# Inequalities in Educational Achievement: Measurement and Evidence from the Programme International of Students' Assessments

Francisco H. G. Ferreira and Jérémie Gignoux[♦]
April, 2009

## 1. Introduction

How large are the observed inequalities in scholastic achievement and to what extent are these inequalities associated with pre-determined circumstance variables such as race, gender, the socio-economic status of one's parents, or place of birth? This paper examines these questions using internationally comparable data on standardized test scores for reading, mathematics and science, for fifteen year-olds in a large number of countries with advanced and developing economies.

While there is a literature that compares inequalities in educational *attainment* across countries, as measured by years of schooling, we are aware of very little work using educational *achievement* data, such as test scores.[1] Yet, it is widely recognized that years of schooling is an unsatisfactory measure of human capital, given the considerable heterogeneity in the quality of education across schools (not to mention across countries).[2] Educational attainment is not the same as educational achievement, and standardized test score measures are generally preferred to years of schooling as an indicator of the latter, since they seek to measure actual learning outcomes, rather than mere enrollment or, at best, attendance. In addition, since the advent of the large international surveys of students' achievements, test scores are substantially more comparable across countries, since the same tests (in different languages) are applied across participating countries.

♦ Development Research Group, The World Bank. Correspondence to fferreira@worldbank.org and jgignoux@worldbank.org.

1 Castelló and Doménech (2002), Thomas, Wang and Fan (2002) and Morrisson and Murtin (2006) examine inequalities in educational attainment across a large number of countries.

2 See, e.g. Pritchett (2004).

One reason why these data have been used relatively seldom to make inequality comparisons is the manner in which they are constructed. Different questions in a test have different levels of difficulty. To infer the adequate weighting scheme that will allow examiners to estimate the student's latent ability from his or her test score, item response theory (IRT) methods are used. These methods yield a distribution of plausible values with an indeterminate metric. The distributions are then standardized, so as to have the same mean and variance. This standardization implies both a translation of the mean, and a rescaling of the dispersion. There is no relative measure of inequality capable of deriving, from the transformed distribution, the inequality in the original distribution.[3] Overall inequality in achievements can therefore not be quantified and inequalities across groups or countries compared cardinally. Moreover, commonly used relative inequality measures, such as the Gini or the generalized entropy indexes, do not even provide a basis for ordinal comparisons (only given absolute inequality measures, such as the variance, allow such comparisons).

In this paper, we propose a solution to this problem. We show that there is a specific measure of inequality between social or spatial groups which satisfies two requirements: (i) because it is a ratio of inequality measures, it is metric-independent; and (ii) given its functional form, the ratio is both translation and scale-invariant. Given that it belongs to a well-defined class of measures of inequality of opportunity, the measure is of intrinsic interest. It provides a lower bound of the share in learning inequality that is attributable to pre-determined and transmitted circumstances lying beyond individual responsibility or genuine ability. Additionally, in the case where the standardization procedure is the same across all units (international surveys of achievement usually apply the same procedure in all countries), some measures of overall inequality can provide a basis for ordinal comparisons across groups and countries.

Another methodological issue is addressed: the selection of the sample of surveyed students due to dropping-out and grade repetition. The international surveys of achievements sample children of a given age currently attending school and not too late at school. The ensuing selection is high in many developing countries and particularly

[3] This follows from Zheng's (1994) impossibility result, namely: there is no meaningful measure of inequality that is both scale and translation-invariant.

problematic for cross-country comparisons of educational inequalities. Non-parametric reweighing procedures are implemented for imputing test scores to unobserved children using two alternative assumptions on the relationships between test scores and observed characteristics. The measures of the shares of inequalities associated to pre-determined circumstances seem quite robust to these corrections.

This paper also has a substantive contribution: it takes advantage of the availability of the large number of comparable surveys of students' achievements of the Programme International of Students' Assessments to provide international comparisons of inequalities in achievements across a large set of countries, including most OECD countries, but also developing countries in Latin American, East and Central Asia, Middle East and North Africa. These data, which were collected in 2006, provide information on achievements in reading, mathematics and science, along with a large set of family and geographic location characteristics. Measures of the shares of learning inequalities associated with predetermined circumstances, and relative measures of the levels of overall learning opportunities, are compared across countries. The application of these measures to policy analysis is also illustrated by relating inequalities in achievements with two basic indicators of the distribution of public expenditure in education and tracking in the educational system respectively: the share of public education expenditure at the primary level and the ratio of vocational attendance at the secondary level.

Our results indicate that a significant share of inequalities in learning achievement can be associated with a set of circumstances variables for parental background, gender and place of residence. This share ranges from 13% to 38% in reading, from 15% to 35%[4] for mathematics, and from 11% to 38% (depending on the country) in science. There is considerable heterogeneity among most regions, but the inequalities associated with pre-determined circumstances are found significantly higher in various continental West-European countries, such as Germany, France or Belgium, and in Latin American countries, notably in Brazil, than in others such as the Canada and the United States, Scandinavian Countries, and most Asian countries. Moreover, no statistically significant relationship is found between the shares or levels in those inequalities and economic

[4] After the exclusion of two outliers (Azerbaijan and Macao) for mathematics.

development measured by per capita growth domestic product. Finally, inequalities in educational opportunities are found to be significantly associated negatively with the share of public educational resources devoted to primary education but positively with early tracking.

The paper is organized as follows. Section 3 is on measurement. It first describes the inequality measures which are suitable for analyzing standardized, IRT-based test-score data, and details the parametric approach we employ to calculate it. It then addresses the coverage issue and proposes a correction procedure. Section 3 describes the PISA data sets, including the test scores and circumstance variables. Section 4 presents the measures of inequality in achievement and provides an international comparison. Section 5 investigates the relationship between those measures and two specific education policies, and Section 6 concludes.

## 2. The measurement of inequality in educational achievements

Two issues must be addressed for the measurement of inequality in educational achievements. The first one concerns the standardization of test scores variables and the second one the incomplete coverage of the samples of surveys of students' achievements. These two methodological issues are subsequently addressed in this section.

### *2.1. Addressing the standardization of test scores*

The first issue arises from the construction of test score variables. There is essentially no scale of scholastic achievements. Test scores are thus provided in arbitrarily standardized scales, and this creates serious constraints on the measurement of inequality in those variables.

Test scores variables are constructs from statistical *item response theory* models. [5] The idea behind item response theory (IRT) is to determine how much of a given unobservable or latent trait (in this case, cognitive skills or achievement) individuals possess. As this trait can not be measured directly, IRT seeks to infer it from a set of responses to test items. IRT methods consist of modeling the item responses as the

[5] See Baker (2001) for an introduction to Item Response Theory, and Martin (1998) or OECD (2006) for a description of how the method was applied to the TIMSS and PISA surveys.

outcome of two sets of independent parameters, one describing the items and the other the examinee's skills. An *item response model* can be written as $p(x|\theta,\alpha)$: the probability of scoring *x* in a given test, given individual latent ability $\theta$ and test item parameters $\alpha$ (such as their difficulty). The approach also requires an assumption about the distribution of the latent variable over the population. This assumption is often simply a normality assumption: $g(\theta) \sim N(\mu,\sigma^2)$. The joint distribution of the latent ability and test items parameters are estimated simultaneously using specific procedures (these procedures are presented in details in Mislevy 1991, Mislevy et al. 1992). A key feature of IRT models for our purpose is that these procedures do not yield a unique metric for the achievement scale. That is, *the mean and the unit of measurement of the estimated achievement are indeterminate*. In practice the model calibration encompasses an anchoring procedure which consists in fixing arbitrarily the metric (or measurement scale) of the estimated achievements. Finally, the achievement estimates presented in cognitive skills survey data are generally standardized.[6] This standardization implies a transformation of the distribution of IRT test scores which involves *both* a scale shift and a translation. The arbitrariness of the scale used for test scores thus proceeds from both the very nature of achievements and the statistical procedures used for their estimation.

This standardization of test scores is not innocuous for inequality measurement purposes. To begin with, note that the standardization makes the values provided by any absolute inequality index useless for measuring cardinality in those variables as those values vary with their rescaling. Only a relative inequality index, satisfying scale invariance (meaning that $I(y)=I(\lambda y), \lambda>0$, for any vector *y* and positive scalar λ)[7], could a priori provide a basis for measurement and cardinal comparisons. Moreover, although most "well-behaved" relative measures of inequality are scale-invariant, Zheng

---

[6] The PISA scores, for instance, were normalized so that the t-scores estimated have a mean of 500 and a standard deviation of 100 for the population of students in the OECD countries surveyed in 2001, using the formula:

$$t_2 = 500 + \frac{100}{sd_1}(t_1 - m_1) \qquad (1)$$

where $t_1$ is the achievement variable obtained from model estimation, and $m_1$ and $sd_1$ are its mean and standard deviation.

[7] For some measures *I()*, *y* must be restricted to be a non-negative vector.

(1994) has shown that there is no meaningful inequality index which can satisfy both scale and translation invariance (implying that $I(y)=I(y+a)$, where $a$ is vector of constants of equal dimension to $y$).[8] In the other hand, most absolute measures of inequality are translation-invariant but not scale-invariant. This implies that any (meaningful) inequality measure defined over the standardized test scores produces a measure of inequality which depends on the arbitrarily chosen scale. Zheng's impossibility result hence means that there is no adequate inequality measure defined over the standardized test scores capable of capturing inequality on a cardinal basis: relative inequality indexes provide no more bases for measuring cardinality in test scores variables than absolute inequality indexes.

The measurement of inequality of achievements must therefore rely on less demanding requirements. A first solution is to restrict the analysis to ordinal comparisons. If one is only interested in learning whether inequality in achievements is higher in a given group a than in another group b (for instance across two countries or two socio-economic groups), some absolute inequality indexes can provide a basis for ordinal comparisons in the sense that, although the values are not, the rankings of two distributions they provide may be scale-invariant. This is in particular the case of the variance. Compare the variance of two distributions of achievements $\{y_{1i}\}$ and $\{y_{2i}\}$. After the imposition of a rescaling, $y_i^s = \frac{1}{\sigma}(y_i - \mu)$, the variance of the distributions is related to the before standardization variance through $Var(y_i^s) = \frac{1}{\sigma}Var(y_i)$, so that the ranking of inequality is preserved across groups. However, many common absolute

[8] A meaningful index of inequality is defined as satisfying three basic properties : a) symmetry which requires that the level of inequality does not change with any permutation of two individuals outcomes, b) simple continuity in any individual income, and c) Pigou-Dalton principle of transfers which requires that inequality decreases (increases) as a result of a progressive (regressive) transfer.

inequality indexes (notably the most common inter-quantiles ratio, the Kuznets or the Kolm-Pollak indexes) do not satisfy this "ranking invariance" property.[9]

If one still wishes to provide cardinal measures of inequality in achievements, a second solution is to reformulate the analysis and to focus, rather than on overall measures of inequalities in achievements, on the extent to which these inequalities are associated with given individual attributes. Indeed, if one is interested in inequality *shares*, rather than levels of overall inequality, there is a straightforward solution to the standardization problem. Consider, for instance, the joint distribution {y, z}, where z is a vector of individual attributes. Now consider a partition of the vector y denoted $\{y_i^k\}$, such that $z_i^k = z^k \Leftrightarrow i \in k$. In other words: $\{y_i^k\}$ is a partition of the population into groups such that the members of each group are identical with respect to all attributes in the vector z.[10] Define a *smoothed distribution* $\{\mu_i^k\}$, corresponding to a particular partition $\{y_i^k\}$, as the distribution that arises from replacing $y_i^k$ with the group-specific

---

[9] The effects of the standardization $y_i^s = \frac{1}{\sigma}(y_i - \mu)$ on the most used relative inequality indexes can be illustrated easily. The Gini can be written $G(y) = \frac{1}{2n^2 \bar{y}} \sum_{i,j} \left| x_i - x_j \right|$, so that the overall inequality of the standardized distribution is given by $G(y^s) = \frac{\bar{y}}{\bar{y} - \mu} G(y)$. The ratio $\frac{\bar{y}}{\bar{y} - \mu}$ depends on the mean of the distribution, so that the application of the same standardization to different distributions (for instance the distributions of achievements in two countries) may change their rankings in terms of overall inequality. For instance, the ranking of the PISA participating countries by Gini of test scores in math differs markedly from the ranking of the variance of the same outcome: Mexico has the 15th higher Gini but only the 44th higher variance, whereas Germany has only the 22nd higher Gini but the 8th higher variance (those re-rankings are consistent with the higher mean score in Germany). For the Theil family indexes, no simple relationship can be found between the levels of overall inequality of the original and standardized distributions. For instance, the mean log deviation can be written $E_0(y) = \frac{1}{n} \sum_i \ln\left( \frac{\bar{y}}{y_i} \right)$, so that the overall inequality of the standardized distribution is given by $E_0(y^{st}) = \frac{1}{n} \sum_i \ln\left( \frac{\bar{y} - \mu}{y_i - \mu} \right)$; no simple relation relates these two expressions. Similarly, one can easily check that the standardization changes the rankings of overall inequality for other measures of inequality such as those from the Kolm-Atkinson family, the inter-quantile ratios, or the coefficient of variation.

[10] The notation $\{y_i^k\}$ is used to refer to a particular partition of all individuals in the distribution. It therefore has two components: the distribution of incomes in the population, and the way it is partitioned.

mean $\mu^k$. It turns out that there exist some indices *I()*, such that the ratio $\Theta_{IB} = I(\{\mu_i^k\}) / I(\{y_i^k\})$ is both scale and translation-invariant. This ratio is, of course, simply the ratio of between-group inequality to total inequality, when the groups are defined as a fine partition by the set *z* of individual attributes.[11]

An index of this form which is both scale- and translation-invariant must satisfy:

$$I(\{\lambda\mu_i^k\}) / I(\{\lambda y_i^k\}) = I(\{\mu_i^k\}) / I(\{y_i^k\}), \lambda > 0 \tag{3}$$

and $I(\{\mu_i^k + a\}) / I(\{y_i^k + a\}) = I(\{\mu_i^k\}) / I(\{y_i^k\}), a > 0$ (4)

simultaneously. Intuitively, this requires an index *I()* that is scale-invariant, and whose absolute counterpart $\mu I()$ (or $\mu^{\alpha} I()$ for any $\alpha>0$) is translation-invariant. If the latter condition holds, adding a constant vector *a* to the vector *y* in both numerator and denominator will not alter the ratio of inequality measures. Two readily available measures that satisfy (3) and (4) are the Gini coefficient and the variance (or its transform in the generalized entropy class, E(2)). Although the Gini is perhaps the best-known of all inequality measures, and although it satisfies a number of desirable properties, it is not perfectly decomposable between population-subgroups: the Gini is not equal to a weighted sum of the inequalities observed within each sub-group and the inequality of the smoothed distribution. There is a third term, which depends on the degree of overlap of conditional distributions. It is therefore ill-suited for our purposes, and we focus on E(2).

