Inequality of Opportunity in Education: A Quantile Regression Approach

Mercy Laita Palamuleni*

Department of Economics and Management, Gustavus Adolphus College

September 26, 2014

Abstract

This paper provides lower-bound estimates of inequality of opportunity in education (IOE) using micro-data from the Programme for International Student Assessment (PISA). The measure is the variation in student mathematics test scores explained by predetermined circumstances (including parental education, gender, and additional community variables). IOE accounts for 10 percent of the variation in test scores for students at the top and bottom of the test score distribution. Three main conclusions are established from using the IOE measure: (1) IOE decreases with an increase in preprimary enrollment rates. Suggesting that improvements in early childhood education might mitigate the effects of IOE factors for some students. (2) IOE increases as overall inequality increases. This indicates the possibility of a more general persistence to inequality factors. Suggesting that equity-based education policies can be a key tool for reducing income inequality. (3) There is evidence of an equity-efficiency tradeoff in education. An implication here is that public education policies aimed at reducing IEO might hinder overall education efficiency, in that it decreases academic achievement for some groups of students.

Keywords: Inequality of Opportunity; Inequality

*Address: 800 W. College Ave., Beck Hall 155, Gustavus Adolphus College, St. Peter, MN 56082, USA. Telephone: 1-704-880-1672. E-mail: *mpalamul@gustavus.edu.*

1 Introduction

Recent attention in the inequality literature has shifted towards examining the inequalities inherent in an individual's access to opportunity. This change of focus is a consequence of recently developing views from egalitarian philosophers, who suggest that distinguishing inequality by source and type might be fruitful. Roemer (1998) is among the most recent influential researchers to interpret the philosophical view of inequality of opportunity in a manner that it can be empirically measured. He suggests dividing inequality into fair and unfair inequalities, as judged by whether or not the inequality is due to environmental conditions that an individual can control. In the human capital context, fair or legitimate inequalities can include, an individual's academic motivation. Conversely, unfair sources of inequality are based on circumstances (whether positive or negative from an academic standpoint) that are beyond a person's control. In this essay, I apply this philosophical view and provide a measure of inequality of opportunity in education (IOE). I implement a modified parametric method for quantifying inequality proposed by Ferreira and Gignoux (2011). After computing IOE, I analyze the role of educational policies and various economic environments in determining IOE. I also examine whether there exist an equity-efficiency tradeoff in education.

I construct an IOE measure using student-level data from the Programme for International Student Assessment (PISA) for the years 2003, 2006, and 2009. For each country and year, I measure IOE as a variation in student achievement explainable by predetermined circumstances in terms of gender, whether a student is a native born, family background (parental education) and school location characteristics.

The measured IOE represents the lower-bound estimate of the actual IOE. Since I use a subset of circumstances that affect a student's achievement; namely those available in the PISA data across time. The inclusion of other relevant variables would improve the magnitude of the measure.

It is important to note that the approach has several advantages. First, in contrast with other inequality measures, the IOE measure in this paper makes more extensive use of available predetermined factors. Second, the reported IOE is a relative measure of inequality, and so satisfies the axioms of inequality measures. Furthermore, since it is a relative measure of inequality this allows meaningful cross-country comparisons.

I contributes to the current education and public policy literatures in several useful ways. Most notably, I quantify IOE by focusing on predetermined characteristics of students and their families, which enables me to use a relatively larger set of circumstances. to quantify IOE. [1] In addition, I also use quantile regression analysis to construct the relative measure of IOE conditional upon a student's test score distribution. This enables an evaluation of how educational systems affect students at different levels of academic achievement. Finally, by taking advantage of several waves of PISA data, I construct a panel dataset at the international level. I use the panel dataset to provide fixed-effect (FE) estimates on the effect of the education system on IOE (which is indirectly the effect of education system on student achieve-

[1]This is in contrast to previous studies that only measure the effect of a single family factor on student outcomes in their representations of IOE. For example, Woessmann (2004) and Shueltz et. al. (2008) measure IOE as an effect of the number of books at a student's home. They defend their choice of measure by explaining that the number of books at home is a robust at explaining international differences in student achievement. It is important to note that this measure is in some ways inadequate, because it excludes other predetermined home factors which might influence a student's achievement.

ment) as well as estimates of the education policy question of equality and efficiency trade-off. One notable disadvantage of the fixed-effect analysis is the inability to estimate the effects of a country's non-invariant education system characteristics, such as student tracking. The panel data FE estimation enables direct control of cross-country heterogeneity, which is impossible in cross-sectional studies.

The main findings are consistent with the findings from other studies that use the PISA dataset. Predetermined circumstances explain differences in student achievement within and between countries, as well as across the student test score distribution. The circumstances accounts for 10% of the variation in test scores of students at the top and bottom of the test score distribution. The FE estimates show that overall economy-wide inequality (as measured by the GINI coefficient) explains cross-country differences in IOE. This has implications in addressing the roles of educational policy and economic activity on academic achievement. They suggest that perhaps policy should focus on combating overall economic disparities as a way of reducing inequality in education. The results also suggest that equity-based education policies can be a key tool for reducing income inequality. The results also demonstrate that access to preprimary education reduces IOE for students at the top of the test score distribution. An implication here is that improvements in early childhood education (such as increasing the average enrollment rate for all children in a country) can mitigate inequality in education opportunity.

Important to education policy is the discussion of the equity-efficiency tradeoff of any existing or planned program. I find a non-robust equity-efficiency tradeoff exists in education sector. The tradeoff is evidenced especially for students at the middle of the test score distribution. This suggests that policies aimed at promoting equality in education opportunity might actually hinder the overall efficiency of a system by decreasing academic achievement for some groups of students. There is also evidence that IOE at the bottom of the test score distribution is positively related to higher average test scores.

The rest of this paper is organized as follows: section 2 includes a brief literature review on the application of inequality of opportunity. Section 3 describes the analytical framework used to compute IOE. Section 4 describes the dataset used. Section 5 reports the findings and investigates potential uses for the IOE measure. Section 6 contains concluding remarks.

