Education Policy Analysis
Archives/Archivos Analíticos de Políticas
Educativas
ISSN: 1068-2341
epaa@alperin.ca
Arizona State University
Estados Unidos

Azevedo Leitão, Ulisses
Social Index of Educational Effectiveness: A New Approach from the Perspective of Promoting Equity
Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas, vol. 23, 2015, pp. 1-26
Arizona State University
Arizona, Estados Unidos

education policy analysis archives

A peer-reviewed, independent,
open access, multilingual journal

Arizona State University

Volume 23 Number 50 May 4th, 2015 ISSN 1068-2341

Social Index of Educational Effectiveness: A New Approach from the Perspective of Promoting Equity

Ulisses Azevedo Leitão
Universidade Federal de Lavras – UFLA
Brazil

Citation: Leitão, U. A., (2015). Social index of educational effectiveness: A new approach from the perspective of promoting equity. *Education Policy Analysis Archives, 23*(50).
http://dx.doi.org/10.14507/epaa.v23.1915

Abstract: In this article, I propose an index, called the Social Index of Educational Effectiveness (SIEE), which allows the establishment of an objective criterion to define the school's profile concerning the promotion of educational equity. It makes it possible to differentiate schools with an "inclusionary profile," (SIEE>0), from those with an "exclusionary profile," (SIEE<0), depending on whether or not the school fosters educational inclusion, and therefore promoting social equity. I have analyzed 272 schools (N=11,732 students) of five urban contexts in Brazil and found behaviors markedly different for the two learning subjects; reading and mathematics, the mean SIEE yielding positive and negative values, respectively. I argue that the SIEE concept is relevant for the analysis of educational effectiveness and should be considered as such by educational researchers and policymakers.
Keywords: large-scale assessment; educational effectiveness; educational accountability policy; efficiency and equity in education; value-added models.

Índice Social de la Eficacia de la Educación: Un Nuevo Enfoque Desde la Perspectiva de Promover la Equidad
Resumen: En este artículo se propone un índice, llamado el Índice Social de la Eficacia de la Educación (SIEE), que permite el establecimiento de un criterio objetivo para definir el perfil de una

Journal website: http://epaa.asu.edu/ojs/
Facebook: /EPAAA
Twitter: @epaa_aape

Manuscript received: 11/24/2014
Revisions received: 4/7/2015
Accepted: 4/7/2015

escuela en relación a la promoción de la equidad educativa. Esto hace que sea posible diferenciar las escuelas con un "perfil de inclusión", (SIEE>0), de los que tienen un "perfil de exclusión", (SIEE <0), dependiendo de si la escuela fomenta la inclusión educativa y, por supuesto, la promoción de la equidad social. Se analizaron 272 escuelas (N = 11.732 estudiantes) de cinco contextos urbanos en el Brasil y encontramos comportamientos marcadamente diferentes de las dúas disciplinas de aprendizaje analizadas, la lectura (Portugués) y la matemática. Se encontró la media SIEE positiva y negativa, respectivamente. Se argumenta que el concepto del SIEE es relevante para el análisis de la eficacia educativa y debe ser considerado por los investigadores educativos y responsables políticos.
Palabras-clave: evaluación a gran escala; efectividad de la educación; política de responsabilización educativa; eficacia educativa y equidad en la educación; modelos de valor agregado.

Índice Social de Eficácia Educacional: Uma Nova Abordagem a Partir da Perspectiva da Promoção da Equidade

Resumo: Neste artigo, propomos um índice, denominado de Índice Social de Eficácia Educacional (SIEE), que permite o estabelecimento de um critério objetivo para definir o perfil da escola em relação à promoção da equidade educacional. Isso torna possível diferenciar as escolas com um "perfil de inclusão", (SIEE> 0), de escolas com um "perfil de exclusão", (SIEE <0), dependendo se a escola promove ou não a inclusão educacional, promovendo, portanto, a equidade social. Foram analisadas 272 escolas de ensino fundamental (N = 11.732 alunos) de cinco contextos urbanos no Brasil. Encontrou-se diferentes comportamentos para as duas disciplinas de aprendizagem analisadas, português e matemática, a média SIEE estimada é positiva e negativa, respectivamente. Argumenta-se que o conceito do SIEE é relevante para a análise da eficácia educacional e deve ser considerado por pesquisadores e formuladores de políticas educacionais.
Palavras-chave: avaliação em larga escala; eficácia educacional; política de responsabilização educacional; eficácia educacional e equidade na educação; modelos de valor agregado.

Introduction

How could education researchers, policymakers, and practitioners verify whether a school is effective in fostering social inclusion and educational equity? The question is relevant and of great interest, especially for countries seeking to implement education policies focused on social development.

Recent initiatives, like Every Child Matters in England, and No Child Left Behind and Race to the Top in the United States suggest a new political commitment in these countries to promote quality education with increasing equity. These initiatives recognize the need for greater investments and better educational strategies, aimed at growing the life chances of groups in situations of social vulnerability.

In Brazil, the National Education Plan (PNE), in particular the current National Education Plan PNE-2011/2020, PL No. 8.035/2010, has established explicit targets to overcome educational inequalities and improve the quality of teaching. The assistance of education to reduce social inequality is generally recognized as intrinsically related to its ability to induce students in social vulnerability situations to advance more than the average student in their learning, aiming to enable equal opportunities for their professional future.

In order to answer the question above, in this article I propose an index, called the Social Index of Educational Effectiveness (SIEE). It is extremely simple to calculate, since it is not limited to a specific statistical modeling, but rather it directly manipulates raw data. It is also easy to interpret and allows the establishment of an objective criterion to define the profile of a school in relation to

the promotion of social equity. The SIEE methodology was then applied in the analysis of the database relative to the Geres Project – a longitudinal study conducted by a consortium of research groups from six Brazilian universities that monitored the proficiency in both subjects; reading (Portuguese) and mathematics of up to 21.529 students, from 303 schools in five cities, between the years 2005 and 2008. In order to avoid missing data, in this study I have focused on 272 schools and 11732 students that participated in the first and fourth assessment cycle. The sample data studied allows the comparison of the growth of learning by measuring the student's achievements in the first four years of elementary education for both the public (local, state and special schools) and the private school system. In spite of the simplicity of the conceptual construction of the SIEE, the proposed analysis enables the identification of important features of the school. In particular, the methodology allows the recognition of both, schools whose Educational Effectiveness has an "exclusionary profile", i.e., schools that are more effective with students who have high levels of previous proficiency (SIEE<0), as well as schools with an "inclusionary profile" that foster social inclusion, i.e., schools that promote educational equity by being more effective with students with low prior proficiency (SIEE>0).

The analysis has allowed us to infer remarkably different behaviors for the two learning subjects analyzed. While the mean SIEE for reading (Portuguese) was found to be strongly positive, SIEE= +0.364, standard deviation +0.268, the mean SIEE for mathematics, by contrast, was found to be negative, SIEE= -0.226, with a remarkably stronger standard deviation +0.481, within the data studied. I investigated the behavior of five school districts and the different school systems in Brazil- i.e. the public and private school systems. The main result of this study is that, throughout all school systems and in all school districts, the process of literacy related to learning reading skills in the investigated schools indeed reduces the differences in proficiency among the students. By contrast, I found that the teaching of mathematics increases these differences. On analyzing the school systems separately, I found a remarkable failure of the public educational system - both the schools belonging to local or state governments -, in reducing the inequalities of proficiency in mathematics, which accounts for the poor educational quality of the public system in the country. On the other hand, schools belonging to the special group of public schools administrated by the federal government and the private schools are almost neutrals, with a SIEE almost null, indicating that the mathematics teaching does not exceed prior educational inequalities at all. The consequences of these empirical findings on the adoption of accountability policies are presented in the final remarks. I argue that the proposed index is relevant for the analysis of Educational Effectiveness and should be considered by policy makers.

This paper is organized as follows: initially, I carry out a brief historical discussion to justify the proposed methodology. Then, I describe the characteristics of the experimental sample here studied; the concept of Social Index of Educational Effectiveness is operationally defined and some real examples of its calculation and interpretation are presented. Finally, I present the results on the application of the proposed methodology to 272 schools enrolling 11,732 students located in five different urban contexts in Brazil and its consequences are analyzed in the concluding remarks.

Research on Educational Effectiveness

Educational Effectiveness Research (EER) has defined effective schools (Sammons, 2007) as schools in which the performance of student´s progress is higher than expected, given the socio-economic and cultural conditions of the family and community context of the student and school. EER has come a long way to arrive at this definition, among heated debate especially in the last two decades. There is some consensus on the importance of analyzing the Educational Effectiveness of

teachers and schools, but there also are strong doubts about the stability and reproducibility of the different statistical models used in its estimation. The main objective of EER is to determine how much of the evolution of student learning can be directly attributed to the school or the teacher. The research also intends to identify an effective school by separating the tangle of variables and individual student characteristics – such as their socio-economic and cultural reality, school infrastructure, teachers' quality, curricula and education policies – from the complex relationships that determine the educational effectiveness.

