Evaluating education systems

Nicolas Gravel[*], Edward Levavasseur[†] and Patrick Moyes[‡]

Preliminary and Incomplete.

May 10th 2017

Abstract

This paper develops and implements dominance criteria for evaluating the performance of *compulsory school systems*.The main criterion that we propose is shown to be the smallest transitive criterion compatible with three elementary principles for evaluating such school systems. The first principle requires that improving the cognitive skill of a children of a given family background is a good thing. The second principle says that the evaluation attached to a given cognitive skill of a children is all the more favorable as the children is coming from a family with an unfavorable background. The third principle says that, for a given distribution of the children cognitive skills and family backgrounds, reducing the correlation between family background and cognitive skill is a good thing. Our dominance criterion considers that school system A is better than school system B if, for any pair of reference family background and cognitive skill, the fraction of the children who have a better cognitive skills and are coming from a worst family background than the reference is weakly larger in A than in B. We then apply our criterion to the ranking of education systems of major OECD countries, taking the standardized PISA scores as the measure of cognitive skills, and considering in turns various indices of the family backgrounds. We show that, albeit incomplete, our criterion enables the comparisons of quite a few educational systems. Educational systems of fast growing asian economies - and in particular Vietnam, appear rather at the top of our rankings while those of wealthy arabic countries such as Qatar or Arab Emirates appear at the bottom. We also consider the possibility of extending our criterion by incorporating some additional value judgement about varying inequality in cognitive skills. The extension gives rise to more discriminatory dominance criteria that we also apply to the ranking of national education systems.

PRELIMINARY AND INCOMPLETE DRAFT

"Schools are remarkably similar in the effect they have on the achievement of their pupils when the socioeconomic background of the students is taken into account."

James Coleman, 1967

[*]Aix-Marseille Univ. (Aix-Marseille School of Economics), CNRS & EHESS, Centre de la Vieille Charité, 2, rue de la Charité, 13002 Marseille Cedex France, Nicolas.gravel@univ-amu.fr

[†]Aix-Marseille Univ. (Aix-Marseille School of Economics), CNRS & EHESS, Centre de la Vieille Charité, 2, rue de la Charité, 13002 Marseille Cedex France, Nicolas.gravel@univ-amu.fr

[‡]CNRS, Université de Bordeaux IV, Patrick.Moyes@orange.fr

1 Introduction

All countries in the world have *compulsory education systems.* These education systems, that are usually made of a mixture of public and private schools that follow specific learning curricula, take children at the age of five, and enroll them in learning programs for about 10 to 12 years, depending upon the country. The result of this enrollment is the acquisition, by the children, of various *cognitive skills* that are of obvious importance for their future welfare. For one things, the cognitive skills are important determinants of the future earnings and employment opportunities of these children (see e.g. Hanushek and Woessmann (2008), Hanushek, Schwerdt, and Woessmann (2015) or Nickell (2004)). But the acquisition of cognitive skills in mathematics, literacy, etc. may also impact individuals' well-being in a way that is not reducible to their pecuniary consequences, however important these may be. As noticed by many (for example Oreopoulos and Salvanes (2011)) cognitive skills may indeed fosters future information acquisition, and help individuals to make better decisions about health, spouse partnership, parental choices, etc. Hence one of the most important standard by which the performance of a compulsory education system can be appraised is through the distribution of cognitive skills acquired by the children at the end of the compulsory curriculum. There are by now a few internationally standardized procedures for measuring these cognitive skills on a regular basis on some suitably chosen samples of children and gathering data on those. One of the most largely commented and discussed such data set is the Programme for International Student Assessment (PISA), which tests math, science, and reading performance of 15-year-olds children on a three-year cycle since 2000. An excellent presentation of these data sources is provided in Hanushek and Woessmann (2011).

It is also widely acknowledged, (see e.g. Mayer (1997), Black, Devereux, and Salvanes (2005), Schutz, Ursprung, and Woessmann (2008) or Dahl and Lochner (2012)) that the children's family background plays a determinant role in this process of cognitive skills acquisition. As one sociologist supposedly put it to the scholar-politician Daniel Patrick Moynihan in reaction to the famous Coleman report - quoted above - on the educational opportunities offered by American schools in the sixties: "Have you heard what Coleman is finding? It's all family." The precise channels through which family backgrounds - as measured by parental income or education or, sometimes (see e. g. Hanushek and Woessmann (2011)) by the number of books at home - affect the acquisition process of the children cognitive skills is still subject to discussion. One of these channels may be genetic. After all, high income - or important books purchasers - parents may also tend to be parents with genetic traits that favor cognitive skill acquisition. Another may be the time and energy spent by the parents in helping the children to acquire those skills. But whatever the channel is, the influence of the family in the children cognitive skill acquisition process must be accounted for when evaluating the global performance of a compulsory school systems. Two school systems who produce the same distribution of cognitive skills can not be considered as equally performing if the distribution of the children family backgrounds differ between the two. Moreover, there is a widely held impression , often developed - again, after the celebrated Coleman report of the sixties - under the heading of "equality of opportunity" (see Schutz, Ursprung, and Woessmann (2008)), that good compulsory school systems are those that succeed somehow in breaking the dependency of the children skill acquisition process upon the family circumstances.

This paper proposes a dominance methodology for comparing alternative compulsory education systems on the basis of a few explicit *elementary principles* that capture what we believe to be common intuitions about what make an education system indisputably performing. The principles that we formulate are applicable to *data* on those education systems (for example, as used in this paper, the PISA data). The approach developed in this paper applies specifically to data where every children observed at the end of the compulsory school system is described by two numbers: one measuring his/her cognitive skill and the other measuring his/her family background. Viewed in this way, the issue of comparing education systems amounts to comparing distributions of pairs of numbers, just as in the traditional multi - actually bi -dimensional normative evaluation developed along the lines of Atkinson and Bourguignon (1982) (see e.g. Atkinson and Bourguignon (1987), Bourguignon (1989), Jenkins and Lambert (1993), Gravel and Moyes (2012),Moyes (2012) for additional theoretical contributions and Duclos, Sahn, and Younger (2006), Gravel, Moyes, and Tarroux (2009), Gravel and Mukhopadhyay (2010) or Hussain, Jorgensen, and Osterdal (2016) for examples of empirical applications). However, the particular nature of the two numbers involved in the description of school systems suggests principles for comparing them that may differ from those considered in the bi-dimensionnal analysis of the Atkinson and Bourguignon (1982) variety.

The *first* principle for comparing education systems that seems hardly disputable is first order dominance in cognitive skills given the family backgrounds. Consider indeed two education systems in which the distribution of the children's family backgrounds is the same. Suppose that, for every conceivable reference level of cognitive skill, the fraction of children with a better cognitive skill than that reference is larger in one education system than in the other. Since family backgrounds are equally distributed in the two education systems, such a first order dominance of the distribution of skills in one education system over that of the other would suggest a clear superiority of the former over the later. Many popular discussions about the relative merit of the different national education systems, notably around the releases of PISA studies, clearly agree with this principle.

The *second* principle concerns the dual situation of two hypothetical educational systems with the same distribution of cognitive skills. Suppose however that the distribution of the children family backgrounds in one education system first order dominates that of the other. This indicates, in a strong sense, that the distribution of family backgrounds is less favorable in one school system than in the other. In such a case, it is arguable to consider that the performance in cognitive skills - observed to be the same in the two systems - is all the *more* impressive as it happens in the system with the *less* favorable distribution of family backgrounds. Put differently, a particular achievement in cognitive skills is more admirable when observed in a child from a low family background than it is when observed in a child with a more favorable one.

The *third* principle reflects the preference alluded to above for education systems who reduce the influence of the family background on the process of cognitive skill acquisition. Consider indeed an education system in which one child with a good family background achieves a high level of cognitive skills while another child with a less favorable family background obtains a lower level of such skills. Consider another school system identical to this one in every respect other than the fact that the high skills child is now coming from the low family background and the low skill one is coming from the high family background. Aversion to correlation would suggest, in a simple case like this one, that the second school system performs better than the first. In effect, the second school system has broken one correlation link between family background and cognitive skill without affecting either the marginal distribution of skills or the marginal distribution of family background in the population. At least all empirical studies - such as Schutz, Ursprung, and Woessmann (2008) - who regress the skills variable over a set of explanatory variables - including of course some that measure family background - and who compare school systems based on the value of the regression coefficient of the family variable would agree with this principle.

In this paper, we show that any transitive application of the combination of the three principles will agree with an easily applicable *dominance criterion* for comparing education systems. The dominance criterion says that one education system dominates another if, for any pair of reference levels of skills and family background, the fraction of the children population with *both* a *better* cognitive skill *and* a *lower* family background than that reference is larger in the dominating system than in the dominated one. This criterion shares with one of the first order criteria of Atkinson and Bourguignon (1982) - when applied to education systems - the agreement with the first and the third principles. However, it differs from the Atkinson and Bourguignon (1982) criterion in considering the second attribute - family status - as having a *negative* impact on the performance of a education system. Atkinson himself (see especially Atkinson (1981a) and Atkinson (1981b)) has applied one of the first order dominance criteria of Atkinson and Bourguignon (1982) to the issue of measuring intergenerational income mobility (see e.g. Fields and Oke (1999) for a survey on income mobility measurement). By so doing, he endorsed the view that improving the distribution of parental status ceteris paribus improves intergenerational mobility. While this view may be defensible for evaluating intergenerational income mobility - at least if one is adopting for that purpose the perspective of Shorrocks (1978) - it is less so when appraisal of education system performance is at stake. We believe, in effect, that a given distribution of cognitive skills is all the more favorably appraised - as an output of the education system - as the children to which it is transmitted are coming from unfavorable family backgrounds.

The three principles just sketched, and the dominance criterion that they characterize, form the bulk of the analysis of this paper. As we show in the empirical analysis, they alone lead to interesting comparisons of national education systems. Yet, since they ride on very consensual principles, they remain fairly incomplete. If one wants to increase the number of education systems that can be compared, it may be appropriate to invoke additional principles. One of them concerns attitude toward inequalities in cognitive skills. Making such an attitude precise requires, when developed in the conventional framework of inequality measurement, that cardinal significance be attached to the measurement of cognitive skills.[1] Provided that we accept this framework, are inequalities in cognitive skills - ceteris paribus of course- a good or a bad thing ?

[1] See e.g. Gravel, Magdalou, and Moyes (2015) for an approach to inequality measurement based on ordinal variables.

While spontaneous intuition - such as that underlying the empirical analysis of Goussé and LeDonné (2015) - seems to favour the second rather than the first answer to this question, a second thought may make one more hesitant. This is at least so if one recognizes, in line with much of the empirical literature (see e.g. Green and Riddell (2003), Heckman, J. Stixrud, and Uzrua (2006) or Barrett (2012)), that income is a convex function of cognitive skills, and that cognitive skill may contribute to well-being in a way that is not reducible to its pecuniary consequence. If this is the case, it is possible that the conversion of cognitive skills into final well being - both through the indirect effect that cognitive skill has on income and its direct intrinsic effect on well-being - be done by a *convex* function. If individual well-being is a convex function of cognitive skills, then a utilitarian ethical observer - or social planner - could favour increasing inequalities - as defined by mean-preserving spread - in cognitive skills, everything else being the same. Following Bazen and Moyes (2012), we call *elitism* the favorable appreciation of mean preserving spreads in cognitive skills when those are performed between children with the same family background. As it happens, when we add elitism to the set of other principles, we characterize an additional dominance criteria that is compatible with the previous one while being much more discriminatory. Unfortunately, and for reasons that we believe to be rather deep and beyond the scope of this paper, we are not capable of characterizing a dominance criterion that respect the three first principles when we replace elitism by the alternative traditional - but nonetheless also defensible - egalitarian view that dislikes such mean preserving spread of cognitive skills. However, as shown below, we can identify dominance criteria which, when it ranks conclusively two education systems, does so in a way that is agreed upon by the three first principles and an aversion to cognitive skills inequality. However we can not establish the converse implication that the failure of this criterion to rank one education system above another implies a conflict between egalitarianism and one of the three aforementioned principles.

We then put our dominance criteria to work by comparing the national education systems based on the 2012 wave of the PISA survey. We specifically compare accross countries the joint distributions of the children's scores in mathematics - as measured by PISA tests- and their parents social status (defined to be the highest International Socio-economic index of the two parents). The most discriminatory criteria - who adds elitism to the three first principles discussed above - are capable of establish a conclusive ranking in about 25% of all pairs of countries. The percentage of clear-cut comparisons obtained from the three core principles alone - without introducing any elitist principles - is only 16%. While these fractions of successful comparisons may be considered small, the robustness of the obtained conclusion is worth emphasizing. Among the noteworthy robust comparisons, one finds that Vietnam has one of the most performant education system in the world. In effect, the Vietnamese education system dominates 33 out of the 45 other countries to which it can be compared, and is dominated by none! To some extent, this reflects the fact that Vietnamese children do very well in their Pisa test even though they come from parents with relatively low status. Among the western developped countries, Japan appears to stand the best against the others.. Its education system dominates indeed that of 16 other countries, and is never dominated. We find also interesting that Finland, often described has having one of the best education system in the world, does not perform particularly well by our criteria. Finland education system dominates those of only four other countries, and is dominated by that of Japan. To some extent, this "average" performance of Finland may be due to the elitist principle on which some of our comparisons are based. Yet, the domination of Finland by Japan is observed even if one focuses on the core three first principles. At the bottom of our rankings, one finds countries such as Jordan and the Arab Emirates. The education systems of Jordan is, in effect, dominated by that of 35 other countries, and dominates none of them. This extremely poor performance is followed closely by that of the Arab Emirates (dominated by 32 countries, without dominating anyone).

The rest of the paper is organized as follows. The next section presents the critera and principles used to compare education systems, and establish the equivalence among them. Section 3 discusses the data and the empirical methodology, section 4 shows and discusses the empirical results and section 5 concludes.

2 Criteria for comparing education systems

2.1 Framework and notation

We are interested in comparing alternative compulsory education systems. Every such system educates a set of n children (with $n \geqq 3$).[2] At the end of the education process, every child acquires a cognitive **s**kill s that is taken from some interval $\mathcal{S} = [\underline{s}, \overline{s}]$ of the set $\mathbb{R}_+$ of real numbers. A given child is also described by his/her family **b**ackground b that is taken from some interval $\mathcal{B} = [\underline{b}, \overline{b}]$ of $\mathbb{R}_+$. For some of the theoretical results below, we find convenient to assume that the numbers $\underline{s}$, $\overline{s}$, $\underline{b}$ and $\overline{b}$ are all rational.