2 Literature review

Does an individual's achievement predominantly depend on effort, or on the predetermined circumstances of that individual's life? This question has challenged researchers, philosophers, and policy makers. For the past 15 years most of the developments in understanding the nature of inequality of opportunity have resurfaced due to works such as Roemer (1998), who first formalized the concept and originally coined it "inequality of opportunity". According to Roemer, the best way to understand inequality of opportunity is to view it as differences in outcome which can be attributed to circumstances beyond one's control. It is important to distinguish inequality of opportunity from inequality of outcome in the education context (which is a common measure of inequality in human capital). Although these concepts might be correlated and both represent inequality, they differ conceptually. The common measurement of inequality of outcome is the variance of the outcome (or other measures of spread).[2] The fundamental focus of inequality of opportunity measure is on quantifying the disparities in opportunities to achieve a goal rather than disparities in outcome.

In general, the concept of inequality of opportunity is motivated by the principles of compensation and natural reward. The principle of compensation attributes inequality of opportunity to the differences in outcomes as a result of factors beyond an individual's control (and therefore, calls for social compensation to address this issue). From this perspective, inequalities of opportunity demand a response such as focused government intervention which can "level the playing field" for those individuals who suffer because of unfortunate personal circumstances. On the other hand, the principle of natural reward states that responsible effort on the part of an individual should be encouraged and rewarded. Even though the concept of inequality of opportunity is established, much of the empirical literature differs on the empirical methodology to measure it.

For the literature following Roemer (1998) philosophy, the primary goal is to decompose the inequality in outcomes into inequality that results from circumstantial factors and inequality resulting from other factors (individual choice, talent, and luck) which is usually called effort. The current literature differs on the usage of parametric

[2]The human capital-growth literature documents both positive and negative effects of human capital inequality on economic growth. For example, Hanushek and Woessmann (2008) show that the variance of student outcome as a measure of inequality of human capital is a strong predictor of economic growth. However, Castello and Domenech, (2002) demonstrate that inequality in human capital hinders economic growth. This implies that inequalities in human capital represent both production efficiencies and production inefficiencies.

approach or non-parametric approach for estimation purposes. Both approaches have their own advantages and limitations. For instance, the main advantage of using non-parametric approaches is that one needs not to specify a direct functional relationship between outcomes and circumstances (or efforts). However as Ferreira and Gignoux (2011) point out, these approaches suffer from data insufficiency when the number of circumstances is large. On the other hand, the parametric approaches are able to include a relatively large set of circumstance, and one can decompose the partial effects of individual circumstances. The major disadvantage of the parametric approach is that one has to assume the functional forms of the relationship between outcomes and circumstances. To this end, the literature generally views the non-parametric and parametric approaches as complementary to each other and not as substitutes for one another (for specific examples, see arguments made by Checchi and Peragine (2010) and Ferreira and Gignoux (2008)). I present a parametric approach.

3 An analytical framework for quantifying IOE

My framework is based on a modified parametric approach to an inequality of opportunity measurement proposed by Ferreira and Gignoux (2011). This method is also closely related to the parametric approaches discussed by Bourguignon et al. (2007), as well as Checchi and Peragine (2010) and Fernando et al. (2012).

The general framework begins with an assumption that each population contains individuals such as the students in my case, which can be indexed by $i \in 1, ..., N$.

Each person's outcome, (each student's test score in mathematics), denoted by T_i, depends on a set of circumstances, C_i, and an amount of effort, E_i, in addition to other environmental factors μ_i, such that:

$$T_i = f(C_i, E_i, \mu_i). \tag{1}$$

In the education context, circumstances are weakly exogenous to a student's test outcome because they are predetermined and cannot be influenced by his or her decisions. However, circumstances can influence a student's effort. For instance, a student's academic effort might depend on her family's social economic status. Thus a reduced form of equation (1) can be expressed as:

$$T_i = f(C_i, E_i(C_i), \mu_i). \tag{2}$$

Once the individual circumstances for a student are defined and identified, students can be partitioned into homogenous groups of circumstance. The most fundamental question in defining inequality is establish the benchmark of equality and measure inequality as overall (or relative) deviations from the equality. In the literature, there are two basic approaches for defining benchmarks of equality in opportunity (thus inversely measuring the inequality of opportunity). An ex-ante approach uses evaluations of the opportunities available to each group, and compares the evaluation to an ideal equality of opportunity that would exist if all sources of IOE were hypothetically eliminated. For example, one can use the mean value of a nation's opportunity set as the standard, then argue that equality of opportunity is achieved when there is no difference in the means across the various subgroups in a that nation (Fleurbaey and Paragine, 2013). Thus inequality in opportunity can be represented by the between-type inequality, (or differences between students who would otherwise have the same characteristics in a given system).

On the other hand, ex-post approaches also offer unique insight into ways of defining the benchmarks of equality of opportunity. The ex-post methodology follows directly from the Roemer (1998, 2001) philosophical body, which argues that equality of opportunity is achieved only when individuals who exert the same effort achieve the same outcome. Thus, inequality of opportunity is expressed as the sum of inequalities within subgroups that exert the same effort. Although the two approaches both approximate inequality of opportunity, they differ in quantifying the degree of inequality. Similar to Ferreira and Gignoux (2011), who build on the models by Bourguignion et al. (2007) and Checchi and Peragine (2010), I adopt the ex-ante approach which allows for computation of the lower bounds of the inequality of opportunity. This inequality of opportunity (IO) can be approximated using non-parametric procedures. However, Ferreira and Gignoux (2011) point out that this can be data-intensive, especially when the vector of circumstances is large (more than three circumstances). Ferreira and Gignoux (2008) present a parametric method based on regression estimates, and use the variance to measure inequality. The procedure involves assuming a relationship between outcome (such as student achievement) and circumstances. Specifically, a simple linear reduced form specification for equation (2) can be expressed as:

$$T_i = C_i'\beta + \varepsilon_i. \tag{3}$$

As it pertains to education literature, equation (3) represents a typical "education production function".[3] In this case a student's test score T_i is regressed on a vector of predetermined circumstances at a student level C_i. I include the following circumstance based on the education production function literature and their comparable availability in the three waves of the PISA dataset: gender, whether a student is a native born, family background (parental education, average index of social economics status), and school location characteristics (whether it is located in a rural area with population of less than 10,000 or urban and average social economic status of students). The last term, ε_i, represents the error term. The coefficient estimates of β in equation (3) capture the cumulative effect of predetermined circumstances, namely the effects which comes directly from circumstances and indirectly through a student's effort. Thus, it does not represent the causal effect of each particular circumstance on student outcome. The predicted value of equation (3) (i.e. $C_i'\hat{\beta}$) where $\hat{\beta}$ is the vector of estimated coefficients of interest) represents the smoothed distribution of student outcomes, drawing from the assumption that students with similar circumstances have the same conditional average test scores.