Pioneering studies carried out in the 1960s and 1970s, with seminal impact on this research subject. Notably, the Coleman Report (Coleman et al., 1966) and the studies by Jencks et al. (1972) have been systematically interpreted with a pessimistic bias in the literature. Several analysts point out that these studies failed to identify a significant relevance of the school and the teacher in the development of students; see, for example, the historical evolution of the EER in Creemers, Kyriakides and Sammons (2010) and Campbell, Kyriakides, Muijs, and Robinson (2004).

In a methodological introduction, the Coleman Report highlights that most of the data needed for the description of equal educational opportunities came from intangible variables. Therefore, the study begins with the innovative but still controversial premise that equality of opportunities should be evaluated by equal results in large-scale assessments. Thus, the researchers collected data on educational resources available in schools for different groups of children, but also on the proficiency of students measured by the results of the learning assessment tests. The main question that the report sought to answer was to determine to what extent and how the schools were able to contribute to overcoming social inequalities of enrolled children. The study showed that variations in school quality had little correlation with the levels of Educational Effectiveness, when students of similar social origins were compared between schools. Moreover, the results of learning assessment were correlated to differences in family background among the students.

It is common to emphasize that the Coleman Report (Coleman et al., 1966) concludes that the quality of the school in which the student actually attended has little influence on his/her final proficiency, compared with individual factors, such as IQ, race, and external factors such as socio-economic status (SES). Studies by Jencks et al. (1972) have also been interpreted as pointing out that family and community environments are significantly more important than school. However, much less attention has been given in literature to the original conclusion of Coleman that the quality of the school and the teacher are more critical precisely for children from families of low SES or who are in a position of vulnerability or social disadvantage. Thus, the authors conclude that:

> The average student's achievement seems to be less affected by the strengths and weaknesses of his school's facilities, curriculum and teachers than is the average minority pupils... The inference might then be made that improving the school of a minority pupil may increase his achievement more than would improving the school of a white child increase his...Similarly, the average minority pupil's achievement may suffer more in a school of low quality than might the average white pupil's…This indicates that it is exactly for most disadvantaged children that improvements in school quality will make the most difference in achievement. (Coleman et al., 1966, p. 55.)

The message behind these early studies was to attribute the causes for the raw achievements of students to their SES, and that these accomplishments should not, therefore, be imputed to the school or the teacher. Nevertheless, these studies have highlighted the crucial importance of the school in promoting educational equity, since they show that the most disadvantaged children are more sensitive to the improvements on school quality in general.

Several criticisms were made to the conclusions of these early studies. Further studies were able to identify, in a much more generally persceptive, the contributions of school (School Effect) and teacher (Teacher Effect) on student performance. The first studies to establish the extent to which the school makes a difference in student education were conducted in the U.S. (Edmonds, 1979) and England (Rutter, Maughan, Mortimore, Ouston, & Smith, 1979). In developing the Tennessee Value-Added Assessment System (TVAAS), Sanders and Rivers (1996) were pioneers inthe concept of establishing teacher effect, as well as the evidence that the teacher effect is long-lasting. In these studies, the authors demonstrated that students with lower proficiency are exactly those who benefit most from more effective teachers. Moreover, they showed that the positive influence of more effective teachers lasts beyond the period in which students and teachers interact.

Despite criticism to the method proposed by the authors (McCaffrey, Lockwood, Koretz, & Hamilton, 2003; McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004), there is a consensus that "one of the key findings from decades of educational effectiveness research is the importance of the classroom level as predictor of pupils outcomes" (Muijs et al., 2014, p. 231).

Studies in Michigan and Chicago (Lee & Bryk, 1989), using multilevel statistics, were able to verify that smaller schools with a positive academic climate – defined as the quality of social interactions in the academic community, in particular between teacher and student – could be associated with a lower variability in the student's achievements, thus fostering greater social equity.

Effectiveness is generally interpreted as the influence of the teacher and/or school on the results of student evaluations, generally large-scale assessments. The result is measured directly by the score obtained by the student in the ratings when using classic statistical tests, or proficiency and skill when using the item response theory (Baker & Kim, 2004; DeMars, 2010; Linder & Hambleton, 1997; Verhelst, 2010a, 2010b). In both cases, the value added by the teacher or the school, the teacher/school effect, is a measure of its effectiveness. However, as the teacher acts within the context of a specific school and the interaction between these two instances is extremely complex, following Cremmers et al. (2010), I will be referring to educational effectiveness in order to identify the integrated and summative effect of both school environment and teacher on student's achievement.

Given the fact that the proficiency of the student is strongly correlated with variables external to the school, the researchers started investigating the evolution of learning to address the gain or the value-added by the school in the educational progress of students. In this context, studies and developments of statistical models to describe the added value from the 1990's have sought to establish hierarchical models using multilevel linear regression modeling (Creemers et al., 2010; McCaffrey et al., 2003, 2004; Sammons, 2007). In this approach, structured data in different units (student, teacher, school, and school system, as well as different disciplines) are grouped in a hierarchical structure, allowing the study of the evolution of the student's learning process, even if he changes school, class or even the school system. There are several approaches for value-added analysis. For a review of different models and the establishment of a generalized one that presents the most important features of the models used in the literature, I refer to McCaffrey et al. (2003, 2004).

The methodological strategy of tracking the value-added by the difference or gain in proficiency and not directly by the raw value of proficiency[1] aims to subtract the factors that are

[1] Unfortunately, despite that this is a consensus that has lasted decades, several initiatives of educational evaluation policies in Brazil are focused on the estimation of student's raw outcomes, without adjustment for SES or any other factor, rather than on standard value-added models. An example is the disclosure of the Índice de Desenvolvimento da Educação Básica (Basic Education Development Index) - IDEB of public

external to the school. In particular, the empirical findings from a strong correlation between initial students' proficiency and their SES disallows a direct association between educational effectiveness with the average proficiency of the school – measured by the school's student of average achievement in large-scale evaluations. Therefore, the determination of the school's effectiveness should be based on the value-added by the school to the educational process of its students.

However, findings of Ballou, Sanders, and Wright (2004) points out a residual correlation between the growth of student proficiency and their family and community context. This is understandable since schools that operate in a reality of low SES are working in a context in which students have less familial support and less incentive to study, which means that the value-added is also lower, but not due to the school's merit. In this sense, Newton, Darling-Hammond, Haertel, and Thomas (2010) argue that, spite of the fact that the value-added models seek to incorporate the influence of SES on student assessment results, the estimated teacher effect is still strongly influenced by different factors. Moreover, the authors criticize the idea that the coherence between the analysis using different statistical models is poor. The work emphasizes the need to take into consideration the whole school context and advocate an improvement on the large-scale evaluations in order to address properly some latent traits, both things considered crucial for estimating the student's proficiency.

Criticism on the use of value-added models in the evaluation of educational effectiveness has gain a new shape with the publication of Rothstein (2010), who raises concerns about the possibility of bias due to non-random assignment of students to teachers. He proposes a falsification test that seems to indicate that the teacher effect is biased in the standard VAM estimates. His proposal started a strong debate about the reliability of VAM using student achievement gains, classroom observations, and student surveys. There have been legitimate concerns raised about the validity of the measures being used (Baker et al., 2010; Rothstein, 2010).

In order to address these concerns, Kane, McCaffrey, Miller, and Staiger (2013) investigated randomly assignments of teachers with different evaluated effectiveness (with appropriate controls for prior student achievement) to different classrooms, and found that a composite measure of effectiveness can identify teachers who produce higher achievement among their students. They conclude that there are causal impacts, estimated with random assignments that are not well different of the predicted by the traditional non-random ones.

In a independent investigation, Kinsler (2012) has shown that the falsification tests employed by Rothstein are inadequate in small samples. Proposing an alternative testing procedure, the authors conclude that, once one accounts for the "smallness" of the data and allows teacher inputs to persist at reasonable rates, it was impossible to reject conditional random assignment.

Nevertheless, although all these advances, the question about the bias caused by non-random assignment of students and teachers had still no end answer until a very important and strongly cited work by Chetty, Friedman & Rockoff (2014). In an impressive analysis, covering 2.5 million childrens in a 20 years time interval spanning from 1989 to 2009, Chetty et al. concludes that value-added models that control for a student's prior-year test scores can provide unbiased forecasts of teachers' causal impacts on student achievement.