This measure is given by $E_2(x) = \frac{1}{2n}\sum_{i=1}^{n}\left[\left(\frac{x_i}{\bar{x}}\right)^2 - 1\right]$, or $E_2(x) = \frac{V(x)}{2\bar{x}^2}$, where $V(x) = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$ is simply the variance. Because E(2) is scale-invariant, and *V(x)* is translation-invariant, $\Theta_{IB}^V = E_2(\{\mu_i^k\}) / E_2(\{y_i^k\})$ (where $E_2$ can be replaced by the variance) is both translation- and scale-invariant. Moreover, because $E_2$ (and the variance) are additively decomposable, this ratio can be interpreted as the share of

[11] $\Theta_{IB}$ corresponds to the RB statistic in Cowell and Jenkins (1995). See also Bourguignon (1979) and Cowell (1980) on inequality decompositions by population subgroups.

between-group inequality other total inequality, the remaining share being within-group inequality.[12]

Moreover, if $z$ is defined appropriately as containing pre-determined circumstances such as race, gender, the socio-economic status of one's parents, or place of birth, the measure of the share of between-group inequality is also one way to compute a measure of inequality of opportunity in scholastic achievements, defined with respect to the observed circumstances $z$.[13] It can correspondingly also be interpreted as a measure of inequality of opportunity in the distribution $y$, with respect to the circumstance vector $z$.

*2.2. Addressing the selection of the sample of examinees*

A second issue must be addressed. Children out of school (or too delayed at school) are not sampled by surveys on learning assessments. For instance, the PISA surveys sample 15 year-olds attending grade 7 or more. The problem is not very serious for most developed countries where the rates of enrollment of 15 years-old are high, but it can be very significant in some developing countries. The samples of the 2006 PISA surveys are representative for less than 80% of the entire population of 15 year-olds in 15 countries. The coverage of the PISA samples is notably low in large countries such as Turkey, Brazil, Mexico, and Indonesia. The measures of both the relative overall inequality and shares of inequality associated with given attributes for a selected sample of students may be different from those that would be obtained from a representative sample of the population.[14] This selection, the extent of which differs across countries, could in particular render international comparisons unreliable.

---

[12] Let us note that for most of the other usual relative indexes used in the literature (which satisfy scale-invariance), the associated absolute inequality index given by $I^a(x) = \bar{x}.I^r(x)$ will not satisfy translation-invariance. Indeed a result of the inequality measurement literature is that it will not be possible to find such a pair for indexes which satisfy a consistency property which requires that decrease (increase) in any sub-population will decrease (increase) total inequality in the total population (Fleurbaey 1996 chapter 5 proposition 4 p.114, and Kolm 1976). The scale-invariance and coherence properties lead to the Kolm-Atkinson relative inequality index, whereas the translation-invariance and coherence properties lead to the Kolm-Pollack absolute inequality index. The Gini satisfies only a weak version of the consistency property, consistency by pair, where one considers only the effect of a change in sub-groups of two observations.

[13] See Ferreira and Gignoux (2008).

[14] Note that the selection needs not lead to underestimate the shares of between-group inequality. The resulting bias depends on the specific selection biases on the estimates of both overall inequality and

There is an essential selection problem here: dropping out of school and repeating grades are likely to be associated with low achievements, but there is no way to learn about the respective roles of an individual's background and his personal characteristics in this selection.[15]

However, one can evaluate the consequences of this selection by making assumptions on the unobserved achievements of out-of-sample children. In this paper, we propose to explore two scenarios regarding the selection process and test the robustness of the measures of learning inequalities to those. The first scenario relies on an assumption of selection on observables; this assumption can be stated in the following way: unobserved children would have the same scores as observed children with similar characteristics. The second scenario makes an opposite assumption that selection is strongly associated with unobservables and can be stated that way: unobserved children would have the smallest score measured for observed children with similar characteristics if they attended school.

These two opposite assumptions are used to impute test scores to out-of-sample children. We use a non-parametric reweighing procedure for implementing these imputations. Both imputation methods consist in using an auxiliary nationally representative household survey in order to estimate the number of missing 15 years-olds with specific attributes. A set x of observed attributes, available and exactly comparable in both the survey of students' achievements and the household survey, is used for this purpose. Observations are added for each out of sample children so that the resulting sample is representative of the national population of 15 years-olds on the basis of the observed attributes x. Test scores are then imputed for the added observations. These imputations are based on the observed conditional distributions of test scores given the attributes x, or $f(y|x)$, for the population covered by the sample of the survey of students' achievements. If the expanded population of 15 year-olds in cell $k$ of this partition $\{s_i^k\}$ in the PISA survey (household survey) is $\phi_{PISA}^k$ ($\phi_{HH}^k$), the imputations

between-group inequality. These biases in turn depend on the population shares and both the between and within components of inequalities among the different groups. We do not provide a formal analysis of these biases in this paper.

[15] Cameron and Heckman (1998) have proposed a statistical model allowing one to estimate the effects of selection in schooling transitions. However those estimates require panel data, which are not available here.

based on the assumption of selection on observables are obtained simply by reweighing the observed conditional distributions $f(y|x)$ by the inverse of the proportion of out-of-sample children, or $\phi_{HH}^{k} / \phi_{PISA}^{k}$ .

The imputations based on the assumption of strong selection on unobservables are obtained by attributing to the added observations with observables z the minimum score observed for the conditional distribution of scores of students with the same attributes. This is done by attributing a proportion $\frac{\phi_{HH}^{k} - \phi_{PISA}^{k}}{\phi_{HH}^{k}}$ of the population in each cell of the parsimonious partition the lowest score in cell k, $\underline{s}^{k}$, after the reweighing process described above.[16]

The correction procedure can therefore be decomposed into three steps. First observations are added into the PISA sample for out of sample 15 years old, using an auxiliary data from a nationally representative household survey, in order to restore the coverage of the sample regarding a selected subset of attributes. Second hypothetical scores are imputed to the added observations according to the two assumptions. Third measures of inequality in achievements are computed for each scenario. Although this procedure does not resolve the information problem concerning the achievements of out of sample children, it should provide some guidance on the robustness of the measures of inequalities in achievement to the selection of the sample.

[16] In this procedure, the specific observations whose scores are modified are chosen randomly within each cell.

## 3. The data

The surveys from the Programme International of Students' Assessments (PISA) offer a unique opportunity to compare inequalities in achievements across a large number of countries. The empirical analyses of this study rely on the data collected in 57 countries for the 2006 round of the PISA surveys. These data were collected between March and November of that year.[17] **Table 1** presents the sample sizes and coverage rates for each countries sorted by regions. Most OECD countries, but also a number of developing countries in Asia, Latin America, North Africa and the Middle East, were surveyed. The sample sizes range from 339 in Luxembourg and 3789 in Iceland to 30971 in Mexico. These samples of examinees are representative for the population of 15 year-olds enrolled in grade 7 or higher in any educational institution. The samples are *not,* therefore, representative of the total population of 15 year- olds in each country, as drop-outs and too delayed students are not covered. [18] The coverage of the total population varies considerably across countries. Although it is almost complete in many OECD countries, it is low in many developing ones: the coverage rates are as low as 47% for Turkey, 53% for Indonesia, 54% for Mexico, and 55% for Brazil. Overall, coverage is less than 80% of the total population of 15 years-olds in 15 countries, including, in addition to those mentioned above, Chile, Colombia, Israel, Jordan, Kyrgyzstan, Macao-China, Portugal, Thailand, Romania, and Uruguay. If not dealt with, this sample selection issue implies that the estimates of inequalities in observed test scores are not assessments of the entire educational system, but only of the achievements of the students who remained in the system. In each country, all children surveyed took three tests in Reading, Mathematics, and Science.[19] The data for achievements in Reading for the

[17] Two precedent rounds of data were collected in 2000/2002 and in 2003 in 43 and 41 countries respectively, and a new round is to be collected in 2009 in 67 countries.

[18] Moreover the sampling procedures induce an under coverage of the total population of enrolled 15 year olds. In particular, listing of schools and weights are established in the year preceding the surveys according to their current enrollment, so that some sampled schools may have closed, while new ones are not included in the sample, and the changes in the enrollment of 15 years olds are not taken into account.

[19] The 2006 round of the survey sought to measure achievements in Science particularly accurately by increasing the number of items devoted to this field.

United States were not issued after a problem occurred during the field operations in that country.

The test scores are based on item response theory and the variables that we use are plausible values. The inference of statistics of unobserved latent variables gives rise to a measurement error problem. Each individual answers a limited number of items so that it is not possible to estimate individual abilities accurately.[20] Plausible values (PVs) are a standard solution for this measurement error problem. The strategy consists in estimating directly the parameters of interest for the population instead of indirectly using estimates of abilities for individuals. The joint distribution of the latent ability and test items parameters is estimated using marginal estimation procedures. A number M of data sets containing the predictions of the marginal distribution of the latent variable are then produced using these estimates. If one is interested in estimating a given statistic $s$, each of these data sets of plausible values is to be used to obtain a set of estimates $\hat{s}_m$ of the statistic, one for each set m=1,…,M of plausible values of test scores. The final estimate $\hat{s}$ of the statistic $s$ is taken to be the average of the M estimates $\hat{s}_m$ obtained with the M sets of plausible values. This strategy allows estimating consistently statistics of the latent trait (achievement here) of the population, such as inequality indexes. Mislevy (1991) presents the details of the procedure.

Note that the variance of the estimate $\hat{s}$ of the parameter of interest is given by the sum of two components, representing respectively the sampling variance of the estimates of the statistic s for each set of plausible values, $\mathrm{var}(\hat{s}_m)$, and the variability of the estimates $\hat{s}_m$ among the different plausible values, that is the measurement error in the estimates unobserved latent trait. This total variance is given by (adapted from Mislevy 1991, equation (8)):

$$\mathrm{var}^{tot}(\hat{s}) = \frac{1}{M}\sum_m \mathrm{var}(\hat{s}_m) + \left(1+\frac{1}{M}\right)\left(\frac{1}{M-1}\right)\sum_j (\hat{s}_m - \hat{s})^2 \qquad (1)$$

[20] In this situation, the distribution of estimates for individual abilities obtained with traditional methods (such as maximum likelihood estimates (MLE)) does not converge to the population distribution of these abilities as the number of examinees increases (Mislevy et al. 1992). These estimates of parameters of this distribution are thus inconsistent (although the asymptotic bias decreases with the number of items per examinee).

where the first term on the right hand side is the mean of the sampling variance of each estimate $\hat{s}_m$, and the second a variance of the estimated plausible values .[21]

If we seek to construct a measure of inequality of educational opportunity, we must also identify "circumstance variables", i.e. individual characteristics that correlate with cognitive skills, but which are economically exogenous to the individual. In other words, predetermined characteristics that are given to the child, and which she can not change; circumstances that she inherits at birth. The PISA data contain information on a number of individual, family and location characteristics that represent such pre-determined circumstances. Ten circumstance variables (or sets of variables) are used in this study: gender, father's and mother's education, father's occupation, language spoken at home, migration status, access to books at home, durables owned by the households, cultural items owned, and area type captured by the location of the attended school.[22]

Parental education is measured by the highest level completed and is coded using ISCED codes into four categories: a) no education or unknown level; b) primary education (ISCED level 1); c) lower secondary education (ISCED level 2), upper secondary (ISCED level 3), or post-secondary non-tertiary education (ISCED level 4); and d) college education (ISCED level 5)). Father's occupation is measured using ISCO codes. We aggregate occupations into three broad categories: a) legislators, senior officials and professionals, technicians and clerks; b) service workers, craft and related trades workers, plant or machine operators and assemblers, and unoccupied individuals;

[21] For PISA, the sampling variance of population parameter estimates can be computed using the Balanced Repeated Replication (BRR) weights provided within the data (PISA 2006). BRR is a replication method for multistage stratified sample designs and is close to the Jackknife. The particular variant of the BRR known as Fay's method was used. For PISA it consisted in forming pairs (called strata) of schools (the primary sampling units) and drawing a number of replicates of the sample (using a so-called Hadamard matrix). 80 replicates were performed. Each of these replicate attributes weight 1.5 to one of the school and weight 0.5 to the other in each strata, the selection being different for each replicate. The BRR weights are then computed as the product of students' original sampling weights and the school weight (1.5 or 0.5) for each particular replication. The variance estimator for the population parameter estimate $\hat{s}$ (computed with the original sampling weights) is then given by: $V_{BRR}(\hat{s}) = 0.05\sum_{t=1}^{80}(\hat{s}_t - \hat{s})^2$ where the estimates $\hat{s}_t$ are computed using each set of BRR weights.

[22] Information on a few additional circumstance variables was available but not used for the analysis because of the likely simultaneous determination of students' achievements and those variables. In particular, mother's occupation and the availability of home educational resources (such as a desk or a computer devoted to home study) were not used. The inclusion of these variables does not change considerably the estimates obtained thereafter (results upon request to the authors).

and c) skilled agricultural and fishery workers, elementary occupations or unknown occupation. The language spoken at home variable is a dummy identifying another language than the language of the test. The migration status variable is a dummy identifying a first or second generation migrant as an individual who was, or whose parents were, born in a foreign country. The number of books at home variable, an indicator of parental human capital, is a categorical variable coded into four categories: a) 0 to 10 books; b) 11 to 25 books; c) 26 to 100 books; and d) more than 100 books. Ownership of durables, an indicator of family wealth, is captured by six dummy variables indicating the ownership of a) a dishwasher; b) a DVD or a VCR player; c) a cell phone; d) a television; e) a computer; f) a car. Ownership of cultural possessions is captured by three dummy variables indicating the ownership of a) books of literature (Shakespeare is mentioned as an example of an author in the formulation of the question); b) books of poetry; and c) works of arts (paintings are mentioned as an example of such works in the formulation of the question). School location is a proxy for the person's inherited spatial endowment and we recode it using three categories: a) villages or small towns (less than 15,000 inhabitants); b) towns (between 15,000 and 100,000 inhabitants); and c) cities (larger than 100,000 inhabitants).[23] School location information was not collected in France, Hong-Kong, and Liechtenstein.

## 4. An international comparison of inequalities in achievements

The inequalities in educational achievements in the countries surveyed by PISA 2006 are now examined using the inequality measures proposed in Section 2. We begin by comparing the relative levels of overall inequalities in achievements, but devote more attention to the measures of between-circumstance groups inequalities, the only ones which can be quantified. The comprehensive set of circumstance variables presented in Section 3 is used to identify circumstance groups. Given the large number of variables, a classic problem of data insufficiency for non-parametric estimation emerges: with the size of cells diminishing, sampling error increases and biases upward the estimates of

[23] PISA also collected information on the mother's occupation. We used this variable in preliminary calculations, and results were unchanged, so we omit them in this analysis.

between group inequalities. The proposed measures are thus implemented using parametric estimates. These parametric estimates rely on functional form assumptions; in particular, by excluding interaction terms, we assume linearity of the effects of circumstances. The statistical relationship between circumstances and test scores are estimated in the regression framework:

$$y_i = z_i \psi + \varepsilon_i \tag{2}$$

Under these functional form assumptions, a parametrically smoothed distribution of achievements is estimated by:

$$\hat{y}_i = \exp\left[z_i \hat{\psi}\right] \tag{3}$$

The share of between group inequalities is then given by:

$$\theta = \mathrm{var}(\hat{y}) / \mathrm{var}(y) \tag{4}$$

and $\mathrm{var}(\{\hat{y}_i\})$ provides an estimate of the levels of inequalities in achievements allowing cross-country comparisons in relative terms (the obtained values depending on the arbitrary standardization). Equation (4) can further be written:

$$\theta = (\mathrm{var}\, y\ )^{-1}\left[\sum_k \psi_k^2 \ \mathrm{var}\, z_k + \frac{1}{2}\sum_k \sum_j \psi_k \psi_j \ \mathrm{cov}(z_k, z_j)\right] \tag{5}$$

The relative levels of overall learning inequalities in the surveyed countries are given by the variances in **Table 1**. Although nothing can be told about the values taken by the variances, a ranking of countries can be established and overall inequalities in achievements are significantly higher in some countries than in others. Among the countries with higher such inequalities are West European countries such as Austria, Belgium, France, Germany, and Italy, East European ones such as Czech Republic and Bulgaria, Latin American countries such as Argentina and Uruguay, but also Israel and Taipei. Among the ones with lower inequalities in achievements are other European countries such as Croatia, Denmark, Estonia, Finland, Ireland, and Latvia, but also Asian countries such as Indonesia, Thailand and Jordan. Countries such as the Great Britain, Japan, or the United States take intermediate rankings. [24]

[24] The inequality measures obtained for Azerbaijan seem particularly small and put the country as an outlier in all the analyses. It is unclear to us how much of this is due to the data collection procedures in this country, but such a different pattern is not likely due to real differences only.