As suggested by Ferreira and Gignoux (2011), it is possible to use variance as the inequality index. In this case, the measure of inequality of education opportunity

[3]This linear specification has been widely used by numerous studies such as Woessmann (2003), Fuch and Woessmann (2007), Hanushek et al. (2011), Woessmann et al. (2009), West and Woessmann (2010), and Woessmann (2011).

can be expressed as:

$$IOE = \frac{Var(C_i'\hat{\beta})}{Var(T_i)}. \tag{4}$$

The index from equation (4) is simply the variation of student achievement, explained by these predetermined circumstances. There are numerous advantages to using equation (4) as a measure of IOE. First from an econometric standpoint, the measure is simply the coefficient of determination or an value from a linear regression model.[4]Thus, it represents lower bounds of the actual measure of IOE, owing to the fact that data limitations restrict the number of circumstances that can be used in the estimation process. In comparison to studies such as Schultz et al. (2008) and Woessmann (2004), this measure contains more information about family background effects. Furthermore, unlike studies that choose a coefficient of circumstance (such as parental education or parental income) to represent the measure of IOE, this measure does not require the justification of a single variable to represent all inequality in opportunity. Finally, this IOE measure can also be used to represent a measure of intergenerational persistence as it applies to education, because it can be decomposed by each individual circumstance (or group of circumstances) that strictly relates to family income.

It is important to note that I obtain my IOE measurements based on the average effects of circumstance on student test scores, and also at different levels of score distribution. In other words, in addition to using the Ordinary Least Squares (OLS)

[4]The coefficient of determination:

$$R^2 = 1 - \frac{\sum_i (T_i - C_i'\hat{\beta})}{\sum_i (T_i - \bar{T})}.$$

estimates of $\hat{\beta}$, I also obtain the estimates of $\hat{\beta}$at different levels of the test score distribution based on the quantile regression analysis proposed by Koenker and Hillock (2001) and Koenker and Bassett (1978). The quantile regression of equation (3) can be expressed as:

$$Q_q(Y_i \Big| C_i) = C_i'\beta_q + \mu_{i,q}.$$

Where $q \in (0,1)$ represents the proportion of a population achieving a test score below the quantile level $1-q$. The estimation process is similar to OLS, with the main assumption being that the error term at each quintile,$\mu_{i,q}$, is independently distributed. The difference here is that instead of minimizing the residual sum of squares to obtain coefficient estimates (as would be done in OLS), the quintile regression attempts to minimize the weighted sum of these residuals.

4 Data

The primary data source for my econometric analysis the 2003 to 2009 waves of Programme for International Student Assessment (PISA) dataset. All country level data comes from educational statistics generated by the World Bank, except for my data on income inequality which is drawn from the United Nations' World Income Inequality Database.

4.1 PISA Dataset

The PISA dataset is organized by the Organization for Economic Cooperation and Development (OECD). PISA tests assess students' skills at the age of fifteen (an age at which most children worldwide are approaching the end of their compulsory education). Unlike other international achievement tests, such as the Trends in International Mathematics and Science Study (TIMSS), the PISA assessment does not focus on a specific type of learning curriculum or grade level. The assessment is in the primary subject areas of mathematics, reading, and science. The OECD has administered the survey triennially since 2000. The number of countries participating in the test has grown over the years, with 65 nations taking part in 2009. Similar to other international achievement tests, the PISA survey uses a complex procedure which follows a two-stage stratified sampling protocol. This allows both private and public schools to serve as the primary sampling unit. It then assesses 35 students from each of the selected schools. In each one of the participating countries approximately 150 schools are sampled, drawing from the total number of fifteen-year-old students in school, regardless of their grade level. The student sample size varies across countries and years, in part because some countries fail to meet the targeted sample size while others take more active advantage of the PISA survey to collect data on their own education systems. The PISA survey contains complimentary questionnaires for the selected schools, selected students, as well as for the parents. Compared to other international achievement tests, PISA provides more detailed information about family background, even addressing aspects such as the highest education level of each parent, and a number of home resource considerations. I take advantage of this family information to quantify IOE. PISA provides five plausible values for each subject area. It is important to note that the plausible value rendered is not the actual score of a student on a particular assessment. Plausible values, rather, are random draws from the distribution of scores that could be reasonably assigned to a student with a specific testing pattern. The primary goal of reporting plausible values is to avoid biases caused by students answering only a subset of questions on a particular test. Because students only answer a fraction of all possible questions, these imputation methods are employed by PISA in order to assess how well students would have performed had they answered all the questions. Thus, instead of reporting a single raw test score, a distribution of scores (with associated probabilities) is generated for each individual student.

4.2 Estimation adjustments while using PISA dataset

Although the reported plausible values are neatly standardized such that the mean test score is 500 and the standard deviation is 100, this standardization also complicates parameter estimates. To obtain an unbiased estimate for any analysis using plausible values, the PISA 2009 manual suggests using all five plausible values for each analysis. Thus, the appropriate statistical estimate is the average of these five. This can be represented as:

$$\beta = \frac{1}{5}\sum_{j=1}^{5}\hat{\beta}_j.$$

Where β is the estimated coefficient of interest and $\hat{\beta}_j$ is the estimated parameter obtained using the j^{th} plausible value. A final weight accounts for the fact that a student from a small school is more likely to be sampled.

5 Empirical results

The first stage of the empirical analysis involves computing IOE measures based on the theoretical model described in section (3).[5] In the second stage of the empirical analysis, I provide evidence of the connections between IOE and relatable factors, such as the overall level of economic development, income inequality, and public education spending (as well as other forms of institutional structure).