In a symposium on EER (National Research Council, 2010, p. 59), one reports some concern by the fact that the school/teacher effect are modeled as an average effect on different students, although it is plausible that the impact of the school over different students may be substantively different. The example highlighted in the debate is that one teacher (or one school)

---

schools. For more details about IDEB see http://www.publicacoes.inep.gov.br/portal/download/503 (Portuguese).

could work more effectively with students of low proficiency but not arouse much enthusiasm – and therefore not being as effective – with students of higher proficiency, and vice versa. One teacher could motivate students of higher proficiency and not achieve the same efficiency with students of low proficiency. In these cases, the standard value-added models are reducing the current analysis of educational effectiveness to a single numerical factor, failing to describe the full complexity of the reality of the educational process within the school.

In conclusion, the question on the reliability and reproducibility of VAM estimates will be in the center of the debate in the next years, but it seems clear now that to control for student's prior-year test scores is crucial for a better estimate of teacher and school effect on a value-added model.

In this article, I propose an index, called Social Index of Educational Effectiveness (SIEE), extremely simple to calculate and easy to interpret, which yields an objective criterion to define the profile of the school, in the mitigation of educational inequality and promoting social equity. The proposed SIEE methodology was applied to the analysis of the Geres Project Database – a longitudinal study carried out by a consortium of research groups from Brazilian universities that monitored the proficiency in reading and mathematics of up to 21,529 students from 303 schools in five cities between 2005 and 2008. The sample studied allows the comparison of the evolution of learning in the first four grades of primary education in the public (local, state and special group) as well as in the private school systems. In a surprisingly simple way, the proposed methodology allows us to identify schools in which their educational effectiveness has an "exclusionary profile," i.e., schools that are more effective with students who have high levels of previous proficiency, as well as schools characterized by an "inclusionary profile," i.e., schools that promote social equity by being more effective with students with low prior proficiency. The data allow us to infer marked differences on the behavior of both subjects analyzed. This study found, in the schools that were investigated, that the process on literacy and reading skills can be characterized as a strongly "inclusionary profile", whereas the teaching of mathematics is extremely exclusionary relative to the distribution of educational effectiveness. Moreover, the methodology made it possible to identify the mean SIEE profile of different cities and school systems. The consequences of this empirical finding in establishing accountability policies focused on social inclusion actions are presented in the final remarks, which argue that the SIEE index represents a clearly important contribution to the analysis of the educational effectiveness distribution, and should be considered as such by policymakers and practitioners.

Brazilian Educational Context

Brazil experienced very dramatic political and social changes in the 1980's and after. The fall of the military regime was engendered in a context of social effervescence. The new Constitution enacted in 1986 established explicitly that Education is "a right that belongs to everybody; the duty of the State and of families, promoted and stimulated with the cooperation of society, with a view to the full development of the individual for the exercise of citizenship and preparation for work".

A federal level, the Law No. 9.394, of December 20, 1996, named National Education Guidelines and Framework Law (Lei das Diretrizes e Bases da Educação, LDB), provides a range of mechanisms aiming improvements of the educational quality. The access to compulsory basic education is set as a subjective public right, so that any citizen can sue the government to demand it.

In order to fulfill this legal statement, a very big effort was made by the Brazilian government in the 1990's. The LDB created mechanisms for administration, organization and financial support for the public education system. One of the main initiatives was the implementation of the FUNDEF (Fund for Maintenance and Development of Elementary

Education) in 1997, expanded as FUNDEB - The Fund for the Development of Basic Education and Valorization of Education Professionals in 2006, which aims to improve the distribution and use of resources inside each state for primary public education. By law, the states and municipalities have to invest at least 25% of their tax and transfers revenues in the public education system. These resources provide financial support to constitute the FUNDEB/FUNDEF. Therefore, in these past few decades, the country almost universalized (Oliveira, 2007) the access of the entire population in elementary school. Moreover, during the 1990s, a substantive portion of previously excluded students also entered in the school system. At the same time, the success rates in primary education increased noticeably.

For this reason, all children are supposed to attend a public school near to their home. The municipal, state or federal governments, in a cooperative framework, financially and administratively support the public schools. The municipalities are supposed to be responsible mainly for the primary education, the state government for the high school and the federal government for the higher education. Since State and Federal Government can provide primary education, by means of cooperation between the different government spheres, the public schools are classified as federal, state or municipal, depending on their respective administrative responsible. However, since the private schools coexist with the public educational system, a child can apply for a public school in the neighborhood or, by choice, attend to a private school.

Nowadays the great challenge for the public education in Brazil is to raise the education quality. The outcomes of the students of public schools are systematically below of the outcomes of the private schools. In this context, the PNE mentioned before is a attempt to improve quality of education of the public educational system. The low performance of the public schools student's is not such a surprise, in face of the different mean SES of attendance of both school systems. In this context, the present article intend to investigate if the public school system is effective in mitigating the educational inequality caused by the low SES of its attendance.

The research on Large-Scale Assessment has evolved rapidly over the past two decades. In Brazil, the systematic use of the Item Response Theory (IRT) (Baker & Kim, 2004; Verhels, 2010a, 2010b) and improvements on both Evaluation Standard Descriptors Matrices, by establishing of appropriate vertical scales, and item elaboration procedures over the last few years have increased the confidence on large-scale assessments. Researchers have identified (Bonamino & Sousa, 2012) at least three phases of Large-Scale Assessment in the country. If during the first phase, assessments simply had an education diagnostic character, in the following phases, they were gradually associated with political accountability, with symbolic or material consequences for educational agents.

Unfortunately, data from large-scale educational assessment available in Brazil have an eminently cross-sectional design, which is inappropriate to the study of value-added (Bonamino & Oliveira, 2013; Creemer et al., 2010; Franco, Brooke, & Alves, 2008), since they do not allow access to the progress on student performance. Because of this, there is growing interest in research with longitudinal design, in order to carry out estimations for the changes in the performance of the same set of students, from one year to another, and, therefore, the estimation of the growth of student's achievement within the school.

As mentioned before, in the last 25 years, Brazil overcame the challenge of granting universal access to primary education. Currently, the main task is to create the appropriate conditions for permanence and progression of students in the school system, which means creating the conditions for quality education and, at the same time, to promote increasing equity. Nowadays, the need for mechanisms to promote social inclusion requires the development of specific indicators that are sensitive to the role of the school in the process of social inclusion and reducing inequalities. In this context, it is extremely important and worthwhile to define ways to distinguish a school that

promotes social inclusion, by mitigating educational inequality, from a school that produces an exclusionary process among its graduating, by increasing the distance between the more proficient and the less proficient students.

Sample and Data sources

This study is based on the analysis of the Geres Project database, available on CD-ROM in the book organized by Nigel Brooke and Alicia Bonamino (2011). In order to prevent problems with missing data, the present investigation has been restricted to the schools for which at least 10 students per classroom have taken part in both the first and the last evaluations.

The Geres Project is a pioneering study in Brazil, with a longitudinal design, carried out between 2005 and 2008, which drew a picture of the evolution of Portuguese (reading) and mathematics learning of students enrolled in the 1st grade of primary school in 2005. Today, these series correspond to the second to fifth years of elementary education, according to the curricular changes introduced by the Ministry of Education in 2007.

Table 1
Data of the Five Municipalities Where This Study Was Performed

Municipalities	State	Inhabitants (2010)*	Enrolled students (2012)*	HDI (2010)**	Schools	Students
Belo Horizonte	Minas Gerais	2,375,151	309,018	0.810	58	2,874
Campinas	São Paulo	1,080,113	131,873	0.805	56	2,796
Campo Grande	Mato Grosso do Sul	786,797	119,310	0.844	47	1,537
Rio de Janeiro	Rio de Janeiro	6,320,446	803,315	0.845	60	2,677
Salvador	Bahia	2,675,656	304,047	0.750	51	1,848
TOTAL					272	11,732

Note. This table shows the inhabitants, the total number of enrolled students in the fundamental school and the Human Development Index of each city. The last two columns show the number of schools and total of students that have taken part in this investigation, respectively. Sources: * IBGE (2014). ** Atlas Brasil (2013).

Initially the project started with 21,529 students from 303 schools, for a representative sample of five Brazilian cities: Belo Horizonte, Rio de Janeiro, Campinas, Campo Grande and Salvador. The sampling was specified so that the point estimate and the variance of the population's social and educational indicators would be equivalent to the larger universe formed by large municipalities in Brazil. Five "assessment cycles" were performed. The first assessment cycle occurred in March 2005, at the beginning of the school year in Brazil. The following assessment cycles occurred always in November, the end of their school year in Brazil, from 2005 to 2008. Due to several mishaps, the schools of the educational district of Salvador did not take part in the fifth assessment cycle. Therefore, only the 11,732 students that participated in the first and in the fourth assessment cycle in all cities are considered for the present study.

Table 1 lists the municipalities, inhabitants, total of enrolled students in the Fundamental School level and Human Development Index of the five cities where the study was carried out. The last two columns show the number of schools and students that took part in this study.

