

## **Measuring inequality of opportunity in education by combining information on coverage and achievement in PISA**

*Márcia de Carvalho*

Departamento de Estatística and Centro de Estudos sobre Desigualdade e Desenvolvimento (CEDE),  
Universidade Federal Fluminense (UFF), Brazil.

*Luis Fernando Gamboa*

Department of Economics, Universidad del Rosario

*Fábio D. Waltenberg*

Departamento de Economia and Centro de Estudos sobre Desigualdade e Desenvolvimento (CEDE), Universidade Universidade Federal Fluminense (UFF)

**Resumo:** Medir quão distantes países ou regiões se encontram do objetivo normativo da igualdade de oportunidades em educação tem sido foco de contribuições recentes na literatura econômica, as quais têm se concentrado, ou em hiatos, entre grupos da população, no acesso a determinados níveis de ensino, ou então em hiatos, entre grupos da população, em resultados em testes de conhecimentos. Argumenta-se neste trabalho que ambos os aspectos – oportunidade ao acesso e ao conhecimento – são importantes, e sobretudo inter-relacionados, de modo que, por consequência, ganham em ser mensurados simultaneamente e combinados num índice composto de igualdade de oportunidades em educação. Aplicando-se o índice a dados recentes provenientes dos exames internacionais do PISA para um conjunto selecionado de países, observam-se mudanças de ordenamento na comparação com rankings estabelecidos com base numa única dimensão. O índice composto pode ser generalizado para incorporar outras dimensões, como nível médio de conhecimento e dissimilaridade no acesso.

**Abstract:** Measuring how far away countries stand from a goal of equality of opportunity in education has been the focus of recent contributions in the economic literature, which have concentrated either on inter-group gaps in access to a given level of studies, or on inter-group gaps in educational achievement. We argue that both aspects are important and interrelated and, as a consequence, should be measured simultaneously and combined in a composite index of equality of opportunity in education. Employing recent PISA data for a selected set of countries, rank reversals are observed with respect to orderings based upon a single dimension. The index can be generalized to include further dimensions, such as average scores or dissimilarity in access.

**Palavras-chaves:** igualdade de oportunidades em educação, mensuração de desigualdade de oportunidades, índices compostos, exame de conhecimentos PISA, América Latina.

**Key words:** equality of opportunity, measurement of inequality of opportunity, composite indices, PISA test scores, Latin America.

**Área ANPEC:** 12 – Economia Social e Demografia Econômica

**JEL:** I24, O54.

# Measuring inequality of opportunity in education by combining information on coverage and achievement in PISA

## 1. Introduction

Measuring how far away countries stand from a goal of equality of opportunity in education has been the focus of recent contributions in the economic literature, which have concentrated either on inter-group gaps in access to a given level of studies, or on inter-group gaps in educational achievement. We argue that both aspects are important and interrelated and, as a consequence, should be measured simultaneously and combined in a composite index of equality of opportunity in education.

Pertaining to a liberal-egalitarian tradition of theories of justice, John Roemer's (1998) concept of equality of opportunity (EOp) states that if a given outcome inequality between a pair of individuals is observed (*ex post*) and if it can be traced back to different predetermined circumstances (*ex ante*), it will be declared that opportunities have not been equal. Equality of opportunities is thus inferred from observing outcome inequalities conditional on circumstances.

If, on the one hand, the fraction of an observed inequality which is due to different circumstances is deemed intolerable, on the other hand, the fraction which is due to deliberate, conscious, choices made by the individuals is acceptable. Different methodologies have been proposed by economists attempting to translate those ideas into measuring procedures, which *grosso modo* try to decompose overall inequality into two components: the fair and the unfair fractions (e.g., Checchi & Peragine, 2010; Dunnzlauf et al. 2010). The larger the calculated share of unfair inequality, the further away from equality of opportunity a given country or region is. Two recent surveys are available documenting the vast literature produced since 1998 (Pignataro, 2012; Ramos & Van de gaer, 2012).

When it comes to assessing equality of opportunities *in education* in particular, at least two major dimensions are important. On the one hand, there is *access*. Being enrolled in a school is a necessary condition for obtaining formal instruction and all the advantages that it provides. If we admit that no kid or teenager should be held responsible for having dropped out, an educational system which fails to retain in schools a considerable portion of a country's pupils' cannot be viewed as one fulfilling EOp in education. If, to the contrary, every kid is enrolled, there is equality of opportunities in the primary, narrow, sense that a minimal threshold of opportunities has been provided to all – a dimension which we denote “access-EOp”.

On the other hand, in a broader, more comprehensive, perspective, access might be viewed as a necessary, but not sufficient condition for EOp in education. All individuals of a given cohort might be enrolled in schools, but the learning conditions faced by different pupils might be so dissimilar – according to, say, their social background or their gender –, that some will not learn adequately, while others will acquire an outstanding level of knowledge. To the extent that such achievement gaps are due to inequalities in predetermined circumstances, there is inequality of opportunities within Roemer's framework. If achievement gaps are not related to predetermined

circumstances, but only to conscious choices – say, more or less time spent doing assigned homework – made by individuals facing equal (or similar) circumstances then equality of opportunity in achievement (or “achievement-EOp” for a shorthand) could be declared.<sup>1</sup> In a way, it can be said that access-EOp precedes achievement-EOp as a requirement for overall EOp, or that the former is necessary and the latter is sufficient. An ideal education system would make sure all kids are enrolled, and would ensure achievement is uncorrelated to pupils’ socioeconomic background.

As mentioned above, measuring how far away a given country stands from equality of opportunity in education has been at the center of recent contributions, which concentrate though on only one of those two important and interrelated dimensions – that is, either on access-EOp (e.g., Paes de Barros et al. 2009; Vega et al. 2010), or on achievement-EOp (e.g., Checchi & Peragine, 2005; Ferreira & Gignoux, 2011; Gamboa & Waltenberg, 2012). In the latter contributions, pupils’ achievement is measured by standardized test scores, such as those made available by OECD’s *Programme for International Student Assessment* (PISA).