5.1 IOE estimates

I estimate IOE (using OLS and quantile regression analysis for quantile values 0.2, 0.5, and 0.8) as a variation in student test scores explainable by a set of circumstances, for each country and for each year. Table 1 and 2 displays the country averages of the estimated results of IOE, and demonstrates substantial cross-country heterogeneity in these values. Territories such as Hon-Kong and Panama have family background information explaining approximately 30% and 37% (respectively) of the difference in student scores on average. In the majority of Scandinavian countries, these same predetermined family factors only explained about 17% of the variation in student test scores. And Azerbaijan had the lowest measure of IOE of around 2%.Table 2 summarizes the estimated IOE measure by year. On average, predetermined

[5]For each country I estimate a linear relationship of the mathematics test score and circumstances in the form of a dummy variable for whether a student's mother has completed high school education (mother has at least high school education =1), parental highest level of education, index of cultural possessions) and the school's average of the index of social economic status. To avoid the results being influenced by gender, school location and whether or not a student is a native born I include, a dummy variable for gender dummy (female =1), whether a student is a native born dummy variable (Native=1), whether a student lives in rural area (with population less than or equal to 10000), These variables were chosen based on the education production literature and on their availability in all three waves of PISA used in the empirical analysis.

circumstances explain about 18.3%, 17.4%, and 17.7% of the variation in student test scores for the years 2003, 2006, and 2009 (respectively). Moreover, family circumstances explain nearly 10% of variation in student test scores at each quantile. I have plotted the IOE measure overtime, to assess the trends in IOE for each country. This is reported in Figure 1 for OECD countries.[6] It shows that the OLS method overstates the IOE in comparison to the quantile regression estimates. no clear trend of IOE exists over time, indicating that the computed measure is stable within each country. Countries such as Greece, Hungary, Turkey, and Luxemburg yield an IOE that declines between 2006 and 2009. Another observation is that measured IOE for students at the top of the test score distribution is greater than the IOE measure for the students at the bottom of the test score distribution. This might suggest intergenerational persistence of some factors in educational context might be greatest for students at the top of the score distribution.

5.2 Application 1: How does institutional structure affect IOE?

The other aim of this study is to find sources of international differences in IOE. Since there is no definitive theory for how educational policies and institutional structure influence IOE, I estimate the relationship on the pooled cross-section of countries using the following:

[6]This enables me to focus only on countries that participated in all waves of the PISA program

$$IOE_{q,ct}{=}X_{ct}^{\prime}\alpha_1+\sum_{t}\tau_t\delta_t+\vartheta_{ct}. \tag{5}$$

The dependent variable, $IEO_{q,ct}$, is the imputed IOE at quantile q, in country$c$, and at time$t$. The vector of presumed determinants of IOE;$X_{ct}^{\prime}$, includes measures of the education system such as average education spending per student (as a percentage of GDP per capital), preprimary enrollment rates, average class size (as measured by the average student teacher ratio), average productivity of the labor force (as measured by GDP per capita at constant PPP), and overall inequality (as measured by a country's average GINI coefficient). I denote α_1 as a vector of the parameters of interest. I also include δ_t as a variable to control for time-fixed effects. The error term is denoted with ϑ. Besides the pooled cross-section, I also estimate the relationship between IOE and its determinants by using a panel model with country fixed-effects, such that:

$$IOE_{q,ct}{=}X_{ct}^{\prime}\alpha_1+\textstyle\sum_{t}\tau_t\delta_t+\displaystyle\sum_{c}\theta Z_c+\xi_{ct}. \tag{6}$$

In equation (6) Z_c represents country specific fixed-effects and ξ_{ct}represents an idiosyncratic error term. The estimation results of equation (5) and equation (6) are in Tables 4 -7. The tables are organized such that the dependent variable for each table is a different measure of IOE. The first regression is the baseline model and it demonstrates the relationship between overall economic inequality and the level of IOE which is conditional on economic development, as measured by average income level per capita. The second regression includes the institutional features. The third regression is similar to the second regression, except that it accounts for the possibility of a non-linear effect from preprimary education. Models (1-3) in each table represent pooled cross-sectional regression estimates, and models (4-6) represent the fixed-effect estimates. The results are best summarized by variable.

Financial resources: All Tables 4 -7 show that financial measures such as expenditures per student and resource endowment (as measured by GDP per capita) do not robustly affect IOE. On the other hand, increasing GDP per capita is associated with higher IOE for students with average test scores or whose scores are in the middle of the test score distribution. However, the association disappears when one analyzes the relationship at the top of the test score distribution. These results are consistent with the findings that the level of economic development explains the cross country differences in student achievement on average. As it pertains to the role of education expenditures, there is no clear support that financial resources impact IOE. These results are in line with the literature, which uses aggregated data and finds no clear relationship between student outcomes and education spending (see e.g. Hanushek and Kimko (2000), Hanushek (2003) and Pritchett (2006)).

Income inequality: Both pooled cross-sections and FE estimates demonstrate that income inequality impacts IOE significantly at all quantile levels, except in the case of FE, at the top of the student test score distribution. On average, 1% increase in the GINI coefficient is associated with about a 0.3% increase in IOE at the top, middle, and bottom of the test score distribution. These results suggest that inequalities from parents can translate to unequal opportunities for students from poor families. Of course, these results do not imply causation; it might also be the case that increased IOE translates to increased overall inequality in society. This is an example of a "vicious circle" theory from economic development.

Class size: There is no consensus regarding how class size affects IOE and hence student test scores. One might argue that small classes should produce higher test scores because students can interact more with their teacher. One might also argue that students in larger classes might outperform students in smaller classes because of other externality effects from their peers. From a policy perspective it is interesting to investigate whether class size reduction reduces IOE. For this reason, I included student-teacher ratio in the analysis. Only FE estimates demonstrate that class size impacts the IOE of average scores and the IOE of scores at the bottom of distribution. The results show that increasing class size by one unit increases IOE by about a tenth of a percentage.