The Geres Project investigated 17 different data strata, defined by the cities and their school systems: private schools, municipal and state public schools, and even included some special schools connected to federal and state universities in Belo Horizonte and Rio de Janeiro, named special group. The research was carried out by a consortium of research groups with experience in large-scale assessments, linked to different universities of the five participating cities and with the support of Federal University of Juiz de Fora. The project was funded by grants from the Ford Foundation, the National Institute of Studies and Projects - INEP, Foundations for Research Support of the State of Rio de Janeiro (FAPERJ) and Minas Gerais (FAPEMIG).

The Geres Project is the largest successful longitudinal study ever carried out in Brazil, its database was recently made available (Brooke and Bonamino, 2011) to allow "further investigations" by external researchers. In addition to data from cognitive assessment, the project collected contextual data from the students, families and schools, thus allowing in-depth studies with different models of statistical analysis, using different explanatory variables.

The project has investigated the evolution of learning skills in reading Portuguese and mathematics, in the first four grades of elementary school. Using the IRT with a logistic model with three parameters in order to calibrate the 116 items and 135 items of Portuguese and mathematics, respectively, the project established its own proficiency scale, which was adjusted to the fourth grade of the Brazilian Assessment System for Basic Education – SAEB scale (Klein, 2003). The correspondence between scale scores and the standard assessments of reading and mathematics skills of the Geres Project's was obtained through a cluster analysis of all items (Brooke & Bonamino, 2011, p. 134), ensuring the appropriate vertical scale for both subjects. In this way, there is the assurance that the proficiency scale covers all the curricular skills of interest in the four years of the elementary education investigated.

Three distinguishing features of the Geres Project are most relevant for the purposes of this study and justify the use of its database. First of all, its longitudinal design allows tracking the student's educational trajectory in the first four years of elementary school, making the analysis of value-added possible as discussed before. Secondly, the existence of a well-established standard-based assessment given by the evaluation of the descriptors matrices, with precise specification and interpretation of skills and abilities to be learned across the first four years of the students' primary education, yields a scale to measure proficiency. These two features are necessary conditions to determine the evolution of the value-added by the school and the definition of SIEE proposed in this paper.

Finally, the fact that the study is focused on the early grades of the primary school, when the cognitive evolution of students is faster and less dependent on their prior school performance (Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2010), provides greater data reliability in the determination of the proficiency's temporal evolution.

In the present study, I have analyzed the 272 schools, both public and private schools distributed across the five cities, aiming at building a statistically significant sample for the reality of each educational district.

Detailed accounts of the research findings of the Geres Project are presented in a number of publications. The reader can check it in (Brooke & Bonamino, 2011). With regard to the development of the student's proficiencies in the early elementary school years, the study dramatically showed that, on average, students of the public schools start elementary school with a lag of 100 points for both subjects – respectively, on the proficiency scale used – compared to the

average proficiency of students from private and special schools. This fact indicates that public education in the country has to work 1 to 1.5 years with their students, so they can first reach the level of initial proficiency of students from private schools. Possibly associated with the difference in mean SES between students of public and private schools systems in Brazil, this fact may still reflect the late literacy and lack of public kindergartens. This was one of the reasons for the introduction of the nine-year elementary school in the country in 2007. Surprisingly, even with this disadvantage of the public education, the Geres research has shown that the value-added by public schools in Brazil is the same as that by private schools. Nevertheless, the evolution of acquired proficiency shows a significant difference in behavior related to the two subjects. While the difference in reading skills falls by approximately 25 %, the lag in mathematics increases by almost 50 %. This fact may indicate that in reading/Portuguese public schools in Brazil reduce inequality, while in mathematics the inequality increases alarmingly. As I will see below, these facts can be best described and analyzed by the SIEE proposed in this work.

Conceptual Definition of the SIEE

I will now define the Social Index of Educational Effectiveness (SIEE) following a conceptual approach. I refer the interested to the Appendix for a more detailed mathematical definition.

First, as discussed before, note that all studies in EER agree about the strong correlation between the student's proficiency and its mean family SES, its social, familial and cultural context. This association is even more dramatic in the early grades of elementary school, when the child lives more intensely his childhood immersed in his familial reality. Therefore, the fundamental idea behind the conceptual development of the SIEE is the establishment of the children's initial proficiency in their first school year as the most relevant predictor or independent variable to assess the non-observable factors that define the relevant student's characteristics at this age.

For a sake of simplicity, suppose I have a classroom with only two students, A and B, with different prior proficiency, $P_A(0)$ and $P_B(0)$, so that $P_A(0) > P_B(0)$. Therefore, the student A represents a student with a higher prior proficiency and student B, one with lower prior proficiency. Let us estimate the proficiency of these students on a time t to be $P_A(t)$ and $P_B(t)$. The gain in proficiency, $V_A(t)$ and $V_B(t)$, is given by the differences in proficiency $V_A(t) = P_A(t) - P_A(0)$ and $V_B(t) = P_B(t) - P_B(0)$, respectively.

The essential feature of the proposed methodology is to figure out if these estimated values of proficiency gain, i.e. value-added by the educational process, show that the educational process is reducing ($V_A(t) < V_B(t)$) or increasing ($V_A(t) > V_B(t)$) the educational inequality. In the first case, a graphical plot of the proficiency gain vs. prior proficiency, $V(t)$ vs. $P(0)$, will be fitted by a decreasing function with negative slope and is then associated with a positive SIEE.

Once a negative slope indicates a positive behavior, that means, indicates that the school is increasing the educational equity, I define conceptually the SIEE as the negative slope that indicate the rate of increasing value-added with decreasing prior proficiencies,

$$SIEE = -\frac{V_A(t) - V_B(t)}{P_A(0) - P_B(0)}. \qquad (1)$$

Of course, for the mathematical definition of SIEE I have to consider a set of students and schools. So, let $P_{\{t,j,i\}}$ be the proficiency of the i-th student in the j-th school at time t. Thus $P_{\{0,j,i\}}$ is its initial proficiency. In the Geres Project database analyzed here, the subscript t stands for the

various assessment cycles. The initial proficiency $P_{\{0,j,i\}}$ corresponds to the estimated proficiency in the first assessment cycle and $P_{\{3,j,i\}}$ means the estimated proficiency in the fourth assessment cycle.

Similarly, the basic assumption of the proposed methodology is that the integral effect due to the set of factors associated with the school, teacher and school system can be accessed by the gain on the proficiency score in a given time interval. Hence, I define

$$V_{\{\Delta t,j,i\}} = P_{\{t+\Delta t,j,i\}} - P_{\{t,j,i\}} \quad (2)$$

as the value-added in proficiency for the i-th student in the j-th school, in the time interval $(t, t + \Delta t)$. In the context of a typical school, a child in a position of vulnerability or social disadvantage will also have an unfavorable learning context. Thus, the average distribution of value-added as a function of the initial proficiency

$$\bar{V}_{\{t,j,i\}}\,(P_{\{0,j,i\}})$$

is expected to be, on average, a monotonic increasing function. This is the key point in my proposition. I believe that the same context (social, familiar, etc.) that determine the prior proficiency will be in charge during the few initial years of the student in primary school. As a rule, the slope for $\bar{V}_{\{t,j,i\}}\,(P_{\{0,j,i\}})$ should be positive, indicating the expectation that a child with an initial proficiency below average will progress slower than a child does with an above average initial one. In this case, the school will be replicating the societal logic of inequality of opportunity. This characteristic is negative in terms of an index that seeks to show the school that overcomes this logic of inequality. The index should thus be positive in the case of a decreasing function $\bar{V}_{\{t,j,i\}}\,(P_{\{0,j,i\}})$, reflecting the fact that, in this case, the school adds more value in terms of proficiency just for students with lower initial proficiency, thus reducing intra-school educational inequality.

I will discuss the consequences of this definition bellow. For now let us only remark that meaningful values of ISEE are between$(-1, +1)$, since for higher values I would have some students with negative learning gains.

Determination and Interpretation of SIEE by School

I now apply the proposed methodology and verify the SIEE to illustrate the process for determining the SIEE for a specific school, as well as interpreting its meaning.

In this analysis I have calculated the value-added simply by the difference found between the fourth assessment cycle evaluation and the first assessment cycle ($t = 0$). So let $V_{\{3,\,j,\,i\}}$ be the value-added in proficiency between the fourth assessment cycle ($t = 3$) and the first assessment cycle, estimated for the i-th student in the j-th school. Figure 1 shows the distribution of the value-added in reading in terms of initial proficiency, that means $V_{\{3,\,j,\,i\}}$ vs. $P_{\{0,\,j,\,i\}}$. The school (identified by code BH013) presents a characteristic profile of social inclusion, which is highlighted by the positive SIEE value. The solid curve refers to the best-fitted linear regression using R (R Core Team, 2013), which allows the estimation of the slope, and therefore SIEE, using the definition given by Eq. 1 (or, more precisely Eq. A2).