Such contributions are important for having improved our knowledge on equality of opportunity in education – that is, on the extent of inter-group gaps in access, as well as of achievement gaps. It must also be acknowledged that the first set of contributions, focusing on the more basic requirement for EOp, which is access-EOp, is somehow “allowed to ignore” the achievement dimension. Nonetheless, an important limitation affects the second set of contributions, dedicated to measuring achievement-EOp only: the more primary requirement of EOp, which is access-EOp, is ignored when it should not be. Countries, or regions within a country, have been compared regarding achievement-EOp, employing samples which are not uniform in terms of *coverage of a given cohort*.<sup>2</sup> For example, if country A shows more calculated achievement-EOp than country B but country A’s sample covers a smaller fraction of its cohort of pupils than does country B’s sample, it will be unclear whether country A should indeed be declared as one providing more education opportunities than country B. The information on coverage (a potential proxy for access, as discussed later) should not be simply disregarded as it has been.

Previous attempts of addressing such problem of disregarding access have been ingeniously performed by Ferreira & Gignoux (2011), who have tried to explicitly reconstruct a full sample before calculating achievement-EOp. Our strategy is of a different nature. Instead of attempting to simulate a full sample, we explicitly acknowledge there are *two different and complementary dimensions of educational opportunity* - (i) access and (ii) achievement conditional on access - and then we aggregate them into a composite index.

Aggregating the dimensions is not trivial, among other reasons because the relative weight attributed to each dimension will certainly vary according to normative preferences or to the objective of each study. Besides choosing weights, another decision concerns the specific functional form of the composite index. We have opted for aggregating the dimensions in a simple, multiplicative way. The procedure provides an index which, while adjustable to different normative preferences, precludes the possibility of perfect substitution of one dimension by the other, which is justified since both are intrinsically important in our view.

---

<sup>1</sup> It should be clear that EOp might coexist with achievement inequalities.

<sup>2</sup> Coverage might be viewed as a proxy for access, a point to which we come back later.

We illustrate our approach for six Latin-American countries and five developed countries with high performance in standardized tests that took part in PISA 2006 and 2009. Depending on the relative weight attributed to each dimension, the index we introduce ranks the countries differently than does a conventional index, exclusively focused on achievement-EOp, and thus is capable of representing a wider spectrum of preferences regarding the relative importance of achievement-EOp with respect to access-EOp.

The paper is organized as follows. Section 2 is devoted to explaining the motivation for this study. In Section 3 we describe our approach. Section 4 reports an application for a selected set of countries. In Section 5 we generalize our method, allowing for further dimensions and weights. Section 6 contains final remarks.

## **2. Achievement-EOp calculations and attempts to circumvent the coverage rate problem**

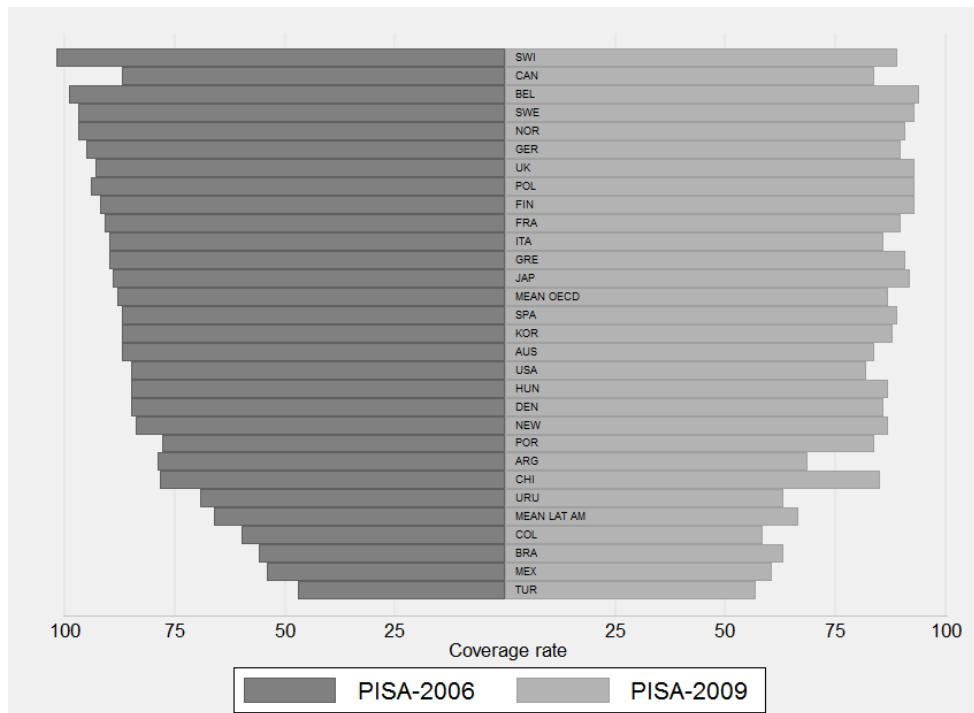
We mentioned before that achievement-EOp calculations have employed standardized test scores from OECD's PISA. Collected every three years since 2000, PISA datasets include test scores of representative samples of students in dozens of countries in three subjects – mathematics, sciences and reading – as well as detailed information on students' background and schools' personnel and functioning conditions. The fourth wave (2009), employed in this paper, contains samples of about 520 thousand students representing pupils of more than 70 countries (OECD, 2012: 25).

PISA samples have a limitation which is rarely emphasized – possibly because it is irrelevant for other purposes but which matters for ours – namely, that the samples do not necessarily represent the national population of 15-year-olds in every participating country. The reasons fall into two categories: (i) some eligible pupils do not take the exams for logistic or fortuitous reasons (e.g., pupils living in a remote region, or pupils who were sick the day the exam took place); local managers of PISA exams might also exclude some pupils for physical or intellectual deficiencies (the accepted cases of which are carefully detailed in PISA manuals), and (ii) individuals who are enrolled in a very low grade (“grade 6” or below<sup>3</sup>) or who are not enrolled in schools are not assessed by PISA, being considered “ineligible”. While the exclusions for fortuitous or technical reasons affect all countries to a certain extent, they are meaningless from the point of view of educational-EOp, since they do not denote any obvious deficiency of the education system. The second category is more problematic. As a consequence, empirical findings based on PISA should not be taken as valid for cohorts of 15-year-old individuals, but for individuals who: (a) have not abandoned the educational system, and (b) have not repeated too many grades. While the actual and the ideal samples do not differ very much in developed countries, they diverge considerably in developing countries.

---

<sup>3</sup> PISA's “grade 6” corresponds with different names in different countries. PISA technical reports (e.g., OECD, 2009 and 2012) provide tables containing country-by-country corresponding labels.