More compactly, an **e**ducation system $\mathbf{e}$ can be described as an $n \times 2$ matrix:

$$\mathbf{e} \equiv (\mathbf{s}; \mathbf{b}) := \begin{bmatrix} s_1 & b_1 \\ \vdots & \vdots \\ s_i & b_i \\ \vdots & \vdots \\ s_n & b_n \end{bmatrix}, \tag{1}$$

where $e_i = (s_i, b_i) \in \mathcal{S} \times \mathcal{B}$ describes child i cognitive ability (s_i) and family background (b_i). We let $\mathcal{E} = (\mathcal{S} \times \mathcal{B})^n$ denote the set of all such possible education systems.

We find useful to represent an education system by a (joint) distribution of the two variables. Specifically, given an education system $\mathbf{e} \equiv (\mathbf{s}; \mathbf{b}) \in \mathcal{E}$, we denote by $N^{\mathbf{e}}(s, b) = \{i \in N : s_i = a \text{ and } b_i = b\}$ the set of children in $\mathbf{e}$ with cognitive skill s and family background b. The (discrete) *joint density function*of $\mathbf{e}$, denoted $f^{\mathbf{e}}$, is defined by:

$$f^{\mathbf{e}}(s, b) = n^{\mathbf{e}}(s, b)/n, \ \forall (s, b) \in \mathcal{S} \times \mathcal{B} \tag{2}$$

[2] The assumption that all school systems educate the same number of children is not essential, provided of course that one adheres the Dalton principle according to which replicating finitely many time a given population of children is a matter of indifference.

where $n^{\mathbf{e}}(s,b) = \#N^{\mathbf{e}}(s,b)$. We also denote by $\boldsymbol{\sigma}(\mathbf{e})$ and $\boldsymbol{\beta}(\mathbf{e})$ the supports of the variables s and b (respectively) in the education system $\mathbf{e}$ defined by:

$$\begin{aligned}\boldsymbol{\sigma}(\mathbf{e}) &= \{s \in \mathcal{S} : \exists i \text{ such that } s_i = s\} \text{ and,}\\ \boldsymbol{\beta}(\mathbf{e}) &= \{b \in \beta : \exists i \text{ such that } b_i = b\}\end{aligned}$$

Hence $\boldsymbol{\sigma}(\mathbf{e})$ and $\boldsymbol{\beta}(\mathbf{e})$ are the set of values of the cognitive skills and family backgrounds that are effectively observed in the education system $\mathbf{e}$. We also assume that $\boldsymbol{\sigma}(\mathbf{e})$ and $\boldsymbol{\beta}(\mathbf{e})$ are made of rational numbers for any education system $\mathbf{e}$[3]. We also indicate by $R^{\mathbf{e}}(s,b) = \{i : s_i \geq s \text{ and } b_i = b\}$ and $Q^{\mathbf{e}}(s,b) = \{i : s_i \geq s \text{ and } b_i \leqq b\}$ the set of children in education system $\mathbf{e}$ who have cognitive skills no smaller than s and a family background equal to b (for $R^{\mathbf{e}}(s,b)$) and no-greater than b (for $Q^{\mathbf{e}}(s,b)$). Finally, for any target s of skill level and b of family background, we denote by $S^{\mathbf{e}}(s,b)$ the **s**uccess relative to s for children with background b defined by:

$$S^{\mathbf{e}}(s,b) = \sum_{i \in R^{\mathbf{e}}(s,b)} (s_i - s)/n$$

This success is defined as the sum, taken over all kids with a better skill than s, of their skill excess over s. This expression will play a key role in the definition of the second dominance criterion based on elitism.

2.2 The dominance approach

We now formulate basic principles that could plausibly underlie comparisons of alternative education systems. Three such principles drive our attention herein. The first of those concerns the favorable appraisal of the impact that an improvement in the cognitive skills of a child can have when this improvement is achieved without any modification in the distribution of income status among parents. Specifically, consider the following definition of an *improvement in a child cognitive skills*.

Definition 1 Improvement in a child cognitive skills. *We say that education system* $\mathbf{e}$ *is obtained from education system* $\mathbf{e}^{\circ}$ *by means of an improvement in a child's cognitive skills if there exists a child* i *such that:*

$$s_i > s_i^{\circ},\ b_i = b_i^{\circ} \text{ and } (s_g, b_g) = (s_g^{\circ}, b_g^{\circ})\ \forall\ g \neq i. \tag{3}$$

In words, an improvement in a child cognitive skills describes the hypothetical situation where one child sees his/her cognitive skills improving, everything else - among other children or among parents - remaining the same. Any such improvement would naturally be considered favorably by an evaluation of an education system.

Consider now the dual situation of two education systems that differ only by the fact that two children with the *same cognitive skill* in the two systems are coming from *different family backgrounds*. In which of the two education systems is this cognitive skill achievement the most remarkable? It would seem that the cognitive skill should be considered more remarkable when observed in

[3] Rationality of the elements of $\boldsymbol{\beta}(\mathbf{e})$ plays no role in the analysis. The assumption that the elements of $\boldsymbol{\sigma}(\mathbf{e})$ are rational is used in the proof of Theorem 2 below.

a child with *low* family background than when observed in a child with a higher one. We define in this spirit as follows the notion of a *deterioration in the family background of a child with a given cognitive skill*.

Definition 2 Deterioration in the family background of a child with a given cognitive skill. *We say that education system* $\mathbf{e}$ *results from system* $\mathbf{e}^{\circ}$ *by means of a deterioration in the family background of a child with a given cognitive skill if there exists a child* i *such that:*

$$s_i = s_i^{\circ},\ b_i < b_i^{\circ} \textit{ and } (s_g, b_g) = (s_g^{\circ}, b_g^{\circ}),\ \forall\ g \neq i. \tag{4}$$

Improvements in a child cognitive skills - given family backgrounds - and Deteriorations in the family background of a child with a given skill are two clear instances of an improvement in the education system. The third one that we consider concerns the extent to which an education system *reduces the correlation* between the child's cognitive skill and the child's family background. To use the terminology of Daniel Patrick Moynihan mentioned earlier, a good education system is one in which "it is not all family". And a better education system than another is one in which it "less family" than in the other. We formulate this reduction in the correlation between cognitive skill and family background in the following minimalist fashion.

Definition 3 Reduction in correlation between cognitive skill and family background. *We say that education system* $\mathbf{e}$ *is obtained from education system* $\mathbf{e}^{\circ}$ *by means of a reduction in correlation between skill and family background if there exist two children* i *and* j $(i \neq j)$ *such that:*

$$s_j = s_i^{\circ} < s_j^{\circ} = s_i;\ b_i = b_i^{\circ} < b_j^{\circ} = b_j \tag{5}$$

and:

$$(s_g, b_g) = (s_g^{\circ}, b_g^{\circ}),\ \forall\ g \neq i, j. \tag{6}$$

Hence, a reduction in correlation between cognitive skill and family background is just switch of cognitive skills between two children who can be ordered by *both* their cognitive skills *and* their family background. We have illustrated in Figure 1 a reduction in correlation between cognitive skill and family background involving two children i and j whose initial and final situations are $(s_i^{\circ}, b_i^{\circ}) = (1, 1)$, $(s_j^{\circ}, b_j^{\circ}) = (2, 2)$, $(s_i, b_i) = (2, 1)$, $(s_j, b_j) = (1, 2)$.

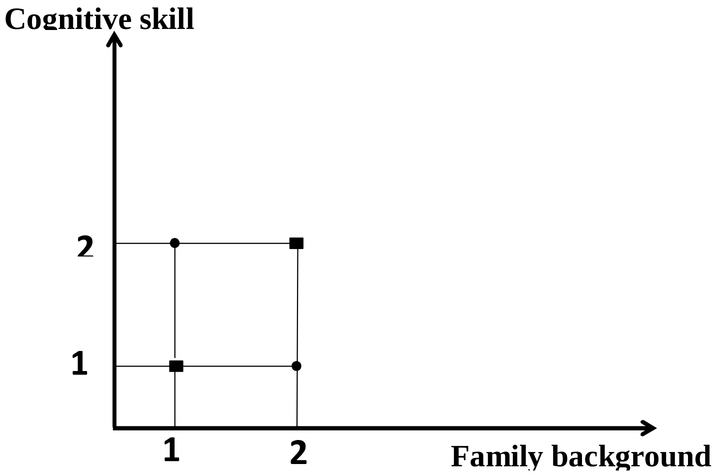

Figure 1: A reduction in correlation between skill and family background.

It is important to notice that this clear reduction in correlation between cognitive skill and family background reduces indeed the positive correlation – or equivalently the positive association – between children skills and family backgrounds without affecting whatsoever the marginal distributions of skills and family backgrounds.

A possible way to compare education systems in $\mathcal{E}$ is resort to the value assigned to these by an *Index* $I(\mathbf{e})$. With such an approach, the statement $I(\mathbf{e}) \geq I(\mathbf{e}')$ means that education system $\mathbf{e}$ performs better than education system $\mathbf{e}'$ in producing cognitive skills in a way that mitigates the influence of family background. What properties could such a function I satisfy so as to serve as a "sensible" way of measuring school system performance ? Obviously, if one agrees that clear improvements in a child's cognitive ability, clear deterioration in the parent status of a given child's cognitive skills and elementary reductions in correlation of between skill and family background are all instances of improvement in the education system, one would then want to require I to reflect this agreement. Formally, this amounts to requiring I to satisfy the following three properties.

Axiom 1 *(Sensitivity to improvements in cognitive skill)* $I(\mathbf{e}) > I(\mathbf{e}')$ *for every two* ***distinct*** *education systems* $\mathbf{e}$ *and* $\mathbf{e}' \in \mathcal{E}$ *such that* $\mathbf{e}$ *has been obtained from* $\mathbf{e}'$ *by means of an improvement of a child"s cognitive ability.*

Axiom 2 *(Sensitivity to deterioration in the family background of a child with a given cognitive skill)* $I(\mathbf{e}) > I(\mathbf{e}')$ *for every two* ***distinct*** *education systems* $\mathbf{e}$ *and* $\mathbf{e}' \in \mathcal{E}$ *such that* $\mathbf{e}$ *has been obtained from* $\mathbf{e}'$ *by means of a deterioration in the family background of a child with a given cognitive skill.*

Axiom 3 *(Sensitivity to reduction in correlation between cognitive skill and family background).* $I(\mathbf{e}) > I(\mathbf{e}')$ *for every two* ***distinct*** *education systems* $\mathbf{e}$ *and* $\mathbf{e}' \in \mathcal{E}$ *such that* $\mathbf{e}$ *has been obtained from* $\mathbf{e}'$ *by means of a reduction in correlation between cognitive skill and family background.*

Beside Axioms 1-3, which reflect the sensitivity of the function I to the three principles mentioned above, we may also require the index to satisfy additional properties. One of them is the *anonymous* requirement that the function I pays no attention to the children's names or other (irrelevant) characteristics when evaluating an education system. The only information on which the evaluation should be based is the distribution of the children skill levels and family backgrounds. Hence, permuting any two pairs of child skill and family background should have no impact on the appraisal of an education system. We formulate precisely this anonymity requirement as follows.

Axiom 4 *(Anonymity) For every school system* $\mathbf{e} \in \mathcal{E}$ *and every* $n \times n$ *permutation matrix* π*, one has* $I(\pi.\mathbf{e}) = I(\mathbf{e})$*.*

The next axiom could be called "*independence with respect to unconcerned children*". It requires indeed the comparisons of any two education systems to be *independent* from the information concerning children cognitive skills and family backgrounds that is common to the two education systems, no matter what these are. The precise statement of this axiom as follows.

Axiom 5 *(Independence with respect to unconcerned children) For all education systems* $\mathbf{e}$, $\mathbf{e}'$, $\mathbf{e}''$ *and* $\mathbf{e}''' \in \mathcal{E}$ *such that, for some group* $G \subset N$ *of children, one has* $(s_g, b_g) = (s'_g, b'_g)$, $(s''_g, b''_g) = (s'''_g, b'''_g)$, $(s_i, b_i) = (s''_i, b''_i)$ *and* $(s'_i, b'_i) = (s'''_i, b'''_i)$ *for all* $g \in G$ *and* $i \in N \backslash G$, $I(\mathbf{e}) \geq I(\mathbf{e}') \Longleftrightarrow I(\mathbf{e}'') \geq I(\mathbf{e}''')$.

It is easy to see that these five axioms have rather strong implications when imposed on a function $I : \mathcal{E} \longrightarrow \mathbb{R}$ that can be taken to be at least once continuously differentiable, as shown in the following proposition.

Proposition 1 *Let* $I : \mathcal{E} \longrightarrow \mathbb{R}$ *be a continuously differentiable function. Then* I *satisfies Axioms* **??**-**??** *if and only if there exists a continuously differentiable function* $\Phi : \mathcal{S} \times \mathcal{B} \longrightarrow \mathbb{R}$ *with partial derivatives* Φ_1 *and* Φ_2 *satisfying, for every* $a \in \mathcal{S}$*, and* $b, c \in \mathcal{B}$ *such that* $b < c$, $\Phi_1(a,b) > \Phi_1(a,c) > 0 > \Phi_2(a,b)$ *for which one has, for every education system* $\mathbf{e} \in \mathcal{E}$*:*

$$I(\mathbf{e}) = h(\sum_{i=1}^{n} \Phi(s_i, b_i)) \tag{7}$$

for some continuously differentiable increasing function $h : IM^{\sum_{i=1}^{n} \Phi}((\mathcal{S} \times \mathcal{B})^n) \longrightarrow \mathbb{R}$.[4]

Proof. *It can be checked that a continuously differentiable function* $I : \mathcal{E} \longrightarrow \mathbb{R}$ *that writes as per (7) for some continuously differentiable function* $\Phi : \mathcal{S} \times \mathcal{B} \longrightarrow \mathbb{R}$ *satisfying, for every* $a \in \mathcal{S}$*, and* b *and* $c \in \mathcal{B}$ *such that* $b < c$, $\Phi_1(a,b) >$

[4] We denote by $IM^g(A)$ the image of the set A under the function g whose domain contains A.