**Table 2** provides our estimates of the total shares and levels of between circumstance groups inequality for achievements in Reading, Math, and Science, for the 56 surveyed countries. These estimates may be interpreted as measures of inequality of learning opportunities. The relative levels of between groups inequality ($\mathrm{var}(\{\hat{y}_i\})$) and the shares $\theta$ of between circumstance groups inequality are provided with the corresponding bootstrap estimates of standard errors.[25] The estimates of the shares of between circumstance groups inequality range between 12.7 and 38.8% of the overall variance of test scores in Reading, between 4.4 (10.2 excluding the outlier Azerbaijan) and 35.1% of the overall variance of test scores in Math, and between 11.1 and 37.9% of the overall variance of test scores in Science.

**Figure 2** provides the same results graphically, for achievements in Math, after ranking the countries both by relative levels (panel a) and shares of between circumstance group inequality (panel b). 95% confidence intervals are presented using the estimates of the standard errors and assuming normal distributions of the estimates, and allow testing the significance of cross-country differences. The rankings by relative levels and shares of between group inequalities are closely related: the Spearman's rank correlation coefficient is 0.90 (the same coefficient is respectively 0.72 and 0.86 for achievements in Reading and Science). Hence one should be able to perform cross-country comparisons on the basis of relative levels or shares of between group inequalities without reaching too different conclusions.

No clear regional pattern of achievements inequalities emerges from the estimates presented in Table 2 and Figure 2. Among the countries with the highest levels or shares of between circumstance groups inequalities, with shares higher than 30%, are West European countries (such as Belgium, France, and Germany) but also East European countries (such as Bulgaria, the Czech Republic, and Hungary), and Latin American countries (such as Argentina, Brazil and Chile). Among the countries with the lowest levels or shares of these inequalities, with shares lower than 20%, are Asian countries (such as Azerbaijan, Macao-China, Hong-Kong, and Japan), Russia, Australia, but also other West European (such as Finland, Iceland, Italy, and Norway) or East European

[25] The relative levels of overall inequality in achievements, $\mathrm{var}(y)$, were given in Table 1.

(such as Estonia and Latvia) countries. Finally, countries such as Great Britain, the United States or Spain, lie in an intermediate position with shares close to 25%. One can use these results to perform specific cross-country comparisons. For instance, focusing on a few developed economies, the shares and levels of these educational inequalities are significantly higher in a few large European countries, such as France and Germany, than in the United States. However these inequalities are significantly lower in Scandinavian countries, such as Finland and Norway, or in Japan and Korea. Regarding developing economies, Latin American countries tend to range in the highest half, while Asian countries, such as Indonesia and Thailand, range in the lowest half, and although the estimates are very imprecise for Indonesia, Thailand exhibits significantly lower inequalities than Latin American countries such as Brazil.

The results in Reading and Science are not discussed in details here. However the estimates of educational inequalities in the three fields are closely related: the Spearman's coefficients of rank correlations for either the estimates of relative levels or shares in Reading, Math and Science, range from 0.75 to 0.92. Therefore the above results should not differ much by field.

**Figure 3** provides an inspection of the relationship between mean test scores and estimates of relative levels (panel a) and shares (panel b) of between group inequality in achievements in the set of countries. The regression line and a 95% confidence interval for the mean are shown on the graphs. No significant relationship is found between mean scores and the inequality in achievements between circumstance groups.

As seen in Section 4, the coverage of the sample of the PISA surveys is far from complete in a number of countries. Coverage is particularly low in developing countries. In order to check the robustness of the previous estimates to the inclusion of out of sample 15 year-olds, the correction procedure presented in Section 2 is now implemented. The procedure is applied to the four countries with the lowest coverage rates, which happen to be very populated emerging economies: Brazil (coverage of 55%), Indonesia (53%), Mexico (54%), and Turkey (47%). These low coverage rates are due to dropping-out and repetition, but also to sampling issues.[26] Four nationally representative

[26] The rates of enrollment of 15 years-olds in those four countries are respectively 89.9 in Brazil, 75.2 in Indonesia, 75.9 in Mexico, and 78.3 in Turkey, while the rates of enrollment in grades 7 or higher are

household surveys are used to estimate the total number of out-of-sample 15 years-old children and a set of their circumstance characteristics. The surveys used are the 2006 round of the national household survey (PNAD) for Brazil, the 2005 round of the national household survey (Susenas) for Indonesia[27], the 2006 round of the national survey of income and expenditure (ENIGH) for Mexico, and the 2006 round of the national household budget survey (HBS) for Turkey.

We partition the population of 15 year-olds in both the household survey and PISA into groups with identical observable circumstances, using a subset of three defining characteristics: the gender, mother's education, and urban-rural status or father's agricultural occupation. The agricultural occupation of the father was preferred to the rural/urban divide in Indonesia because the definition of rural areas in the Susenas survey is not based uniquely on population and does not match the PISA definition. This variable was also preferred for Mexico because it is much more associated to under-coverage than the urban/rural distinction.[28] The three variables used for each country are defined identically in the two surveys, so that the partitions should be strictly comparable.[29] Following the two scenarios presented in Section 2, observations are then imputed in the PISA data set for unobserved children. The measures of the relative levels of overall inequalities and shares of between group inequalities in achievements are finally computed in the two cases.

**Figure 4** shows the kernel density functions for the standardized PISA test scores in reading in 2006, under three different scenarios, in the case of Turkey[30]. The top panel depicts the observed sample distribution, with no correction for selection. The middle panel depicts the counterfactual distribution with the "selection on observables"

---

respectively 71.1 in Brazil, 71.6 in Indonesia, 74.6 in Mexico, and 61.7 in Turkey. Brazil and Turkey thus have higher enrollment rates but large numbers of delayed students. Sampling issues account from 15 to 20% of the under-coverage of the national populations of 15 years-old in those four countries.

[27] The Susenas 2006 was not available.

[28] Sampling issues could have been more important in urban areas in this country making one unable to distinguish the selection due to dropping-out with the one due to sampling.

[29] A finer partition would be possible, but would generate statistically imprecise estimates of population weights in the household survey, given the sample size.

[30] Similar results for the other three countries are available upon request to the authors.

correction for selection. The bottom panel depicts the counterfactual distribution with the “strong selection on unobservables” correction.

**Table 3** provides the estimates of overall and between circumstance groups inequalities in achievements obtained after applying both corrections. For each of the four countries, for each of the three distributions of test scores (no correction, lower alternative and higher alternative correction for selection) and for each subject (reading, mathematics and science), we report both the total variance and the share of between group inequalities of test scores. The “selection on observables” correction has limited effects on the overall variance for the distribution of test scores. If the variance increases for test scores in reading in Turkey and decreases in all three fields in Indonesia, those changes are modest (about 10%), the correction has almost no effect on the other variances. The “strong selection on unobservables” correction, on the other hand (and as could be expected from an inspection of Figure 4), increases the variances between two- and fourfold (threefold in the case of Turkey). The shares of between circumstance groups inequality, on the other hand, do not vary much across subjects and turn out to be relatively insensitive to the alternative corrections for selection. With no selection correction, 27 to 32% of the variance of scores is associated to the set of circumstances in Brazil, 22 to 24% in Indonesia, 26 to 28% in Mexico, and 24 to 25% in Turkey. These shares remain almost unchanged or decrease slightly under the more conservative selection correction procedure to 26-31% of the variance of scores in Brazil, 18-22% in Indonesia, 24-27% in Mexico, and 24-25% in Turkey. Under the higher alternative correction procedure, the estimate of the share of between group inequality in educational achievement rises to some 38-40% in Brazil, 26-27% in Indonesia, and 32-33% in Turkey, and remains almost unchanged to 23-25% in Mexico. The estimated shares of between circumstance groups of inequalities appear thus relatively robust to corrections for sample selection. Although we applied the correction procedure to a limited set of four countries, the selection of a sample of attendees seems to lead to an underestimation of these shares, but this bias is not very large.

The bottom panel of the **Table 4** reports the partial shares of inequalities in achievements in Math associated with individual circumstances, namely gender, father’s education, mother’s education, father’s occupation, type of area, language spoken at home, immigration status, number of books, ownership of durables, and cultural possessions.[31] These partial shares are given by:

$$\theta^{PJ}\left(\left\{y_i^k\right\}\right)=\left(\operatorname{var} y\right)^{-1}\left[\psi_j^2 \operatorname{var} z_j+\frac{1}{2}\sum_k \psi_k \psi_j \operatorname{cov}\left(z_k, z_j\right)\right] \tag{6}$$

These partial shares sum up to the overall parametric estimate of between-group inequality given by (5) and also satisfy a decomposition path-independence property (Foster and Shneyerov 2001). [32] The partial shares are only reported for the regression without any correction for sample selection, because the partition $\left\{s_i^k\right\}$ used for that correction is based on some but not all of the independent variables in the regression, which would tend to bias the estimates of the shares of inequality associated with each variable.

A causal interpretation of these results is problematic. However the family educational and cultural resources seem to be associated with the largest share of inequality of learning achievement. Mother's and father's education combined account for a mean of 3.7 and a maximum of 9.2 (in Hungary) percentage points of the overall shares of explained inequality in the set of 57 countries which take the mean of 24.7. The number of books at home accounts for a mean of 7.2 and a maximum of 14.4 (in Austria) percentage points in these shares. Add parental education, language at home, numbers of books, and cultural possessions, and this set of "educational and cultural variables" add up to a mean of 15.0 points. Family economic resources also appear has an important source of learning inequalities. Father's occupation and the three "asset" indicators durables account respectively for means of 3.6 and 3.8 and maximums 7.2 (in Luxembourg) of 18.4 (in Brazil) percentage points of the overall shares of explained inequality. With immigration status, the set of "economic variables" explains a mean of 7.8 points.[33] [34] Finally, the area type where schools are located accounts for a mean of 1.6

[31] Similar decompositions for achievements in reading and science can be obtained from the authors.

[32] We show in a related paper (Ferreira et al. 2009) that equation (9) is the simple average between the direct and residual estimates of the partial shares which correspond to a smoothed and a standardized distributions respectively.

[33] Immigration status can be considered as both a cultural and an economic resource variable.

[34] There are interesting cross-country and regional variations in these partial shares of learning inequalities.

and a maximum of 10.7 (in Kyrgyzstan) points of the overall shares, whereas the student's gender accounts for a rather limited mean of 0.6 and a maximum of 2.1 (in Chile) points of the overall shares.

## 5. Application: exploring the relationship between economic development, specific education policies, and the distribution learning opportunities

Equity is an important concern of educational policies, and numerous empirical studies have sought to estimate the impacts of such policies on the distributions of access to education and achievement of given educational attainments. Recognizing that quality of education is a major determinant of individuals' achievements, recent studies tend to focus on the impacts of policies on inequalities in achievements, measured by test scores. However, the lack of satisfactory measures of those inequalities is impeding progress in such policy analysis. The methodology proposed above can remove these limitations as it addresses the two main obstacles for the measurement of learning inequalities: it is both robust to the standardization of test scores and it allows exploring the implications of the selection of samples of attendees. In order to illustrate their potential applications, the previous measures of learning inequalities are now used in an exploratory analysis of the distributional outcomes of two specific educational policies: the distribution of the educational public resources among the various levels of education and the tracking of pupils between general and vocational schools or classes.

The distributional incidence of public expenditure in education and the allocation of financial resources among the different segments of the education system have been examined by various studies (Al-Samarrai and Zaman 2007, Birdsall 1996, Castro-Leal et al. 1999, Jimenez 1986, Van de Walle and Nead 1995, Yaqub 1999). Given that children with disadvantaged backgrounds tend to drop-out earlier than other children, the allocation of larger shares of financial resources to the primary level of schooling is more likely to have redistributive effects, whereas their allocation to higher levels, in particular

---

For instance, the share of learning inequality associated with the set of educational and cultural family resources has a higher mean in Western and Eastern European countries than in countries of other regions, whereas the share of learning inequality associated with the set of economic resources has a higher mean in Latin American countries.

the tertiary level, should be much less redistributive or even anti-distributive. [35]

The impacts of tracking policies on the efficiency and equity in educational systems are another example of education policies having received attention in recent studies (Ariga et al. 2006, Brunello and Checchi 2007, Brunello et al. 2006, Hanushek and Woessman 2006, Manning and Pisckhe 2006). The theoretical arguments do not provide clear-cut predictions of the effects of early tracking on educational achievements: while homogenous classrooms, and the associated specialization of teaching and curricula to students' needs and abilities, could allow efficiency gains, least performing groups could at the opposite be slowed down even more by unfavorable allocations of resources, including less well endowed schools, teacher sorting, peer effects, or differences in curricula[36]. Moreover, a frequent concern has been that, given that much of early inequality and therefore track placements depend on parental resources, tracking might reinforce the effects of family background on educational achievements.

Let us illustrate how the distributional impacts of those policies can be explored using international comparisons and our measures of learning inequalities. The data used for these exploratory analyses are from the UNESCO Institute for Statistics (UIS). These data are provided by the UNESCO member states through an annual data collection performed by UIS.[37] Our indicator of the distribution of public education resources is the share of educational expenditure in primary (the first ISCED level corresponding to grades 1 to 6) among total public educational expenditure, while our indicator of tracking is the share of technical/vocational enrollment at the secondary level (including lower and upper secondary or the second and third ISCED levels, usually corresponding to grades 7 to 12) among total enrollment at that level. The information on the distribution of education expenditure across levels is missing for six countries (Canada, Montenegro, Qatar, Russia, Serbia and Taipei) and the information on the share of technical and vocational enrollment at the secondary level is missing in five countries (Latvia,

[35] The allocation of resources between different types of schools or kinds of expenditure should also have distributive implications.

[36] Too early tracking can also have some costs in terms of the misallocation of students to tracks and in terms of foregone versatility in the production of skills (Brunello and Checchi 2006).

[37] The data for 2006 correspond to the school year 2005-06 for countries where the school year laps over two calendar years.

Montenegro, Serbia, Taipei and the United States). Moreover two countries are excluded from those analyses: Liechtenstein and Luxembourg. The number of observations for Liechtenstein (339 examinees) makes the estimates of learning inequalities unreliable and Luxemburg appears too much as an outlier in terms of GDP per capita in 2006 (at about 69.000 US dollars, the second highest GDP being observed for the United States at 44.000 US dollars).