**Figure 1: Coverage rates at PISA 2006-2009**



The proportion of the cohort of 15-year-old individuals which has been excluded is not uniform, neither across countries, nor over time. As shown in Figure 1, differences can be substantial casting doubts on the reliability of cross-country comparisons. The coverage rates<sup>4</sup> in 2006 and 2009 in a selected set of Latin American countries are a good example: Argentina (79%, 69%), Brazil (55%, 63%), Chile (78%, 85%), Colombia (60%, 59%), Mexico (54%, 61%) and Uruguay (69%, 63%). These figures reveal that: (i) coverage rates are not high on average, (ii) although all countries come from the same region, cross-country dispersion is substantial, with a range of around 25 percentage points in each year, (iii) there are within-country oscillations across waves (6.5 percentage points on average).

For example, if Mexico turns out to show a better achievement-EOp figure than Argentina in 2006, it will be unclear whether that indeed reflects less EOp in the latter than in the former, or whether such result is driven by a more homogeneous sample in Mexico (lowest coverage rate in that year) than in Argentina (highest rate). In other words, had all 15-year-old Mexicans taken the exam, would Mexico’s good EOp record as compared to Argentina’s remain?

To address coverage rate problems – in PISA and possibly in other test scores datasets – , we view at least three paths. The first is simply to be cautious when interpreting results of studies employing PISA for developing countries. That is simple, but risky, since readers (and policymakers) might basically overlook the call for caution. Gamboa & Waltenberg (2012) opt for that path, presenting results based on PISA limited samples and emphasizing that caution is necessary in their interpretation.

<sup>4</sup> Along this paper, we shall call “coverage rate” the ratio in a given country and year between the weighted estimate of *eligible non-excluded 15-year-olds from the student sample* and the *national 15-year-old population*.

A second path consists of explicitly reconstructing full samples. Recently Ferreira & Gignoux (2011) have done so for four countries: Brazil, Indonesia, Mexico, and Turkey. Due to the absence of information in PISA samples about non-participant pupils, it is not possible to perform a correction such as Heckman's procedure (except through appropriate instrumental variables and complementary datasets). They have turned to ancillary datasets (i.e. household surveys), which, however, do not contain information on test scores, and they have then imposed some assumptions in order to undertake two kinds of simulation. The first consists of re-weighting test scores observations in PISA using information taken from the ancillary databases on the fraction of different types of individuals in the population.<sup>5</sup> The second one consists of imputing into the dataset pupils who were not evaluated, ascribing to them scores equal to the lowest score obtained by individuals very similar to them according to observable characteristics.

The first simulation, which relies on more conventional assumptions, provides results almost equal to the original results, both in terms of overall inequality and of unfair inequality. While such somewhat unexpected finding might offer relief for those employing PISA datasets, applying the procedure to other countries/years could lead to more substantial changes. The second simulation leads to sizeable differences with respect to the naive calculations, but as Ferreira & Gignoux (2011) themselves acknowledge, their assumptions are “admittedly extreme”.

We would like to view our approach as complementary to those two we have just described. Yet, we believe the former one is subject to misinterpretation of results. And as for the latter, sample corrections still present high costs (especially being data-intensive) and limited benefits (results based on strong assumptions). For these reasons we tend to favor a different path.

### 3. A composite index that combines access and achievement

We explicitly assume there are *two dimensions of educational opportunity* – namely, (i) access, and (ii) achievement conditional on access – each of which might rely on a different branch of the EOp literature and thus be measured according to a specific procedure before being aggregated into a composite index.

Since we have used as the motivation for this paper shortcomings in the calculations of achievement-EOp based on PISA, we shall emphasize the way we can circumvent them also based on PISA. But the procedure we discuss might be adapted and employed with other datasets and in different contexts. Let us now describe how each dimension could be calculated before turning to aggregation issues.

1. Inequality of opportunity in achievement. We restrict the calculation of inequality of opportunity in achievement to the available PISA samples employing Ferreira & Gignoux's (2011) regression-based index of inequality of educational opportunity (hereafter:  $IO_{FG}$ ). This is calculated as the proportion of the variance of test scores that is explained by a set of circumstances, ranging from  $IO_{FG} = 0$  (perfect equality of opportunity) to  $IO_{FG} = 1$  (perfect inequality of opportunity). The exercise undertaken is essentially a static decomposition of inequality (as expressed by the

---

<sup>5</sup> They adapt the methodology proposed by DiNardo et al. (1996).

variance) into unfair inequality (the R-squared) and fair – more precisely, “residual” – inequality (1 – R-squared).<sup>6</sup>

2. Inequality of opportunity in access. To the best of our knowledge, no consensus has emerged concerning how to adequately measure inequality of opportunity in terms of access. We propose employing countries’ coverage rates available in PISA technical reports, which might range from 0 (no coverage) to 1 (full coverage). Following Paes de Barros et al. (2009) and Vega et al. (2010) the overall coverage rate is denoted  $\bar{p}$ , with  $0 \leq \bar{p} \leq 1$ . The rationale is to sanction each country according to the extent to which its educational system excludes individuals from PISA exams.<sup>7</sup>

Employing the coverage rate or the enrolment rate as a measure of access might be defended by those who view access exclusively as a function of circumstances, such that any deviation from full coverage should be sanctioned, irrespectively of which groups are excluded, in which case the greater the distance from full coverage ( $\bar{p} = 1$ ) the larger the inequality of opportunity in that country.

One might object that what would matter for Roemer’s concept of EOp would not be the overall coverage rate, but rather the extent to which different groups are excluded from PISA. That would require penalizing a country whose coverage rate varies across groups. In such scenario, an index as the so-called ‘Human Opportunity Index’ (Paes de Barros et al., 2009) could be a candidate, since one of its component terms is a dissimilarity index, expressing cross-types inequalities in access. However, a powerful objection to employing *HOI* is that it implicitly assumes that observed access inequalities are partly fair, which is inconsistent with EOp theory in that being out of schools (or PISA) is not a choice for which teenagers should be held responsible.<sup>8</sup>

Information on PISA’s coverage rate has another advantage over the alternative mentioned above, namely, the fact that it can be easily gathered, since they are readily available in PISA reports. Not requiring the handling of further datasets might be an attractive feature for both researchers and policy makers willing to make international or intertemporal comparisons relatively swiftly. Finally, since the coverage rate is provided in PISA technical reports, the data employed in both dimensions (achievement and access) would come from the same source, which means more clarity and uniformity in definitions, as compared to the alternative which would involve different sources of data for each country.

