterpillar plots of teacher effectiveness scores in reading

<Insert Figure 1>

Note: “Bottom teacher” positions are located on the left side of the figure, while “top teacher” positions are on the right side.

Source: Authors’ own calculations from PIRLS and TIMSS 2011 data.

Figure 2. Caterpillar plots of teacher effectiveness scores in mathematics

<Insert Figure 2>

Note: “Bottom teacher” positions are located on the left side of the figure, while “top teacher” positions are on the right side.

Source: Authors’ own calculations from PIRLS and TIMSS 2011 data.

⁷ This “overlapping” can be easily observed if we focus on the horizontal line departing from the “0” value –the average of all BLUP scores–. Those teachers placed in the bottom positions of the ranking –left side of the figures– and whose upper bounds of their confidence intervals do not touch the average line can be considered as significantly different from “average teachers”, so they are denoted as “bottom teachers”. Similarly, those teachers in the high ranking positions –right side of the figures– whose lower bounds of their confidence intervals do not touch the average line can be considered as significantly different from “average teachers”, and hence as “top teachers”.

Conclusions

The purpose of this research was to delimit, both theoretically and empirically, the factors, teaching methodologies and efficient practices affecting teachers' effectiveness, focusing on Spanish students. In order to achieve this aim, a novel combination of techniques to measure teachers' efficiency and effectiveness was employed.

In the case of teachers' efficiency, a relevant conclusion is that the proper use of the available resources that this concept entails is essential to engage students in lessons and, thus, increase teachers' effectiveness. In this sense, the availability of sufficient teaching materials may be necessary for the proper development of lessons, so direct communication regarding this subject between head teachers and teachers should be achieved. Overcrowded classrooms seem to be a problem in mathematics lessons, so the number of students per class should be controlled and the alternative of splitting classrooms into groups may be advisable. Finally, an increase in the number of teaching hours may be beneficial for students, up to a certain threshold –5 hours a week–, a new practice which has been employed in Spain with the application of the new education act –LOMCE (*Ley Orgánica de Mejora de la Calidad Educativa*)–.

From our results on teachers' effectiveness it can be inferred that interactions between teachers on how to explain a concept in a lesson could be positive, while visiting other teachers' classrooms is not. This might mean that teachers should manage the contents of their lessons with the help and participation of other teachers, but from the moment the class is in the hands of the teacher in charge of the classroom they should take charge of their students without the intervention of another teacher. Conclusions on homework are also relevant, as results may indicate that effective reading teachers usually monitor the completion of their students' homework and discuss its content in order to provide clarifications. These actions, together with making students take a written quiz in mathematics every or almost every lesson and putting some emphasis on national achievement tests should be included in teachers' curriculum, as they have been found to be effective teaching practices.

Furthermore, making students work on problems on their own without paying enough attention or not guiding them during the process could be detrimental for their engagement, as students may lose interest in the task. In addition, reading aloud to the class could provoke a decrease in students' attention to the reading, as they are not able to read at their own pace and understand the content and sense of the text.

As efficiency is integrated in the concept of effectiveness, putting more stress on making teachers manage their resources in a proper way would also help them to make their students become more interested and exert an effort in learning.

In conclusion, this research has provided an insight into the determination of the characteristics, teaching procedures and efficient use of resources which can make a teacher effective. To the extent that these effective teachers are able to improve their students' engagement, thanks to their effective methods and independently of their students' socio-economic backgrounds, they are fostering the improvement of social mobility.