Before turning to education policies, Figure 5 provides a very basic description of the relationship between economic development, measured by GDP per capita[38], and measures of inequality in learning given by both the relative levels (panel a) and the shares of between circumstance group inequality (panel b) for achievements in Math[39]. The regression line and a 95% confidence interval for the mean are shown on the graphs. No significant statistically relationship is found between GDP per capita and either measure of learning inequalities. In order to test whether outliers such as Azerbaijan or Macao-China drive the statistical relationship, the procedure proposed by Besley, Kuh and Welsch (1980) is implemented to identify outliers and the test of a linear relationship is performed again after the exclusion of the corresponding observations. These estimates confirm the absence of a significant relationship between our measures of learning inequalities and GDP per capita.[40]

Turning to education policies, we begin by considering the distribution of education expenditure across levels. The share of expenditure allocated to the primary level of education within the set of 49 countries is quite heterogeneous: while the mean share is 25.8%, the lowest share is observed in Romania at 13.8% and the lowest in Colombia at 41.8% (the 0.25 quantile is at 20.2% and the 0.75 quantile at 30.4%). Figure 6 provides a graphical inspection of the relationship between the share of expenditure allocated to the primary level and both the levels and shares of learning inequalities

[38] GDP per capita is measured in purchasing power parity in 2006 US dollars; the data are from the World Development Indicators (WDI) database.

[39] Similar results are found for achievements in reading and science.

[40] Although these estimates indicate negative but insignificant relationships between GDP per capita and the shares of between group inequalities (the relationship is significant at 10% only for Math), a positive relationship is found between the levels of between-group inequalities for test scores in Science and insignificant ones for the same relationship in Reading and Math.

between circumstance groups. Here again the regression line and a 95% confidence interval for the mean are shown on the graphs. **Table 5** gives the tests of significance of this relationship both without any controls (in row 1) and controlling for per capita GDP and public education expenditure per pupil (row 2). Significant negative associations are found between the shares of resources allocated to the primary level and between circumstance group inequalities in achievements, in both reading and science and without or with controls for per capita GDP and expenditure per pupil, but also for achievements in Math with controls. These correlations are slightly more significant for inequalities measured in shares and the coefficients for these relations lie between -0.001 and -0.003 indicating that an increase of 10 points of the share of resources allocated to the primary level is associated with decreases of 1 to 3 points in the share of between circumstance groups inequalities in learning.

Considering now the relations between our indicator of tracking and learning inequalities, the shares of technical and vocational enrollment at secondary levels take a mean of 19.8 percent and are comprised between 0.9 in Qatar and 51.4 percent in the Netherlands (the .25 quantile is at 12.9 and the .75 quantile at 31.2). As above, Figure 7 provides a graphical inspection of these shares of technical and vocational enrollment at secondary levels and both the levels and shares of learning inequalities between circumstance groups, while **Table 6** gives the tests of significance of this relationship both without any controls (row 1) and controlling for per capita GDP and public education expenditure per pupil (row 2). A pattern of significant positive relationships, consistent across fields and specifications, is found. Higher shares or levels of inequalities in learning outcomes between circumstance groups tend to be associated with higher shares of technical and vocational enrollment. The coefficients of correlations lie between 0.001 and 0.002 for learning inequalities measured in shares, indicating that an increase of 10 points of the share of technical or vocational enrollments is associated with an increase of 1 to 2 points of the shares of between circumstance groups learning inequalities.

These preliminary analyses thus indicate that our measures of learning inequalities are associated negatively with the allocation of educational resources to primary level of education and positively with tracking into general or technical/vocational schooling at the secondary level. These results seem in line with and extend those of studies devoted to these relationships. For instance, while Hanushek and Woessman (2006) find tracking to be associated with higher levels of overall inequality in test scores, our results suggest it also tends to come with higher levels of inequality of learning opportunities.[41] These analyses remain descriptive in nature, and do nothing for controlling for the heterogeneity in the education systems or populations of pupils. However they illustrate the potential use of robust indicators of learning inequalities in studying the distributive impacts of education policies. Extensions of these analysis – notably the use of panel data (some international surveys of students' assessment have now collected various round of data) - could allow to test for causality in those relationships.

## 6. Conclusion

Measures of inequality in learning achievements are needed both for studying the distributional impacts of educational policies and for understanding the formation of economic inequalities. However two methodological issues have been preventing the development of robust measures of these inequalities: the arbitrary standardization of test scores variables and the selection of the samples of surveys for the assessment of students based on school attendance. The first objective of this paper has thus been to propose measures of learning inequalities taking into account these two constraints. Its second objective was to perform international comparisons of the levels of learning inequalities and of their association with a large set of family and geographic pre-determined circumstances. These measures can be viewed as measures of inequality of learning opportunities. A third objective was to illustrate the use of these measures for studying the distributional impacts of education policies.

Our first contribution is therefore methodological. Using some results on the measurement of inequality (notably Zheng 1994), we have argued that the standardization

[41] However, the long term effects of early tracking remain debated. For instance, Brunello and Checchi (2007) find that although it tends to increase the link between family background and educational attainments by diverting some individuals to progress to the tertiary level of education, it seems to reduce the impacts of family background on adult literacy and promote further on-the-job training by offering more effective curricula to less well performing students.

of test scores variables makes it impossible to provide cardinal measures of overall inequalities in achievements. However, we show first that the levels of overall inequalities can be compared in relative terms (using ordinal measures) when using a data sets in which test scores are obtained using a common standardization procedure, and second that cardinal measures can be provided for the shares of learning inequalities associated with given attributes using between-group decompositions and the variance inequality index: these measures are robust to the standardization of test scores. These measures can be applied for the analysis of inequalities in other fields in which the outcomes variables do not have a proper scale, an example of such variables being the wealth indexes constructed using data analysis methods on the basis of information on durables, or any other such index. Besides, when the attributes used to perform these decompositions are pre-determined circumstances variables, such as gender, ethnicity or family background, these measures can be interpreted as indicators of the inequality of learning opportunities. Another methodological contribution has consisted in tests of the robustness of these measures of learning inequalities to the inclusion of out-of-sample children following two opposite assumptions on the role of unobserved heterogeneity in sample selection.

Our contribution is also substantive. It consisted primarily in comparing the relative levels and patterns of learning inequalities across the 57 countries which participated in the PISA 2006 international assessment of 15 years-old pupils. We found that significant cross-country differences exist in the relative levels and shares of learning inequalities associated with a broad set of circumstance variables. In particular, these inequalities tend to be higher in European and Latin American countries, intermediate in others such as the United States and other European countries, and lower in a set of countries comprising notably Asian or Scandinavian countries. These measures are found to be relatively robust to the inclusion of out-of-sample children for the four countries with the lowest rates of coverage.

In another substantive contribution, and in order to illustrate the use of these measures in policy analysis, we documented the relationships between our measures of learning inequalities and indicators of two specific education policies: the distribution of resources across education cycles and tracking into technical/vocational at the secondary level. In line with the results of recent empirical studies, we found that while resources allocated to the primary cycle are associated with fewer inequalities in achievement between circumstance groups, the reverse was true for tracking.

The analyses presented in this paper might be developed into two directions. In a methodological perspective, it would be useful to provide a complete characterization of the inequality measures which prove independent to the arbitrary determination of scales. We have focused on additionally decomposable measures, but some studies have extended the notion of decomposability to alternative "representative income" such as the geometric mean (Foster and Shneyerov 1999). These studies might provide other satisfying measures of standardized variables such as test scores. In a substantive perspective, the measures of learning inequalities proposed in this study could be used to study the distributive impacts of various education policies, or the relationships between educational and economic inequalities.

## References


Ariga, K., G. Brunello, R. Iwahashi, and L. Rocco (2006): "On the Efficiency Costs of De-Tracking Secondary Schools". IZA Discussion Paper No. 2534.

Baker, F. (2001): *The Basics of Item Response Theory*. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, MD.

Besley D.,E. Kuh and R. Welsch (1980): *Regression Diagnostics: Identifying Influential Data and Sources of Colinearity*, New York, Wiley.

Birdsall, N. (1996): "Public Spending on Higher Education in Developing Countries: Too Much or Too Little?", *Economics of Education Review*, 15(4): 407-19

Bourguignon, F. (1979): "Decomposable Income Inequality Measures", *Econometrica*, 47(4): 901-20.

Brunello. G. and D. Checchi (2007): "Does School Tracking Affect Equality of Opportunity? New International Evidence", *Economic Policy*, 22: 781-861.

Brunello, G., K. Ariga and M. Giannini (2006): "The Optimal Timing of School Tracking", forthcoming in P. Peterson and L.Wößmann, (eds), *Schools and the Equal Opportunity Problem*, MIT Press, Cambridge MA.

Cameron S.V. and J.J. Heckman (1998): "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males", *Journal of Political Economy*, University of Chicago Press, 106(2): 262-333.

Castello, A. and R. Domenech (2002): "Human Capital Inequality and Economic Growth: Some New Evidence", *The Economic Journal*, 112:C187-200.

Castro-Leal, F., J. Dayton, L. Demery, and K. Mehra, (1999): "Public Social Spending in Africa: Do the Poor Benefit?", *World Bank Research Observer*, 14(1): 49-72.

Cowell, F. A. (1980): "On the Structure of Additive Inequality Measures", *Review of Economic Studies*, 47: 521-31.

Cowell, F.A. and S.P. Jenkins (1995): "How much inequality can we explain? A methodology and an application to the USA", *The Economic Journal*, 105: 421-30.

Ferreira, F.H.G. and J. Gignoux (2008): "The Measurement of Inequality of Opportunity: Theory and an application to Latin America", Policy Research Working Paper #4659, The World Bank.

Ferreira F.H.G., J. Gignoux, and M. Aran (2009): "Inequality of Economic Opportunity in Turkey: An assessment using asset indicators", mimeo.

Fleurbaey M. (1996), *Theories de la justice*, Economica.

Foster, J.E. and A.A. Shneyerov (1999): "A General Class of Additively Decomposable Inequality Measures", *Economic Theory*, 14:89-111.

Hanushek, E. and L.Wößmann (2006): "Does Educational Tracking Affect Performance and Inequality? Differences-In-Differences Evidence Across Countries", *Economic Journal* 116: C63-C76.

Jimenez, E. (1986): "The Public Subsidization of Education and Health in Developing Countries: a Review of Equity and Efficiency", *World Bank Research Observer*, 1(1).

Kolm, S.C., (1976): "Unequal Inequalities", *Journal of Economic Theory,* 12: 416-442.

Manning, A. and J.S. Pisckhe (2006): "Comprehensive versus Selective Schooling in England in Wales: What Do We Know?", IZA Discussion Paper 2072.

Martin, K. (eds.) (1998), *TIMSS Technical Report Volume II: Implementation and Analysis* , Primary and Middle School Years.

Mislevy, R. (1991): "Randomization Based Inference about Examinees in the Estimation of Item Parameters", *Psychometrika*, 56:177-196.

Mislevy, R., A. Beaton, B. Kaplan and K. Sheehan (1992): "Estimating Population Characteristics from Sparse Matrix Samples of Item Responses", *Journal of Educational Measurement*, 29(2): 133-161.

Morrisson, Ch. and F., Murtin (2007): "Education inequalities and the Kuznets curves: a global perspective since 1870", PSE Working Paper 2007-12.

OECD (2006), *PISA 2006 technical report*.

Prittchet, L. (2004): "Does Learning to Add Up Add Up?: The returns to schooling in aggregate data", BREAD Working Paper 053, Cambridge, MA.

Al-Samarrai, S. and H. Zaman (2007): "Abolishing School Fees in Malawi: The Impact on Education Access and Equity", *Education Economics*, 15(3): 359-375.

Thomas V., Y. Wang and X. Fan (2002): "A New Dataset on Inequality in Education: Gini and Theil Indices of Schooling for 140 countries, 1960-2000", unpublished manuscript, The World Bank, Washington, DC.

Van de Walle, D. and K. Nead (1995): *Public Spending and the Poor: Theory and Evidence*, Johns Hopkins and World Bank, Washington DC.

Yaqub, S. (1999): "How Equitable is Public Spending on Health and Education?", background paper to the World Development Report 2000/01, The World Bank, Washington DC.

Zheng, B. (1994): "Can a Poverty Index be Both Relative and Absolute?", *Econometrica*, 62(6): 1453-1458.

**Table 1: Sample sizes, coverage rates, mean scores, and overall inequality in achievements, in Reading, Math, and Science**