The remaining step is to aggregate such dimensions into one composite index. Being aware that different views on the relative importance of each dimension will never be

---

<sup>6</sup> Since the R-squared never goes down when new variables are added to the regression, it can be interpreted in the present context as a lower bound for achievement-InEOp. For details see Ferreira & Gignoux (2011).

<sup>7</sup> A similar strategy would consist of employing enrollment rates of 15-year-olds, which also range from 0 to 1, sanctioning each country according to the extent to which its educational system excludes individuals from schools (generally). To that rate could be added information, if available, on the proportion of 15-year-olds who have repeated too many grades, assuming that multiple grade repetition is not only a barrier to being eligible to PISA exams, but also a predictor of restricted life opportunities. For the sake of simplicity, we employ PISA coverage rates.

<sup>8</sup> Sill, in section 5, the more general version of the composite index, will include a measure of inter-group inequality of access, satisfying those who believe that would be an important ingredient in measuring EOp in education.

consensually settled, we have opted for aggregating the dimensions by means of a Cobb-Douglas specification allowing different choices of relative weights to each dimension.<sup>9</sup>

Although there is no consensus about the best aggregation procedure Cobb-Douglas's has been frequently employed above all due to its simplicity. With such functional form, the attributes cannot be perfect substitutes as they would be with an additive form, which makes sense since both dimensions are important and interrelated in our view. Moreover, it respects two properties, which are well-known in microeconomics: (i) mononicity: an increment in the value of either attribute always increases the index's value; and (ii) convexity: a moderate improvement in both dimensions results in a more substantial increment of the index value than a (relatively) large improvement in only one dimension. Mononicity is important because we would like to reward improvements observed in each dimension; convexity is useful because we would like to reward more those countries which care about both dimensions, and do not, so to speak, "specialize" in one of them. The Cobb-Douglas specification is also very flexible, allowing the inclusion of new variables and the use of weights for each dimension reflecting the relative importance attributed to each component.

The "Bidimensional Composite Index of Educational EOp" (or *BIE*) we propose takes then the following simple form:

$$BIE = \underbrace{\bar{p}^a}_{\text{access-EOp}} \cdot \underbrace{(1 - IO_{FG})^b}_{\text{achievement-EOp}} \quad (3.1)$$

with:  $0 \leq \bar{p} \leq 1$ ,  $0 \leq IO_{FG} \leq 1$ ,  $0 \leq BIE \leq 1$ ,  $0 \leq a \leq 1$ ,  $0 \leq b \leq 1$ .

The parameters  $a$  and  $b$  reflect the relative importance of each component, and, for simplicity, it is common to assume  $a+b=1$ .<sup>10</sup> The index is increasing in  $\bar{p}$  and decreasing in  $IO_{FG}$  as would be desirable. Since  $IO_{FG}$  ranges in the interval  $[0,1]$ , with 0 standing for perfect equality of opportunity in achievement,  $(1 - IO_{FG})$  will equal 1 in the case where circumstances are completely unrelated to outcomes for pupils who have taken PISA exams. In such limiting case, *BIE* will depend solely upon the coverage rate: the higher it is, the larger the educational opportunities offered to 15-year-olds of a given country will be.

Conversely, in a case of full coverage, where  $\bar{p} = 1$ , the index *BIE* will depend solely upon inequality of opportunity in achievement – a situation that might make sense for analyses of educational EOp in developed countries, where  $\bar{p}$  does actually tend to 1.

Developing countries more often face the challenge of having to improve simultaneously in both dimensions, facing limited budgets that need to be allocated based on priorities (e.g., long vs short run). In some cases coverage might be easier to tackle than actual learning – but such policy choice will impact less in the index than

---

<sup>9</sup> In the course of this research, we also employed fuzzy sets transformations. As results were qualitatively equal to those reported here, we opted for omitting those based on fuzzy sets.

<sup>10</sup> Cobb Douglas functions are homogeneous which implies that if we multiply each component by a constant, the outcome will be that constant raised to the sum of exponents.



policy choices which conduct to improvements in both dimensions (for reasons already explained).

#### 4. Application

We compare rankings of educational-EOp for selected countries that took part in PISA 2006 and 2009 using  $BIE_1$  and a ranking obtained from a conventional index that only takes into account achievement-EOp.

In calculating the achievement dimension, we follow Ferreira & Gignoux (2011), employing PISA's "plausible values" for Mathematics as a measure of test scores and a set of circumstances including: mother and father education, father occupation, stock of educational capital, city size and ownership of specific durables. In this first application, the weights have been set to  $a = b$ . Results are reported in Table 1.