References

- Afshartous, D., & Wolf, M. (2007). Avoiding 'data snooping' in multilevel and mixed effects models. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170(4), 1035–1059. <http://dx.doi.org/10.1111/j.1467-985X.2007.00494.x>
- Arpino, B., & Varriale, R. (2010). Assessing the quality of institutions' rankings obtained through multilevel linear regression models. *Journal of Applied Economic Sciences*, 11(1), 7–22. Retrieved from http://cesmaa.eu/journals/jaes/files/VolumeV_Issue1_11__Spring2010_fulltext.pdf
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Battese, G. E., & Coelli, T. J. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics*, 38(3), 387–399. [http://dx.doi.org/10.1016/0304-4076\(88\)90053-X](http://dx.doi.org/10.1016/0304-4076(88)90053-X)
- Blatchford, P., & Martin, C. (1998). The effects of class size on classroom processes: 'It's a bit like a treadmill—working hard and getting nowhere fast'. *British Journal of Educational Studies*, 46(2), 118–137. <http://dx.doi.org/10.1111/1467-8527.00074>
- Bonesrønning, H. (2004). Can effective teacher behavior be identified? *Economics of Education Review*, 23(3), 237–247. <http://dx.doi.org/10.1016/j.econedurev.2003.07.002>
- Borich, G. (1996). *Effective teaching methods*. New York: Macmillan.
- Braga, M., Paccagnella, M., & Pellizzari, M. (2014). Evaluating students' evaluations of professors. *Economics of Education Review*, 41, 71–88. <http://dx.doi.org/10.1016/j.econedurev.2014.04.002>
- Brophy, J. E., & Good, T. L. (1986). Teacher behaviour and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 328–375). New York: MacMillan.
- Calero, J., Choy, A., & Waisgrais, S. (2010). Determinantes del riesgo de fracaso escolar en España: una aproximación a través de un análisis logístico multinivel aplicado a PISA-2006. *Revista de Educación, Extraordinary*, 225–256. Retrieved from http://www.revistaeducacion.educacion.es/re2010/re2010_09.pdf
- Carabaña, J. (2013). Repetición de curso y puntuaciones pisa ¿cuál causa cuál? In INEE (Ed.), *PISA 2012: Programa para la evaluación internacional de los alumnos. Informe español. Volumen II: Análisis secundario* (pp. 32–66). Madrid: Ministerio de Educación, Cultura y Deporte, Instituto Nacional de Evaluación Educativa.
- Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the Linkages. *Research in Higher Education*, 47(1), 1–32. <http://dx.doi.org/10.1007/s11162-005-8150-9>
- Chakraborty, K., Biswas, B., & Lewis, W. C. (2001). Measurement of Technical Efficiency in Public Education: A Stochastic and Nonstochastic Production Function Approach. *Southern Economic Journal*, 67(4), 889–905. <http://dx.doi.org/10.2307/1061576>
- Creemers, B. P. M. (1994). *The effective classroom*. London: Cassell.
- Creemers, B. P. M., & Reezigt, G. J. (1996). School level conditions affecting the effectiveness of instruction. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, 7(3), 197–228. <http://dx.doi.org/10.1080/0924345960070301>
- Elbaum, B., Arguelles, M. E., Campbell, Y., & Saleh, M. B. (2004). Effects of a Student-Reads-Aloud Accommodation on the Performance of Students With and Without Learning Disabilities on a Test of Reading Comprehension. *Exceptionality: A Special Education Journal*, 12(2), 71–87. http://dx.doi.org/10.1207/s15327035ex1202_2

- Foy, P., & O'Dwyer, L. M. (2013). *School Effectiveness Models and Analyses*. In M. O. Martin, & I. V. S. Mullis (Eds.), *TIMSS and PIRLS 2011 Relationships report* (Technical Appendix B). United States: Boston College.
- Fullan, M. (2002). The Role of Leadership in the Promotion of Knowledge Management in Schools. *Teachers and Teaching: theory and practice*, 8(3), 409–419.
<http://dx.doi.org/10.1080/135406002100000530>
- Gagne, R. M., Yekovich, C. W., & Yekovich, F. R. (1993). *The cognitive psychology of school learning*. New York: HarperCollins.
- Goldstein, H., & Healy, M. J. R. (1995). The graphical presentation of a collection of means. *Journal of the Royal Statistical Society: Series A (Statistics and Society)*, 158(1), 175–177.
<http://dx.doi.org/10.2307/2983411>
- Grilli, L., Pennoni, F., Rampichini, C., & Romero, I. (2014). *Exploiting TIMSS and PIRLS combined data: multivariate multilevel modelling of student achievement*. Paper presented at the VI European Congress of Methodology of the European Association of Methodology, Utrecht, Netherlands.
- Grossman, P., Wineburg, S., & Woolworth, S. (2001). Toward a theory of teacher community. *Teachers College record*, 103(6), 942–1012. Retrieved from
<http://www.tcrecord.org/content.asp?contentid=10833>
- Jürges, H., & Schneider, K. (2007). Fair ranking of teachers. *Empirical Economics*, 32(2), 411–431. <http://dx.doi.org/10.1007/s00181-006-0112-3>
- Klassen, R. M., & Chiu, M. M. (2010). Effects on Teachers' Self-Efficacy and Job Satisfaction: Teacher Gender, Years of Experience, and Job Stress. *Journal of Educational Psychology*, 102(3), 741–756. <http://dx.doi.org/10.1037/a0019237>
- Kyriakides, L. (2008). Testing the validity of the comprehensive model of educational effectiveness: a step towards the development of a dynamic model of effectiveness. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, 19(4), 429–446. <http://dx.doi.org/10.1080/09243450802535208>
- Leckie, G., & Goldstein, H. (2009). The limitations of using school league tables to inform school choice. *Journal of the Royal Statistical Society: Series A (Statistics and Society)*, 172(4), 835–851. <http://dx.doi.org/10.1111/j.1467-985X.2009.00597.x>
- Lenkeit, J. (2013). Effectiveness measures for cross-sectional studies: a comparison of value-added models and contextualized attainment models. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, 24(1), 1–25.
<http://dx.doi.org/10.1080/09243453.2012.680892>
- Leung, M., Chan, J. K. W., & Wang, Z. (2006). Impact of school facilities on working behavior of teachers. *International Journal of Strategic Property Management*, 10(2), 79–91. Retrieved from <http://dx.doi.org/10.1080/1648715X.2006.9637546>
- Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, 19(6), 527–537.
<http://dx.doi.org/10.1016/j.learninstruc.2008.11.001>
- Lockheed, M., & Hanushek, E. (1994). Concepts of Educational Efficiency and Effectiveness. In R. Husén, & T. N. Postlethwaite (Eds.), *International Encyclopedia of Education* (pp. 1779–1784). Oxford: Pergamon.
- McEwan, P. J., & Carnoy, M. (2000). The Effectiveness and Efficiency of Private Schools in Chile's Voucher System. *Educational Evaluation and Policy Analysis*, 22(3), 213–239.
<http://dx.doi.org/10.3102/01623737022003213>