$\Phi_1(a,c) > 0 > \Phi_2(a,b)$ and some continuously differentiable increasing function $h : IM^{\sum_{i=1}^{n}\Phi}((\mathcal{S}\times\mathcal{B})^n) \longrightarrow \mathbb{R}$ verifies Axioms 1-5. As for the other direction, suppose that $I : \mathcal{E} \longrightarrow \mathbb{R}$ is a continuously differentiable function satisfying Axioms 1-5. Define the ordering $\succsim$ on $\mathcal{E}$ by $\mathbf{e} \succsim \mathbf{e}' \iff I(\mathbf{e}) \geq I(\mathbf{e}')$. Since I is continuous, the ordering $\succsim$ is continuous in the sense of Debreu (1954). Moreover the set $\mathcal{E} = (\mathcal{S}\times\mathcal{B})^n$ is a Cartesian product of n connected and separable topological spaces. Since I satisfies Axioms 1 and 2, each of the n elements or this Cartesian product is essential in the sense of Debreu (1960) (definition 4). Since I satisfies axiom 5, the ordering $\succsim$ satisfies Debreu (1960)'s property of independence (definition 4 again) across elements of the Cartesian product. Hence, by virtue of theorem 3 of Debreu (1960), the ordering $\succsim$ can be numerically represented (in the sense of Debreu (1954)) by a continuous function $\Psi : (\mathcal{S}\times\mathcal{B})^n \longrightarrow \mathbb{R}$ that writes, for every $\mathbf{e} \in (\mathcal{S}\times\mathcal{B})^n$:

$$\Psi(\mathbf{e}) = \sum_{i=1}^{n} \Phi_i(s_i, b_i) \tag{8}$$

for n continuous functions $\Phi_i : (\mathcal{S}\times\mathcal{B}) \longrightarrow \mathbb{R}$. Since the ordering $\succsim$ is numerically represented equally well by the function I that satisfies the anonymity Axiom 4, one must have $\Phi_i(a,b) = \Phi_j(a,b) = \Phi_j(a,b)$ for all i, j and all $(a,b) \in \mathcal{S}\times\mathcal{B}$ for some function $\Phi : (\mathcal{S}\times\mathcal{B}) \longrightarrow \mathbb{R}$. Since the functions I and Ψ numerically represent the same ordering $\succsim^E$, one must have:

$$\begin{aligned} I^E(\mathbf{e}) &= h(\Psi(\mathbf{e})) \\ &= h(\sum_{i=1}^{n} \Phi(s_i, b_i)) \end{aligned}$$

for some increasing function $h : IM^{\sum_{i=1}^{n}\Phi}((\mathcal{S}\times\mathcal{B})^n) \longrightarrow \mathbb{R}$. Since I is continuously differentiable, so must be both h and Φ. The requirement that I be sensitive to improvements in children cognitive skill requires Φ to be increasing with respect to its first argument. Analogously,the requirement for I to be sensitive to deteriorations in the family background of a child with given cognitive skill requires Φ to be decreasing with respect to its second argument. Hence, under differentiability of Φ, one has $\Phi_1(a,b) > 0 > \Phi_2(a,b)$ for all $(a,b) \in (\mathcal{S}\times\mathcal{B})$. The proof that Φ must satisfy $\Phi_1(a,b) > \Phi_1(a,c)$ for all $a \in \mathcal{S}$ and b, $c \in \mathcal{B}$ such that $b < c$ if the ordering $\succsim$ (or the function I) satisfies Axiom 3 results from lemma 4.1 in Gravel and Moyes (2012). ■

Proposition 1 thus provides a specific way of comparing education systems. It says indeed that the performance of any education system *can be* measured as a sum, taken over all children educated through this system, of a function of both the child's cognitive skill and family background (the same function for all children). Let us call *additive* any index that writes as per expression (7) for some functions Φ and some increasing function h. Proposition 1 also imposes specific properties on the function Φ, that result from the requirement imposed on the index to satisfy Axioms 1-3. Yet Proposition 1 still leaves one with a wealth of possible ways of generating these comparisons, as many ways in fact as there are functions Φ satisfying $\Phi_1(s,b) > \Phi_1(s,b') > 0 > \Phi_2(s,b)$ for all $s \in$ S and b, $b' \in \mathcal{B}$ such that $b < b'$.

It is of interest to contrast the comparisons of education systems made by additive indices of the kind identified by Proposition 1 with those often performed often in the empirical literature that are based on regression coefficients of cognitive skills on family background, as used for example in Schutz, Ursprung, and Woessmann (2008). Specifically, many authors estimate, for every education system $\mathbf{e}$, the following regression model (abstracting from the additional "control" variables that are often considered in the regressions, in addition to the family background):

$$s_i = \alpha^{\mathbf{e}} + \beta^{\mathbf{e}} b_i + \varepsilon_i^{\mathbf{e}}$$

where $\varepsilon_i^{\mathbf{e}}$ is the regression error term observed on child i of system $\mathbf{e}$ and $\alpha^{\mathbf{e}}$ and $\beta^{\mathbf{e}}$ are the (theoretical) constant and regression coefficient of this linear model applied to system $\mathbf{e}$. It is well-known from elementary econometrics that the least-square estimate of $\beta^{\mathbf{e}}$, denoted $\widehat{\beta}(\mathbf{e})$, is defined by:

$$\begin{aligned} \widehat{\beta}(\mathbf{e}) &= \frac{cov(s_i, b_i)}{var(b_i)} \\ &= \frac{\sum_{i=1}^{n} s_i b_i - n\overline{s}(\mathbf{e})\overline{b}(\mathbf{e})}{\sum_{i=1}^{n} b_i^2 - n\overline{s}\overline{b}(\mathbf{e})} \end{aligned} \tag{9}$$

where $\overline{s}(\mathbf{e})$ and $\overline{b}(\mathbf{e})$ denote the average skill and family background (respectively) observed in the education system $\mathbf{e}$. $\widehat{\beta}(\mathbf{e})$ is clearly an index that enables the comparison of any two such systems. This index is usually used as a negative measure of the performance of a school system (the lower the index, the better is the school system). When used this way, this index, that is clearly continuously differentiable, satisfies Axiom 3 and Axiom 4, as the reader can easily verify. However, the index $\widehat{\beta}(\mathbf{e})$ violates Axioms 1, 2 and 5. As is apparent from Expression (9), $\widehat{\beta}(\mathbf{e})$ can not be written by an additive expression such as (7) for some function Φ applied to every child i. The fact that the index $\widehat{\beta}(\mathbf{e})$ violates Axiom 1 strikes us as a serious limitation for measuring the performance of education system. Indeed, it is easy to see that a reduction in the cognitive skill of a child coming from a below average background reduces the value of $\widehat{\beta}(\mathbf{e})$ and, therefore, improves the performance of the school system. We find this conclusion not very appealing from a policy point of view. Moreover, we consider the additivity of an index to be a convenient property for computation and manipulation.

We therefore stick to the class of additive indices that can be written as per expression (7) for some function Φ satisfying $\Phi_1(s,b) > \Phi_1(s,b') > 0 > \Phi_2(s,b)$ for all $s \in$ S and b, $b' \in \mathcal{B}$ such that $b < b'$. Denote by $\mathbf{\Phi}$ the class of all functions. Since we do not have reason to favour one of these functions rather than another, we find safe to seek for a ranking of education systems that commands agreement of all such functions. This give rise the following notion of "additive dominance".

Definition 4 *(additive dominance) We say that the education system* $\mathbf{e}$ *addi-*

tively dominates the system $\mathbf{e}'$ *for the class* Φ *of functions if one has:*

$$\sum_{i=1}^{n} \Phi\left(s_i, b_i\right) \geqq \sum_{i=1}^{n} \Phi\left(s_i', b_i'\right), \ \forall\ \Phi \in \Phi \tag{10}$$

The following theorem establishes a tight link between the notion of additive dominance - as per Definition 4 - on the one hand, and the notions of clear improvement in a child cognitive skill, clear deterioration in a child family backgrounds and clear reduction in correlation between cognitive skill and family background on the other. It also identifies an empirical test that enables one to verify, given any two education systems, whether or not one additively dominates the other for the class Φ. The empirical test is easy to implement. It amounts to verifying, for any possible reference pair of child's cognitive skill and family background, if the fraction of children with *both* a *worse* family background *and* a *better* skill than the reference is larger in one education system than another. The formal statement of the result is as follows. Proof of parts of the equivalence is also provided.

Theorem 1 *Let* $\mathbf{e}, \mathbf{e}^{\circ} \in \mathcal{E}$*. Then, statements (a), (b) and (c) below are equivalent:*

(a) $\mathbf{e}$ *is obtained from* $\mathbf{e}^{\circ}$ *by means of a finite sequence of clear improvements in child cognitive skill, clear deterioration in the family background of children with given skill and/or clear reduction in correlation between cognitive skill and family background.*

(b) $\mathbf{e}$ *additively dominates* $\mathbf{e}^{\circ}$ *for any additive index based on the class* Φ *of functions.*

(c-1) $\#Q^{\mathbf{e}}(s,b) \geq \#Q^{\mathbf{e}^{\circ}}(s,b)$ *for all* $(s,b) \in [\boldsymbol{\sigma}(\mathbf{e}) \cup \boldsymbol{\sigma}(\mathbf{e}^{\circ})] \times [\boldsymbol{\beta}(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})]$.

Proof.
Statement (a) implies Statement (b). *The proof of this implication has already been provided in the proof of Proposition 1.*
Statement (b) implies Statement (c) *Assume that* $\mathbf{e}$ *is at least as good as* $\mathbf{e}^{\circ}$ *for any additive index based on the class* Φ *of functions. Hence, one has:*

$$\sum_{i=1}^{n} \Phi(s_i, b_i) \geq \sum_{i=1}^{n} \Phi(s_i^{\circ}, b^{\circ}{}_i) \tag{11}$$

for all functions Φ *in the class* Φ*. In particular therefore, inequality (11) holds for the function* Φ^{sb} *defined, for any* $(s,b) \in \mathcal{S} \times \mathcal{B}$*, by:*

$$\begin{aligned} \Phi^{sb}(x,y) &= 1 \text{ if } x \geq s \text{ and } y \leq b \\ &= 0 \text{ otherwise} \end{aligned} \tag{12}$$

This function is not continuously differentiable. Yet, it can be approximated by a continuously differentiable function (see e.g. Fishburn and Vickson (1978)). However it satisfies all the other properties of functions in the class Φ*. It is clearly increasing in its first argument, decreasing in its second argument, and is such that "the rate of increase" of the function with respect to its first argument* *is decreasing with respect to the second argument. Since Inequality (11) holds for any such function* Φ^{sb}*, one has, for every* $(s,b) \in \mathcal{S} \times \mathcal{B}$*:*

$$\begin{array}{rcl} \sum_{i=1}^{n} \Phi^{sb}(s_i, b_i) & \geq & \sum_{i=1}^{n} \Phi^{sb}(s_i^{\circ}, b^{\circ}{}_i) \\ & \Longleftrightarrow & \\ \#Q^{\mathbf{s}}(s,b) & \geq & \#Q^{\mathbf{s}^{\circ}}(s,b) \end{array}$$

as required by (c).

Statement (c) implies Statement (b). *Assume that (c) holds. Let us show that this implies that inequality (11) holds for all functions* Φ *in the class* Φ*. Since the bounds of the intervals* $[\underline{s}, \overline{s}]$ *and* $[\underline{b}, \overline{b}]$ *and the elements of the sets* $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\mathbf{e}) \cup \boldsymbol{\sigma}(\mathbf{e}^{\circ})$ *and* $\boldsymbol{\beta} = \boldsymbol{\beta}(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})$ *are all rational numbers, one can subdivide the intervals* $[\underline{s}, \overline{s}]$ *and* $[\underline{b}, \overline{b}]$ *into* l *and* m *(respectively) subintervals* $[s_h, s_{h+1}]$ *and* $[b_k, b_{k+1}]$ *of equal length* $\Delta^s = s_{h+1} - s_h$ *and* $\Delta^b = b_{k+1} - b_h$ *(for* $h = 0, ..., l-1$ *and* $k = 0, ..., m-1$*) in such a way that:*
(1) $\underline{s} = s_0$*,* $\overline{s} = s_l$*,* $\underline{b} = b_0$*,* $\overline{b} = b_m$
(2) for every $s \in \boldsymbol{\sigma}$ *and* $b \in \boldsymbol{\beta}$*, there are* $h \in \{0, ..., l\}$ *and* $k \in \{0, ..., m\}$ *such that* $s_h = s$ *and* $b_k = b$*.,*
Taking such a subdivision, one can write Inequality (11) as:

$$\sum_{h=1}^{l} \sum_{k=1}^{m} \Delta f(s_h, b_k) \Phi(s_h, b_k) \geq 0 \tag{13}$$

where, for every $(s,b) \in [\underline{s}, \overline{s}] \times [\underline{b}, \overline{b}]$*:*

$$\Delta f(s,b) = f^{\mathbf{e}}(s,b) - f^{\mathbf{e}^{\circ}}(s,b)$$

Of course the possibility that $\Delta f(s,b) = 0$ *for some* $(s,b) \in [\underline{s}, \overline{s}] \times [\underline{b}, \overline{b}]$ *is not ruled out. Inverting the order of summation of the skill variable, we can write alternatively Inequality (13) as:*

$$\sum_{h=l}^{1} \sum_{k=1}^{m} \Delta f(s_h, b_k) \Phi(s_h, b_k) \geq 0 \tag{14}$$