**Table 1: Estimations using PISA 2006 and 2009**

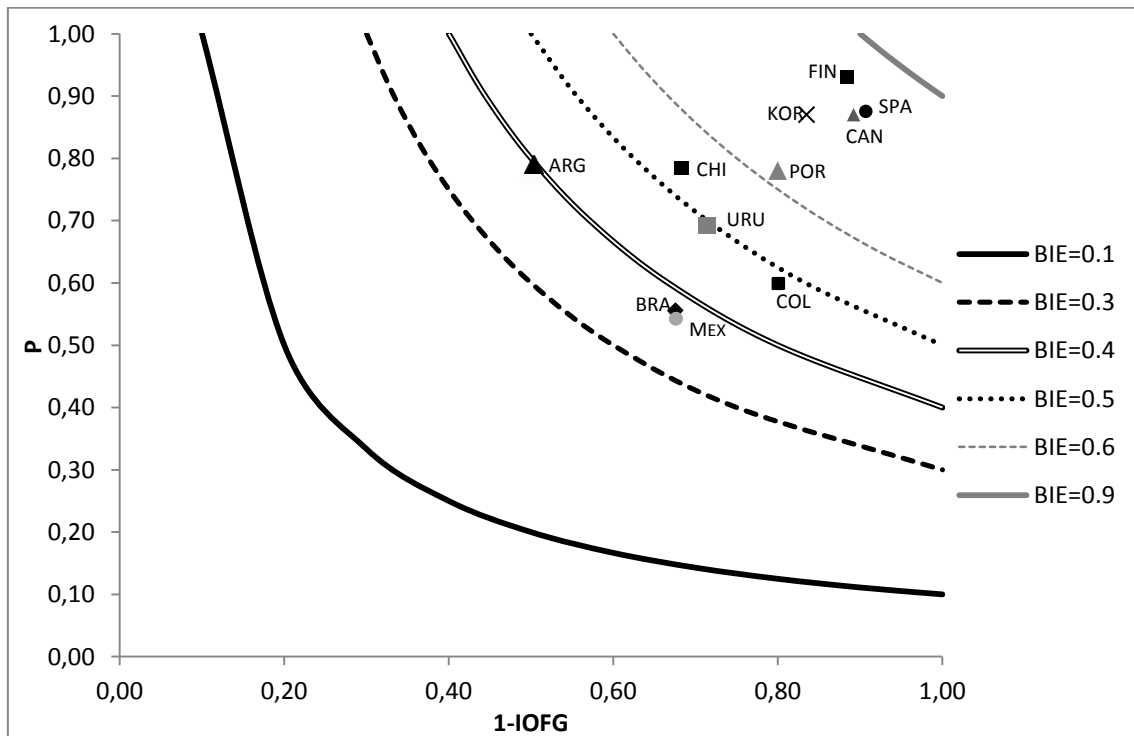
<b>Panel A: 2006</b>					
<b>Country</b>	$\bar{p}$	$IO_{FG}$	$Sd (IO)$	$(1-IO_{FG})$	$BIE$
Argentina	0,79	0,497	0,001	<b>0,503</b>	<b>0,397</b>
Brazil	0,56	0,325	0,001	<b>0,675</b>	<b>0,375</b>
Chile	0,78	0,318	0,001	<b>0,683</b>	<b>0,535</b>
Colombia	0,60	0,199	0,001	<b>0,801</b>	<b>0,480</b>
Mexico	0,54	0,324	0,001	<b>0,676</b>	<b>0,367</b>
Uruguay	0,69	0,286	0,000	<b>0,714</b>	<b>0,494</b>
Portugal	0,78	0,200	0,008	<b>0,800</b>	<b>0,624</b>
Spain	0,87	0,099	0,006	<b>0,901</b>	<b>0,784</b>
Finland	0,93	0,116	0,009	<b>0,884</b>	<b>0,822</b>
Korea	0,87	0,165	0,006	<b>0,835</b>	<b>0,727</b>
Canada	0,87	0,108	0,004	<b>0,892</b>	<b>0,776</b>
<b>Panel B: 2009</b>					
<b>Country</b>	$\bar{p}$	$IO_{FG}$	$Sd (IO)$	$(1-IO_{FG})$	$BIE$
Argentina	0,69	0,55	0,0008	<b>0,449</b>	<b>0,308</b>
Brazil	0,63	0,238	0,0006	<b>0,762</b>	<b>0,481</b>
Chile	0,85	0,269	0,0006	<b>0,731</b>	<b>0,623</b>
Colombia	0,59	0,302	0,0004	<b>0,697</b>	<b>0,408</b>
Mexico	0,61	0,308	0,0006	<b>0,692</b>	<b>0,419</b>
Uruguay	0,63	0,387	0,0006	<b>0,613</b>	<b>0,387</b>
Portugal	0,84	0,206	0,010	<b>0,794</b>	<b>0,667</b>
Spain	0,89	0,076	0,005	<b>0,924</b>	<b>0,822</b>
Finland	0,93	0,122	0,009	<b>0,878</b>	<b>0,817</b>
Korea	0,88	0,156	0,006	<b>0,844</b>	<b>0,743</b>
Canada	0,84	0,134	0,005	<b>0,866</b>	<b>0,727</b>

BIE: equality of opportunity (achievement and access)  
Sd (IO) Standard Error

Both in 2006 and in 2009 rank reversals are observed when switching from an EOp index that focuses exclusively on achievement ( $1 - IO_{FG}$ ) to an index that encompasses both achievement-EOp and in access-EOp ( $BIE$ ) and weighs them equally.<sup>11</sup> Argentina is the most opportunity-unequal country in 2006 in terms of achievement (lowest value in column ( $1 - IO_{FG}$ )), but after taking into account its relatively good coverage rate, it moves up two positions (third-lowest value in column  $BIE$ ). Chile also moves up two positions. Colombia and Mexico do the opposite movement (5<sup>th</sup> to 8<sup>th</sup>, and 9<sup>th</sup> to 11<sup>th</sup>, respectively). In 2009, Brazil and Chile, for example, exchange their positions as we move from column ( $1 - IO_{FG}$ ) to column  $BIE$ .