| | Observations | Coverage rate | Mean in Reading | Variance in Reading | | Mean in Math | Variance in Math | | Mean in Science | Variance in Science | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Asia & North Africa* | | | | | | | | | | | |
| Azerbaijan | 5184 | 0.88 | 355.0 | **4935.8** | 296.3 | 476.8 | **2300.5** | 157.6 | 385.3 | **3100.3** | 213.8 |
| Hong Kong-China | 4645 | 0.97 | 538.9 | **6689.3** | 314.6 | 551.4 | **8721.8** | 432.9 | 546.1 | **8410.8** | 351.6 |
| Indonesia | 10647 | 0.53 | 383.9 | **5593.2** | 357.5 | 380.7 | **6401.9** | 508.5 | 384.8 | **4908.9** | 457.4 |
| Israel | 4584 | 0.76 | 441.3 | **14242.3** | 665.2 | 443.3 | **11519.5** | 687.1 | 455.6 | **12421.4** | 428.9 |
| Japan | 5952 | 0.89 | 409.5 | **10483.2** | 480.4 | 389.2 | **8283.9** | 374.5 | 427.1 | **10024.8** | 403.0 |
| Jordan | 6509 | 0.65 | 500.2 | **8853.7** | 422.2 | 525.6 | **7006.7** | 326.8 | 533.7 | **8075.3** | 340.4 |
| Korea | 5176 | 0.87 | 290.5 | **7795.6** | 472.0 | 315.9 | **8572.4** | 576.8 | 326.3 | **8110.5** | 423.3 |
| Kyrgyzstan | 5904 | 0.63 | 556.1 | **10425.4** | 511.8 | 547.2 | **7566.1** | 351.7 | 521.9 | **7033.4** | 338.5 |
| Macao-China | 4760 | 0.73 | 490.6 | **5830.6** | 348.4 | 524.4 | **7038.8** | 255.5 | 509.5 | **6058.1** | 247.4 |
| Qatar | 6265 | 0.90 | 312.5 | **11690.6** | 248.9 | 317.7 | **8143.2** | 249.4 | 349.1 | **6937.7** | 227.9 |
| Russian Federation | 5799 | 0.81 | 442.4 | **8691.4** | 348.1 | 478.7 | **8016.0** | 283.3 | 481.5 | **8023.8** | 238.6 |
| Chinese Taipei | 8812 | 0.88 | 506.7 | **7119.8** | 292.3 | 562.7 | **10632.6** | 444.3 | 543.7 | **8920.7** | 307.1 |
| Thailand | 6192 | 0.72 | 425.2 | **6699.2** | 283.0 | 425.5 | **6631.0** | 255.1 | 429.7 | **5955.5** | 223.9 |
| Tunisia | 4640 | 0.90 | 379.0 | **9467.7** | 483.5 | 363.9 | **8454.5** | 430.6 | 384.2 | **6787.3** | 336.4 |
| Turkey | 4942 | 0.47 | 452.9 | **8631.1** | 510.5 | 428.2 | **8693.9** | 803.1 | 427.6 | **6923.2** | 521.8 |
| *Latin America* | | | | | | | | | | | |
| Argentina | 4339 | 0.79 | 383.9 | **15431.1** | 899.7 | 388.1 | **10230.3** | 192.6 | 398.3 | **10250.1** | 531.2 |
| Brazil | 9295 | 0.55 | 389.2 | **10497.6** | 685.4 | 365.6 | **8467.3** | 60.7 | 385.3 | **7970.9** | 344.2 |
| Chile | 5233 | 0.78 | 447.9 | **10658.3** | 503.6 | 417.1 | **7645.4** | 115.6 | 443.1 | **8405.1** | 316.5 |
| Colombia | 4478 | 0.60 | 390.3 | **11628.4** | 511.4 | 373.8 | **7750.9** | 118.5 | 391.9 | **7192.5** | 306.5 |
| Mexico | 30971 | 0.54 | 427.4 | **9154.5** | 433.9 | 420.7 | **7270.6** | 58.6 | 422.6 | **6513.4** | 237.3 |
| Uruguay | 4839 | 0.69 | 424.7 | **14694.7** | 491.4 | 435.5 | **9861.0** | 184.2 | 437.7 | **8919.0** | 326.9 |
| *North America & Oceania* | | | | | | | | | | | |
| Australia | 22646 | 0.87 | 508.7 | **9263.7** | 276.9 | 516.3 | **7359.7** | 176.2 | 523.1 | **8871.4** | 214.2 |
| Canada | 14170 | 0.87 | 512.3 | **8795.8** | 187.2 | 517.4 | **7749.0** | 191.2 | 522.5 | **10045.9** | 204.0 |
| New Zealand | 4823 | 0.84 | 522.7 | **11069.3** | 331.7 | 523.8 | **8700.0** | 224.1 | 532.7 | **11513.0** | 290.8 |
| United States | missing | 0.85 | | | | 474.7 | **8055.6** | 341.7 | 488.3 | **11251.7** | 355.5 |
| *Eastern Europe* | | | | | | | | | | | |
| Bulgaria | 4498 | 0.83 | 406.8 | **13807.7** | 939.5 | 417.4 | **10221.9** | 736.9 | 439.1 | **11389.6** | 680.9 |
| Czech Republic | 5932 | 1.01 | 509.6 | **12368.4** | 643.7 | 536.0 | **10637.7** | 429.8 | 537.6 | **9685.0** | 394.0 |
| Estonia | 4865 | 0.94 | 502.4 | **7257.9** | 318.5 | 516.8 | **6509.3** | 247.9 | 533.7 | **7013.6** | 182.8 |
| Croatia | 5213 | 0.85 | 477.6 | **7890.4** | 376.7 | 467.3 | **6940.8** | 250.9 | 493.7 | **7347.1** | 247.5 |
| Hungary | 4490 | 0.85 | 488.1 | **8910.4** | 448.0 | 496.2 | **8288.2** | 351.8 | 508.7 | **7779.3** | 269.3 |
| Lithuania | 4744 | 0.93 | 469.3 | **9128.7** | 287.2 | 485.6 | **8064.5** | 309.5 | 486.5 | **8097.7** | 273.4 |
| Latvia | 4719 | 0.85 | 484.9 | **8226.9** | 306.0 | 491.2 | **6857.7** | 251.0 | 493.8 | **7119.4** | 218.7 |
| Montenegro | 4455 | 0.84 | 388.2 | **7994.3** | 295.4 | 395.8 | **7131.9** | 305.2 | 408.8 | **6350.7** | 191.2 |
| Poland | 5547 | 0.94 | 512.6 | **10043.6** | 296.2 | 500.9 | **7485.7** | 195.8 | 503.3 | **8077.1** | 199.1 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Romania | 5118 | 0.66 | 392.0 | **8438.8** | 539.4 | 415.0 | **7051.7** | 479.5 | 416.6 | **6586.6** | 385.3 |
| Serbia | 4798 | 0.83 | 402.9 | **8434.5** | 309.6 | 436.6 | **8419.1** | 323.8 | 436.9 | **7249.9** | 265.4 |
| Slovak Republic | 4731 | 0.95 | 470.6 | **11042.1** | 526.1 | 495.1 | **8936.3** | 466.3 | 491.2 | **8676.5** | 332.2 |
| Slovenia | 6595 | 0.88 | 468.6 | **7738.2** | 438.2 | 482.2 | **7965.2** | 243.2 | 494.2 | **9625.5** | 265.7 |
| *Western Europe* | | | | | | | | | | | |
| Austria | 4927 | 0.92 | 494.0 | **11698.0** | 686.2 | 509.5 | **9616.3** | 449.5 | 513.9 | **9570.0** | 471.2 |
| Belgium | 8857 | 0.99 | 507.1 | **12104.5** | 616.6 | 526.9 | **11263.4** | 703.3 | 516.3 | **9941.0** | 399.0 |
| Switzerland | 12192 | 1.02 | 496.6 | **8849.4** | 321.9 | 528.3 | **9495.2** | 311.6 | 508.0 | **9863.4** | 320.1 |
| Germany | 4891 | 0.95 | 496.5 | **12532.1** | 596.9 | 504.3 | **9817.2** | 501.6 | 516.2 | **9996.9** | 397.3 |
| Denmark | 4532 | 0.85 | 493.8 | **7975.3** | 290.1 | 512.2 | **7199.9** | 258.5 | 494.7 | **8673.2** | 264.0 |
| Spain | 19604 | 0.87 | 479.5 | **7892.9** | 203.1 | 501.7 | **7906.8** | 193.2 | 504.5 | **8197.2** | 176.4 |
| Finland | 4714 | 0.93 | 547.1 | **6597.8** | 175.8 | 549.0 | **6539.5** | 164.0 | 563.4 | **7331.6** | 171.1 |
| France | 4716 | 0.91 | 488.7 | **10805.1** | 571.5 | 496.4 | **9134.9** | 375.6 | 496.1 | **10317.6** | 426.2 |
| United Kingdom | 13152 | 0.94 | 495.6 | **10387.2** | 343.9 | 497.3 | **7906.7** | 233.8 | 514.3 | **11403.9** | 321.4 |
| Greece | 4873 | 0.90 | 461.9 | **10528.9** | 597.5 | 462.0 | **8518.9** | 437.1 | 476.6 | **8487.2** | 373.2 |
| Ireland | 4585 | 0.94 | 518.6 | **8536.7** | 343.2 | 502.3 | **6721.8** | 245.5 | 509.5 | **8902.1** | 284.1 |
| Iceland | 3789 | 0.96 | 485.0 | **9426.6** | 238.3 | 505.6 | **7758.6** | 156.7 | 491.0 | **9382.9** | 183.7 |
| Italy | 21773 | 0.90 | 477.0 | **11829.3** | 378.3 | 473.6 | **9181.5** | 318.7 | 487.2 | **9131.8** | 250.1 |
| Liechtenstein | 339 | 0.84 | 510.7 | **9054.6** | 556.2 | 524.9 | **8659.3** | 404.2 | 522.3 | **9401.1** | 406.4 |
| Luxembourg | 4567 | 1.03 | 480.1 | **9969.3** | 144.5 | 490.5 | **8677.1** | 135.6 | 486.8 | **9317.4** | 128.8 |
| Netherlands | 4871 | 0.96 | 513.9 | **9335.4** | 476.8 | 537.4 | **7850.6** | 385.7 | 530.8 | **9146.0** | 313.5 |
| Norway | 4692 | 0.97 | 484.4 | **11056.4** | 403.3 | 489.8 | **8386.6** | 253.3 | 486.9 | **9238.9** | 380.7 |
| Portugal | 5109 | 0.78 | 476.8 | **9765.1** | 449.7 | 470.9 | **8218.2** | 358.3 | 479.0 | **7843.2** | 303.5 |
| Sweden | 4443 | 0.97 | 509.0 | **9646.1** | 347.9 | 503.2 | **8038.8** | 246.4 | 504.2 | **8875.4** | 264.0 |

**Table 2: Estimates of the shares and levels of between circumstance groups inequality in achievements in Reading, Math, and Science**

| | Between group inequality in Reading | | Share of between group inequality in Reading | | Between group inequality in Math | | Share of between group inequality in Math | | Between group inequality in Science | | Share of between group inequality in Science | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Asia & North Africa* | | | | | | | | | | | | |
| Azerbaijan | **852.7** | 165.2 | **0.173** | 0.028 | **101.8** | 29.3 | **0.044** | 0.012 | **346.7** | 71.8 | **0.112** | 0.024 |
| Hong Kong-China | **1186.1** | 146.3 | **0.177** | 0.016 | **1340.5** | 184.4 | **0.154** | 0.016 | **1394.9** | 196.6 | **0.166** | 0.018 |
| Indonesia | **1399.7** | 283.0 | **0.250** | 0.038 | **1515.3** | 378.9 | **0.237** | 0.042 | **1079.3** | 315.0 | **0.220** | 0.045 |
| Israel | **2800.4** | 326.8 | **0.197** | 0.018 | **2373.9** | 290.5 | **0.206** | 0.019 | **2426.2** | 245.1 | **0.195** | 0.016 |
| Japan | **2157.0** | 251.2 | **0.206** | 0.017 | **1678.1** | 213.9 | **0.203** | 0.020 | **1899.5** | 205.3 | **0.189** | 0.016 |
| Jordan | **3065.9** | 331.6 | **0.346** | 0.024 | **1908.0** | 236.8 | **0.272** | 0.024 | **2190.7** | 216.4 | **0.271** | 0.019 |
| Korea | **1671.3** | 253.5 | **0.214** | 0.022 | **1795.8** | 289.4 | **0.209** | 0.021 | **1405.5** | 217.2 | **0.173** | 0.019 |
| Kyrgyzstan | **3273.4** | 359.4 | **0.314** | 0.023 | **2316.9** | 286.1 | **0.306** | 0.027 | **1892.0** | 226.3 | **0.269** | 0.023 |
| Macao-China | **737.9** | 110.9 | **0.127** | 0.012 | **718.9** | 77.8 | **0.102** | 0.009 | **675.4** | 61.6 | **0.111** | 0.008 |
| Qatar | **3609.0** | 176.1 | **0.309** | 0.010 | **2072.6** | 111.8 | **0.254** | 0.009 | **1830.6** | 107.9 | **0.264** | 0.009 |
| Russian Federation | **2066.6** | 242.9 | **0.238** | 0.021 | **1318.6** | 168.8 | **0.165** | 0.020 | **1469.8** | 187.5 | **0.183** | 0.020 |
| Chinese Taipei | **2132.7** | 186.2 | **0.300** | 0.017 | **2922.2** | 333.3 | **0.275** | 0.022 | **2504.9** | 236.3 | **0.281** | 0.019 |
| Thailand | **2174.7** | 207.3 | **0.325** | 0.023 | **1524.1** | 178.9 | **0.230** | 0.021 | **1575.5** | 172.7 | **0.265** | 0.022 |
| Tunisia | **2035.0** | 284.0 | **0.215** | 0.024 | **2309.5** | 353.9 | **0.273** | 0.031 | **1293.6** | 226.7 | **0.191** | 0.026 |
| Turkey | **2163.5** | 313.0 | **0.251** | 0.026 | **2099.6** | 457.4 | **0.241** | 0.033 | **1721.0** | 329.1 | **0.249** | 0.032 |
| *Latin America* | | | | | | | | | | | | |
| Argentina | **4462.0** | 514.9 | **0.289** | 0.024 | **3220.2** | 76.6 | **0.315** | 0.007 | **3199.4** | 375.5 | **0.312** | 0.026 |
| Brazil | **2817.5** | 366.6 | **0.268** | 0.020 | **2694.6** | 30.4 | **0.318** | 0.005 | **2280.5** | 246.5 | **0.286** | 0.021 |
| Chile | **2646.2** | 326.2 | **0.248** | 0.022 | **2520.3** | 43.4 | **0.330** | 0.001 | **2514.5** | 256.5 | **0.299** | 0.021 |
| Colombia | **2100.7** | 263.1 | **0.181** | 0.018 | **1676.3** | 72.6 | **0.216** | 0.007 | **1390.0** | 165.2 | **0.193** | 0.018 |
| Mexico | **2549.4** | 321.3 | **0.278** | 0.024 | **1899.7** | 24.7 | **0.261** | 0.002 | **1765.7** | 206.1 | **0.271** | 0.024 |
| Uruguay | **3245.0** | 281.1 | **0.221** | 0.015 | **2419.9** | 47.7 | **0.245** | 0.004 | **2208.8** | 150.3 | **0.248** | 0.012 |
| *North America & Oceania* | | | | | | | | | | | | |
| Australia | **1842.7** | 127.3 | **0.199** | 0.010 | **1129.4** | 81.6 | **0.153** | 0.009 | **1458.5** | 102.3 | **0.164** | 0.009 |
| Canada | **2128.2** | 121.6 | **0.242** | 0.011 | **1635.0** | 116.4 | **0.211** | 0.011 | **2079.8** | 122.4 | **0.207** | 0.010 |
| New Zealand | **3060.1** | 200.0 | **0.276** | 0.013 | **2093.6** | 137.7 | **0.241** | 0.012 | **3098.9** | 192.6 | **0.269** | 0.013 |
| United States | **missing** | | | | **2244.3** | 236.0 | **0.279** | 0.020 | **3176.4** | 296.0 | **0.282** | 0.019 |
| *Eastern Europe* | | | | | | | | | | | | |
| Bulgaria | **5211.5** | 697.4 | **0.377** | 0.028 | **3380.8** | 513.0 | **0.331** | 0.030 | **4141.3** | 554.4 | **0.364** | 0.030 |
| Czech Republic | **3665.1** | 411.9 | **0.296** | 0.021 | **2853.3** | 277.1 | **0.268** | 0.019 | **2701.5** | 275.4 | **0.279** | 0.020 |
| Estonia | **1965.3** | 140.3 | **0.271** | 0.013 | **1343.6** | 116.0 | **0.206** | 0.013 | **1457.6** | 109.7 | **0.208** | 0.012 |
| Croatia | **2343.6** | 211.5 | **0.297** | 0.017 | **1542.3** | 135.7 | **0.222** | 0.015 | **1759.0** | 143.2 | **0.239** | 0.014 |
| Hungary | **3072.7** | 316.9 | **0.345** | 0.023 | **2704.0** | 268.1 | **0.326** | 0.022 | **2535.1** | 207.5 | **0.326** | 0.019 |
| Lithuania | **2902.5** | 219.2 | **0.318** | 0.017 | **2253.9** | 186.0 | **0.279** | 0.017 | **2120.6** | 183.6 | **0.262** | 0.016 |
| Latvia | **2088.1** | 187.6 | **0.254** | 0.017 | **1375.9** | 171.7 | **0.201** | 0.020 | **1331.9** | 138.9 | **0.187** | 0.016 |
| Montenegro | **2014.7** | 155.5 | **0.252** | 0.013 | **1587.8** | 131.2 | **0.223** | 0.012 | **1253.8** | 95.6 | **0.197** | 0.011 |
| Poland | **2766.9** | 197.4 | **0.275** | 0.014 | **1802.2** | 123.8 | **0.241** | 0.013 | **1946.9** | 146.4 | **0.241** | 0.014 |
| Romania | **2536.8** | 347.6 | **0.301** | 0.026 | **2204.6** | 323.6 | **0.313** | 0.028 | **2042.5** | 277.3 | **0.310** | 0.027 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Serbia | **2626.1** | 220.8 | **0.311** | 0.018 | **2322.0** | 206.8 | **0.276** | 0.017 | **1850.7** | 162.4 | **0.255** | 0.016 |
| Slovak Republic | **3227.8** | 397.7 | **0.292** | 0.026 | **2828.7** | 387.9 | **0.317** | 0.030 | **2574.1** | 284.8 | **0.297** | 0.024 |
| Slovenia | **2602.4** | 265.9 | **0.336** | 0.018 | **2096.7** | 167.0 | **0.263** | 0.016 | **2578.5** | 184.4 | **0.268** | 0.014 |
| | | | | | | | | | | | | |
| *Western Europe* | | | | | | | | | | | | |
| Austria | **3463.3** | 338.9 | **0.296** | 0.019 | **2883.9** | 302.1 | **0.300** | 0.020 | **3100.8** | 336.7 | **0.324** | 0.022 |
| Belgium | **4059.3** | 334.7 | **0.335** | 0.015 | **3709.9** | 412.9 | **0.329** | 0.018 | **3359.3** | 253.2 | **0.338** | 0.015 |
| Switzerland | **2767.7** | 180.5 | **0.313** | 0.013 | **2681.3** | 168.9 | **0.282** | 0.013 | **3172.0** | 182.6 | **0.322** | 0.012 |
| Germany | **4606.8** | 396.0 | **0.368** | 0.021 | **3446.1** | 302.2 | **0.351** | 0.018 | **3519.9** | 303.5 | **0.352** | 0.019 |
| Denmark | **1825.1** | 157.7 | **0.229** | 0.015 | **1580.0** | 129.5 | **0.219** | 0.014 | **2156.3** | 187.7 | **0.249** | 0.017 |
| Spain | **1919.2** | 137.6 | **0.243** | 0.013 | **1890.8** | 120.5 | **0.239** | 0.012 | **2112.7** | 134.3 | **0.258** | 0.013 |
| Finland | **1628.0** | 108.9 | **0.247** | 0.014 | **1173.4** | 78.5 | **0.179** | 0.010 | **1225.0** | 89.6 | **0.167** | 0.011 |
| France | **3291.7** | 330.5 | **0.305** | 0.019 | **3059.9** | 273.3 | **0.335** | 0.019 | **3559.7** | 299.9 | **0.345** | 0.018 |
| United Kingdom | **2847.1** | 204.7 | **0.274** | 0.014 | **2043.1** | 127.8 | **0.258** | 0.012 | **3135.9** | 186.8 | **0.275** | 0.012 |
| Greece | **2748.0** | 372.4 | **0.261** | 0.023 | **1942.2** | 262.8 | **0.228** | 0.022 | **2080.0** | 235.2 | **0.245** | 0.019 |
| Ireland | **2214.7** | 215.2 | **0.259** | 0.018 | **1579.3** | 159.2 | **0.235** | 0.017 | **2137.3** | 193.2 | **0.240** | 0.016 |
| Iceland | **2203.1** | 116.1 | **0.234** | 0.009 | **1298.0** | 76.4 | **0.167** | 0.009 | **1724.9** | 102.2 | **0.184** | 0.009 |
| Italy | **2449.5** | 229.5 | **0.207** | 0.015 | **1634.4** | 156.4 | **0.178** | 0.014 | **1884.9** | 168.6 | **0.206** | 0.014 |
| Liechtenstein | **3516.8** | 399.9 | **0.388** | 0.031 | **2797.6** | 348.0 | **0.323** | 0.034 | **3560.9** | 355.5 | **0.379** | 0.030 |
| Luxembourg | **3434.2** | 104.6 | **0.344** | 0.008 | **2521.6** | 82.6 | **0.291** | 0.008 | **3056.8** | 95.1 | **0.328** | 0.009 |
| Netherlands | **2308.6** | 255.0 | **0.247** | 0.022 | **2126.7** | 237.3 | **0.271** | 0.023 | **2588.5** | 269.1 | **0.283** | 0.023 |
| Norway | **2993.8** | 257.1 | **0.271** | 0.016 | **1634.0** | 149.1 | **0.195** | 0.014 | **2032.0** | 237.6 | **0.220** | 0.018 |
| Portugal | **2962.8** | 322.2 | **0.303** | 0.021 | **2251.4** | 240.4 | **0.274** | 0.019 | **2096.3** | 224.0 | **0.267** | 0.020 |
| Sweden | **2561.0** | 191.0 | **0.265** | 0.014 | **1873.7** | 128.9 | **0.233** | 0.012 | **2217.5** | 148.0 | **0.250** | 0.013 |