Preprimary enrollment rate: Schuetz et al. (2008) emphasize the role of preprimary education in influencing the effect of family background on student achievement. In general, it is ambiguous whether accessibility of preprimary education influences IOE. On one hand, preprimary education can level the playing field of students if it is made accessible to students coming from disadvantaged families. On the other hand, it can increase IOE if the accessibility to preprimary education is dependent on economic status. This can cause the students from advantaged families to attend preprimary education more exclusively, thus exacerbating IOE. Schuetz et al. (2008) provides a theoretical model (and empirical evidence) that shows a non-linear relationship between preprimary education and IOE. More specifically, using cross-sectional data from TIMSS, they find an inverted-U relationship to be present between these factors. In this case, increasing enrollment rates increases student outcomes initially, but eventually the especially high enrollment rates result in lower student outcomes. I include the preprimary enrollment rate in models (2) and (5) of Tables 4-7. To test for the possibility of non-linearity, I also include the square of this in models (3) and (6) of Tables 4-7. Consistent with the findings of Schuetz et al. (2008), my models predict that there is an inverted-U relationship between these variables. However, these results are only statistically significant for the pooled-regression results of IOE at the top of the student score distribution. This suggests that accessibility to preprimary education initially increases IOE, but then eventually lowers IOE overall, most particularly for the students at the top of the test score distribution. The lack of clear evidence from FE estimates makes interpreting results from the cross-sectional (international) dataset unsubstantial.

Robustness check: These results are not robust for the samples from the OECD and non-OECD countries in Tables 8-11and Tables 12-15 respectively. For example, both the pooled cross-section regression and panel data FE estimates from Tables 8-11 show that financial resources do not predict IOE at various student levels of test score distribution. However, it does seem that increased income inequality leads generally to increased IOE.

5.3 Application 2: Is there equity-efficiency tradeoff in education exist?

Relevant to education policy discussions is an evaluation of whether attempts to improve education efficiency (such as efforts to increase student test scores) come at a cost of unintentionally exacerbating the existing inequalities in education. Education policy evaluations also involves analysis of the extent to which policies intended to reduce IOE might in turn decrease overall student achievement. Therefore, I investigate whether there is an equity-efficiency tradeoff in education. I report the pooled correlation between the average test score in mathematics and associated IOE measurements which are conditional on a student's test score in Figure 2.

The first graph of Figure 2 represents the relationship between mean mathematics test scores and the average IOE. The other graphs represent the relationship between mean mathematics test scores and IOE at the 20th, 50th, and 80th quintiles. The scatter plots in Figure 2 show no definitive tradeoff between equality and efficiency in education, but the fitted correlation at the top of student score distribution is suggestive of a possible tradeoff in some cases. One possible explanation of the absence of the equity-efficiency tradeoff is that the plots do not control for features that might influence IOE (such as the level of economic development). I proceed with estimating the equity-efficiency tradeoff through a pooled cross-section model in the form of:

$$\bar{T}_{ct} = \varphi_1 ln(GDPpercapita)_{ct} + \sum_q \varphi_q IOE_{q,ct} + \sum_t \tau_t \delta_t + u_{ct}. \tag{7}$$

Where $\bar{T}_{ct}$denote the average student test score in mathematics for country c in year$t$. The natural log of GDP per capita in constant PPP is included in order to control for the level of economic development. The imputed inequality measure at each quantile $IEO_{q,ct}$, is included to capture and measure the tradeoff, and thus φ_q is the parameter of interest. I also include time dummies to control for time fixed effects that are common to all countries, denoted with δ_t. The last term, u_{ct}, is the error term. Besides reporting results from the pooled cross-section model presented in equation (7), I also employ a panel fixed-effect model expressed as:

$$\bar{T}_{ct} = \varphi_1 ln(GDPpercapita)_{ct} + \sum\nolimits_q \varphi_q IOE_{q,ct} + \sum\nolimits_t \tau_t \delta_t + \sum\nolimits_c \theta_c Z_c + \mu_{ct.} \qquad (8)$$

.Equation (8) differs from equation (7) in that it includes country dummies, Z_c, which control for country specific fixed-effects. I present pooled-regression and panel fixed-effect results in Table 16.

Table 2-16 show results from the pooled regressions (models 1-3) and for the fixed-effect estimates (models 4-6). The first regression in Table 16 is a baseline model for the pooled regression. It includes the logarithm of GDP per capita in order to control for average productivity. The second regression includes the average measure of IOE, and the third regression includes the disaggregated measures of IOE at the top, middle, and bottom of the test score distribution. The estimated results show that, on average, a single percentage increase in GDP per capita increases average mathematics test scores by 56 points, and that the results are statistically significant for all of the pooled regressions.

As for the relationship between IOE and student achievement, the results don't show robust equity-efficiency tradeoff at either levels of the student test score distribution. This is in line with the findings of Woessmann (2004), who used the TIMSS dataset and found no clear evidence of tradeoff between a country's mean test scores and inequality. However, model 3 shows that a percentage increase in IOE at the top of student test score distribution lowers the average math achievement by 6.515 points. An implication from the pooled regressions is that the equity-efficiency tradeoff exists for some groups of students.

The fixed-effect estimates which accounts for the cross country heterogeneity are presented in models (4-6) of Table 16. They show a different outlook of the equity-efficiency tradeoff. First, they show no evidence that initial economic conditions matter for the average student test score. Furthermore, model (6) of Table 2-16 demonstrates that an increase in IOE for a student at the bottom of the score distribution actually improves the average mathematics test scores, and that the results are statistically significant at a 1% level. On the other hand, model (6) also demonstrates that the equity-efficiency tradeoff for IOE exists at the middle of the test score distribution. The results show that at lower levels of the score distribution, high levels of IOE are associated with higher levels of math test scores. In contrast to Woessmann et al. (2008) and Woessmann (2004), which find no clear evidence of the tradeoff on average, the fixed-effect estimates suggest that the tradeoff exists when one considers the inequality measure at different points of the test score distribution, and especially for students at the middle of the distribution.

5.4 Data limitations and future research

This study has several limitations. Data collection issues arising from the fact that extremely poor countries, such as those in sub-Saharan Africa, did not participate in any waves of PISA testing. Therefore, the students sampled were not fully representative of all fifteen-year-old students in the world, and the country-level panel dataset in this study only describes middle and higher-income countries. It would be interesting to analyze education policy factors in developing countries, and to monitor their influence on inequality of opportunity in the education sector.

In addition, data availability restricts the extent of family background information that can be included in approximations for the effect that these factors have on achievement. As a result, the estimates of IOE are lower bounds values, and the regression results used to derive IOE should be interpreted with care. Additionally, there is still no consensus in the literature as to what determines IOE, or how predetermined family background variables affect student achievement. To this end, I have not considered any nonlinearity that might arise while defining the effect of predetermined circumstance on student outcomes.