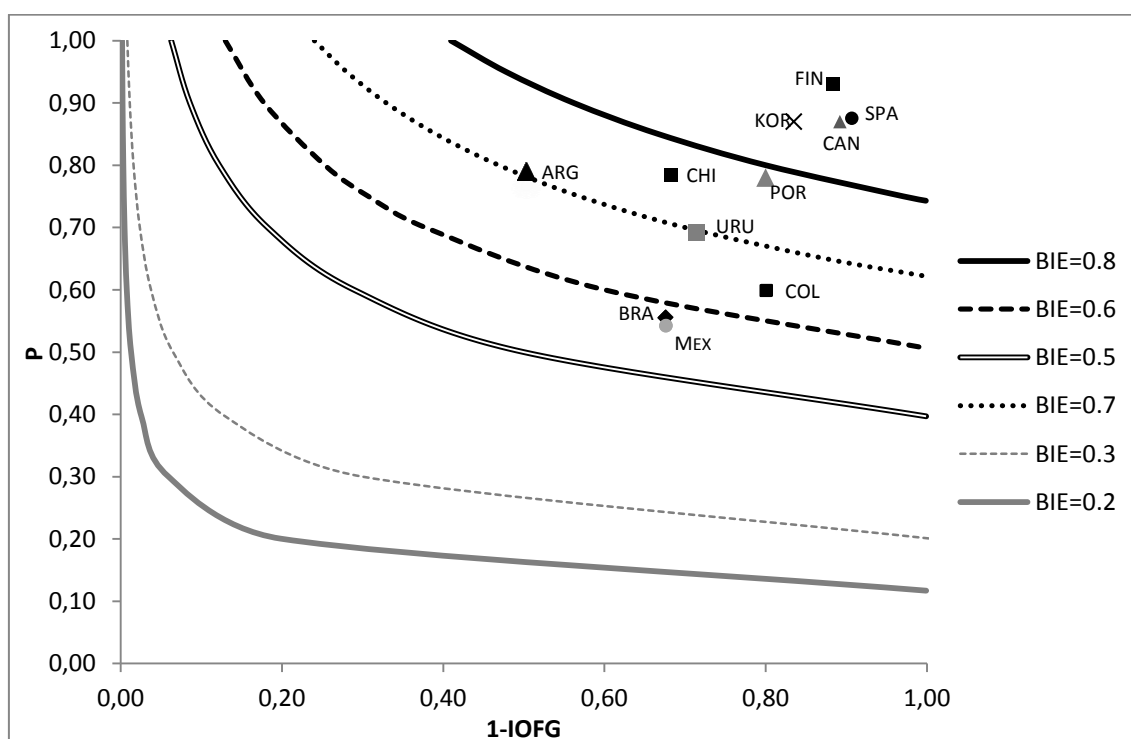
Our results can be plotted over a map of so-called “iso-opportunity curves”, as do Barros et al. (2009), simultaneously showing where a country stands in either dimension (in each axis), and in terms of overall educational EOp (the level curves). Figure 2a contains iso-opportunity curves for PISA 2006. Points representing Brazil and Mexico, which lie below the curve “ $BIE=0,4$ ”, are to be contrasted with Chile’s, which is above the curve “ $BIE=0,5$ ”, or Portugal, above “ $BIE=0,5$ ”. Other developed countries cluster way higher. It is interesting to compare Argentina and Colombia: while the former performs relatively well in the access-EOp dimension and poorly in the achievement-EOp dimension, the latter presents the opposite situation. With equal weights, Colombia reaches a higher iso-opportunity curve.

**Figure 2a. Iso-opportunity curves and EOp in selected countries employing  $BIE$ , with weights  $a = b$ . Source: PISA 2006.**



<sup>11</sup> Bootstrapped standard errors indicate that the difference between the countries’ calculated indices are statistically significant. Results are not reported here, but are available upon request.

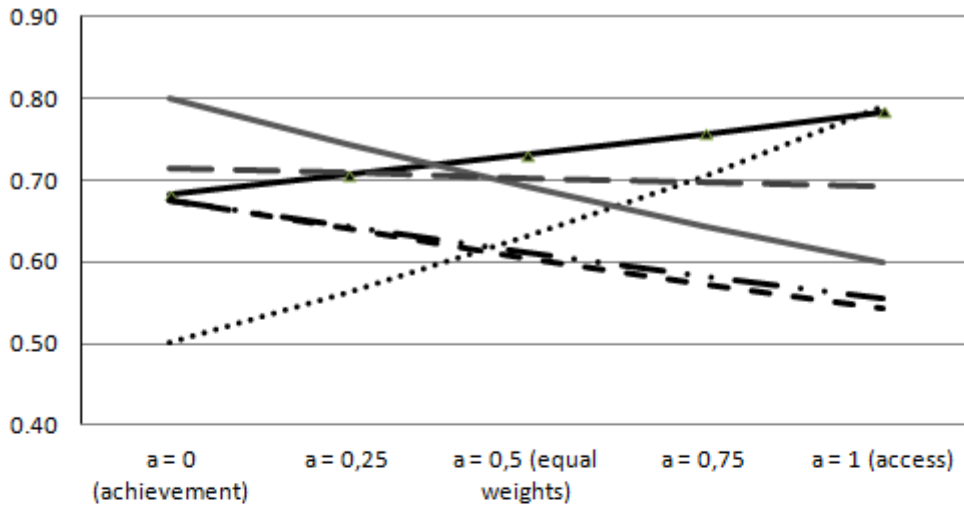
**Figure 2b. Iso-oppportunity curves and EOp in selected countries employing *BIE*, with weights  $a = 3/4$ ;  $b = 1/4$ . Source: PISA 2006.**



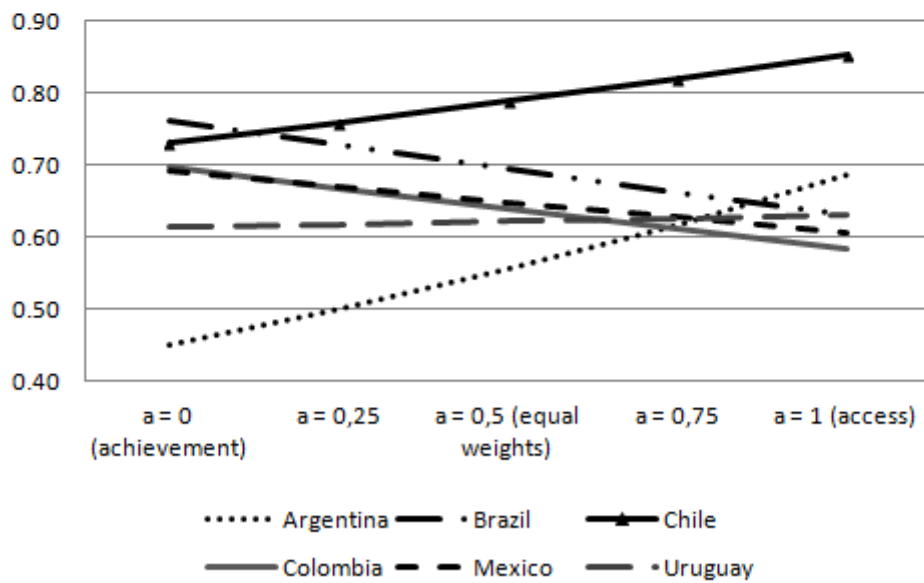
In Figure 2b, we employ different weights for each dimension, privileging the access-EOp dimension ( $a = 3/4$ ;  $b = 1/4$ ). Uruguay and Argentina, which were in different iso-oppportunity curves in Figure 2a, now share the same iso-oppportunity curve, since Argentina performs well in access in comparison to other countries. With the new weights, Argentina's iso-oppportunity curve is higher than Colombia's (while they were tied before). Developed countries in Figure 2b cluster again above the curve "*BIE*=0.8", with the exception of Portugal, which is close to Chile.