- McLaughlin, M., McGrath, D. J., Burian-Fitzgerald, M. A., Lanahan, L., Scotchmer, M., Enyeart, C., & Salganik, L. (2005). *Student content engagement as a construct for the measurement of effective classroom instruction and teacher knowledge*. Paper presented at the annual meeting of the American Educational Researchers Association, Montreal, Canada.
- Muijs, D., Kyriakides, L., Van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art - teacher effectiveness and professional learning. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, 25(2), 231–256. <http://dx.doi.org/10.1080/09243453.2014.885451>
- Muijs, D., & Reynolds, D. (2000). School Effectiveness and Teacher Effectiveness in Mathematics: Some Preliminary Findings from the Evaluation of the Mathematics Enhancement Programme (Primary). *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, 11(3), 273–303. [http://dx.doi.org/10.1076/0924-3453\(200009\)11:3;1-G;FT273](http://dx.doi.org/10.1076/0924-3453(200009)11:3;1-G;FT273)
- Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*. Chestnut Hill, Massachusetts: TIMSS & PIRLS International Study Center, Boston College.
- Raudenbush, S. W., & Willms, J. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20(4), 307–335. <http://dx.doi.org/10.3102/10769986020004307>
- Reynolds, D., Sammons, P., Stoll, L., Barber, M., & Hillman, J. (1996). School effectiveness and school improvement in the United Kingdom. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, 7(2), 133–158. <http://dx.doi.org/10.1080/0924345960070203>
- Rothstein, R., Ladd, H. F., Ravitch, D., Baker, E. L., Barton, P. E., Darling-Hammond, L.,... Shepard, L. A. (2010). *Problems with the use of student test scores to evaluate teachers*. Washington, DC: Economic Policy Institute.
- Ryan, K., & Cooper, J. (2007). *Those who can, teach*. New York: Houghton Mifflin.
- Schneider, M. (2002). *Do school facilities affect academic outcomes?* Washington, DC: National Clearinghouse for Educational Facilities.
- Shirvani, H. (2009). Examining an Assessment Strategy on High School Mathematics Achievement: Daily Quizzes Vs. Weekly Tests. *American Secondary Education*, 38(1), 34–45. Retrieved from http://www.jstor.org/stable/41406065?seq=1#page_scan_tab_contents
- Slater, H., Davies, N. M., & Burgess, S. (2012). Do Teachers Matter? Measuring the Variation in Teacher Effectiveness in England. *Oxford Bulletin of Economics and Statistics*, 74(5), 629–645. <http://dx.doi.org/10.1111/j.1468-0084.2011.00666.x>
- Taylor, B. M., Pearson, P. D., Clark, K., & Walpole, S. (1999). *Beating the odds in teaching all children to read: Lessons from effective schools and exemplary teachers*. Michigan: Center for the Improvement of Early Reading Achievement.
- Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition without control: When a little frontal lobe goes a long way. *Current directions in Psychological Science*, 18(5), 259–263. <http://dx.doi.org/10.1111/j.1467-8721.2009.01648.x>
- Timothy R. B., Rupley, W. H., & Nichols, W. D. (2007). The Effective Teacher of Reading: Considering the ‘What’ and ‘How’ of Instruction. *The Reading Teacher*, 60(5), 432–438. <http://dx.doi.org/10.1598/RT.60.5.3>
- Tindal, G., Heath, B., Hollenbeck, K., Almond, P., & Harniss, M. (1998). Accommodating students with disabilities on large-scale tests: An experimental study. *Exceptional Children*, 64(4), 439–450. <http://dx.doi.org/10.1177/001440299806400401>