Note, first, that the data show great variability and dispersion, in view of the students' different realities and abilities. The class is by no means homogeneous. However, the graphic shows a linear bias, more clearly revealed by the linear regression line with the setting data. Note also that the negative slope in this line highlights the fact that, on average, students of low prior proficiency (leftmost points on the horizontal axis) advanced more rapidly in their learning than students with higher initial proficiency (points over to the right on the horizontal axis).

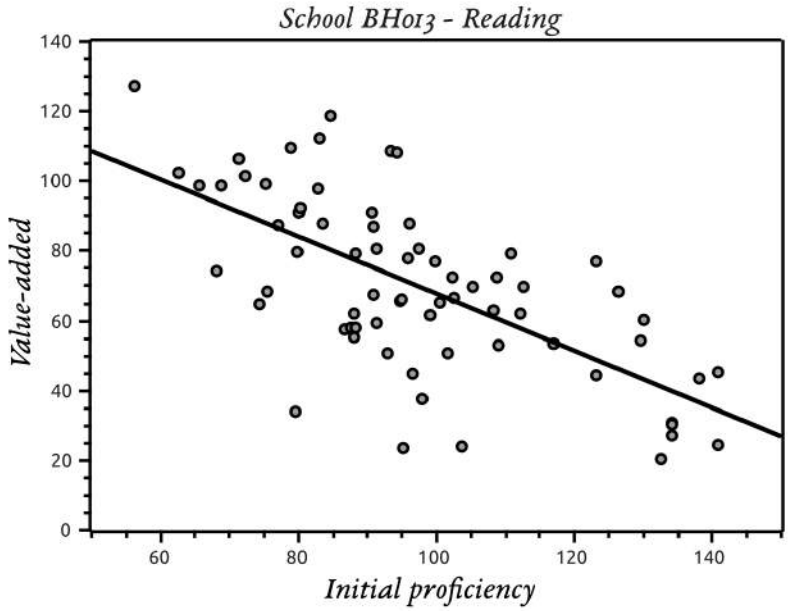

Figure 1. Example of profile distribution of added value on the initial proficiency - School BH013, city of Belo Horizonte, Brazil, a state public school. Discipline: Reading / Portuguese. Straight line shows best-fit yielding SIEE= .82. The data shows a typical behavior of educational effectiveness with an inclusionary profile.

Typically in this school, for students presenting low initial proficiency an average value-added of about 100 points was estimated, based on the standard scale of the Geres project for reading skills, whereas for students of high initial proficiency an average value of around 40 was estimated. This indicates a drastic reduction on the huge initial intra-school inequality. This fact is shown by the strong positive value yielding SIEE= .82.

Figure 2 shows the distribution $V_{\{3,\,j,\,i\}}$ vs. $P_{\{0,\,j,\,i\}}$, corresponding to the subject mathematics at the same school, involving the same students presented in Figure 1. Note that the linear regression shows a neutral profile, characterized by a SIEE value close to zero, SIEE= .07. In this case, there is no improvement in the initial low proficiency students.

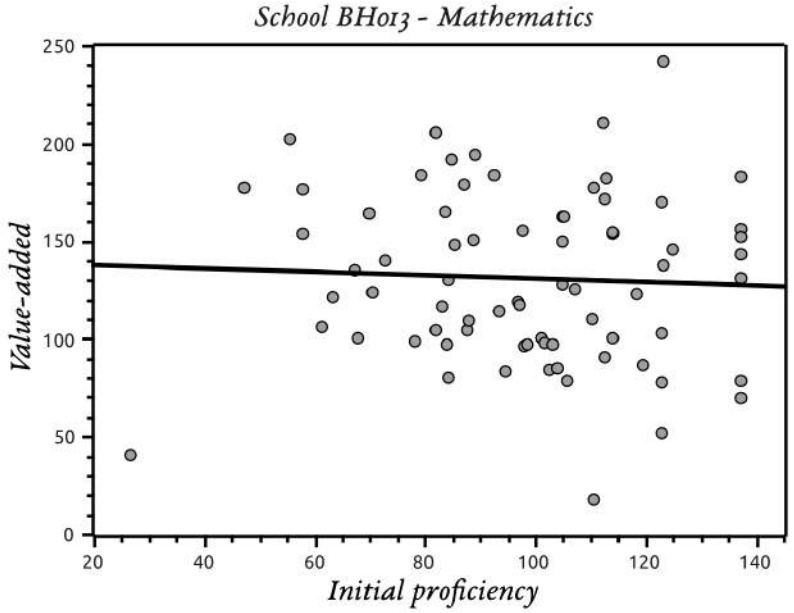

Figure 2. Example of profile distribution of value-added on the initial proficiency - School BH013, city of Belo Horizonte, Brazil, state public school. Discipline: Mathematics. Straight line shows best-fit yielding $\mathrm{SIEE} = .07$. Characteristic behavior is a neutral profile.

Figure 3 shows the distribution $V_{\{3,\,j,\,i\}}$ vs. $P_{\{0,\,j,\,i\}}$, also corresponding to mathematics teaching, at a school - code BH016 - with a strongly exclusionary profile. Note in this case that only the students with a higher initial proficiency have progressed more in their mathematics learning, which leads to an increase in inequality.

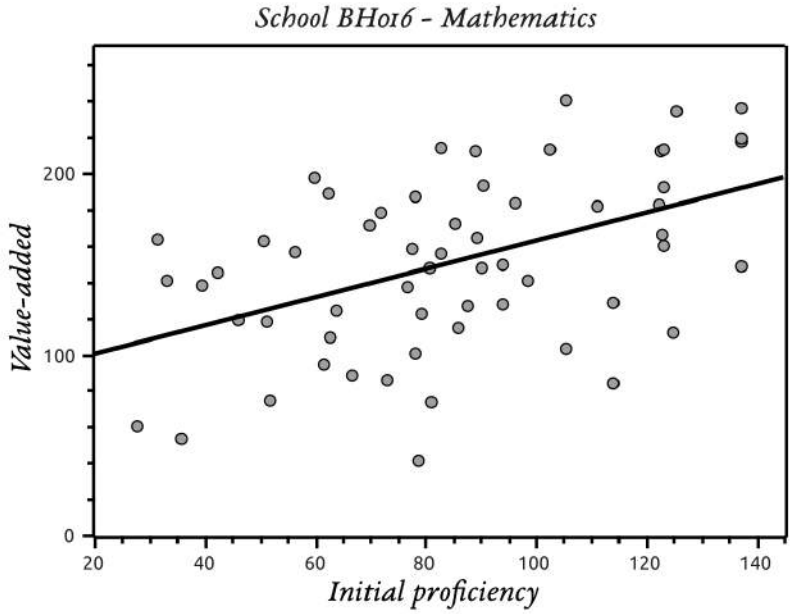

Figure 3. Example of profile distribution of value-added on the initial proficiency - School BH016, city of Belo Horizonte, Brazil, state public education. Discipline: Mathematics. Straight line shows best-fit yielding $\mathrm{SIEE} = -.78$. Characteristic behavior of educational effectiveness is an exclusionary profile.

Also note that both schools shown in Figures 02 and 03 have an equivalent average pupils' achievements in mathematics, i.e., about 150 points in the Geres Project standard proficiency scale for mathematics. The difference between the averages SES of the schools is not significant (see Table 1), hence, no matter which statistical model is used to determine the value-added by both schools, it is expected that both schools will also have similar values for educational effectiveness. However, the strongly negative value of SIEE for the second school (code BH016, Figure 3), means that the average difference in student's proficiency increases. It grows approximately from 100 points to 200 points, which shows a sharp increase of intra-school educational inequalities. Apparently, the second school is much less effective at promoting educational equity, a fact that is highlighted by its negative SIEE.

Results

Table 2 summarizes the estimated results of the SIEE and its variance, according to the methodology proposed in this paper, for both subjects, Portuguese (reading) and mathematics, as well as the correlation between SIEE and SES estimated. The SIEE of only 272 schools – of the 303 participating in the project - were estimated because the data of schools with a small number of students participating in the assessments did not support a reliable linear regression and were discarded. The data were analyzed as a function of the school district and school network.

An overall observation of the estimated values in Table 2 is sufficient to show the drastic difference in behavior between the two disciplines. The values for the SIEE Portuguese (reading skill) are systematically positive, ranging between $.31$ and $.52$, for both, the different education districts and different school networks. For mathematics, in contrast, the SIEE is negative, ranging between $.00$ and $-.38$, with a remarkable variance.