Doing a discrete, or Abelian (see e.g. Fishburn and Vickson (1978) eq. 2.49), sum by part of expression (13) yields (after exploiting the fact that $\sum_{h=l}^{1} \sum_{k=1}^{m} \Delta f(s_h, b_k) = 0$*):*

$$\begin{aligned} & \sum_{h=l}^{2} \sum_{k=1}^{m} \Delta f(s_h, b_k) \Phi_1(s_{h-1}, b_m) \\ & - \sum_{h=l}^{1} \sum_{k=1}^{m-1} \Delta f(s_h, b_k) \Phi_2(s_l, b_k) \\ & - \sum_{h=l}^{2} \sum_{g=l}^{h} \sum_{k=1}^{m-1} \sum_{j=1}^{k} \Delta f(s_g, b_j) \Phi_{12}(s_{h-1}, b_k) \\ \geq \; & 0 \end{aligned} \tag{15}$$

where, for every $h \in \{1,...,l\}$ and $k \in \{1,...,m\}$:

$$\begin{aligned}\Phi_1(s_h, b_k) &= \Phi(s_{h+1}, b_k) - \Phi(s_h, b_k) \\ \Phi_2(s_h, b_k) &= \Phi(s_h, b_{k+1}) - \Phi(s_h, b_k)\end{aligned}$$

denote the partial discrete (to the right) difference of Φ with respect to cognitive skill and to family background respectively and where:

$$\Phi_{12}(s_h, b_k) = \Phi(s_{h+1}, b_{k+1}) - \Phi(s_h, b_{k+1}) - \Phi(s_{h+1}, b_k) + \Phi(s_h, b_k) \tag{16}$$

denote the discrete (to the right) difference of difference of Φ first with respect to skill and second with respect to family background. Hence, a sufficient condition for (15) to hold for all functions Φ satisfying $\Phi_1 \geq 0 \geq \Phi_2$ and $\Phi_{12} \leq 0$ is to have:

$$\sum_{g=l}^{h} \sum_{j=1}^{k} \Delta f(s_g, b_j) \geq 0$$

for all $h \in \{1,...,l\}$ and $k \in \{1,...,m\}$ of, equivalently, to have:

$$\#Q^{\mathbf{s}}(s_h, b_k) \geq \#Q^{\mathbf{s}^\circ}(s_h, b_k)$$

for all $(s_h, s_k) \in [\boldsymbol{\sigma}(\mathbf{e}) \cup \boldsymbol{\sigma}(\mathbf{e}^\circ)] \times [\boldsymbol{\beta}(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^\circ)]$.
Statement (c) implies Statement (a) *The argument follows from an adaptation of theorem 4.1 in Gravel and Moyes (2012) and other results in the literature on uni- or multi-dimensionnal measurement such as Lehmann (1955), Quirk and Saposnik (1962) and Osterdal (2010).* ■

2.3 Attitudes toward inequalities in cognitive skills

Theorem 1 provides a simple test for checking whether an education system results from another by a finite sequence of clear improvements in cognitive skills, clear deterioration of family backgrounds and/or clear reduction in correlation between family background and cognitive skill. This criterion consists in verifying if, for any pair of skill and family background levels, the fraction of children with *both* a better skill *and* a lower family background than those levels is larger in one system than another. Yet, the large consensus over which the criterion stands - actually all rankings of education systems who agree with Axioms 1-3 - is likely to make this criterion highly incomplete. As usual in dominance analysis, any gain in discriminatory power comes at the cost of requiring the criterion to satisfy additional principles. What could these be ?

One concerns attitude toward inequalities in cognitive skills, given family background. Consider indeed two alternative hypothetical situations concerning two children with the same family background who end up, at the end of their compulsory schooling curriculum, with different cognitive skills. Assume that the average cognitive skill - calculated over the two children - is the same in the two situations but that the two skill levels are more spread out in one situation than in the other. By "more spread out", we mean, as is standard in inequality analysis, "resulting from a mean preserving spread" or, equivalently, by a mean preserving "regressive transfer". Of course the very notion of a mean preserving spread of cognitive skill rests on the faith that cognitive skills are measured on a cardinal scale so that the very notion of a "mean" is, without playing with words, meaningful.[5] Suppose we have this faith. Is a more spread-out distribution of skill better or worse than a less spread-out one ? Intuition coming from conventional attitude toward income inequality could suggest that a less spread out - or a more equal - distribution of skills is better than a more spread-out one. Yet, such an intuition may be misleading for an attribute such a skill. This is at least so if one adopts the welfarist perspective according to which the "social goodness" of alternative states of affairs depends only upon the distribution of individual well-being that they generate. As documented in the literature, there seems to be two channels by which cognitive skill affects well-being. One of them is the income that the skills enables the individual to earn on the market. If one restricts attention to skills as the unique determinant of income, one can denote by $y(s)$ the income that an individual with skill s can earn on the market. The empirical evidence (see e.g. Green and Riddell (2003), Heckman, J. Stixrud, and Uzrua (2006) or Barrett (2012)) on the functional relation y connecting skill to income is that it is *increasing* and *convex*. That is, the *gain* in earning capacity brought about by an increase in skill is *increasing* with skill. The other channel by which cognitive skill affects well-being is of course a direct one. Cognitive skill helps the individual to make better decisions (in the choice of his/her partner, career profile, medical treatments, etc.) and better use of the information. This suggests that the individual well-being u is an *intrinsic* function of two variables: skills (s) and income (y):

$$u = U(s, y)$$

As conventionally considered in economics, the (utility) function U that associates well-being to every combination of skill and income would be increasing in both variables and concave with respect to income (marginal utility of income is decreasing with respect to income at any skill level). Yet, economic theory, empirical evidence, and introspection do not provide clear evidence about the concavity or convexity of U with respect to skill. Convexity - e.g. the fact that, given income, the marginal utility of an increase in skill is increasing with skill - is not implausible. A similar lack of *a priori* intuition concerns the relation between the (positive) marginal utility of income and the skill level. Are skill and income complement, or substitute, for the achievement of a given level of well-being ? It is not implausible to believe that they are complement, so that the marginal utility of income is increasing (at least weakly) with respect to skill. To sum up, it is plausible to assume that U satisfies (assuming differentiability):

$$\begin{aligned} U_j(.) &\geq 0 \text{ for } j = s, y && (17) \\ U_{yy}(.) &\leq 0 && (18) \\ U_{ss}(.) &\geq 0 && (19) \\ U_{ys}(.) &\geq 0 && (20) \end{aligned}$$

where, for every i, $j \in \{s, y\}$, $U_j(.)$ and $U_{ij}(.)$ denote, respectively, the partial derivative of U with respect to j and the second derivative first with respect to

[5] See Allison and Foster (2004) for a discussion of the difficulty of applying conventional concepts of inequality measurement to distribution of an ordinal variable and Gravel, Magdalou, and Moyes (2015) for a dominance approach to the issue.

i and second with respect to j. Under these assumptions, one can then consider the "final" functional relation connecting any skill level s to utility u through both indirect and direct effects:

$$\begin{aligned} u &= U(s, y(s)) \\ &= \Psi(s) \end{aligned}$$

Hence $\Psi(s)$ gives the level of well being produced by a level of skill s through both its direct and indirect (pecuniary) effect. Assuming $y(.)$ to be differentiable, one can see that:

$$\begin{aligned} \Psi_s(.) &= U_s(.) + U_y(.) y_s(.) \\ &\geq 0 \end{aligned}$$

if $U(.)$ satisfies (17) and the function y is increasing in skills. More importantly for our purpose, one can also see that, under assumptions (17)-(20) and the convexity of the function y:

$$\Psi_{ss}(.) = \underbrace{U_{ss}(.)}_{+} + \underbrace{2U_{sy}(.) y_s(.)}_{+} + \underbrace{U_{yy}(.) [y_s(.)]^2}_{-} + \underbrace{U_y(.) y_{ss}(.)}_{+}$$

Hence, while the sign of $\Psi_{ss}(.)$ is *a priori* ambiguous, it is not implausible that the sum of the three positive terms overweights the negative one.

We therefore consider this possibility for skill to have a positive and convex final effect on well-being. If one does that, it becomes possible to defend the view that, for a population of children with the same family background, a more unequal - or spread out - distribution of skills of a given mean is better than a less-spread out one. This approach, which can be viewed as the opposite to egalitarianism, has been called elitism by Bazen and Moyes (2012). The main elementary operation that describes the notion of dispersion in the cognitive skills that is considered favorably by elitism is the following.

Definition 5 Elementary dispersion in children cognitive skills. *We say that education system* $\mathbf{e}$ *is obtained from education system* $\mathbf{e}^{\circ}$ *by means of a clear dispersion in children cognitive skills if there are two children* h *and* i *and a real number* Δ *such that:*

$$s_i^{\circ} = s_i - \Delta \geq s_h + \Delta = s_h^{\circ}, \;\; b_h = b_h^{\circ} = \;\; b_i = b_i^{\circ} \text{ and } (s_g, b_g) = (s_g^{\circ}, b_g^{\circ}) \; \forall \; g \neq i.$$

We now formulate as follows the requirement, for an index $I : \mathcal{E} \longrightarrow \mathbb{R}$ that measures the performance of an education system to be sensitive to such an elementary dispersion in the children cognitive skills.

Axiom 6 *Sensitivity to Elementary dispersion in children cognitive skills.* $I(\mathbf{e}) > I(\mathbf{e}')$ *for every two* ***distinct*** *education systems* $\mathbf{e}$ *and* $\mathbf{e}' \in \mathcal{E}$ *such that* $\mathbf{e}$ *has been obtained from* $\mathbf{e}'$ *by means of a clear dispersion in children cognitive skills.*

Requiring the continuously differentiable index I to satisfy Axiom 6, in addition to Axioms 1-4, renders the index additive with respect to a class $\mathbf{\Phi}^{**}$ of functions Φ in Expression 7 that is more restricted than Φ. Indeed, in addition to the properties that define the class Φ, the functions in the class $\mathbf{\Phi}^{**}$

will be **convex** with respect to their first argument. Specifically, any function Φ in $\mathbf{\Phi}^{**}$ will satisfy the requirement that $\Phi_1(s,b) \geq \Phi_1(s',b) > 0$ for every s, $s' \in \mathcal{S}$ satisfying $s > s'$ and every $b \in \mathcal{B}$. While the class of functions $\mathbf{\Phi}^{**}$ is a (significantly) proper subset of the class Φ, it is still a somewhat large class of functions. Without additional reason to choose one function rather than another, it may seem safe to look for rankings of education systems that command unanimity over all functions in this class. The following Theorem provides an exact empirical test that enables one to check for such a ranking.

Theorem 2 *Let $\mathbf{e}, \mathbf{e}^{\circ} \in \mathcal{E}$. Consider the following statements.*

1. *$\mathbf{e}$ is obtained from $\mathbf{e}^{\circ}$ by means of a finite sequence of improvements in children cognitive skill, deteriorations in the family background of children with a given skill, reduction in correlation between cognitive skill and family background and/or elementary dispersions in children cognitive skills*

2. *$\mathbf{e}$ additively dominates $\mathbf{e}^{\circ}$ for any additive index based on the class $\mathbf{\Phi}^{**}$ of functions.*

3. $\sum_{b \in \beta(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})} S^{\mathbf{e}}(s(b), b) \geq \sum_{b \in \beta(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})} S^{\mathbf{e}^{\circ}}(s(b), b)$ *for all increasing functions $s : \boldsymbol{\beta}(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ}) \to \boldsymbol{\sigma}(\mathbf{e}) \cup \boldsymbol{\sigma}(\mathbf{e}^{\circ})$ and $\#Q^{\mathbf{e}}(\underline{s}, b) \geq \#Q^{\mathbf{e}^{\circ}}(\underline{s}, b)$ for every $b \in \boldsymbol{\beta}(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})$.*
   Then, statement 1 implies statement 2 and statements 2 and 3 are equivalent.

Proof.
Statement 1 implies Statement 2: *This is an immediate consequence of the fact that additive rankings of education systems based on the class $\mathbf{\Phi}^{**}$ of functions are all sensitive to improvements in children cognitive skill, deteriorations in the family background of children with a given skill, reduction in correlation between cognitive skill and family background and elementary dispersions in children cognitive skills.*
Statement 2 implies statement 3:*Assume that $\mathbf{e}$ is at least as good as $\mathbf{e}^{\circ}$ for any additive index based on the class $\mathbf{\Phi}^{**}$ of functions. Hence, one has Inequality (11) for all functions Φ in the class $\mathbf{\Phi}^{**}$. In particular therefore, inequality (11) holds for the function $\Phi^{s(.)}$ defined, for any increasing function $s : \mathcal{B} \to \boldsymbol{\sigma}(\mathbf{e}) \cup \boldsymbol{\sigma}(\mathbf{e}^{\circ})$, by:*

$$\Phi^{s(.)}(x,y) = \max[x - s(y), 0]$$

Just like in the proof of Theorem 1, this function is not continuously differentiable. However, it can be approximated by a continuously differentiable function following the arguments of Fishburn and Vickson (1978). $\Phi^{s(I)}$ is clearly increasing (weakly) with respect to its first argument (for a given y). It is also weakly decreasing with respect to y if s is increasing. Moreover, the "rate of increase of $\Phi^{s(.)}$" with respect to x is decreasing with respect to y if s is increasing. Finally, this function $\Phi^{s(.)}$ is convex with respect to x for any given y. Hence Inequality

(11) must hold for any such function $\Phi^{s(.)}$. Hence, one has, for every $(s,b) \in \mathcal{S} \times \mathcal{B}$:

$$\begin{array}{rcl} \sum_{i=1}^{n} \Phi^{s(.)}(s_i, b_i) & \geq & \sum_{i=1}^{n} \Phi^{s(.)}(s_i^{\circ}, b^{\circ}{}_i) \\ & \Longleftrightarrow & \\ \sum_{b \in \beta(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})} S^{\mathbf{e}}(s(b), b) & \geq & \sum_{b \in \beta(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})} S^{\mathbf{e}^{\circ}}(s(b), b) \end{array}$$

as required by the first part of statement (3). As for the second part of statement (3), one can simply observe that the function Φ^b defined by:

$$\begin{array}{rcl} \Phi^b(x,y) & = & 1 \text{ if } y \geq b \\ & = & 0 \text{ otherwise} \end{array}$$

belongs also to the class $\boldsymbol{\Phi}^{**}$ for any $b \in \boldsymbol{\beta}(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})$. Hence inequality (11) must hold for the function Φ^b as well and, as a result, one must have $\#Q^{\mathbf{e}}(\underline{s}, b) \geq \#Q^{\mathbf{e}^{\circ}}(\underline{s}, b)$ for every $b \in \boldsymbol{\beta}(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})$.