- Willms, J. D. (2003). *Student engagement at school. A sense of belonging and participation. Results from PISA 2000*. OECD Publishing. <http://dx.doi.org/10.1787/9789264018938-en>
- Wolters, C. A., & Daugherty, S. G. (2007). Goal Structures and Teachers' Sense of Efficacy: Their Relation and Association to Teaching Experience and Academic Level. *Journal of Educational Psychology*, 99(1), 181–193. <http://dx.doi.org/10.1037/0022-0663.99.1.181>
- Zuzovsky, R. (2009). Teachers' qualifications and their impact on student achievement: Findings from TIMSS 2003 data for Israel. In M. von Davier, & D. Hastedt (Eds.), *IERI Monograph Series Issues and Methodologies, Vol. 2* (pp. 37–62). Hamburg/Princeton, Germany/USA: IERI Institute.

Table 1. Teacher efficiency estimates for reading and mathematics

Variables	Ln(Reading Scores)	Ln(Mathematics Scores)
Student level variables		
Ln (Home Resources for Learning index)	0.223*** (0.012)	0.282*** (0.013)
Student with diglossia (Reference: Student without diglossia)	-0.040*** (0.007)	-0.033*** (0.007)
Teacher level variables		
Classrooms are overcrowded (Reference: Serious problem)		
Not a problem	0.002 (0.012)	0.040*** (0.015)
Minor problem	-0.005 (0.012)	0.034* (0.018)
Moderate problem	0.007 (0.013)	0.038** (0.015)
Teachers do not have adequate instructional materials and supplies (Reference: Serious problem)		
Not a problem	0.062** (0.024)	0.058 (0.041)
Minor problem	0.067*** (0.024)	0.048 (0.042)
Moderate problem	0.061** (0.024)	0.061 (0.041)
Minutes per week teaching reading/mathematics (Reference: 240 or less)		
More than 240 to 300	0.023*** (0.009)	0.009 (0.009)
More than 300	0.013 (0.008)	0.012 (0.013)
Constant	5.706*** (0.039)	5.526*** (0.058)
Observations: Students	3,438	3,432
Groups: Teachers	184	184
Wald test	416.61***	517.19***

Source: Authors' own calculations from PIRLS and TIMSS 2011 data.

Dependent variable: Ln(Students achievement in reading/mathematics in PIRLS/TIMSS). Range of the variables: Ln(Reading Scores): 5.72 – 6.53 (Original variable range: 304 – 687); Ln(Mathematics Scores): 5.50 – 6.52 (Original variable range: 245 – 677).

Estimation method: Time-invariant stochastic frontier. The efficiency term is supposed to follow a truncated normal distribution.

Standard errors in parentheses.

*** denotes variable significant to level 1%; ** to 5%; * to 10%.

Table 2. Teacher effectiveness estimates for reading and mathematics

Variables	Engagement in Reading	Engagement in Mathematics
Student level variables		
Female student (Reference: Male student)	0.431*** (0.066)	0.038 (0.068)
Teacher level variables		
Female teacher (Reference: Male teacher)	0.281* (0.149)	0.097 (0.138)
Years of experience	0.011** (0.005)	0.011** (0.005)
Efficiency	0.809*** (0.290)	0.766*** (0.275)
Discussions with other teachers about how to teach a concept (Reference: Never)		
Daily or almost daily	0.502** (0.246)	1.002*** (0.214)
1 to 3 times a week	0.106 (0.221)	0.460** (0.184)
2 or 3 times a month	0.131 (0.218)	0.546*** (0.180)
Visit another teacher's classroom (Reference: Never)		
Daily or almost daily	-0.589 (0.572)	-1.571*** (0.454)
1 to 3 times a week	-0.291 (0.203)	-0.216 (0.182)
2 or 3 times a month	0.020 (0.159)	0.018 (0.150)
Monitor whether or not the homework was completed: Always or almost always (Reference: Sometimes, Never or almost never)		
	0.659** (0.282)	-0.333 (0.258)
Discuss the homework in class (Reference: In reading: Never or almost never; In mathematics: Sometimes, Never or almost never)		
Always or almost always	1.095** (0.436)	0.064 (0.135)
Sometimes	1.304*** (0.435)	- -
National achievement tests (Reference: Little or no emphasis)		
Major emphasis	0.275 (0.233)	-0.134 (0.195)
Some emphasis	0.118 (0.122)	0.286** (0.115)
Take a written test or quiz (Reference: In reading: Never; In mathematics: Some lessons, Never)		
Every or almost every lesson	0.147 (0.418)	0.388* (0.222)
About half the lessons	-0.161 (0.255)	0.448** (0.186)
Some lessons	0.042 (0.243)	- -
Work on problems (individually or with peers) while I am occupied with other tasks (Reference: Never)		
Every or almost every lesson	- -	-0.008 (0.164)
About half the lessons	- -	-0.099 (0.171)
Some lessons	- -	-0.281**