On a more technical note, the PISA dataset does not include individuals who repeated grades such that they are not in grade 6 or higher by the age of fifteen, nor does it include dropouts. The reported measure should be interpreted only for students who did not leave school or repeat multiple grades. The reported IOE should not be interpreted, therefore, as representing IOE for all 15-year-old students in a country. The provided IOE does, however, usefully demonstrate how the effect of family background information varies across countries. Finally, there is no underlying and established theory behind the determinants of IOE. For this reason, omission of unconsidered variables might worsen the endogeneity problem in the second stage of my estimates. However, the results from fixed-effect estimates (which account for heterogeneity between countries) help to control this endogeneity problem. A notable disadvantage of the fixed-effect analysis is that reliable estimations relating to a country's non-invariant education system characteristics (such as student tracking) cannot be generated. An appropriate response to this issue would be to use a difference-in-difference approach. My focus, however, is on more conscious measures of the education system, and so I reserve such investigation for future works.

6 Concluding remarks

The results from the Program for International Student Assessment (PISA) have triggered serious debate about the efficacy of various educational systems. One important finding is that inequality of education opportunity (IOE), as measured by the effect of a student's family background on test scores, is a demonstrable force in influencing educational outcomes. Numerous hypotheses seek to explain this phenomenon, with most policy makers primarily showing interest in understanding the extent to which a nation's education system affects IOE. To this end, I investigate the role of both education policy and implemented educational systems towards explaining international differences in IOE.

The availability of international microeconomic data on student achievement in the past decade enables a much deeper investigation of cross-country differences in IOE. A few studies have recently used panel data techniques to analyze the role of education systems in influencing IOE at different levels of students' test score distribution. I add to this body of research by examining the specific role of policy with regards to cross-country differences in the presence of IOE.

I use quantile regressions analysis to construct measures of IOE for countries that have participated in the PISA assessment since 2003. I construct the measure of IOE as a relative variation in student test scores that can be explained by predetermined circumstances. The main advantage of this measure is that it allows for simple parametric estimations, and that it makes more substantial use of family background information. This is in contrast to studies such as Woessmann (2004), and Schuetz et al. (2008), which only examine family background according to the value of a single metric. The constructed IOE varies greatly across international borders (and sometimes even within countries) at different points in the students' score distribution. I show that overall inequality in society strongly predicts IOE. Additionally, increasing preprimary enrollment rates also seems to reduce the IOE measure at the top of the test score distributions. One implication for these findings is that improvements in early childhood education can mitigate the deleterious effects of IOE for some students. Additionally there is an equity-efficiency tradeoff in education, especially for students at the middle of the test score distribution. Policies aimed at reducing inequality of education opportunity also risk harming efficiency in other ways, such as by lowering academic achievement for certain groups of students.

One acknowledged shortfall associated with using fixed-effect estimation is that one cannot approximate the effects of non-invariant education system characteristics, such as student tracking in schools. An appropriate method to mitigate this issue would be to use a difference-in-difference approach. Analysis in this vein may prove productive for future projects.

This paper can also be extended in several ways. The provided measure of education inequality may be used to critically analyze the role of IOE on economic growth and earnings. Additionally, one could disaggregate inequality of opportunity in human capital development, and analyze the effects. Finally, with an increased availability of data from developing countries (such as those in sub-Saharan Africa), it is now possible to analyze the roles of aid policy, trade, and financial remittance in explaining IOE. It may even be fruitful to deconstruct IOE by source, and thereby obtain even more accurate estimations of intergenerational persistence in IOE.

Acknowledgment

I am grateful to seminar participants at Kansas State University Department of Economics and my dissertation committee members for helpful comments.

References

Bourguignon, F., F. H. G. Ferreira, and M. Menendez.2007. Inequality of opportunity in Brazil. Review of Income and Wealth 53 (4): 585-618.

Castello, A. and R. Domenech. 2002. Human capital inequality and economic growth: some new evidence. The Economic Journal 112 (478): 187-200.

Checchi, D., and V. Peragine. 2010. Inequality of opportunity in Italy. Journal of Economic Inequality 8 (4): 429-450.

Corak, M. 2013. Income Inequality, Equality of Opportunity, and Intergenerational Mobility. Journal of Economic Perspectives, 27(3): 79-10

Fernando G.L., and F.D. Waltenberg. 2012. Inequality of opportunity for educational achievement in Latin America: Evidence from PISA 2006-2009. Economics of Education Review 31 (5) (OCT 2012): 694-708.

Ferreira, F. and J. Gignoux, 2008. The measurement of inequality of opportunity: theory and an application to Latin America. World Bank Policy Research Working Paper 4659.

———.2010. The measurement of inequality of opportunity: Theory and an application to Latin America. Review of Income and Wealth 57 (4) (DEC 2011): 622-657.

Fleurbaey M., and V. Peragine. 2013. Ex-ante versus ex-post equality of opportunity. Economica 80 (317): 118–130.

Fuchs, T., and L. Woessmann. 2007. What accounts for international differences in student performance? A re-examination using PISA data. Empirical Economics 32 (2-3): 433-464.

Hanushek, E. A., and D. D. Kimko. 2000. Schooling, labor-force quality, and the growth of nations. American Economic Review 90 (5): 1184-1208.

Hanushek E. A., and L Woessmann. 2011. The economics of international differences in educational achievement. Handbook of the Economics of Education 3: 89-200.

Hanushek, E. A., and M.E. Raymond. 2005. Does school accountability lead to improved student performance? Journal of Policy Analysis and Management 24 (2): 297-327.

Koenker, R., and G. Bassett. 1978. Regression quantiles. Econometrica 46 (1): 33-50.

Koenker, R., and K. F. Hallock. 2001. Quantile regression. Journal of Economic Perspectives 15 (4): 143-56.

Lefranc, A., N. Pistolesi, and A. Trannoy. 2008. Inequality of opportunities vs. inequality of outcomes: Are western societies all alike. Review of Income and Wealth. 54(4): 513-546.

Marrero, G.A., and J. G. Rodriguez. 2013. Inequality of opportunity and growth. Journal of Development Economics 104: 107-122. OECD. 2004. "Education and equity-policy brief.