Statement 3 implies statement 2

We start from Inequality (11) that we write, just like in inequality (14), as follows (after inverting the order of summation) :

$$\sum_{k=1}^{m} \sum_{h=l}^{1} \Delta f(s_h, b_k) \Phi(s_h, b_k) \geq 0 \tag{21}$$

Abel decomposing the inner term of Inequality (21) yields:

$$\begin{array}{rl} & \sum_{k=1}^{m} \sum_{h=l}^{1} \Delta f(s_h, b_k) \Phi(s_1, b_k) \\ + \sum_{k=1}^{m} \sum_{h=l}^{2} \sum_{g=l}^{h} \Delta f(s_g, b_k) \Phi_1(s_{h-1}, b_k) & \geq \ 0 \end{array} \tag{22}$$

where, as in the proof of Theorem 1, $\Phi_1(s_h, b_k)$ is defined by:

$$\Phi_1(s_h, b_k) = \Phi(s_{h+1}, b_k) - \Phi(s_h, b_k)$$

for any $s_h \in \{s_1, ..., s_{\sigma-1}\}$ and $b_k \in \{b_1, ..., b_{\beta}\}$. Paralleling the ingenious approach proposed by Bourguignon (1989) for his ordered poverty gap criterion, one can observe that the second term of the left hand side of Inequality (22) can be written as:

$$\begin{array}{rcl} \sum_{k=1}^{m} \sum_{h=l}^{2} \sum_{g=l}^{h} \Delta f(s_g, b_k) \Phi_1(s_{h-1}, b_k) & = & \sum_{k=1}^{m} \sum_{h=l}^{2} \sum_{g=l}^{h} [\Delta f(s_g, b_k) + g_{k-1}^h - g_k^h] \Phi_1(s_{h-1}, b_k) \\ & & - \sum_{k=2}^{m} \sum_{h=\sigma}^{2} \Phi_{12}(s_{h-1}, b_{k-1}) g_{k-1}^h \end{array} \tag{23}$$

for any list of $l \times m$ non-negative numbers g_{k-1}^h (for $k = 1, ..., m$ and $h = 1, ..., l$) satisfying $g_0^h = g_m^h = 0$ for every $h \in \{1, ..., l\}$, with the discrete cross derivative

$\Phi_{12}(.)$ *defined as per expression (16) of Theorem 1. Doing an additional Abel decomposition of the first term of the right hand side of Expression (23) enables one to write this expression as:*

$$\sum_{k=1}^{m}\sum_{h=l}^{2}\sum_{g=l}^{h}\Delta f(s_g,b_k)\Phi_1(s_{h-1},b_k) =$$

$$\begin{aligned}
= \quad & \sum_{k=1}^{m}[\sum_{h=l}^{2}\sum_{g=l}^{h}\Delta f(s_g,b_k)+\sum_{h=l}^{2}g_{k-1}^{h}-\sum_{h=l}^{2}g_{k}^{h}]\Phi_1(s_1,b_k) \\
& -\sum_{k=1}^{m}\sum_{i=l}^{2}[\sum_{h=l}^{i}\sum_{g=l}^{h}\Delta f(s_g,b_k)+\sum_{h=\sigma}^{i}g_{k-1}^{h}-\sum_{h=\sigma}^{i}g_{k}^{h}][\Phi_1(s_{i-1},b_k)-\Phi_1(s_i,b_k)] \\
& -\sum_{k=2}^{m}\sum_{h=l}^{2}\Phi_{12}(s_{h-1},b_{k-1})g_{k-1}^{h} \qquad (24)
\end{aligned}$$

Substituting (24) back into inequality (22) enables one to write this latter inequality as follows (after performing an Abel decomposition of the first term of the right hand side of this inequality this time with respect to the k indexed summation):

$$\begin{aligned}
& -\sum_{k=1}^{m}\sum_{j=1}^{k}\sum_{h=l}^{1}\Delta f(s_h,b_j)\Phi_2(s_1,b_k) \\
& +\sum_{k=1}^{m}[\sum_{h=l}^{2}\sum_{g=l}^{h}\Delta f(s_g,b_k)+\sum_{h=l}^{2}g_{k-1}^{h}-\sum_{h=l}^{2}g_{k}^{h}]\Phi_1(s_1,b_k) \\
& +\sum_{k=1}^{m}\sum_{i=l}^{2}[\sum_{h=l}^{i}\sum_{g=l}^{h}\Delta f(s_g,b_k)+\sum_{h=l}^{i}g_{k-1}^{h}-\sum_{h=l}^{i}g_{k}^{h}][\Phi_{11}(s_{i-1},b_k)] \\
& -\sum_{k=2}^{m}\sum_{h=l}^{2}\Phi_{12}(s_{h-1},b_{k-1})g_{k-1}^{h} \\
\geq \quad & 0 \qquad (25)
\end{aligned}$$

where $\Phi_{11}(s_{i-1},b_k)=\Phi_1(s_i,b_k)-\Phi_1(s_{i-1},b_k)$. *We now observe that, for every* $i\in\{1,...,l\}$ *and* $k\in\{1,...,m\}$*:*

$$\begin{aligned}
\sum_{g=l}^{i}(s_i-s_g)\Delta f(s_g,b_k) &= s_i\sum_{g=l}^{i}\Delta f(s_g,b_k)-\sum_{g=l}^{i}s_g\Delta f(s_g,b_k) \\
&= s_i\sum_{g=l}^{i}\Delta f(s_g,b_k)-s_i\sum_{g=l}^{i}\Delta f(s_g,b_k)+\sum_{h=l}^{i+1}[s_{h-1}-s_h]\sum_{g=l}^{h}\Delta f(s_g,b_k) \\
&= -\Delta^s\sum_{h=l}^{i+1}\sum_{g=l}^{h}\Delta f(s_g,b_k) \qquad (26)
\end{aligned}$$

Using this expression, one can write Inequality (25) as:

$$
\begin{aligned}
& -\sum_{k=1}^{m}\sum_{j=1}^{k}\sum_{h=l}^{1}\Delta f(s_h,b_j)\Phi_2(s_1,b_k) \\
& +\sum_{k=1}^{m}[\sum_{g=l}^{1}(s_g-s_1)\Delta f(s_g,b_k)/\Delta^s+\sum_{h=l}^{2}g_{k-1}^h-\sum_{h=l}^{2}g_k^h]\Phi_1(s_1,b_k) \\
& +\sum_{k=1}^{m}\sum_{i=l}^{2}[\sum_{g=l}^{i-1}(s_g-s_{i-1})\Delta f(s_g,b_k)/\Delta^s+\sum_{h=l}^{i}g_{k-1}^h-\sum_{h=l}^{i}g_k^h][\Phi_{11}(s_{i-1},b_k)] \\
& -\sum_{k=2}^{m}\sum_{h=l}^{2}\Phi_{12}(s_{h-1},b_{k-1})g_{k-1}^h \\
\geq\ & 0 \qquad (27)
\end{aligned}
$$

As is clear, a sufficient condition for Inequality (27) to hold for all functions Φ such that $\Phi_2\leq 0$, $\Phi_{12}\leq 0$, $\Phi_1\geq 0$ and $\Phi_{11}\geq 0$ is to have:

$$\sum_{j=1}^{k}\sum_{h=l}^{1}\Delta f(s_h,b_j) \geq 0 \qquad (28)$$

$$\sum_{g=l}^{i-1}(s_g-s_{i-1})\Delta f(s_g,b_k)/\Delta^s+\sum_{h=l}^{i}g_{k-1}^h-\sum_{h=l}^{i}g_k^h \geq 0 \qquad (29)$$

for all $k=1,...,m$ and $i=1,...,l$ for some list of $l\times m$ non-negative numbers g_{k-1}^h (for $k=1,...,m$ and $h=1,...,l$) satisfying $g_0^h=g_m^h=0$ for every $h\in\{1,...,l\}$. We observe that having Condition (28) satisfied for every k is equivalent to having $\#Q^{\mathbf{e}}(s_1,b_k)\geq\#Q^{\mathbf{e}^\circ}(s_1,b_k)$ for every k. We now establish that a sufficient condition for the existence of a list of $l\times m$ non-negative numbers g_{k-1}^h (for $k=1,...,m$ and $h=1,...,l$) satisfying $g_0^h=g_m^h=0$ for every $h\in\{1,...,l\}$ for which (29) holds for every $k\in\{1,...,m\}$ and $i=1,...,l$ is to have:

$$\sum_{g:s^i\leq s_g\leq s_l}(s_g-s^i)\Delta f(s_g,b_k)/\Delta^s\geq 0 \qquad (30)$$

for every $(s^1,...,s^m)\in[\boldsymbol{\sigma}(\mathbf{e})\cup\boldsymbol{\sigma}(\mathbf{e}^\circ)]^m$ such $s^1\leq...\leq s^m$. To see this, suppose that Inequality (30) holds for all $(s^1,...,s^m)\in[\boldsymbol{\sigma}(\mathbf{e})\cup\boldsymbol{\sigma}(\mathbf{e}^\circ)]^m$ such $s^1\leq...\leq s^m$. Define the non-negative numbers $\sum_{h=l}^{i}g_{k-1}^h$, for every $i\in\{1,...,l\}$ and $k\in\{1,...,m\}$ recursively as follows:

$$\sum_{h=l}^{i}g_{k-1}^h=\min_{s\geq s^k}\sum_{g=l}^{i}(s_g-s)\Delta f(s_g,b_k)/\Delta^s+\sum_{h=l}^{i}g_k^h$$

starting from:

$$\sum_{h=l}^{i}g_{m-1}^h=\min_{s\geq s^m}\sum_{g=l}^{i}(s_g-s)\Delta f(s_g,b_m)/\Delta^s$$

and setting of course $\sum_{h=l}^{i} g_0^h = 0$ *for every* i*. Hence* $\sum_{h=l}^{i} g_{k-1}^h$ *is the minimal difference in success between education systems* $\mathbf{e}$ *and* $\mathbf{e}^{\circ}$ *for all family background above* k *and all success line superior to* s^k*. These number are clearly non-negative and decreasing, as required, and this completes the proof.* ■

Theorem 2 provides an exact empirical test of the additive dominance for all functions in $\mathbf{\Phi}^{**}$. For any two education systems $\mathbf{e}$ and $\mathbf{e}^{\circ}$, the test works as follows. One first assigns, to every value of the family background observed in either of the two education systems, a target of cognitive skill that is increasing with respect to family background. That is to say, children from high family backgrounds are assigned a higher target than those of lower backgrounds. One then sums, over all levels of family background observed in the two systems, the excess of cognitive skills of the children over their respective target. An education system who exhibits a larger sum of excess of cognitive skills of children over their target than another for every assignment of targets to family background that are increasing with respect to the family backgrounds is then said to dominate that other system. Observe that when the marginal distributions of the family background differ between the two education systems, the test for additive dominance for all functions in $\mathbf{\Phi}^{**}$ requires also that the marginal distribution of family backgrounds in the dominating system be worse - as per usual first order dominance - than that of the dominated one.

One can certainly hesitate in adhering to the elitist value judgement that cognitive skills inequalities among children of a given family background are, ceteris paribus, a good thing. For reasons that are not completely clear to us (and to others, see the discussions of similar difficulties made by Muller and Trannoy (2012) on p. 138-139), there does not seem to be exact empirical test of additive dominance for all functions in the class Φ that are concave with respect to cognitive skills and, therefore, exhibit aversion toward inequalities in those skills. However, if one adheres to the elitist point of view that inequalities in cognitive skills among children of a given family background are, everything else the same, welfare improving, it is possible to develop a more discriminatory dominance criterion that incorporates an additional value judgement that the love for elitism, which underlies the requirement that the function Φ be convex with respect to skill, be itself decreasing with respect to family background. That is, elitism has more favorable consequence when it takes place among children of low family backgrounds than when it takes place among children of more favorable backgrounds. Incorporating such a value judgement in the analysis amounts to restricting the class $\mathbf{\Phi}^{**}$ of functions Φ over which additive dominance is seek to the class $\mathbf{\Phi}^{***}$ of all functions in $\mathbf{\Phi}^{**}$ that satisfy the requirement that $\Phi_1(s, b') - \Phi_1(s', b') \geq \Phi_1(s, b) - \Phi_1(s', b)$ for every two levels of skills s and s' and family backgrounds b and b' such that $s > s'$ and every $b > b'$. The following theorem provides an exact empirical test that enables one to check for additive dominance over all functions Φ in the class $\mathbf{\Phi}^{***}$.

Theorem 3 *Let* $\mathbf{e}, \mathbf{e}^{\circ} \in \mathcal{E}$*. Then the following two statements are equivalent.*

1. $\mathbf{e}$ *additively dominates* $\mathbf{e}^{\circ}$ *for any additive index based on the class* $\mathbf{\Phi}^{***}$ *of functions.*

2. $\sum_{b \in \beta(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ}) : b \leq \overline{b}} S^{\mathbf{e}}(s,b) \geq \sum_{b \in \beta(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ}) : b \leq \overline{b}} S^{\mathbf{e}^{\circ}}(s,b)$ *for every success target* $s \in \boldsymbol{\sigma}(\mathbf{e}) \cup \boldsymbol{\sigma}(\mathbf{e}^{\circ})$, *and every family background* $\overline{b} \in \boldsymbol{\beta}(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})$.