In Figure 3, we present a wider range of views on the relative importance of access-EOp and achievement-EOp, by ranking now only our six Latin American countries according to multiple weights. As expected, as we diminish the relative weight attributed to achievement, we approach rankings based on access alone. Figure 3 also allows a direct visual inspection of strengths and weaknesses of different countries in each dimension of educational EOp. For example, in both years Argentina and Chile increase their level of educational EOp as more emphasis is put on access-EOp (going from left to right in the graph). Brazil, Mexico and Colombia follow an opposite pattern. Uruguay's calculated level of educational EOp is roughly flat across different values for the parameter  $a$ .

**Figure 3. Equality of educational opportunity in six Latin American countries as measured by *BIE*, with varying weights. Sources: PISA 2006 and 2009.**



**PANEL A: 2006**



**PANEL B: 2009**

To some extent Figure 3 summarizes all that has been discussed so far, in three senses. Firstly, it confirms the need to take access into account when dealing with PISA-based analyses encompassing developing countries. Secondly, it shows that different observers might rank differently a given set of countries and thus reach different conclusions regarding educational EOp according to the relative weight attributed to each dimension. Thirdly, it attests that conventional rankings of educational EOp – either based on access alone or on achievement alone – can be viewed as particular cases of *BIE*.

## 5. Generalization

To employ the composite indices defined in previous sections means working with social welfare functions – or, more precisely, “equality-of-opportunity functions” – based upon certain attributes,  $BIE = f(\text{access-EOp, achievement-EOp})$ . However, other attributes could be incorporated, as well as sub-attributes. Over this aggregation process, each dimension might vary its relative importance providing different interpretations and rankings. More generally, the index could be composite, but with multiple components.

Average scores, for example, might be a relevant dimension of opportunities. Education is certainly a “positional good”, in the sense that various advantages that derive from having more of it depend on the relative level a person has as compared to those living in the same city, or province, or country. For most jobs, for example, an employer will choose the best candidate - e.g., the one with the best education record – among a pool of local competitors. In that sense, what matters is only the educational achievement of a person as compared to those around her. But it might also be true that for some jobs – in an increasingly globalized economy – competition is at an international level. If this is so, it will be useless to have the highest achievement in your locality if that is low at a more global level. Even if we ignore international jobs (which arguably are not the most frequent), it still is possible to argue that higher average scores are better than lower average scores. It might be better to live in a city, or province, or country, where the overall level of achievement is higher than elsewhere. There might be positive spillovers from education, such that, for example, more jobs are created in more dynamic places.

**Table 2: Composite index with multiple dimensions**

Panel A: 2006						
Country	$\bar{p}$	$IO_{FG}$	$BIE$	Average score	$\bar{s}$	MIE
Argentina	0,79	0,497	0,397	381	0,544	0,216
Brazil	0,555	0,325	0,375	370	0,529	0,198
Chile	0,784	0,317	0,535	411	0,587	0,314
Colombia	0,599	0,199	0,480	370	0,529	0,254
Mexico	0,542	0,324	0,366	406	0,580	0,213
Uruguay	0,692	0,286	0,494	427	0,610	0,301
Portugal	0,78	0,200	0,624	466	0,666	0,415
Spain	0,87	0,099	0,784	480	0,686	0,538
Finland	0,93	0,116	0,822	548	0,783	0,644
Korea	0,87	0,165	0,727	547	0,781	0,568
Canada	0,87	0,108	0,776	527	0,753	0,584

In Table 2, we show both for 2006 and 2009 the calculated value of multiplying  $BIE$  reported in Table 1 by  $(\bar{s})$ , defined as ratio of the average score in each country to the maximum possible score in PISA (700). Mexico and Brazil exchange positions in this new ranking as compared to Table 1’s ranking due to Brazil’s extremely low average scores. The same happens in 2009 between Colombia and Uruguay. With this new

index, figures obtained from Finland, Korea, and Canada are much higher than those obtained in the Latin American countries, implying another layer of advantage concerning the provision of opportunities of the former countries over the latter ones.

**Table 2 (cont.): Composite index with multiple dimensions**

Panel B: 2009						
Country	$\bar{p}$	$IO_{FG}$	$BIE$	Average score	$\bar{s}$	MIE
Argentina	0,686	0,55	0,309	388	0,554	0,171
Brazil	0,632	0,238	0,482	386	0,551	0,266
Chile	0,852	0,268	0,624	421	0,601	0,375
Colombia	0,585	0,303	0,408	381	0,544	0,222
Mexico	0,607	0,308	0,420	419	0,599	0,251
Uruguay	0,631	0,387	0,387	427	0,610	0,236
Portugal	0,84	0,206	0,667	487	0,696	0,464
Spain	0,89	0,076	0,822	483	0,690	0,567
Finland	0,93	0,122	0,817	541	0,773	0,631
Korea	0,88	0,156	0,743	546	0,780	0,579
Canada	0,84	0,134	0,727	527	0,753	0,548

Not only might the achievement dimension of the index be enriched – for example, by taking into account average scores as mentioned above. The access dimension can also be sophisticated, for example, in order to account for cross-groups inequalities in access, as measured, for instance, by the dissimilarity index – already alluded to in this paper.<sup>12</sup> We could then indeed have four components in a meaningful multidimensional composite index of equality of educational opportunity (*MIE*), which would now include two new terms: cross-types dissimilarity with respect to access ( $1 - D$ ), and a country's average score ( $\bar{s}$ ), suitably transformed to fit the  $[0,1]$  interval.

$$MIE = \underbrace{[\underbrace{\bar{p}^\alpha}_{\text{overall across-types}} \cdot \underbrace{(1-D)}_{\text{access-EOp}}]}_{\text{access-EOp}} \cdot \underbrace{[\underbrace{\bar{s}^\gamma}_{\text{overall across-types}} \cdot \underbrace{(1-IO_{FG})^\delta}_{\text{achievement-EOp}}]}_{\text{achievement-EOp}} \quad (3.2)$$

where:  $0 \leq \bar{p} \leq 1$  and  $0 \leq IO_{FG} \leq 1$  are defined as in Equation 3.1;  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  are (normative) weights, all of them nonnegative and normalized to sum 1;  $0 \leq D \leq 1$  stands for a dissimilarity index, which aggregates the distance between the average coverage rate and each type's coverage rate weighted by the relative frequency of each type in the population;  $\bar{s}$  represents average scores transformed to fit the  $[0,1]$  interval.