Note in Table 2 that the values of Pearson correlation coefficient between SIEE and SES are widespread and depend strongly on the educational district as well as the school system. The results show a strong positive correlation between SIEE and SES for the special public schools in reading, and for mathematics. This is not a desired characteristic, since this would mean that the schools that are more inclusionary are just the schools with positive SES. That means that the more positive the mean SES of the school, the more the school is inclusionary. I will discuss this fact later. It should be noted that the only case with a negative correlation is the school district of Campo Grande in mathematics, but the mean SIEE value is strongly negative, which is also a negative aspect.

Table 2
Average, Variance and SIEE – SES Correlation of the SIEE for Proficiency in Reading and Mathematics in the Studied Schools, Stratified by School Districts and School Systems.

Portuguese (reading)			
School District	Mean SIEE	Variance	Correlation SIEE x SES
Belo Horizonte	.31	.08	-.12
Campinas	.36	.17	.11
Campo Grande	.45	.13	-.23
Rio de Janeiro	.40	.06	.22
Salvador	.31	.09	.04
Schools System	Mean SIEE	Variance	Correlation SIEE x SES
Special Public Schools	.52	.02	.83
State Government Schools	.43	.14	-.04
Local Government Schools	.32	.04	.30
Private Schools	.34	.17	-.03
Matemathics			
School District	Mean SIEE	Variance	Correlation SIEE x SES
Belo Horizonte	-.25	.20	.78
Campinas	-.32	.24	.26
Campo Grande	-.23	.33	-.47
Rio de Janeiro	-.11	.20	.42
Salvador	-.22	.19	.04
School System	Mean SIEE	Variance	Correlation SIEE x SES
Special Public Schools	.00	.10	-.12
State Government Schools	-.23	.23	-.04
Local Government Schools	-.38	.14	.04
Private Schools	-.05	.33	.16

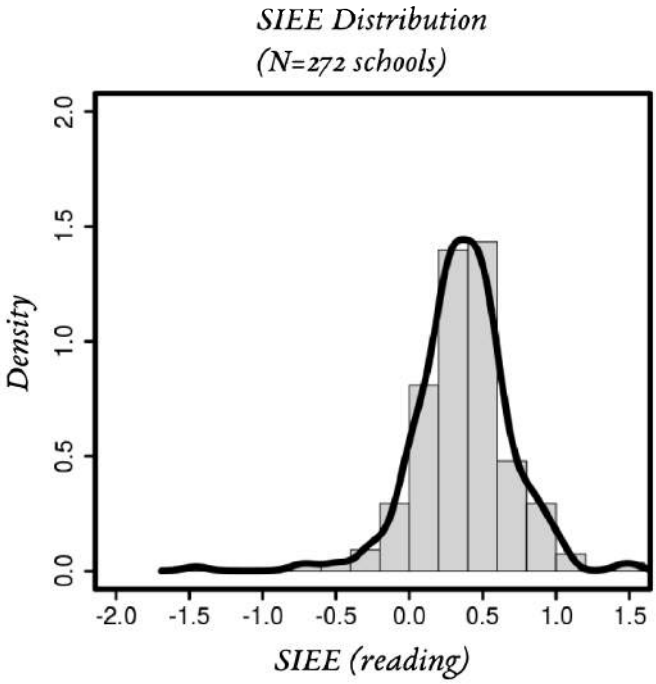

Figure 4. Histogram of frequency distribution of the SIEE of reading skill. Curve is best-fitted by fast Fourier Transform using R.

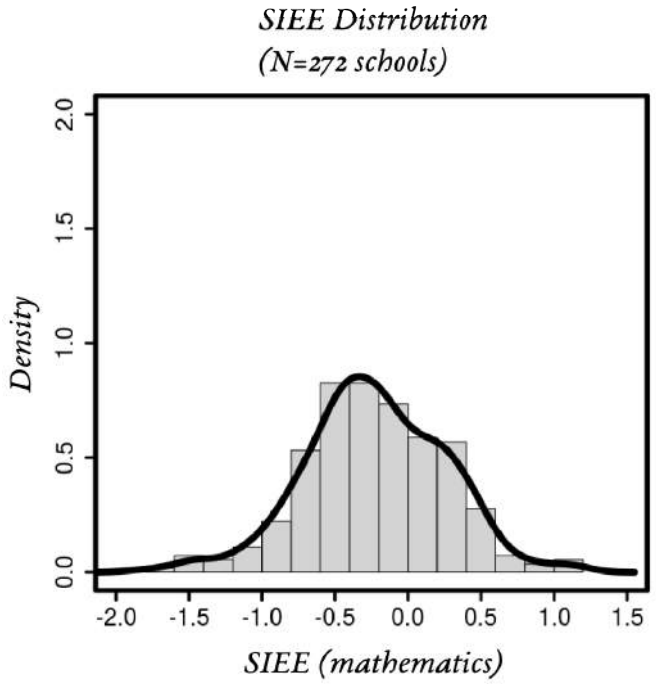

Figure 5. Histogram of frequency distribution of the SIEE of mathematics skill. The graphic is plotted in the same scale used in Figure 4. Curve is best-fitted by fast Fourier Transform using R.

The SIEE histograms make dramatically explicit the different behavior for these two subjects to their profile of educational effectiveness, especially with regard to three different aspects.

First, note that the histogram of the SIEE for teaching mathematics shows great dispersion, fully covering the interval $(-2, +1)$. It is noteworthy that this fact provides an additional indication of the anomalous behavior observed in the teaching of mathematics, since, as mentioned before, realistic values of SIEE should be in the range $(-1, +1)$. For reading skills, however, there is a strong concentration of the SIEE index in the region where the SIEE has positive values in the range $(0, +1)$. Both facts seem to indicate that in the teaching of mathematics there is not the same didactic, pedagogical and methodological uniformity among the different schools as I were able to observe in the teaching of Portuguese (reading).

Secondly, the frequency distribution shown in the histograms demonstrates that the average value of SIEE is markedly different in both subjects, as already seen in Table 2. While the average value of the SIEE for Portuguese is clearly positive, the average SIEE for the teaching of mathematics is negative.

Finally it is worth noting that, unlike the frequency distribution of the SIEE for mathematics (Figure 5), the frequency distribution of the SIEE for Portuguese (Figure 4) holds no similarity to the frequency distribution of the average SES of the schools (Figure 6). This fact is corroborated by the contrast in the calculated correlation coefficients SIEE – SES of the two disciplines.

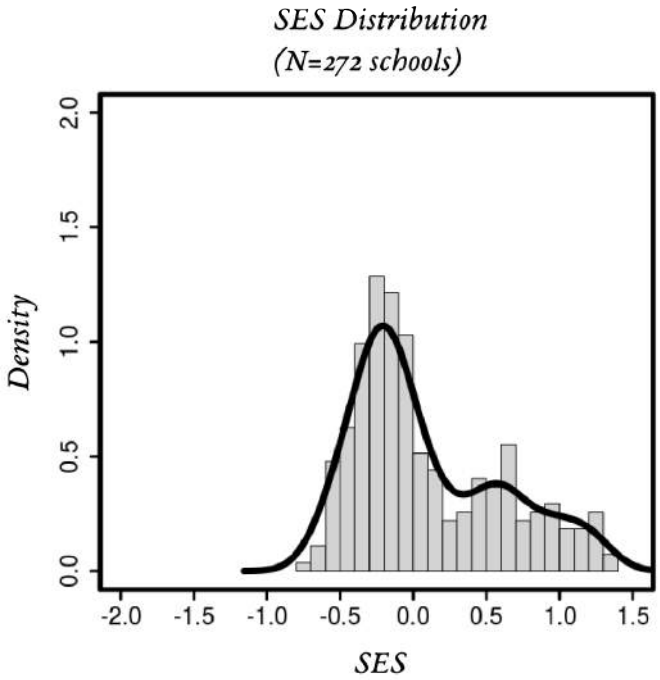

Figure 6. Histogram of frequency distribution of the average SES of the schools studied. Curve is best-fitted by fast Fourier Transform using R.

In summary, by using the methodology proposed in this paper, I found there is a marked difference between the teaching of Portuguese and mathematics in the schools investigated. The average of reading skills SIEE is positive, average $SIEE_{Reading} = .364$, and has a lower variability in its frequency distribution. In contrast, the average of skills on mathematics SIEE is negative, average $SIEE_{Math} = -.226$. Moreover, there is no correlation between the average school SES and SIEE for reading (Portuguese), which shows that, on average, these schools are able to overcome the logic of inequality in this discipline. This fact is not observed in relation to teaching of mathematics.

Discussion

There are several and consistent criticisms directed at educational effectiveness studies. A broad debate was held some years ago in a special issue of the journal School Effectiveness and School Improvement (SESI, 2001). On the one hand, there were complaints that the studies in the EER have been too pragmatic, preoccupied only to identify what works, rather than having a greater concern with the theoretical issues and the decisive influence of social structure on the results analyzed. There is still strong criticism on promoting the culture of "tarnishing" the image of schools by using an index for accountability that shows students' proficiency in a cross-section design. Nearly forty years of research in educational effectiveness have systematically revealed that this is an unfair approach, since the average proficiency of a school will always be strongly determined by the social, economic and cultural reality of the community in which it operates, that means, its context.