**Table 3: Tests of robustness of the estimates of the shares and levels of between circumstance groups inequality in achievements in Reading, Math, and Science to the inclusion of out of sample 15 year olds**

| | PISA population without any correction | | | Correction assuming *selection on observables* | | | Correction assuming *strong selection on unobservables* | | |
|---|---|---|---|---|---|---|---|---|---|
| | Reading | Math | Science | Reading | Math | Science | Reading | Math | Science |
| TURKEY | | | | | | | | | |
| Total inequality | **8631.1** | **8693.9** | **6923.2** | **9678.0** | **8360.1** | **6819.9** | **24231.6** | **17965.5** | **14790.0** |
| | 510.5 | 803.1 | 521.8 | | | | | | |
| Parametric | **0.251** | **0.241** | **0.249** | **0.250** | **0.236** | **0.250** | **0.327** | **0.320** | **0.326** |
| | 0.026 | 0.033 | 0.032 | | | | | | |
| BRAZIL | | | | | | | | | |
| Total inequality | **10497.6** | **8467.3** | **7970.9** | **10579.2** | **8179.8** | **7525.6** | **32336.8** | **21514.8** | **21366.9** |
| | 685.4 | 60.7 | 344.2 | | | | | | |
| Parametric | **0.268** | **0.318** | **0.286** | **0.265** | **0.309** | **0.262** | **0.404** | **0.404** | **0.385** |
| | 0.020 | 0.005 | 0.021 | | | | | | |
| MEXICO | | | | | | | | | |
| Total inequality | **9154.5** | **7270.6** | **6513.4** | **9144.9** | **7228.4** | **6269.4** | **38749.9** | **26500.0** | **18766.3** |
| | 433.9 | 58.6 | 237.3 | | | | | | |
| Parametric | **0.278** | **0.261** | **0.271** | **0.267** | **0.242** | **0.255** | **0.256** | **0.250** | **0.228** |
| | 0.024 | 0.002 | 0.024 | | | | | | |
| INDONESIA | | | | | | | | | |
| Total inequality | **5593.2** | **6401.9** | **4908.9** | **5045.1** | **5816.8** | **4322.3** | **17045.9** | **18466.6** | **12722.4** |
| | 357.5 | 508.5 | 457.4 | | | | | | |
| Parametric | **0.250** | **0.237** | **0.220** | **0.218** | **0.200** | **0.181** | **0.274** | **0.261** | **0.261** |
| | 0.038 | 0.042 | 0.045 | | | | | | |

**Table 4: Estimates of the partial shares of between circumstance groups inequality in achievements in Reading, Math, and Science associated with individual circumstance variables**

| | Total | Gender | Father's education | Mother's education | Father's occupation | Area type | Language at home | Immigration status | Number of books | Durables | Cultural possessions |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Asia & North Africa* | | | | | | | | | | | |
| Azerbaijan | 0.044 | 0.000 | 0.000 | 0.000 | 0.001 | 0.003 | 0.000 | 0.006 | 0.017 | 0.008 | 0.010 |
| Hong Kong-China | 0.154 | 0.009 | 0.012 | 0.007 | 0.026 | 0.000 | 0.000 | 0.013 | 0.062 | 0.009 | 0.018 |
| Indonesia | 0.237 | 0.009 | 0.009 | 0.005 | 0.018 | 0.072 | 0.002 | 0.000 | 0.025 | 0.096 | 0.009 |
| Israel | 0.206 | 0.004 | 0.002 | 0.039 | 0.057 | 0.006 | 0.001 | 0.000 | 0.065 | 0.003 | 0.030 |
| Japan | 0.203 | 0.012 | 0.042 | 0.027 | 0.025 | 0.005 | 0.000 | 0.004 | 0.032 | 0.013 | 0.044 |
| Jordan | 0.272 | 0.001 | 0.030 | 0.029 | 0.043 | 0.022 | 0.007 | 0.000 | 0.021 | 0.103 | 0.016 |
| Korea | 0.209 | 0.004 | 0.017 | 0.011 | 0.000 | 0.019 | 0.000 | 0.001 | 0.086 | 0.014 | 0.061 |
| Kyrgyzstan | 0.306 | 0.000 | 0.002 | 0.012 | 0.014 | 0.107 | 0.008 | 0.007 | 0.066 | 0.053 | 0.037 |
| Macao-China | 0.102 | 0.006 | 0.008 | 0.001 | 0.007 | 0.003 | 0.005 | 0.003 | 0.010 | 0.021 | 0.039 |
| Qatar | 0.254 | 0.010 | 0.011 | 0.005 | 0.052 | 0.035 | 0.079 | 0.016 | 0.018 | 0.012 | 0.017 |
| Russian Federation | 0.165 | 0.001 | 0.001 | 0.009 | 0.030 | 0.009 | 0.004 | 0.003 | 0.046 | 0.037 | 0.024 |
| Chinese Taipei | 0.275 | 0.005 | 0.029 | 0.015 | 0.031 | 0.026 | 0.000 | 0.008 | 0.088 | 0.018 | 0.054 |
| Thailand | 0.230 | 0.001 | 0.023 | 0.026 | 0.048 | 0.028 | 0.001 | 0.000 | 0.024 | 0.079 | 0.000 |
| Tunisia | 0.273 | 0.009 | 0.001 | 0.000 | 0.072 | 0.032 | 0.005 | 0.000 | 0.046 | 0.077 | 0.034 |
| Turkey | 0.241 | 0.003 | 0.042 | 0.041 | 0.007 | 0.018 | 0.000 | 0.001 | 0.051 | 0.045 | 0.034 |
| *Latin America* | | | | | | | | | | | |
| Argentina | 0.315 | 0.004 | 0.014 | 0.026 | 0.024 | 0.022 | 0.000 | 0.003 | 0.079 | 0.114 | 0.029 |
| Brazil | 0.318 | 0.009 | 0.019 | 0.024 | 0.027 | 0.014 | 0.005 | 0.001 | 0.025 | 0.184 | 0.011 |
| Chile | 0.330 | 0.021 | 0.016 | 0.055 | 0.050 | 0.026 | 0.001 | 0.000 | 0.068 | 0.060 | 0.033 |
| Colombia | 0.216 | 0.017 | 0.009 | 0.015 | 0.014 | 0.014 | 0.003 | 0.000 | 0.049 | 0.085 | 0.010 |
| Mexico | 0.261 | 0.003 | 0.001 | 0.025 | 0.018 | 0.074 | 0.014 | 0.002 | 0.033 | 0.077 | 0.014 |
| Uruguay | 0.245 | 0.005 | 0.013 | 0.047 | 0.029 | 0.006 | 0.000 | 0.000 | 0.056 | 0.059 | 0.030 |
| *North America & Oceania* | | | | | | | | | | | |
| Australia | 0.153 | 0.008 | 0.007 | 0.009 | 0.044 | 0.002 | 0.000 | 0.000 | 0.055 | 0.011 | 0.016 |
| Canada | 0.211 | 0.008 | 0.029 | 0.011 | 0.035 | 0.017 | 0.003 | 0.000 | 0.078 | 0.013 | 0.018 |
| New Zealand | 0.241 | 0.005 | 0.036 | 0.016 | 0.036 | 0.003 | 0.000 | 0.000 | 0.074 | 0.034 | 0.037 |
| United States | 0.279 | 0.004 | 0.014 | 0.018 | 0.062 | 0.013 | 0.000 | 0.003 | 0.122 | 0.036 | 0.010 |
| *Eastern Europe* | | | | | | | | | | | |
| Bulgaria | 0.331 | 0.000 | 0.005 | 0.020 | 0.052 | 0.032 | 0.001 | 0.012 | 0.102 | 0.048 | 0.060 |
| Czech Republic | 0.268 | 0.004 | 0.010 | 0.035 | 0.045 | 0.007 | 0.001 | 0.001 | 0.089 | 0.052 | 0.024 |
| Estonia | 0.206 | 0.000 | 0.000 | 0.019 | 0.061 | 0.003 | 0.007 | 0.000 | 0.080 | 0.012 | 0.028 |
| Croatia | 0.222 | 0.011 | 0.006 | 0.000 | 0.041 | 0.007 | 0.000 | 0.004 | 0.060 | 0.046 | 0.048 |
| Hungary | 0.326 | 0.005 | 0.038 | 0.054 | 0.038 | 0.016 | 0.000 | 0.002 | 0.099 | 0.034 | 0.042 |
| Lithuania | 0.279 | 0.001 | 0.007 | 0.023 | 0.030 | 0.024 | 0.001 | 0.002 | 0.080 | 0.061 | 0.051 |
| Latvia | 0.201 | 0.002 | 0.000 | 0.025 | 0.028 | 0.007 | 0.000 | 0.000 | 0.069 | 0.048 | 0.024 |
| Montenegro | 0.223 | 0.006 | 0.000 | 0.014 | 0.025 | 0.002 | 0.001 | 0.007 | 0.071 | 0.021 | 0.081 |
| Poland | 0.241 | 0.004 | 0.014 | 0.035 | 0.019 | 0.008 | 0.000 | 0.000 | 0.078 | 0.030 | 0.051 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Romania | 0.313 | 0.004 | 0.000 | 0.006 | 0.057 | 0.022 | 0.000 | 0.001 | 0.084 | 0.062 | 0.078 |
| Serbia | 0.276 | 0.003 | 0.006 | 0.011 | 0.034 | 0.020 | 0.003 | 0.000 | 0.086 | 0.063 | 0.050 |
| Slovak Republic | 0.317 | 0.008 | 0.030 | 0.027 | 0.033 | 0.004 | 0.001 | 0.014 | 0.137 | 0.054 | 0.009 |
| Slovenia | 0.263 | 0.002 | 0.022 | 0.043 | 0.044 | 0.003 | 0.000 | 0.006 | 0.105 | 0.003 | 0.038 |
| *Western Europe* | | | | | | | | | | | |
| Austria | 0.300 | 0.017 | 0.003 | 0.017 | 0.026 | 0.006 | 0.018 | 0.008 | 0.144 | 0.017 | 0.044 |
| Belgium | 0.329 | 0.002 | 0.029 | 0.049 | 0.056 | 0.009 | 0.053 | 0.000 | 0.065 | 0.030 | 0.040 |
| Switzerland | 0.282 | 0.006 | 0.024 | 0.019 | 0.028 | 0.012 | 0.050 | 0.006 | 0.104 | 0.012 | 0.021 |
| Germany | 0.351 | 0.012 | 0.019 | 0.050 | 0.047 | 0.007 | 0.014 | 0.012 | 0.131 | 0.010 | 0.049 |
| Denmark | 0.219 | 0.005 | 0.018 | 0.020 | 0.028 | 0.002 | 0.015 | 0.013 | 0.064 | 0.008 | 0.047 |
| Spain | 0.239 | 0.004 | 0.014 | 0.026 | 0.028 | 0.002 | 0.010 | 0.001 | 0.103 | 0.032 | 0.020 |
| Finland | 0.179 | 0.008 | 0.011 | 0.018 | 0.019 | 0.000 | 0.009 | 0.004 | 0.073 | 0.006 | 0.033 |
| France | 0.335 | 0.002 | 0.034 | 0.025 | 0.059 | 0.000 | 0.007 | 0.008 | 0.104 | 0.028 | 0.069 |
| United Kingdom | 0.258 | 0.010 | 0.027 | 0.021 | 0.051 | 0.002 | 0.000 | 0.004 | 0.113 | 0.010 | 0.019 |
| Greece | 0.228 | 0.001 | 0.040 | 0.024 | 0.036 | 0.008 | 0.003 | 0.003 | 0.059 | 0.037 | 0.017 |
| Ireland | 0.235 | 0.006 | 0.011 | 0.024 | 0.025 | 0.001 | 0.001 | 0.006 | 0.103 | 0.017 | 0.040 |
| Iceland | 0.167 | 0.001 | 0.014 | 0.049 | 0.027 | 0.001 | 0.004 | 0.003 | 0.061 | 0.000 | 0.012 |
| Italy | 0.178 | 0.008 | 0.006 | 0.011 | 0.016 | 0.024 | 0.003 | 0.000 | 0.061 | 0.028 | 0.023 |
| Liechtenstein | 0.323 | 0.001 | 0.058 | 0.008 | 0.033 | 0.000 | 0.020 | 0.029 | 0.050 | 0.049 | 0.076 |
| Luxembourg | 0.291 | 0.010 | 0.007 | 0.011 | 0.072 | 0.009 | 0.018 | 0.007 | 0.102 | 0.013 | 0.041 |
| Netherlands | 0.271 | 0.006 | 0.009 | 0.020 | 0.065 | 0.010 | 0.018 | 0.004 | 0.111 | 0.004 | 0.024 |
| Norway | 0.195 | 0.002 | 0.010 | 0.013 | 0.050 | 0.000 | 0.006 | 0.003 | 0.063 | 0.006 | 0.041 |
| Portugal | 0.274 | 0.007 | 0.000 | 0.029 | 0.056 | 0.009 | 0.013 | 0.000 | 0.072 | 0.051 | 0.042 |
| Sweden | 0.233 | 0.001 | 0.002 | 0.020 | 0.052 | 0.004 | 0.011 | 0.004 | 0.095 | 0.009 | 0.034 |