<sup>12</sup> Although in that point we rejected the use of the human opportunity index or the dissimilarity index, in the basis of conceptual and practical disadvantages, and defended the use of PISA's coverage rate, it might be the case that other researchers or policymakers have a different view or find it worth it to invest time in gathering and ancillary, national, datasets. This is why we include the dissimilarity index in this version of the index.

In Equation 3.2, the term  $(\bar{s})$  is within the achievement-EOp dimension the analogue of  $(\bar{p})$  within the access-EOp dimension, in the sense that both reflect overall educational opportunities available in the country (average scores and overall access, respectively). Similarly the remaining terms,  $(1 - D)$  and  $(1 - IO_{FG})$ , are analogous since both indicate the way the available opportunities (for access and for achievement, respectively) are distributed across groups.

As argued before, different persons might value differently each dimension of equality of educational opportunity. If someone thinks average scores might have nothing to do with equality of opportunity, it simply requires setting  $\gamma = 0$ . The opposite extreme view is one which would ignore the distributional issues captured by  $\bar{p}$ ,  $D$  and  $IO_{FG}$  and care only about average scores, which would only require setting  $\gamma = 1$  and  $\alpha = \beta = \delta = 0$ . All those cases can be viewed as particular cases of the more general *MIE*.

## 6. Final remarks

Measuring educational-EOp has been the focus of recent contributions in the economic literature, which concentrate either on opportunity for access to a given level of studies, or on opportunity in terms of educational achievement. In this paper, we combine both concerns. The method introduced here points out the inappropriateness of EOp indices that overlook participation especially in developing countries where many young individuals still abandon the educational system in early years of their lives, temporarily or for good – in either case, becoming ineligible to PISA and probably facing the prospects of limited life opportunities.

The computation of any multidimensional index requires solving the difficult issues of assigning weights to each dimension in the aggregation step and the choice of an appropriate aggregation method. Our application to selected countries shows how rankings approach unidimensional-EOp orderings as relative weights tend to limiting values (as would be expected). For a sufficiently large weight attributed to the access dimension, countries are ranked very differently than they are by EOp indices regarding achievement only, confirming the need to take access into account when dealing with PISA-based analyses encompassing developing countries.

It is important to notice that, while motivated by PISA's problem, the index offered here is applicable with due adaptations to analyses that employ other datasets presenting similar problems. In a way, any computation of achievement-EOp which ignores the access dimension will miscalculate the actual education opportunities provided in a country.

In addition to incorporating further dimensions, as proposed in section 5 through *MIE*, there are other paths for deepening or extending what has been done here. Regarding the achievement-EOp dimension, the sensitivity of results to different sets of circumstances included in the regression-based estimations could be tested. As for the access-EOp dimension, the results presented here could be compared to those relying on other definitions of access-EOp (enrolment rate, for example). It would also be interesting to study the effect of employing alternative aggregating procedures. Finally it might be useful to decompose the contribution to educational-EOp of different dimensions and sub dimensions, and to observe how they evolve over time, possibly connecting the observed evolution to policies or exogenous shocks affecting pupils and their families.

## References

- Barr, N. (2012), *Economics of the welfare state*, 5<sup>th</sup> edition, Oxford: Oxford Univ. Press.
- Cecchi, D. ; V. Peragine (2010), "Inequality of Opportunity in Italy". *Journal of Economic Inequality*, 8, pp. 429-450
- Cecchi, D., V. Peragine (2005), "Regional Disparities and Inequality of Opportunity: The Case of Italy". IZA Discussion Papers 1874, Institute for the Study of Labor (IZA)
- DiNardo, J., Fortin, N. and Lemieux, T. (1996), "Labor market institutions and the distribution of wages 1973–1992: a semiparametric approach". *Econometrica* 64 (5): pp. 1001–1044.
- Dunzlauf, L., Neumann, D., Niehues, J., Peichl, A. (2010), "Equality of opportunity and redistribution in Europe". IZA Discussion Papers 5375, Institute for the Study of Labor (IZA).
- Ferreira, F., Gignoux, J. (2011), "The measurement of educational inequality: Achievement and opportunity". *Working Papers 240*, ECINEQ, Society for the Study of Economic Inequality.
- Gamboa, LF.; F.D. Waltenberg (2012). "Inequality of opportunity in educational achievement in Latin America: Evidence from PISA 2006-2009". *Economics of Education Review*, 31(5), pp. 694-708, October
- OECD (2009), *PISA 2006 Technical Report*, Paris: OECD. Retrieved from [www.oecd.org/dataoecd/0/47/42025182.pdf](http://www.oecd.org/dataoecd/0/47/42025182.pdf).
- OECD (2012), *PISA 2009 Technical Report*, Paris: OECD Publishing <http://dx.doi.org/10.1787/9789264167872-en>.
- Paes de Barros R., Ferreira F., Molinas J., and Saavedra J. (2009), *Measuring Inequality of Opportunities in Latin America and the Caribbean*. Washington, DC: World Bank.
- Peragine, V.: Review of "Measuring inequality of opportunities in Latin America and the Caribbean", by Paes de Barros, R., Ferreira, F., Molinas, J., Saavedra, J.:World Bank and Palgrave Macmillan, 2009, *Journal of Economic Inequality* (2011) 9:137–143. DOI 10.1007/s10888-010-9151-2 (2010)
- Pignataro, G. (2012), "Equality of opportunity: policy and measurement paradigms". *Journal of Economic Surveys*. 26 (5), 800-834. doi: 10.1111/j.1467-6419.2011.00679.x.
- Ramos, X., Van de Gaer, D. (2012), "Empirical approaches to inequality of opportunity: Principles, measures, and evidence." *Working Papers 259*, ECINEQ, Society for the Study of Economic Inequality
- Roemer, J. (1998), *Equality of opportunity*. Cambridge, MA: Harvard University Press.
- Vega, J. R. M., Paes de Barros, R.P. de, Saavedra, J., Guibale, M. (2010), *Do our children have a chance? The 2010 Human Opportunity Report for Latin América and the Caribbean*. World Bank, Washington, DC, 176p
- Yalonetzky, G. (2012). "A dissimilarity index of multidimensional inequality of opportunity". *Journal of Economic Inequality*, 10(3), pp. 343-373. DOI 10.1007/s10888-010-9162-z September