Proof. Statement 1 implies statement 2: *Assume that* $\mathbf{e}$ *is at least as good as* $\mathbf{e}^{\circ}$ *for any additive index based on the class* $\boldsymbol{\Phi}^{***}$ *of functions. Hence, one has Inequality (11) for all functions* Φ *in the class* $\boldsymbol{\Phi}^{***}$. *In particular therefore, inequality (11) holds for the function* $\Phi^{\overline{b}s}$ *defined, for any* $s \in \boldsymbol{\sigma}(\mathbf{e}) \cup \boldsymbol{\sigma}(\mathbf{e}^{\circ})$ *and* $\overline{b} \in \boldsymbol{\beta}(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})$, *by :*

$$\begin{aligned} \Phi^{\overline{b}s}(x,y) &= \max[x-s,0] \text{ if } y \leq \overline{b} \\ &= 0 \text{ otherwise} \end{aligned}$$

Like in the proofs of the two previous theorems, this function is not continuously differentiable, but it can be approximated by a continuously differentiable function following the arguments of Fishburn and Vickson (1978). $\Phi^{\overline{b}s}$ *is clearly increasing (weakly) with respect to its first argument (for a given* y*). It is also weakly decreasing with respect to* y. *Moreover, the "rate of increase of* $\Phi^{\overline{b}s}$*" with respect to* x *is decreasing with respect to* y *if* s *is increasing. The function* $\Phi^{\overline{b}s}$ *is clearly convex with respect to* x *for any given* y. *As is clear, this convexity is weakly decreasing with respect to* y *(as the function becomes constant when* y *becomes strictly larger than* $\overline{b}$. *Hence Inequality (11) must hold for any such function* $\Phi^{\overline{b}s}$ *so that one must have, for every* $(s,\overline{b}) \in \mathcal{S} \times \mathcal{B}$:

$$\begin{aligned} \sum_{i=1}^{n} \Phi^{\overline{b}s}(s_i,b_i) &\geq \sum_{i=1}^{n} \Phi^{\overline{b}s}(s_i^{\circ},b^{\circ}{}_i) \\ &\Longleftrightarrow \\ \sum_{b \in \beta(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ}) : b \leq \overline{b}} S^{\mathbf{e}}(s,b) &\geq \sum_{b \in \beta(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ}) : b \leq \overline{b}} S^{\mathbf{e}^{\circ}}(s,b) \end{aligned}$$

as required by the first part of statement (3). As for the second part of statement (3), one can simply observe that the function Φ^{b} *defined by:*

$$\begin{aligned} \Phi^{b}(x,y) &= 1 \text{ if } y \geq b \\ &= 0 \text{ otherwise} \end{aligned}$$

belongs also to the class $\boldsymbol{\Phi}^{***}$ *for any* $b \in \boldsymbol{\beta}(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})$. *Hence, inequality (11) must hold for the function* Φ^{b} *as well and, as a result, one must have* $\#Q^{\mathbf{e}}(\underline{s},b) \geq \#Q^{\mathbf{e}^{\circ}}(\underline{s},b)$ *for every* $b \in \boldsymbol{\beta}(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})$.

Statement 3 implies statement 2

We start from Inequality (15). If we Abel decompose this Inequality once more *with respect to the cognitive skill sum operator, we obtain:*

$$\begin{aligned}
& \sum_{h=l}^{2}\sum_{g=l}^{h}\sum_{k=1}^{m}\Delta f(s_g,b_k)\Phi_{11}(s_{h-1},b_m) \\
& -\sum_{h=l}^{1}\sum_{k=1}^{m-1}\Delta f(s_h,b_k)\Phi_2(s_1,b_k) \\
& -\sum_{h=l}^{2}\sum_{g=l}^{h}\sum_{k=1}^{m-1}\Delta f(s_g,b_k)\Phi_{21}(s_{h-1},b_k) \\
& -\sum_{h=l}^{2}\sum_{g=l}^{h}\sum_{k=1}^{m-1}\sum_{j=1}^{k}\Delta f(s_g,b_j)\Phi_{12}(s_1,b_k) \\
& +\sum_{h=l}^{3}\sum_{g=l}^{h}\sum_{e=l}^{g}\sum_{k=1}^{m-1}\sum_{j=1}^{k}\Delta f(s_e,b_j)[\Phi_{12}(s_{h-2},b_k)-\Phi_{12}(s_{h-1},b_k)] \\
\geq\ & 0
\end{aligned}$$

or, recalling that

$$\sum_{g=l}^{i}(s_g-s_i)\Delta f(s_g,b_k)/\Delta^s=\sum_{h=l}^{i+1}\sum_{g=l}^{h}\Delta f(s_g,b_k)$$

for every $i\in\{1,...,l\}$ *and* $k\in\{1,...,m\}$ *:*

$$\begin{aligned}
& -\sum_{k=1}^{m}\sum_{g=l}^{1}(s_g-s_1)\Delta f(s_g,b_k)\Phi_{11}(s_{g-1},b_m)/\Delta^s \\
& -\sum_{k=1}^{m-1}\sum_{g=l}^{1}(s_g-s_1)\Delta f(s_g,b_k)\Phi_{21}(s_{g-1},b_k)/\Delta^s \\
& -\sum_{k=1}^{m-1}\sum_{j=1}^{k}\sum_{g=l}^{1}(s_g-s_1)\Delta f(s_g,b_j)\Phi_{12}(s_1,b_k)/\Delta^s \\
& +\sum_{h=l}^{3}\sum_{k=1}^{m-1}\sum_{g=l}^{h-1}(s_g-s_{h-1})\sum_{j=1}^{k}\Delta f(s_g,b_j)[\Phi_{12}(s_{h-2},b_k)-\Phi_{12}(s_{h-1},b_k)]/\Delta^s \\
\geq\ & 0 \qquad (31)
\end{aligned}$$

Recognizing that

$$\begin{aligned}
[\Phi_{12}(s_{h-2},b_k)-\Phi_{12}(s_{h-1},b_k)] &= [[\Phi(s_{h-1},b_{k+1})-\Phi(s_{h-2},b_{k+1})]-[\Phi(s_{h-1},b_k)-\Phi(s_{h-2},b_k)]] \\
&\quad -[\Phi(s_h,b_{k+1})-\Phi(s_{h-1},b_{k+1})]-[\Phi(s_h,b_k)-\Phi(s_{h-1},b_k)]] \\
&= [[\Phi(s_{h-1},b_{k+1})-\Phi(s_{h-2},b_{k+1})]-[\Phi(s_h,b_{k+1})-\Phi(s_{h-1},b_{k+1})]] \\
&\quad -[[\Phi(s_{h-1},b_k)-\Phi(s_{h-2},b_k)]-[\Phi(s_h,b_k)-\Phi(s_{h-1},b_k)]] \\
&= [\Phi_1(s_{h-2},b_{k+1})-\Phi(s_{h-1},b_{k+1})]-[\Phi_1(s_{h-2},b_k)-\Phi_1(s_{h-1},b_k)] \\
&= \Phi_{11}(s_{h-2},b_k)-\Phi_{11}(s_{h-2},b_{k+1}) \\
&\geq 0
\end{aligned}$$

for every $\Phi \in \mathbf{\Phi}^{***}$, one can see that a sufficient condition for Inequality (31) to hold for all functions $\Phi \in \mathbf{\Phi}^{***}$ is to have:

$$\sum_{g=l}^{h}(s_g - s_h)\sum_{j=1}^{k}\Delta f(s_g, b_j) \geq 0$$

for every $h \in \{0, ..., l\}$ and every $k \in \{1, ..., m\}$. But this is nothing else than requiring $\sum_{b\in\beta(\mathbf{e})\cup\boldsymbol{\beta}(\mathbf{e}^{\circ}):b\leq\overline{b}} S^{\mathbf{e}}(s,b) \geq \sum_{b\in\boldsymbol{\beta}(\mathbf{e})\cup\boldsymbol{\beta}(\mathbf{e}^{\circ}):b\leq\overline{b}} S^{\mathbf{e}^{\circ}}(s,b)$ to hold for every success target $s \in \boldsymbol{\sigma}(\mathbf{e}) \cup \boldsymbol{\sigma}(\mathbf{e}^{\circ})$, and every family background $\overline{b} \in \boldsymbol{\beta}(\mathbf{e}) \cup \boldsymbol{\beta}(\mathbf{e}^{\circ})$. ■

Theorem 3 thus provides an additional criterion, consistent with, but more discriminatory then, the two previous ones, for comparing alternative school systems. The test works in a very simple fashion. One first assigns, to every value of the family background below some threshold, a target of cognitive skill. One then sums, over all children with family background below the threshold, the excess of cognitive skills of the children who exhibit a better performance than the cognitive target. An education system who exhibits a larger sum of excess of cognitive skills of children over their target than another for every choice of cognitive target and every threshold of family background is then considered better than the other. It is intuitively clear that this test is more discriminatory than the previous one. In effect, it can be viewed as a special case of the previous test for which the family background-dependant cognitive target is constant below some background threshold, and then increased to the maximal possible level for all family background above the threshold.

3 Empirical Analysis

3.1 Data and variables

We base our empirical analysis on the 2012 wave of the OECD Program for International School Assessment (PISA) survey. This survey assesses the skills of 510 000 students - aged 15 - across 65 countries from the OECD. It also provides information on their parents and family environment that comes from the children' parents themselves. The sample of children in each country is based on a random selection of a sample of schools from which, in a second step, a random selection of the pupils is performed. In the last step, individuals are weighted in such a way as to make the sample representative of the actual population of the country. It is important to notice that despite these corrections, the samples of pupils evaluated in the PISA survey are not totally representative of the population of interest (children aged 15). Indeed, children who are not enrolled at school, or who are enrolled in very low grades for their age, or who do not go to school because of physical or intellectual deficiencies are excluded from the sample. While this limitation in the coverage of children is small in developed countries (where more than 90% of the children are represented by the sample), the fraction covered is significantly lower for developing countries, and is definitely not uniform across countries (see e.g. Carvalho, Gamboa, and Waltenberg (2012)). The cognitive skills evaluated by the PISA survey concern 3 different subjects: Mathematics, Reading and Science. We focus herein on the Mathematics test, which seems to be less culturally biased, and less correlated with family background. The results achieved by children on the test are standardized by the PISA team through a somewhat complex Item Response methods described in PISA (2014). At the end of the procedure, the pupils' performance at the test are standardized at the OECD mean of 500 with a standard deviation of 100, as shown in the following figure, that depicts the world distribution of this standardized score. We emphasizes that this standardization - based on the normal distribution - should make one somewhat hesitant in attaching cardinal significance to the information conveyed by the score.

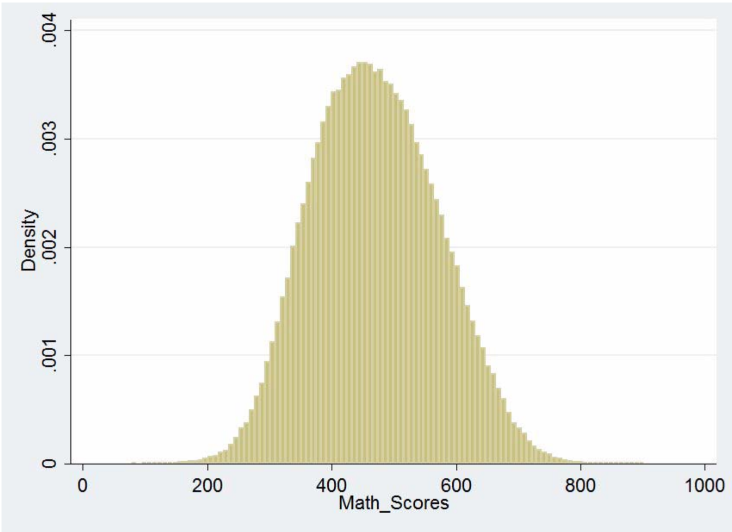

Figure 2: Distribution of standardized scores in mathematics at the world level.

While the distribution of these scores is by construction normal (Gaussian) at the world level, this normality is not reproduced at the level of every country. In this paper, we focus attention on the 46 "full" countries that have more than 5 millions inhabitants. This excludes cities such as Shanghai or Hong Kong, and Taipei which are only parts of larger countries. We believe indeed that it would not have been appropriate to treat Shanghai - say - as a national education system as a whole that could be compared with that of, say, the US. The following table shows the ranking of the 46 countries as per their average math score.

Rank	Country	Maths PISA Score	Rank	Country	Maths PISA Score
1	Singapore	573	24	USA	481
2	Korea	554	25	Sweden	478
3	Japan	536	26	Hungary	477
4	Switzerland	531	27	Israel	466
5	Netherlands	523	27	Greece	453
6	Finland	519	28	Serbia	449
7	Canada	518	29	Turkey	448
8	Poland	518	30	Romania	445
9	Belgium	515	31	Bulgaria	439
10	Germany	514	32	Arab Emirates	434
11	Vietnam	511	33	Kazakhstan	432
12	Austria	506	34	Thailand	427
13	Australia	504	35	Chile	423
14	Denmark	500	36	Mexico	413
15	Czech Republic	499	36	Uruguay	409
16	France	495	37	Brazil	391
17	UK	494	38	Argentina	388
18	Norway	489	39	Tunisia	388
19	Portugal	487	40	Jordan	386
20	Italy	485	41	Colombia	376
21	Spain	484	42	Qatar	376
22	Russia	482	43	Indonesia	375
23	Slovakia	482	44	Peru	368

Table 1: 46 OECD countries ranked by Average PISA Maths Score

As for the measurement of family background, the PISA 2012 database provides two pieces of information. The first one is the *highest educational attainment of the parents*, defined as the highest ISCED (International Standard Classification of Education) level of the two parents. The ISCED comprises seven discrete categories, ranging from 0 (no education) to 6 (second stage of tertiary education). The PISA database contains also information on the *highest occupational status* of the parents, as defined by the International Socioeconomic Index (ISEI) of Occupational Status of Ganzeboom, Graaf, and Treiman (1992)). The ISEI can be described as a weighted average of the parent education and their profession, the later being itself ranked by the average level of income associated to that profession within the country. In the PISA data base, the largest ISEI of the two parents appears as a continuous variable ranging from 0 to 100. While we have done the analysis for either of the two indicators of the family background, we focus in what follows on the ISEI. We justify this choice by the fact that this indicator is sensitive to both the income and the education of the children parents, and that it is somewhat well documented that both parents education and parents income contribute to the children human capital. To that extent, we feel that ISEI is a more comprehensive summary indicator of the family background of a given child than the more educationally focused ISCED one. Figure 4 below shows the world distribution of ISEI in the PISA data base.

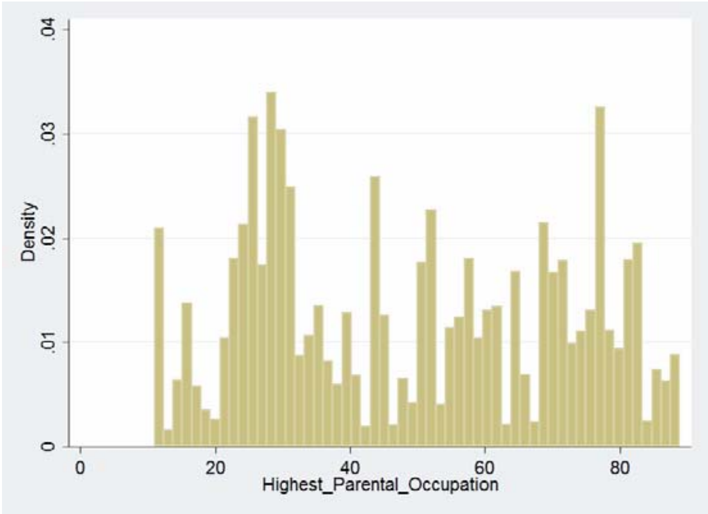

Figure 4: Distribution of the occupational status in the 2012 PISA data base.