	-	(0.137)
Read aloud to the class (Reference: Once or twice a month, Never or almost never)		
Every or almost every day	-0.675*** (0.252)	-
Once or twice a week	-0.647** (0.282)	-
Constant	7.289*** (0.570)	9.027*** (0.371)
Observations: Students	3,191	3,201
Groups: Teachers	156	158
Wald test	105.08***	78.66***
Snijders/Bosker R-squared student level	0.061	0.053
Snijders/Bosker R-squared teacher level	0.274	0.313
ICC	0.086***	0.063***
LR test for linear regression	121.28***	71.65***

Source: Authors' own calculations from PIRLS and TIMSS 2011 data.

Dependent variable: reading/mathematics lessons engagement index –index data from students– in PIRLS/TIMSS. Range of the variables: Engagement in Reading: 2.07 – 14.32; Engagement in Mathematics: 2.68 – 13.27.

Estimation method: Multilevel regression.

Standard errors in parentheses.

*** denotes variable significant to level 1%; ** to 5%; * to 10%.

Figure captions

Figure 1. Caterpillar plots of teacher effectiveness scores in reading

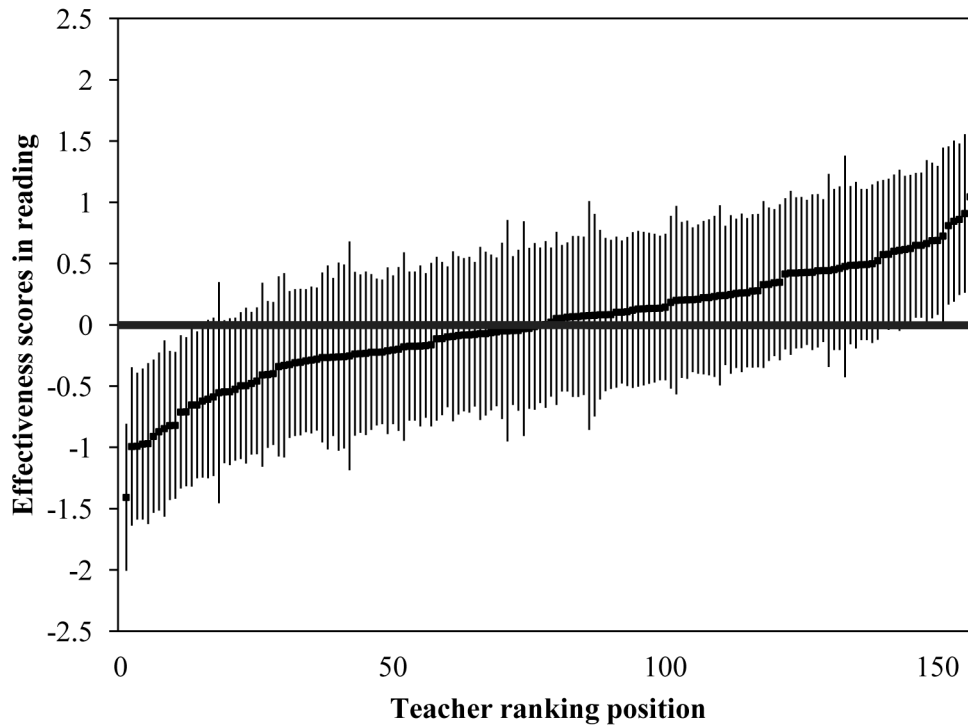


Figure 2. Caterpillar plots of teacher effectiveness scores in mathematics