OECD. (2003). PISA 2006: Technical report. Retrieved May 2013 from www.pisa.oecd.org

——— 2009. PISA 2006: Technical report. Retrieved May 2012. URL:www.pisa.oecd.org.

——— 2009a. Manual for the PISA 2009 Database. Paris: OECD.

———.2010. Pisa 2009 results: Overcoming social background-equity in learning opportunities and outcomes. Vol. II. —— .2010a. PISA 2009 results: What students know and can do student performance in reading, mathematics and science (Volume I). URL: http://dx.doi.org/10.1787/9789264091450-en

———. 2010b. Education at a glance: OECD Indicators 2010. Paris. Roemer, J.E,1998. Equality of opportunity. Harvard University Press, Cambridge.

Pritchett L., 2006. Does Learning to add up add up? The returns to schooling in aggregate data. Handbook of the Economics of Education 1: 635–695.

Schuetz, G., H. W. Ursprung, and L. Woessmann. 2008. Education policy and equality of opportunity. Kyklos 61 (2): 279-308. UNU-WIDER. 2008. World Income Inequality Database, Accessed, May 2013. URL:http://www.wider.unu.edu/research/Database/enGB/database/ .

Woessmann, L. 2003. Schooling resources, educational institutions and student performance: The international evidence. Oxford Bulletin of Economics and Statistics 65 (2): 117-170.

Woessmann, L. 2004. How equal are educational opportunities? Family background and student achievement in Europe and the US. IZA DP Paper1284.

Woessmann, L. 2010. Families, schools and primary-school learning: Evidence for Argentina and Colombia in an international perspective. Applied Economics 42 (21): 2645-2665.

World Bank. 2011a. Education Statistics. Washington, DC: World Bank Producer and Distributor. URL: http://databank.worldbank.org

——.2011a. World Development Indicators. Washington, DC: World Bank Producer and Distributor. URL: http://databank.worldbank.org

Woessmann, L. 2003. Schooling resources, educational institutions and student performance: The international evidence. Oxford Bulletin of Economics and Statistics 65 (2): 117-170.