As for the math score, we also provide, in the following tables, the ranking of our 46 countries in terms of the children's parents ISEI.

(TO BE PROVIDED BY EDWARD)

As our criteria are somewhat sensitive to the correlation between cognitive skill - as measured by the math score - and family background, it may be appropriate to provide empirical evidence on this. The two following tables provides a ranking of the countries in terms of the correlation that they exhibit between children PISA scores in mathematics and family backgrounds. We use in turns two measures of correlation: the usual Spearman coefficient (covariance divided by the product of the standard deviation of each of the variable), and the regression coefficient $\widehat{\beta}(\mathbf{e})$ discussed above. At the world level, the correlation between math cognitive skill and family backgrounds is significant (0.3327), but not outlandish.

Figure 5 shows the correlation between the two variables measured at the country level. Again, the correlation is visible, but not astonishing.

Rank	Country	Elasticity	Rank	Country	Elasticity
1	Kazakhstan	0.475	24	Turkey	1.402
2	Indonesia	0.785	25	United States	1.405
3	Jordan	0.814	26	United Kingdom	1.415
4	Mexico	0.889	27	Netherlands	1.418
5	Japan	1.016	28	Serbia	1.507
6	Canada	1.141	29	Peru	1.537
7	Brazil	1.172	30	Romania	1.577
8	Colombia	1.172	31	Poland	1.587
9	Norway	1.174	32	Arab Emirates	1.632
10	Italy	1.228	33	Thailand	1.658
11	Korea	1.233	34	Singapore	1.730
12	Finland	1.239	35	Bulgaria	1.730
13	Russia	1.274	36	France	1.759
14	Vietnam	1.286	37	Portugal	1.780
15	Sweden	1.310	38	Qatar	1.781
16	Denmark	1.320	39	Hungary	1.831
17	Argentina	1.335	40	Belgium	1.853
18	Spain	1.347	41	Germany	1.860
19	Switzerland	1.359	42	Uruguay	1.895
20	Greece	1.362	43	Chile	1.923
21	Tunisia	1.370	44	Israel	1.938
22	Australia	1.376	45	Czech Republic	1.962
23	Austria	1.379	46	Slovakia	2.059

Table 2: 46 OECD countries ranked by Elasticity of Parental Social Status (ISEI) on their Children's Maths PISA Score

Rank	Country	Correlation	Rank	Country	Correlation
1	Kazakhstan	0.162	24	Spain	0.342
2	Japan	0.232	25	Brazil	0.345
3	Korea	0.233	26	United States	0.345
4	Jordan	0.257	27	Serbia	0.359
5	Indonesia	0.260	28	Denmark	0.363
6	Norway	0.265	29	Greece	0.374
7	Arab Emirates	0.279	30	Tunisia	0.386
8	Mexico	0.281	31	Poland	0.396
9	Canada	0.286	32	Argentina	0.397
10	Italy	0.294	33	France	0.407
11	Finland	0.302	34	Germany	0.409
12	Australia	0.310	35	Belgium	0.415
13	Turkey	0.311	36	Czech Republic	0.417
14	Vietnam	0.315	37	Thailand	0.417
15	Switzerland	0.317	38	Israel	0.418
16	Sweden	0.318	39	Portugal	0.419
17	Netherlands	0.321	40	Bulgaria	0.422
18	Qatar	0.324	41	Peru	0.423
19	Singapore	0.325	42	Hungary	0.434
20	Russia	0.326	43	Slovakia	0.435
21	Austria	0.330	44	Romania	0.444
22	Colombia	0.335	45	Uruguay	0.472
23	UK	0.338	46	Chile	0.540

Table 3: 46 OECD countries ranked by Correlation of Parental Social Status (ISEI) with their Children's Maths PISA Score

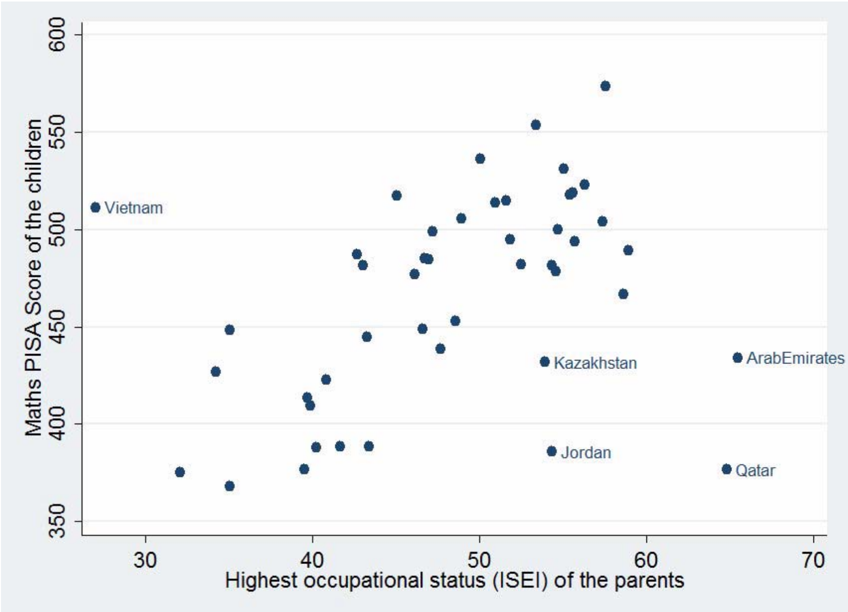

Figure 5: Cross country relation between children cognitive skill and family background.

The somewhat atypical position of Vietnam, with very good average performance in cognitive skill and low average family background, and countries such as Qatar or Arab Emirates (with good average family backgrounds and somewhat modest performance in math scores should be noticed and kept in mind for the understanding of the results, to the presentation of which we now turn.

3.2 Results

In order to apply the criteria discussed in the previous section, we find appropriate to categorize a bit the children in terms of their Math Scores parental ISEI so as to avoid an excessive dependency of the conclusion of the analysis upon the measurement of the skill or family backgrounds provided by our indicators. Our criteria are, indeed, incomplete - they do not always lead to clear-cut comparisons of education systems - and they are more likely to be incomplete when applied to school systems that differ by a large number of distinct pairs of numbers. Hence, reducing the number of possible pairs of distinct numbers by categorizing them increases the likelihood of obtaining clear-cut comparisons.

Concerning math scores, we have adopted the following categorization of the variable into eight groups, defined as follows:

- Category 1 : $\textit{PISA Math Score} \leq 300$
- Category 2 : $300 \leq \textit{PISA Math Score} \leq 350$
- Category 3 : $350 \leq \textit{PISA Math Score} \leq 400$

- Category 4 : $400 \leq$ *PISA Math Score* ≤ 450
- Category 5 : $450 \leq$ *PISA Math Score* ≤ 500
- Category 6 : $500 \leq$ *PISA Math Score* ≤ 550
- Category 7 : $550 \leq$ *PISA Math Score* ≤ 600
- Category 8 : *PISA Math Score* ≥ 600

This corresponds to an equal spacing (by 50) of the interval of scores [300,600] (which concentrates 95% of the world sample population) to which are added the bottom (less than 300) and above (more than 600) part of the [0,1000] interval.

As for the parent's ISEI score, we have categorize it based on the quintiles of world distribution of this variable, as depicted on Figure 4 above. This generates the following 5 categories:

- Category 1 : $ISEI \leq 25.95$
- Category 2 : $25.95 < ISEI \leq 36.35$
- Category 3 : $36.35 < ISEI \leq 56.98$
- Category 4 : $56.98 < ISEI \leq 72.94$
- Category 5 : $ISEI \geq 72.94$

We start by comparing countries on the basis of first-order stochastic-dominance applied only to the distribution of the children's Maths PISA scores separated into 8 categories, as just described. The results are presented in the Hasse diagram shown on Figure 1. First, countries are ordinally plotted by mean Maths PISA score from bottom to top, and then links are drawn between them to indicate the possible presence of inverted first-order stochastic-dominance. For instance, one can see that Korea - with the second highest mean Maths PISA score - dominates Japan and Switzerland among others, but does not dominate the Netherlands despite having a higher mean Maths PISA score. Similarly, we have applied first-order stochastic-dominance to the distribution of the Parents' highest ISEI, separated - as mentioned above - into 5 categories. The results are shown in figure 2. Again, the countries are ordinally plotted by mean highest ISEI of the two parents, and linked to one other to signal the possible presence of first-order stochastic dominance. By comparing figures 1 and 2, one can note that the Arab Emirates who is the most dominating country in terms of highest parental ISEI (figure 2), is dominated by a great number of countries in Maths PISA scores (figure 1). In contrast, one may also note that Vietnam - which is the most dominated country in terms of highest parental ISEI - dominates a large portion of the countries in Maths PISA Scores.

Insert Hasse Diagram with Children Maths scores

Insert Hasse Diagram with Parents HISEI

The two rankings in figures 1 and 2 are also helpful to understand the results that have been generated by our criteria, and that we present in the in the next section. Indeed the first-order criterion that we have developed in this paper (make reference), requires - as necessary conditions - that there be first -order Inverse Stochastic Dominance for the children, and that there be first-order stochastic dominance for the parents. In other words, for our first-order criterion to rank a country A as better than B, one needs at least for:

- the children's marginal distribution of Maths PISA scores in A to dominate that in B, with first-order inverse stochastic dominance
- the parents' marginal distribution of highest ISEI in B to dominate that in A, with first-order stochastic dominance

It follows that a country which dominates another in figure 1, and who is dominated by the the same country in figure 2 is likely to be ranked better by our criterion. For instance, one can see that the Finnish children dominate the Norwegian children with first-order inverse Stochastic dominance (figure 1), and that the Finnish parents are dominated by the Norwegian parents with first-order stochastic dominance (figure 2). Thus, our criterion might rank Finland as better than Norway.

The dominance conditions applied to the marginal distributions of parents and children, being necessary conditions, if a country A is not ranked better than B in figure 1, or not ranked worse than B in figure 2, then it will be impossible for A to be ranked better than B in our rankings. Such is the case with Singapore and Vietnam, since Singapore both has better ranked children and better ranked parents than Vietnam.

Furthermore, one must remember that our criterion also takes into account the correlation between the parents and their children. Although two countries satisfy the dominance conditions discussed above, they might not be ranked better than one another if the correlation between parents and children is too high. In the next section, before we unveil the results of our criterion - with its 3 variants - we will first discuss the statistical inference methodology used to improve the reliability of our results.We provide on Figure 6 the Hasse diagram that describes the incomplete ranking of the 46 countries provided by our first criterion (associated to the unanimity of all additive indices that agree with Axioms 1-3). Any continuous (without "hole") combination of horizontal and at least one vertical lines connecting two countries indicates a dominance of the "above" country over the "below" one. As can be seen, there are many pairs of countries that can not be compared by this very demanding (but very robust) dominance criterion. Finland, Germany, Japan, Singapore and Switzerland are the only countries that can not be compared with any others. They are, in this sense, outliers. At the top of the tree are countries that dominate several countries but who are dominated by none. These countries are Austria, Canada, Belgium, Czechia, Netherlands Poland, Romania, Russia, Mexico, Spain, Thailand, Turkey, Uruguay and, of course, Vietnam. This country, who dominates twelve others, is clearly the major player in this top group. This prominence is explained by the fact that Vietnam pupils are by large coming from family with very unfavorable family background. Since these pupils are doing rather well at PISA test, it appears that Vietnam education system outperforms many others. Another very good performer in this "top club" is Austria, who dominates 8 countries. Minor players in this top club are Belgium (who dominates only Israel) and Mexico (who dominates Argentina and Jordan). At the bottom of the trees, one finds countries that are dominated by several others, but dominate none. Members of this "bottom" pools are Arab Emirates, Argentina, Brazil, Bulgaria, Columbia, Denmark, Greece, Israel, Indonesia, Italy, Jordan, Kazakhstan, Malaysia, Peru, Qatar and the US. Particularly noticeable members of this clubs are the Arab Emirates and Qatar (who are each dominated by eight countries). The main reason for this is the polar extreme opposite of Vietnam.

Indeed, the children trained in the education systems of these countries are coming from family with good background (as measured by ISEI) even though they end up performing poorly in PISA Math tests. While the performance of Denmark and the US pupils at Math tests is a bit better than that of Arab Emirates and Qatar, those two countries also appear at the bottom of the list because of their extremely favorable distribution of family backgrounds.

Overall, the criterion is somewhat incomplete. Indeed, among all the possible pairs of countries that could be compared, only 7.6% are ranked conclusively by this criterion. While this may be seen as disappointing, we prefer viewing the 7.6% of the ranking obtained as being extremely robust. To insist on the Vietnam case, our ranking concludes indeed in an outstanding performance of the education system of this country. We view this as something that does not emerge spontaneously from a casual look at PISA ranking, in which the average Math Score of Vietnam (511) makes this country only a bit above the OECD average. However, considering that the average social status of Vietnamese parents (ISEI) is by far the lowest of all countries, the Math attainments of their children appear in fact to be outstanding. To highlight just how impressive the attainments of Vietnamese children are at the Math PISA test, one can compare them to the attainments of children from countries with similar average parental social-statuses. Peru and Indonesia for instance whose average parental social status are comparable to the Vietnamese, respectively score an average of 368 and 375 at the Math PISA tests, making them the 2 lowest ranked countries of the official PISA ranking.

TO BE CONTINUED.

A way to generate more comparison is to pay the price of accepting additional value judgements. One of them is the idea of elitism, which considers that, among children with the same family background and the same average score, it is a good thing to increase the spread of the distribution. Figure 6 provides a similar Hasse ranking of the countries that adds this elitist principle to the other three principles underlying Axioms **??**-**??**. We emphasize that this value judgement requires that cardinal significance be attached to the test score categories, an assumption that is not innocuous, given the rather important normalization done by the PISA investigators.