**Table 5: Regression coefficients for the relationship between the share of education public expenditure at the primary level and the share and level of between circumstance group inequality in achievements in Reading, Math, and Science**

Without controls

| | Reading | | Math | | Science | |
|---|---|---|---|---|---|---|
| Share all countries | -0.00217*** | (0.00092) | -0.00077 | (0.00112) | -0.00152 | (0.00105) |
| Share excluding outliers | -0.00300*** | (0.00078) | -0.00113 | (0.00101) | -0.00172* | (0.00101) |
| Level all countries | -20.98 | (15.77) | -10.77 | (12.68) | -20.98 | (15.77) |
| Level excluding outliers | -21.27* | (12.60) | -17.16 | (11.51) | -21.27* | (12.60) |

*Notes: regression estimates of the relationship between the share of public expenditure in education allocated to the primary level and the estimated share and level of between circumstance group inequality in achievements with and without outliers; outliers are identified using the method proposed by Besley, Kuh and Welsch (1980); source: UNESCO Institute for Statistics database; ***/**/*: significant at 1/5/10%.*

Controlling for GDP and public expenditure in education per pupil

| | Reading | | Math | | Science | |
|---|---|---|---|---|---|---|
| Share all countries | -0.00197** | (0.00087) | -0.00013 | (0.00120) | -0.00103 | (0.00113) |
| Share excluding outliers | -0.00184*** | (0.00072) | -0.00181* | (0.00102) | -0.00185* | (0.00108) |
| Level all countries | -16.27 | (16.97) | -3.81 | (13.79) | -10.57 | (15.17) |
| Level excluding outliers | -29.18** | (13.12) | -19.57* | (11.32) | -22.25* | (13.07) |

*Notes: regression estimates controlling for per capita GDP and public expenditure in education per pupil; ***/**/*: significant at 1/5/10%.*

**Table 6: Regression coefficients for the relationship between tracking (measured by the share of technical and vocational enrollment at the secondary level) and the shares and levels of between circumstance groups inequality in achievements in Reading, Math, and Science**

Without controls

| | Reading | | Math | | Science | |
|---|---|---|---|---|---|---|
| Share all countries | 0.00106* | (0.00059) | 0.00130* | (0.00070) | 0.00179*** | (0.00063) |
| Share excluding outliers | 0.00158** | (0.00060) | 0.00109* | (0.00062) | 0.00160*** | (0.00059) |
| Level all countries | 23.00*** | (9.59) | 19.48*** | (7.59) | 28.10*** | (7.70) |
| Level excluding outliers | 21.45*** | (9.17) | 17.84*** | (7.27) | 26.09*** | (7.13) |

*Notes: regression estimates of the relationship between the share of technical and vocational enrollment at the secondary level and the estimated share and level of between circumstance group inequality in achievements with and without outliers; outliers are identified using the method proposed by Besley, Kuh and Welsch (1980); source: UNESCO Institute for Statistics database; ***/**/*: significant at 1/5/10%.*

Controlling for GDP and public expenditure in education per pupil

| | Reading | | Math | | Science | |
|---|---|---|---|---|---|---|
| Share all countries | 0.00148*** | (0.00057) | 0.00173*** | (0.00074) | 0.00214*** | (0.00068) |
| Share excluding outliers | 0.00090* | (0.00047) | 0.00175*** | (0.00065) | 0.00205*** | (0.00067) |
| Level all countries | 32.15*** | (10.21) | 22.12*** | (8.55) | 18.90*** | (7.64) |
| Level excluding outliers | 22.00*** | (8.84) | 18.90*** | (7.64) | 21.36*** | (8.62) |

*Notes: regression estimates controlling for per capita GDP and public expenditure in education per pupil; ***/**/*: significant at 1/5/10%.*

**Figure 1: Ranking of countries by estimated relative levels of the variance of achievements in Math**

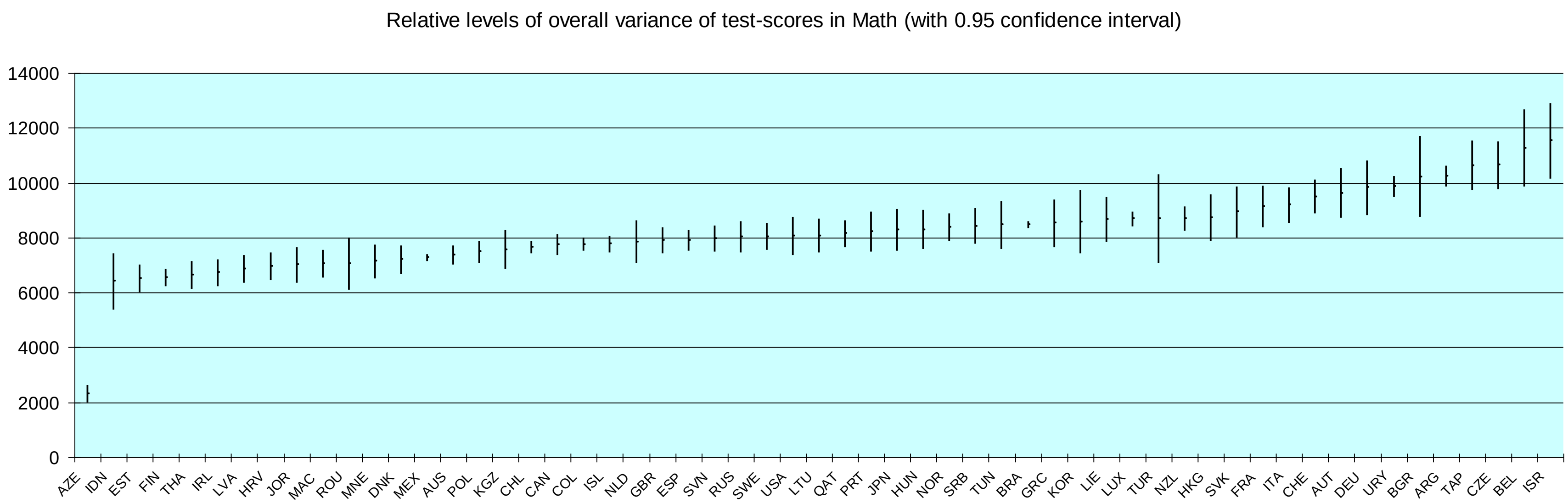
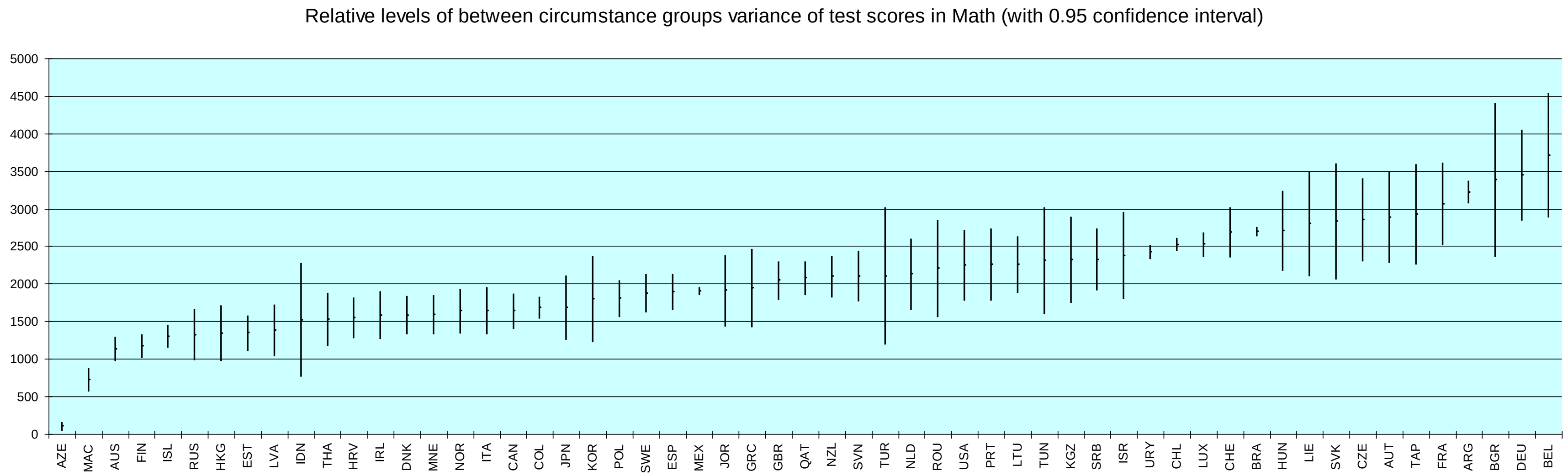


**Figure 2: Ranking of countries by estimated relative levels (panel a) and shares (panel b) of between circumstance groups variance for achievements in Math**

Panel a



Panel b

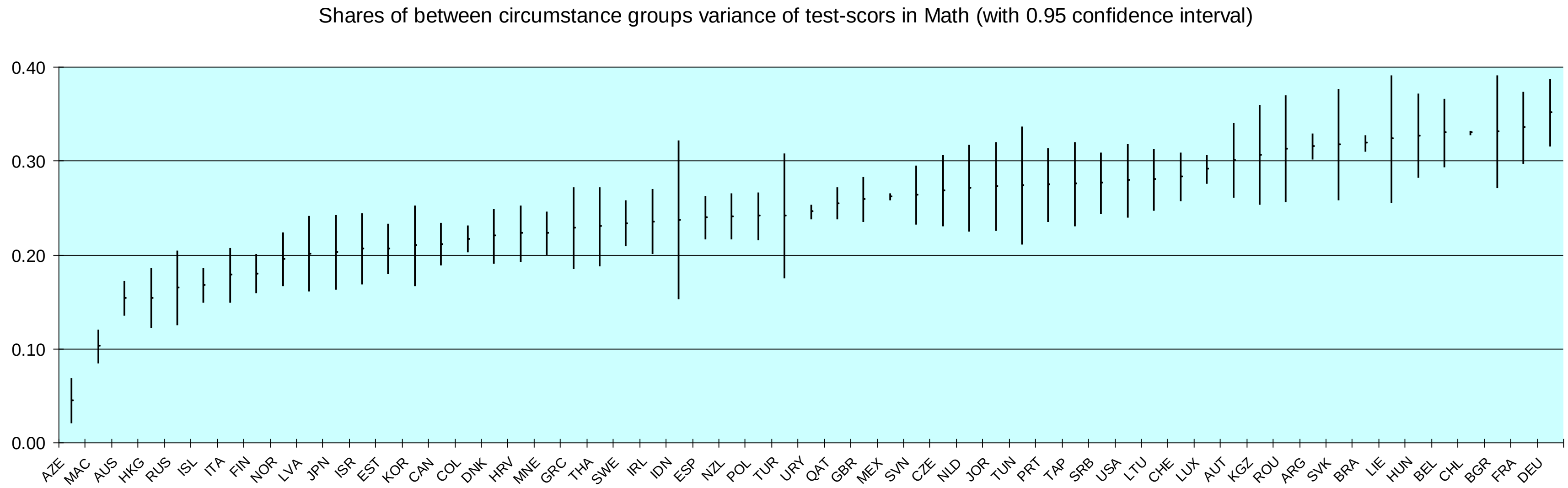
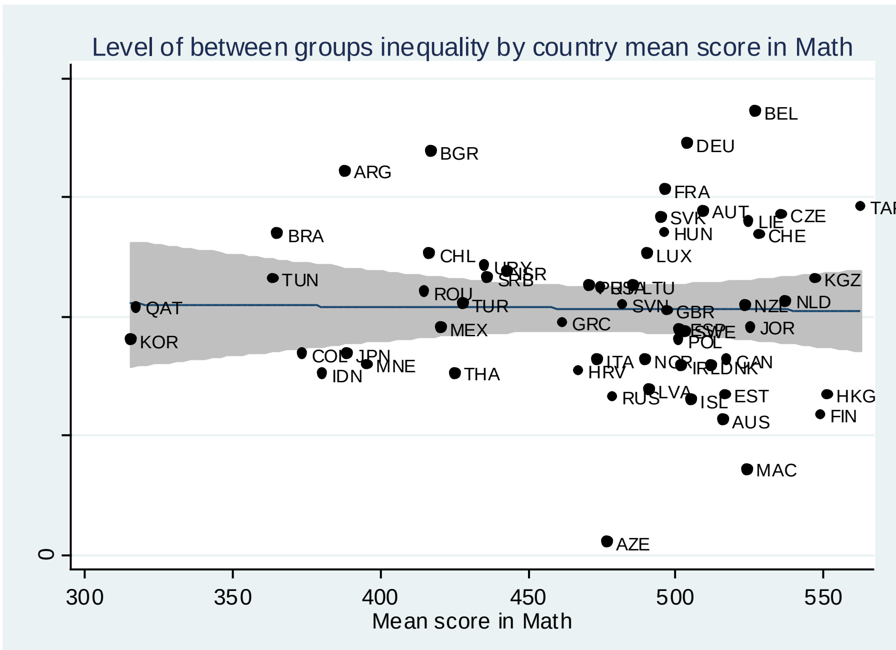

Shares of between circumstance groups variance of test-scors in Math (with 0.95 confidence interval)
0.40
0.30
0.20
0.10
0.00
AZE
MAC
AUS
HKG
RUS
ISL
ITA
FIN
NOR
LVA
JPN
ISR
EST
KOR
CAN
COL
DNK
HRV
MNE
GRC
THA
SWE
IRL
IDN
ESP
NZL
POL
TUR
URY
QAT
GBR
MEX
SVN
CZE
NLD
JOR
TUN
PRT
TAP
SRB
USA
LTU
CHE
LUX
AUT
KGZ
ROU
ARG
SVK
BRA
LIE
HUN
BEL
CHL
BGR
FRA
DEU

**Figure 3: Relative levels (panel a) and shares (panel b) of between group inequality by country mean scores for achievements in Math**