Moreover, indicators based on the concept of value-added by the school may also be, to a lesser extent, distorted by this social context, given that families with higher income and greater cultural capital offer on average more favorable conditions to their children's learning (Bourdieu, 1984). In this context, the school effectiveness also has correlations with the average school´s SES.

This paper presents an alternative methodology to assess the effectiveness of a school. The Social Index of Educational Effectiveness, defined by Eqs. (1) and (A2), shows the distribution profile of value-added by the school according to the student's initial proficiency. A positive SIEE index shows a school that is acting to reduce educational inequality in its specific context. On the other hand, a negative SIEE index shows a school that is increasing intra-school educational inequalities.

A very high positive value of SIEE can mean an anomalous school that cares only for students with low initial proficiency, abandoning or discouraging students that are more proficient. This fact, of course, cannot be considered a positive factor and certainly will be associated with null or negative learning gains for those having greater prior proficiency. Thus, the average value-added of the school remains an important factor to be considered in the analysis of educational effectiveness of schools.

Together with the analysis carried out using value-added and SIEE indexes, this study emphasizes the important role played by the SIEE – SES correlation. As seen in results presented in Table 2, the SIEE – SES correlations are strongly dispersed and depend strongly on educational district as well as the school system. This fact is due to the great sensibility of the SIEE – SES correlations to the profile distribution of SIEE within an Educational District or a School System. For example, if the average SIEE is positive and the correlation is negative, the school system has an inclusionary profile, and is more inclusive just in schools placed in regions with a low SES. In the Geres data investigated in this paper, this is the case found in Campo Grande. The mean SIEE of $.45$ with a SIEE – SES correlation of $-.23$ achieved by this municipality in reading indicates an Educational District with a strong commitment to the construction of educational equity, and that is more effective in the right region, that means, in regions with low SES. This situation contrasts strongly with what is observed in Belo Horizonte regarding to the mathematics teaching. The mean ISEE in this case is fairly negative (-0.25) and the ISEE – SES correlation is strongly positive ($+0.78$), indicating the worst scenario: an educational district that does not promote educational equity and, the best results of to the SIEE are just for schools in regions with high SES.

It´s worth mentioning that the classes for the primary school in Brazil are held, generally, by only one teacher, who is responsible for both subjects, reading and mathematics. There are not many studies that investigate directly the question of whether value-added measures are correlated

across subjects or not. In a recent review, Goldhaber, Cowan and Walch (2012) highlight that the correlations of math and reading value-added estimated for elementary and middle school teachers found in the literature range from .35-.65. In agreement with these results that yielded differential teaching effectiveness across subjects, my results suggest the possibility that specializing elementary teachers could indeed result in improvements in student outcomes in the context of the Brazilian primary school.

Using the concepts of mean value-added, the index SIEE that was proposed in this paper and its SES correlation, in Table 3 I propose a ranking of the profile of Educational Effectiveness, on eight classes in descending order of relevance to the quality education policy to promote social justice.

Table 3
Classification of the School Depending on the Values of Mean Value-Added and the SIEE.

Profile	SIEE	Value-Added	Correlation SIEE x SES	Educational Effectiveness Profile
A1	Positive	High	Negative	Effective school that promotes social inclusion, belonging to a school system that is more inclusive in regions with low SES
A2	Positive	High	Positive	Effective school that promotes social inclusion, but the school system is more inclusive in regions with higher SES
A3	Positive	Low	Negative	Less effective school that promotes social inclusion, belonging to a school system that is more inclusive in regions with low SES
A4	Positive	Low	Positive	Less effective school that promotes social inclusion, but the school system is more inclusive in regions with higher SES
B1	Negative	High	Negative	Effective school that does not promotes social inclusion, belonging to a school system that is more inclusive in regions with low SES
B2	Negative	High	Positive	Effective school that does not promotes social inclusion, but the school system is more inclusive in regions with higher SES
B3	Negative	Low	Negative	Less effective school that does not promotes social inclusion, belonging to a school system that is more inclusive in regions with low SES
B4	Negative	Low	Positive	Less effective school that does not promotes social inclusion, but the school system is more inclusive in regions with higher SES

The ideal is always that the school can get an average value-added above the mean values of the schools of a certain group. But, in order to contribute to overcoming social inequalities, the school should also have a positive SIEE, Profile A1 to A4. However, schools that operate in the region of social vulnerability and for this reason have a value-added below the average, must

necessarily have a positive SIEE index, A3 and A4 profiles, aiming to contribute to overcoming these inequalities.

These profiles should be considered of more value in a policy of fairness than the B1 to B4 profiles, characteristics of schools that have a negative index SIEE - indicating they are raising inequality. Evidently, profile B4 depicts the worst scenario: schools that offer little value-added and their value-added is poorly distributed.

The proposed methodology was applied to classroom estimates of SIEE, but the SIEE can be also applied for teacher estimates. The SIEE is not based on any statistical model that preclude some hidden variables of the students and schools context. However, this fact is also the weakness of the proposed methodology. Since I are not aware of the causal factors that determine why a school is found to be inclusionary or exclusionary.

More detailed studies are needed to investigate the causes and extent of the anomaly observed in the SIEE mathematics in the present study. It would be very interesting to see in other databases of countries that have well-established longitudinal data, whether the exclusionary nature of mathematics teaching in early elementary school years found in this study is a widespread phenomenon or if it is a localized challenge to the teaching of mathematics in Brazil. Another relevant point for future research is the question of if and how teachers' training courses for the lower grades may be contributing to this anomaly. It is worth investigating methodological alternatives to overcome it. Note that the differences observed in teaching Portuguese (reading) and mathematics could be assigned to different levels of pedagogical skill formation during the teacher formation courses.

Therefore, I conclude that the present study demonstrates that the proposed Social Index of Educational Effectiveness has strong relevance for the analysis and identification of the role of the school in fostering social inclusion and in reducing inequalities. Educational evaluation and accountability policy by public policymakers that are willing to face the challenge of social inclusion and equity should be taken it into account on their assessments, assumptions and work agenda. The proposed approach does not discard traditional indices, such as average proficiency rates and average value-added by the school, but should be considered along with them for the establishment of a fair criterion for determining Educational Effectiveness considering the role of schools in promoting educational inclusion.

Authors' Note

The author would like to thank the staff of Geres Project, through its organizers Nigel Brooke and Alicia Bonamino, for generously make the database of the Geres Project publicly avaiable. Thanks are due to R.M. Fontes Martins and S. Abranches for critical reading of the draft.

Declaration of Conflict of Interests

The author declares no potential conflict of interests with respect to the research, authorship, and/or publication of this article.

Funding

The research reported here was conducted as part of the Basic Education Research Agreement CAPES-FAPEMIG. Grant APQ-03366-12 supports partially this research.

References

Atlas Brasil. (2013). *Atlas of Human Development in Brazil 2013* [Atlas do Desenvolvimento Humano no Brasil]. [S.l]: PNUD, 2013. Retrieved September 15, 2014 from: http://www.atlasbrasil.org.br/2013/en/.

Baker, F. B., & Kim, S. (2004). *Item response theory: parameter estimation techniques*. New York, NY: Marcel Dekker.

Baker, E.L, Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Ravitch, D., Rothstein, R., Shavelson, R. J. and Shepardet, L. A. (2010). Problems with the use of student test scores to evaluate teachers. *Economic Policy Institute Briefing Paper #278*. Retrieved March 12, 2015, from: http://www.epi.org/publication/bp278/

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, *29*, 37-65. http://dx.doi.org/10.3102/10769986029001037

Bonamino, A. M. C., & Oliveira, L. H. G. (2013). Estudos longitudinais e pesquisa na educação básica [Longitudinal studies and research in basic education]. *Linhas Críticas* (Online), *19*, 33-50. Retrieved April 18, 2014, from: http://seer.bce.unb.br/index.php/linhascriticas/article/view/8918

Bonamino, A. M. C, & Souza, S. Z. L. (2012). Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola [Three generations of basic education assessments – Brazil: interfaces with the curriculum of/in the school]. *Educação e Pesquisa* (*USP*), *38*, 373-388. http://dx.doi.org/10.1590/S1517-97022012005000006

Bourdieu, P. (1984). *Distinction*. Cambridge, MA: Harvard University Press.

Brooke, N., & Bonamino, A. M. C. (Org.). (2011). *Geres 2005: Razões e resultados de uma pesquisa longitudinal* [*Geres 2005: Reasons and results of a longitudinal research on school effectiveness*]. 1a. Ed., Rio de Janeiro: WalPrint.