It is visually apparent that this diagram is significantly more "vertical" than the previous one. This visual impression is comforted by the observation that this new dominance criterion enables one to rank conclusively 34.9% of all possible comparisons. Hence, it seems that elitism is a principle that contributes a great deal to the explanation of the performance of the various national education systems. For instance, the usual champion in PISA studies - Finland - could not be compared with any other countries with our first criterion. It can now be ranked above Norway and Sweden, and below Japan. Since having a worse distribution of family background - as per first order stochastic dominance - is a necessary condition for having one distribution ranked above another, the ranking of Finland above Norway and Sweden implies that the distribution of family backgrounds is less favorable in Finland than in each of these two countries as per first order dominance. As Finland could not be compared with Norway and Sweden by our first criterion, one must conclude that this failure was coming from the fact that the distribution of scores in mathematics among Finland pupils did not dominate that of either Norway or Sweden at the first order (even though Finland has a better average sore than these two countries). The fact that Finland dominates Norway and Sweden when one introduces elitism thus implies that the average score of the best performing children in Finland is better than the average score of their Norwegian or Swedish counterparts no matter how one defines the threshold of excellence.

4 Conclusion

In this paper we have developed two robust criteria for comparing education systems on the basis of their ability to equalize the opportunities of children from all backgrounds, whilst maintaining best possible level of education. This criterion requires that everything else being equal, the improvement of a child's cognitive abilities is considered positively. It also requires that all else being equal, a child's achievement is all the more remarkable as his background is unfavorable to his success. The third condition requires that the correlation between the cognitive success of a child and their parents educational success be as low as possible. The criterion that we implement coincides with the unanimity of all comparisons of education systems that agree with these three principles.

TO BE COMPLETED AND DETAILED.

References

Allison, R. A., and J. E. Foster (2004): "Measuring health Inequality Using Qualitative Data," *Journal of Health Economics*, 23, 505–524.

Atkinson, A. B. (1981a): "The Measurement of Economic Mobility," in *Inkommens Verdeling en Openbard Financien*, ed. by P. Eggelshoven, and L. vanGemerden. Het Spectrum.

——— (1981b): "On Intergenerationnal Income Mobility in Britain," *Journal of Post-Keynesian Economics*, 3, 194–218.

Atkinson, A. B., and F. Bourguignon (1982): "The Comparison of Multi-dimensioned Distribution of Economic Status," *Review of Economic Studies*, 49, 183–201.

——— (1987): "Income Distributions and Differences in Needs," in *Arrow and the Foundation of the Theory of Economic Policy*, ed. by G. R. Feiwel. Macmillan, London.

Barrett, G. J. (2012): "The Returns to Cognitive Skills in the Australian Labour Market," *The Economic Record*, 88, 1–17.

Bazen, S., and P. Moyes (2012): "Elitism and Stochastic Dominance," *Social Choice and Welfare*, 39, 207–251.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005): "Why the Apple doesn't fall far: understanding intergenerational transmission of human capital," *American Economic Review*, 95, 437–439.

Bourguignon, F. (1989): "Family Size and Social Utility: Income Distribution Dominance Criteria," *Journal of Econometrics*, 42, 67–80.

Carvalho, M. D., L. F. Gamboa, and F. D. Waltenberg (2012): "Equality of Educational Opportunity Employing Pisa Data: Taking Both Achievement and Access into Account," ECINEQ working paper no. 2012-277.

Dahl, G. B., and L. Lochner (2012): "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax," *American Economic Review*, 102, 1927–1956.

Debreu, G. (1954): "Representation of a Preference Ordering by a Numerical Function," in *Decision Processes*, ed. by R. L. D. R. M. Thrall, C. H. Coombs, pp. 159–165. Wiley, New York.

——— (1960): "Topological methods in cardinal utility theory," in *Mathematical methods in the social sciences, 1959; proceedings*, ed. by S. K. K. J. Arrow, and P. Suppes, pp. 16–26. Stanford University Press, Stanford.

Duclos, P. Y., D. Sahn, and S. D. Younger (2006): "Robust Multidimensional Poverty Comparisons," *The Economic Journal*, 116, 943–968.

Fields, J. S., and E. A. Oke (1999): "The measurement of income mobility: an introduction to the literature," in *Handbook on Income Inequality Measurement*, ed. by J. Silber, pp. 557–596. Kluwer, Norwell, MA.

Fishburn, P. C., and R. G. Vickson (1978): "Theoretical Foundations of Stochastic Dominance," in *Stochastic Dominance*, ed. by G. A. Withmore, and M. C. Findlay. Lexington Books.

Ganzeboom, H. B. G., P. D. Graaf, and D. J. Treiman (1992): "A standard international socio-economic index of occupational status," *Social Science Research*, 21, 1–56.

Goussé, M., and N. LeDonné (2015): "Why do inequality in 15-year-old cognitive skill increases so much in France between 2000 and 2009 ?," Mimeo, Université Laval.

Gravel, N., B. Magdalou, and P. Moyes (2015): "Ranking Distributions of an Ordinal Attribute," AMSE working paper, no 2015-50.

Gravel, N., and P. Moyes (2012): "Ethically Robust Comparisons of bidimensionnal Distributions with an Ordinal Attribute," *Journal of Economic Theory*, 147, 1384–1426.

Gravel, N., P. Moyes, and B. Tarroux (2009): "Robust International Comparisons of Distributions of Disposable Income and Access to Regional Public Goods," *Economica*, 76, 432–461.

Gravel, N., and A. Mukhopadhyay (2010): "Is India better off now than fifteen years ago ? A robust multidimensional answer," *Journal of Economic Inequality*, 8, 173–195.

Green, D. A., and W. C. Riddell (2003): "Literacy and Earnings: An Investigation of the Interaction of Cognitive and Unobserved Skills in Earnings Generation," *Labour Economics*, 10, 165–184.

Hanushek, E. A., G. Schwerdt, and L. Woessmann (2015): "Returns to Skills around the World: Evidence from PIAAC," *European Economic Review*, 73, 103–130.

Hanushek, E. A., and L. Woessmann (2008): "The Role of Cognitive Skills in Economic Development," *Journal of Economic Literature*, 46, 607–668.

——— (2011): "The Economics of International Differences in Educational Achievement," in *Handbook of the Economics of Education, Vol. 3*, ed. by E. A. Hanushek, S. Machin, and L. Woessmann. North Holland, Amsterdam.

Heckman, J. J., J. J. Stixrud, and S. Uzrua (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 24, 411–482.

Hussain, M. A., M. M. Jorgensen, and L. P. Osterdal (2016): "Refining Population Health Comparisons: A Multidimensional First Order Dominance Approach," *Social Indicators Research*, forthcoming.

Jenkins, S. P., and P. J. Lambert (1993): "Ranking Income Distributions when Needs Differ," *Review of Income and Wealth*, 39, 337–356.

Lehmann, E. L. (1955): "Ordered Family of Distributions," *The Annals of Mathematics and Statistics*, 26, 399–419.

Mayer, S. A. (1997): *What Money Can't Buy: Family Income and Children's Life Chances*. Harvard University Press, Cambridge, MA.

Moyes, P. (2012): "Comparaisons of Heterogeneous Distributions and Dominance Criteria," *Journal of Economic Theory*, 147, 1351–1383.

Muller, C., and A. Trannoy (2012): "Multidimensionnal Inequality Comparisons: A Compensation Perspective," *Journal of Economic Theory*, 147, 1427–1449.

Nickell, S. (2004): "Poverty and Worklessness in Britain," *Economic Journal*, 114, C1–C25.

Oreopoulos, P., and K. G. Salvanes (2011): "Priceless: the Non-Pecuniary Benefit of Education," *Journal of Economic Perspectives*, 25, 159–184.

Osterdal, L. P. (2010): "The Mass Transfer Approach to Multivariate Discrete First Order Stochastic Dominance: Direct Proof and Implications," *Mathematical Social Sciences*, 46, 1222–1228.

PISA (2014): "PISA 2012 Technical Report," Paris, OECD, 476 pages.

Quirk, J. D., and R. Saposnik (1962): "Admissibility and Measurable Utility Functions," *Review of Economic Studies*, 29, 140–146.

Schutz, G., H. W. Ursprung, and L. Woessmann (2008): "Education Policy and Equality of Opportunity," *Kiklos*, 61, 279–308.

Shorrocks, A. F. (1978): "The Measurement of Mobility," *Econometrica*, 46, 1013–1024.

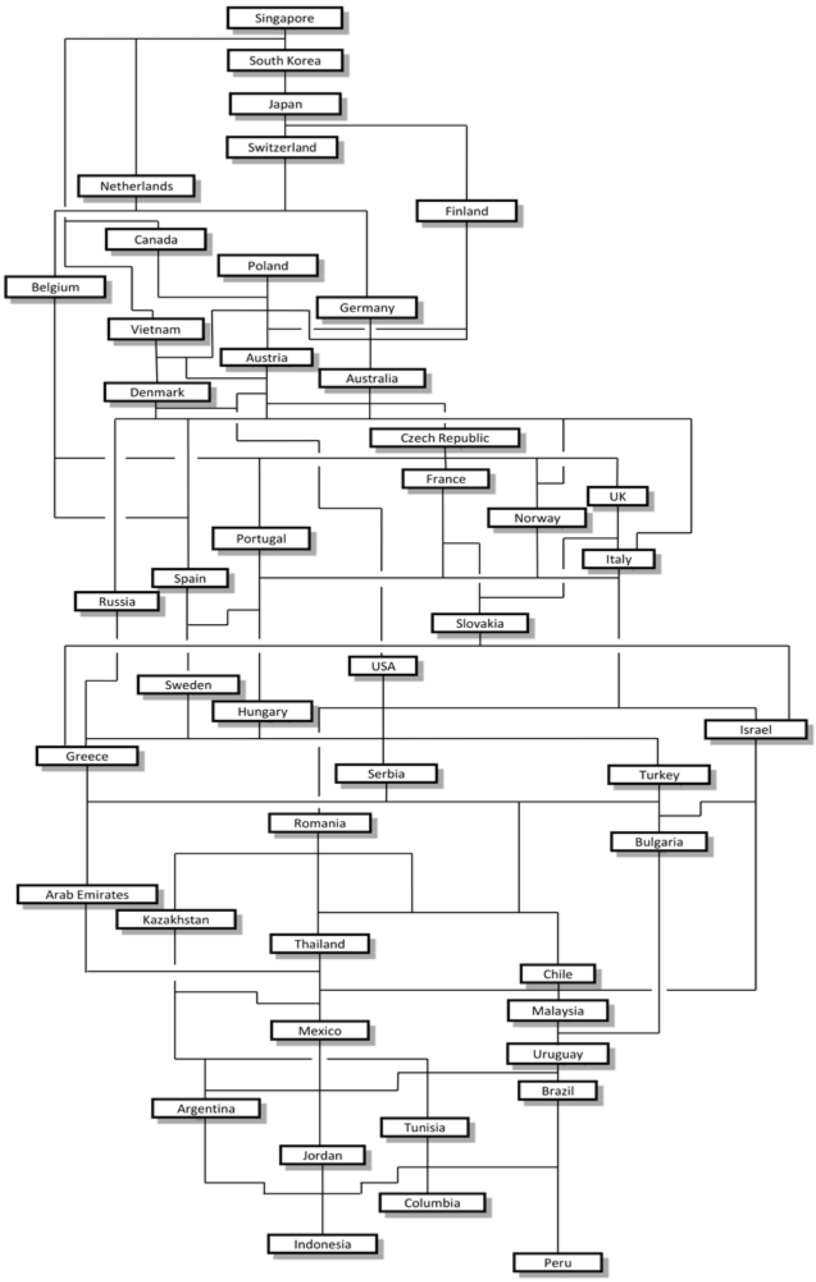

Figure 3: 1st order dominance chart of the distribution of Pisa Scores in Mathematics

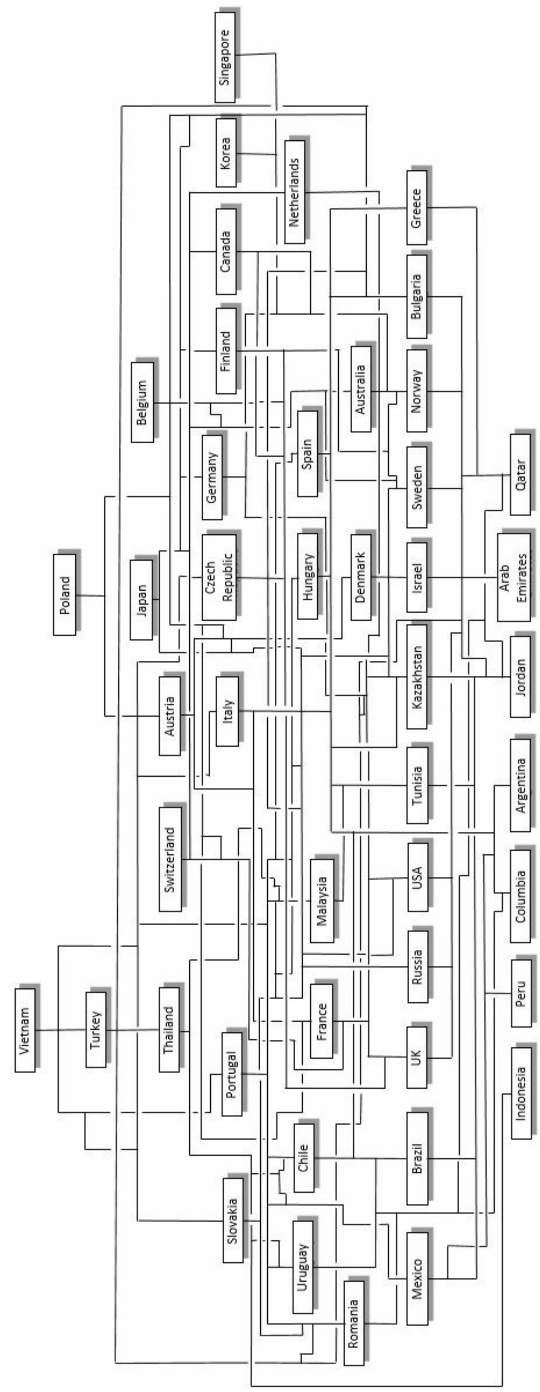

Figure 6: Hasse diagram of dominance, first criterion.