Panel a




Panel b

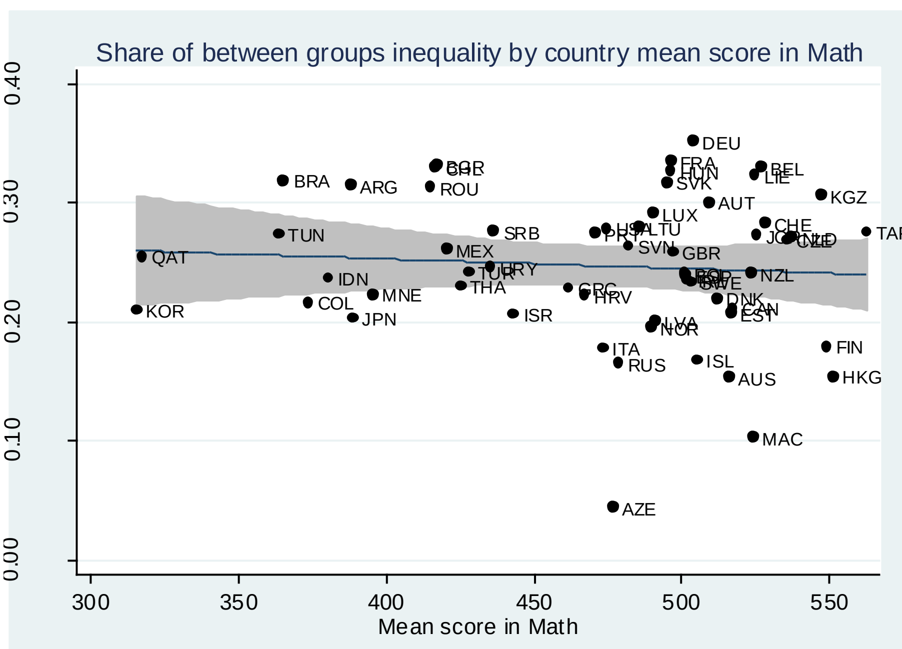


**Figure 4: Distribution of standardized Turkish reading test scores under three alternative assumptions about selection into PISA participation**

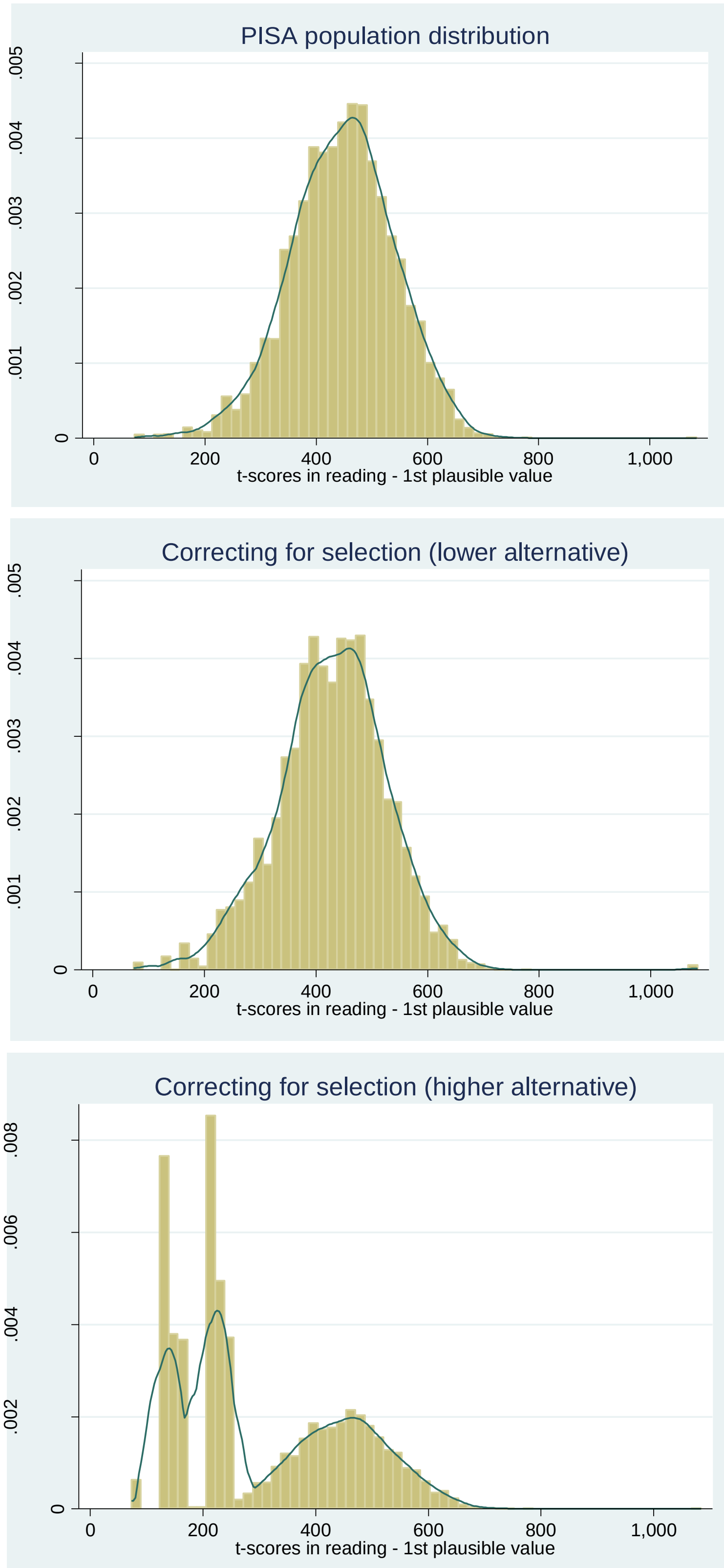
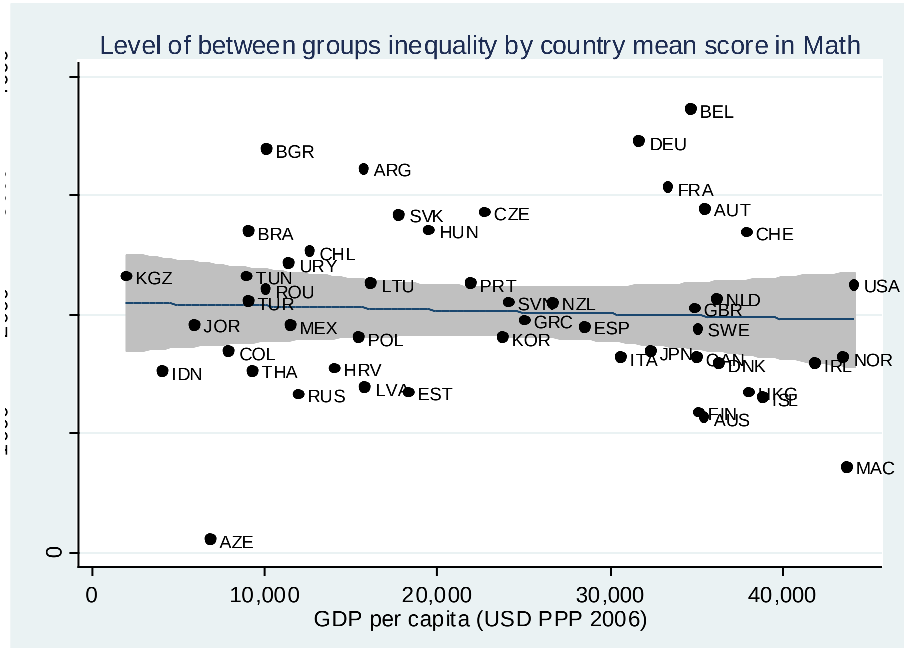


**Figure 5: measures of between circumstance groups inequalities (levels in panel a and shares in panel b) in learning and GDP per capita in 2006**

Panel a




Panel b

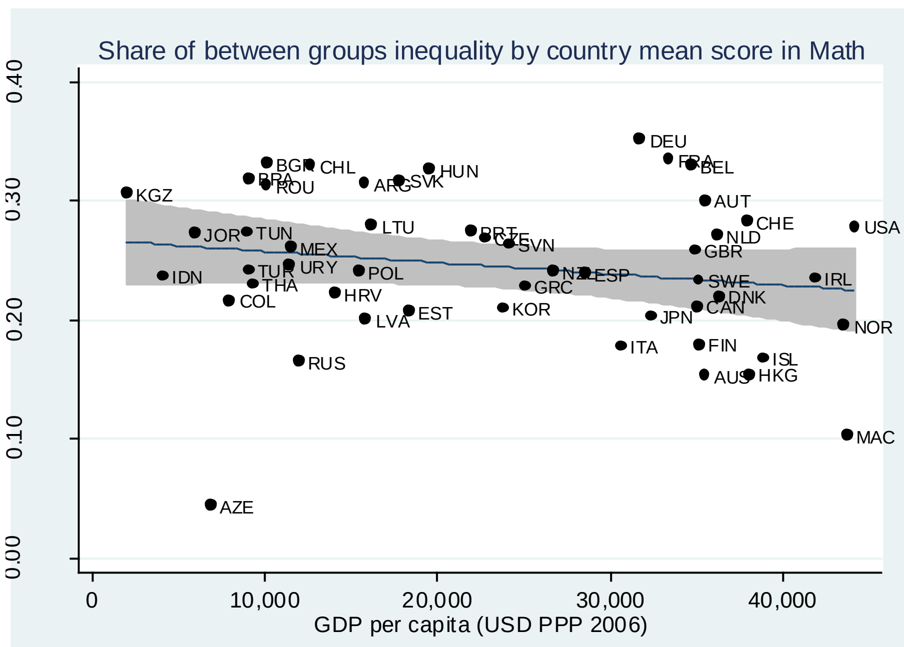
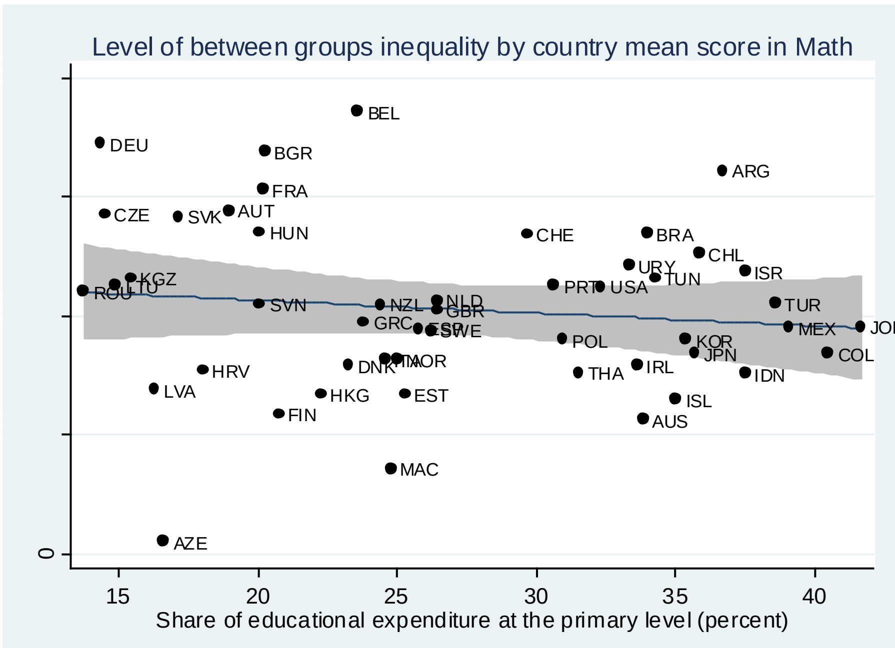


**Figure 6: measures of between circumstance groups inequalities in learning (levels in panel a and shares in panel b) and share of education public expenditure at the primary level**

Panel a




Panel b

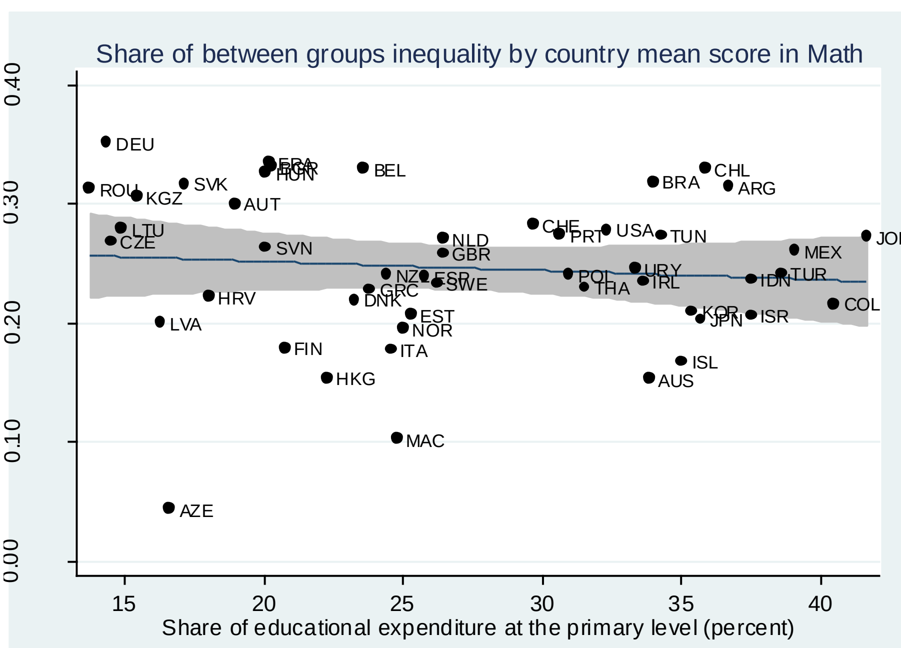
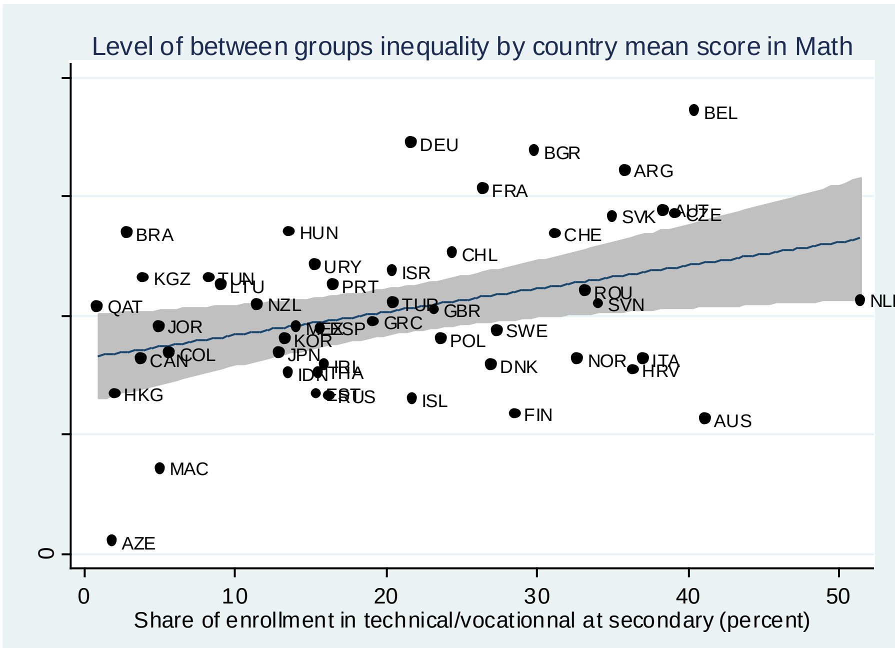


**Figure 7: measures of between circumstance groups inequalities in learning (levels in panel a and shares in panel b) and shares of enrollment in technical or vocational curricula at the secondary level**

Panel a




Panel b