Campbell, J., Kyriakides, L., Muijs, D., & Robinson, W. (2004). Assessing teacher effectiveness: Developing a differentiated model. New York, NY: Routledge Falmer. http://dx.doi.org/10.4324/9780203403709

Chetty, Raj, Friedman, J. N., & Rockoff, J. E. (2014) Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104*(9), 2593-2632. http://dx.doi.org/10.1257/aer.104.9.2593

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). Equality of educational opportunity. Washington, D.C.: U.S. Government Printing Office.

Creemers, B., Kyriakides, L., & Sammons, P. (Eds.). (2010). *Methodological advances in educational effectiveness research*. London, UK: Routledge Taylor Francis.

DeMars, C. (2010). *Item response theory*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195377033.001.0001

Edmonds, R. R. (1979). Effective schools for the urban poor. *Educational Leadership*, *37*(10), 15–24.

Franco, C., Brooke, N., & Alves, F. (2008). Longitudinal study of quality and equity in the Brazilian Elementary Education: GERES 2005 (In Portuguese). *Ensaio: Avaliação e Políticas Públicas em Educação*, *16* (61), 625-638. Retrieved April 18, 2014 from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-40362008000400008&lng=en &tlng=. 10.1590/S0104-40362008000400008.

Goldhaber, D., Cowan, J., & Walch, J. (2012). *Is a good elementary teacher always good? Assessing teacher performance estimates across subjects*. CEDR Working Paper 2012-7.2. Seattle, WA: University of Washington.

IBGE – Brazilian Institute of Geography and Statistics. (2014). Information on the Brazilian municipalities. Retrieved September 15, 2014 from: http://www.cidades.ibge.gov.br/xtras/home.php?lang=_EN.

Jencks, C., Smith, M., Acland, H., Bane, M.J., Cohen, D., Gintis, H., Heyns, B., & Michelson, S. (1972). *Inequality: A reassessment of the effects of family and schooling in America*. New York, NY: Basic Books.

Kane, T. J., McCaffrey, D. F., Miller, T., & Staiger, D. O. (2013). *Have I identified effective teachers? Validating measures of effective teaching using random assignment*. Seattle, WA: Bill & Melinda Gates Foundation.

Klein, R. (2003). Utilização da teoria de resposta ao item no Sistema Nacional de Avaliação da Educação Básica (SAEB) [Using the item response theory in the national assessment system for the basic education (SAEB)]. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, *11*(40), 283-296.

Lee, V. E., & Bryk, A.S. (1989). A multilevel model of the social distribution of high school achievement. *Sociology of Education*, *62* (Julho), 172-192. http://dx.doi.org/10.2307/2112866

Linder, W. J. van der & Hambleton, R. K. (Eds.). (1997) *Handbook of Modern Item Response Theory*. Spring-Verlag, New York, NY. http://dx.doi.org/10.1007/978-1-4757-2691-6

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND Corporation. http://dx.doi.org/10.1037/e658712010-001

McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistic*, *29*(1), 67-101. http://dx.doi.org/10.3102/10769986029001067

National Research Council. (2010). *Getting value out of value-added: Report of a workshop*. Retrieved March 15, 2014 from: http://216.78.200.159/Documents/RandD/Other/Getting%20Value%20out%20of%20Value-Added.pdf.

Newton, X., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Education Policy Analysis Archives*, *18* (23). DOI: http://dx.doi.org/10.14507/epaa.v18n23.2010

Oliveira, R. P. de. (2007). From the universalization of teaching to the challenge of quality: A historical analysis. *Educ. Soc – Campinas*, *28* (100 – Special), 661-690. Retrieved September 15, 2014 from: http://www.cedes.unicamp.br

R Core Team. (2013). *R: A language and environment for statistical computing*. Retrieved September 1, 2014 from R Foundation for Statistical Computing: http://www.R-project.org.

Radhakrishna Rao, C., Toutenburg, H., Shalabh, & Heumann, C. (2008). *Linear models and generalizations*. 3rd ed. New York, NY: Rao Springer.

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *Quarterly Journal of Economics, 125*(1), 175-214. http://dx.doi.org/10.1162/qjec.2010.125.1.175

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. Cambridge, MA: Harvard University Press.

Sammons, P. (2007). *School effectiveness and equity: Making connections. A review of school effectiveness and improvement research and its implications for practitioners and policymakers, report commissioned by CFBT*, Retrieved March, 14, 2014 from: http://www.cfbt.com/evidenceforeducation/pdf/91168_SchEffectiveness_FullLiteratureReview.pdf.

Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (Ed.) (2010). *Early childhood matters: Evidence from the effective pre-school and primary education project.* London, UK: Routledge Taylor Francis.

UNESCO. (2010). *World data on education,* 7th ed. Retrieved March 14, 2015 from: http://www.ibe.unesco.org/fileadmin/user_upload/Publications/WDE/2010/ pdf-versions/Brazil.pdf .

Verhelst, N. (2010a). Using item response theory to measure outcomes and factors. In B Creemers, L. Kyriakides & P. Sammons (Eds.), *Methodological advances in educational effectiveness research* (pp. 153-182). London, UK: Routledge Taylor Francis.

Verhelst,N. (2010b). IRT modles. In B Creemers, L. Kyriakides & P. Sammons (Eds.), *Methodological advances in educational effectiveness research* (pp. 183-218). London, UK: Routledge Taylor Francis.

Appendix
Mathematical Definition of the SIEE

Since the slope of $\overline{V}_{\{t,j,i\}}$ $(P_{\{0,j,i\}})$ should be negative for a school with positive ISEE, the SIEE index sign is contrary to the sign of the slope or the rate of variation of $\overline{V}_{\{t,j,i\}}$ as a function of $P_{\{0,j,i\}}$. Hence, I define

$$SIEE_{\{t,j\}} \equiv lim_{\Delta P_{\{0,j,i\}} \to 0} \left(\frac{\overline{V}_{\{t,j,i\}} (P_{\{0,j,i\}} + \Delta P_{\{0,j,i\}}) - \overline{V}_{\{t,j,i\}} (P_{\{0,j,i\}})}{\Delta P_{\{0,j,i\}}} \right) . \quad \text{(Eq. A1)}$$

Graphically, the SIEE is the slope of the curve $\overline{V}_{\{t,j,i\}}$ $\left(P_{\{0,j,i\}}\right)$ and can be determined by the slope of the linear regression of $\overline{V}_{\{t,j,i\}}$ as function of $P_{\{0,j,i\}}$. Thus, following the linear regression method (Radhakrishna Rao, Toutenburg, Shalabh & Heumann, 2008), the Social Index Educational Effectiveness regarding the j-th school, due to the value-added in the period between the first evaluation and the evaluation made at time t, can be estimated by:

$$ISEE_{\{t,j\}} = - \frac{\Sigma i[(P_{\{0,j,i\}} - \bar{P}_{\{0,j,i\}})(P_{\{t,j,i\}} - P_{\{0,j,i\}})]}{\Sigma i\left[(P_{\{0,j,i\}} - \bar{P}_{\{0,j,i\}})^2\right]} \quad \text{(Eq. A2)}$$

with a standard error given by:

$$\square ISEE_{\{t,j\}} = - \frac{S}{\sqrt{\Sigma i\left[(P_{\{0,j,i\}} - \bar{P}_{\{0,j,i\}})^2\right]}}. \quad \text{(Eq. A3)}$$

In this expression, S is the variance of $V_{\{t,j,i\}}$,

$$S_j^2 = \frac{\Sigma i[(P_{\{t,j,i\}} - P_{\{0,j,i\}}) - a(P_{\{0,j,i\}}) - b]^2}{N_j - 2}. \quad \text{(Eq. A4)}$$

In determining the variance S above, the parameter a = $ISEE_{\{t,j\}}$ is the slope, b= $\sum_i \left[\left(P_{\{t,j,i\}} - P_{\{0,j,i\}} \right) - a\left(P_{\{0,j,i\}} \right) \right]$ is the intercept and N_j is the number of students from the j-th school that participated in the study.

About the Author

Ulisses A. Leitão
Exact Science Department at the Federal University of Lavras (UFLA)
ulisses@ufla.br
Twitter: @ualeitao
Dr. rer. nat. Ulisses A. Leitão (Germany, 1988) is Adjunct Professor of the Exact Science Department at the Federal University of Lavras (UFLA) since 2009. He takes part as full master degree advisor in the National Master Program on Physics Education, promoted by the Brazilian Physics Society and CAPES. His interests include inquiring-based physics and mathematics education, applications of Item Response Theory on large-scale evaluation for improvements in quality education.

The study reported in this paper was conducted as part of the Basic Education Research Agreement CAPES-FAPEMIG. Grant APQ-03366-12 supports partially this research.

education policy analysis archives

Volume 23 Number 50 May 4th, 2015 ISSN 1068-2341