Figure 1: IOE trends for OECD countries from 2003 to 2009

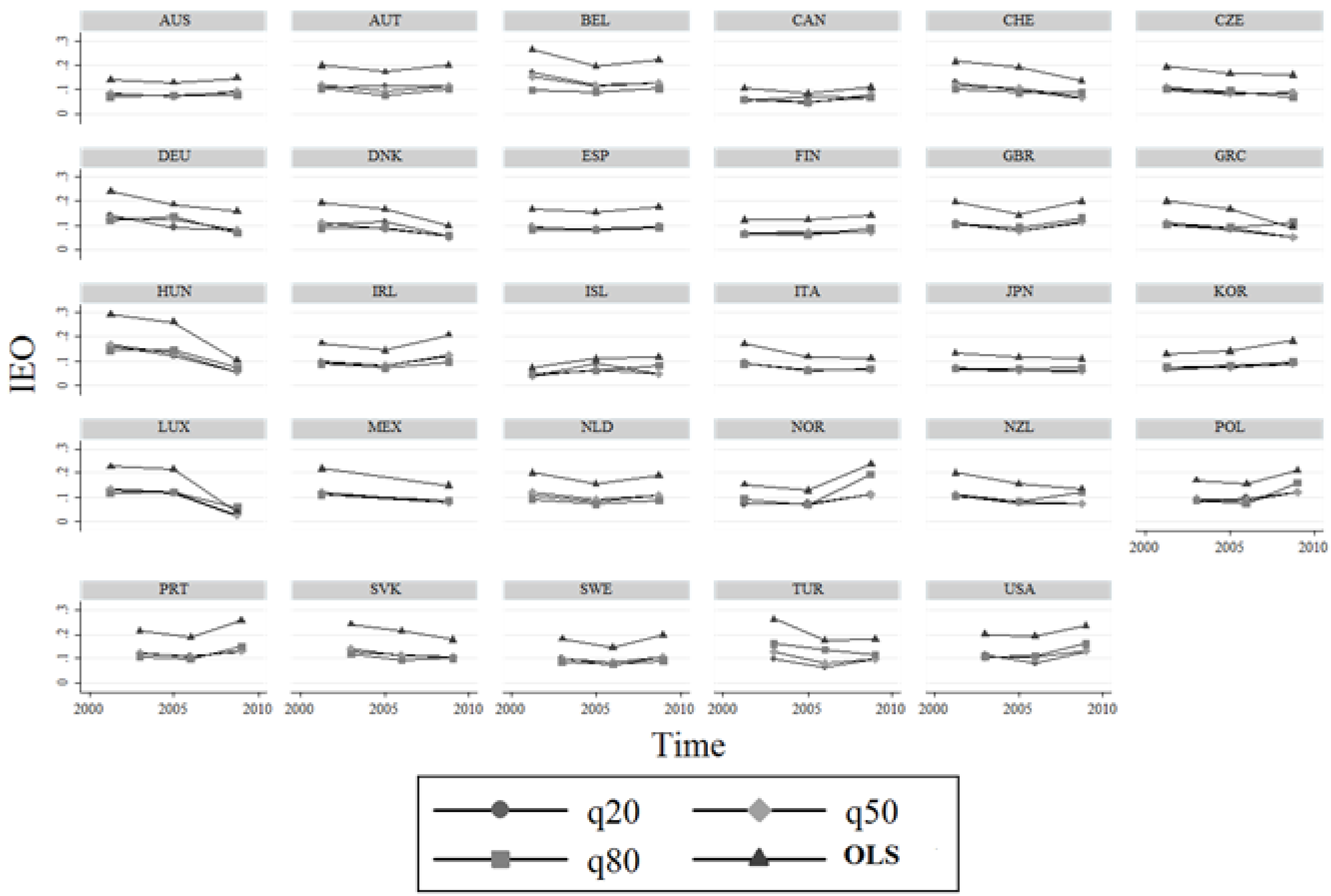

Figure 2: Equity-efficiency tradeoff in education

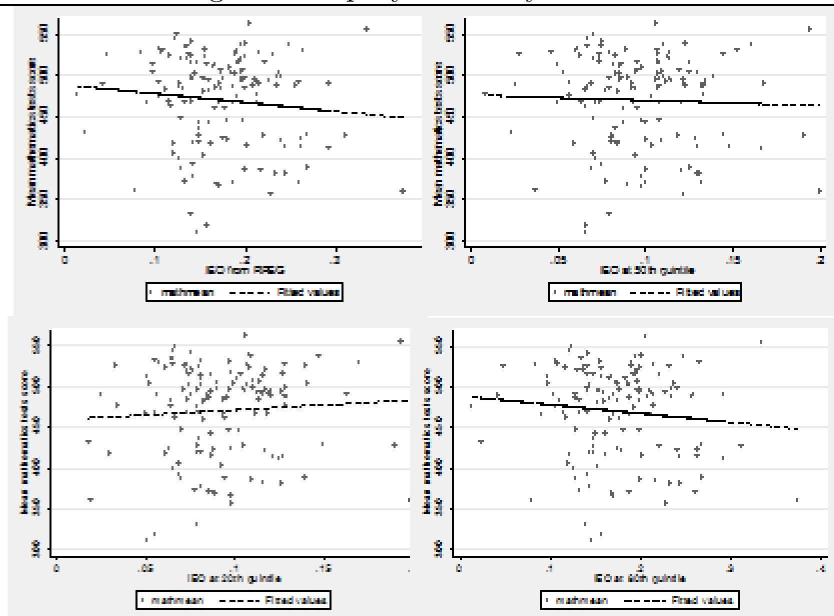

Table 1: Decade averages of IOE as measured by variation in mathematics test scores explained by predetermined circumstances in non-OECD countries.

Non-OECD	IEO-OLS	IEO-q50	IEO-q20	IEO-q80
Argentina	25.150	12.839	12.796	14.913
Azerbaijan	1.933	1.639	2.615	2.496
Brazil	22.993	11.876	9.384	15.697
Bulgaria	26.600	14.688	13.622	15.925
Colombia	21.633	11.54	10.628	12.168
Costa Rica	13.00	7.388	7.264	8.843
Croatia	15.567	8.375	8.724	8.291
Georgia	20.500	12.88	12.74	9.218
Hong-Kong China	33.500	19.48	19.327	11.603
India	11.200	5.546	5.401	7.71
Indonesia	11.878	6.354	5.553	9.326
Jordan	12.100	7.040	6.883	7.158
Kazakhstan	14.633	8.078	7.517	9.851
Kyrgyzstan	14.367	7.290	6.512	8.926
Latvia	14.047	7.481	8.080	7.837
Republic of Moldova	16.35	7.827	7.279	10.155
Romania	17.55	9.29	8.354	10.375
Russian Federation	12.116	6.472	6.456	6.856
Lithuania	20.083	12.080	11.262	9.070
Macao-China	9.300	5.091	5.26	5.729
Malaysia	17.400	10.274	10.106	10.18
Malta	18.500	10.406	10.137	10.685
Mauritius	17.400	9.278	9.142	9.370
Panama	37.500	20.02	19.887	16.100
Peru	17.100	9.83	9.798	11.188
Qatar	16.300	8.978	8.865	8.623
Serbia	14.967	8.559	8.356	8.558
Thailand	14.296	7.581	5.753	11.692
Uruguay	23.791	13.665	12.80	11.854

Note: The IEO is measured as R2*100.

Table 2: Decade averages of IOE as measured by variation in mathematics test scores explained by predetermined circumstances in OECD countries.

OECD	IEO-OLS	IEO-q50	IEO-q20	IEO-q80
Austria	19.044	10.979	11.423	9.491
Belgium	22.807	13.093	13.836	9.794
Chile	26.700	13.868	12.163	16.288
Czech Republic	17.329	9.297	8.873	8.778
Estonia	11.967	6.219	6.256	7.526
Finland	12.989	7.155	6.981	7.252
France	14.100	9.046	9.039	6.753
Greece	15.273	8.233	7.861	10.278
Hungary	21.916	11.995	11.306	12.035
Italy	13.004	7.502	7.524	7.362
Japan	12.002	6.613	6.179	6.96
Mexico	18.33	10.082	9.470	9.941
New Zealand	16.484	9.011	8.606	10.348
Poland	17.971	10.041	10.081	10.73
Portugal	21.96	12.08	11.988	11.865
Slovak Republic	21.069	12.00	11.621	10.564
Slovenia	19.267	10.88	9.906	10.075
Spain	16.54	9.093	9.092	8.599
Sweden	17.500	9.665	9.449	8.496
Turkey	20.669	10.281	8.623	13.717
United Kingdom	17.973	10.131	9.945	10.974
United States	20.933	11.503	10.77	12.747

Note: The IEO is measured as R2*100.

Table 3: Average IEO values for 2003, 2006, and 2009.

Year 2003				
	Mean	Std. Dev	Min	Max
IEO-OLS	18.315	6.341	1.933	37.5
IEO-q50	10.124	3.547	1.639	20.02
IEO-q20	9.61	3.608	1.942	19.887
IEO-q80	10.035	2.809	2.496	16.288
Year 2006				
	Mean	Std. Dev	Min	Max
IEO-OLS	17.368	6.159	1.467	37.5
IEO-q50	9.493	3.475	0.851	20.02
IEO-q20	8.937	3.094	3.305	19.887
IEO-q80	9.8	3.069	3.053	17.85
Year 2009				
	Mean	Std. Dev	Min	Max
IEO-OLS	17.727	6.362	2.4	37.5
IEO-q50	9.832	3.523	2.426	20.02
IEO-q20	9.776	3.57	1.811	19.887
IEO-q80	10.221	3.136	1.256	17.025

Note: The IEO is measured as R2*100.

Table 4: The determinants of IEO on average.

	Pooled-OLS			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP per capita)	2.167**	2.208*	1.683	5.384**	4.112*	4.687*
	(0.920)	(1.099)	(1.134)	(2.151)	(2.340)	(2.397)
GINI	0.251***	0.310***	0.252**	0.188***	0.197**	0.197**
	(0.093)	(0.104)	(0.112)	(0.069)	(0.082)	(0.083)
ln(Expenditure)		3.178	3.168		3.944	3.815
		(2.660)	(2.625)		(2.713)	(2.690)
Class-Size		-0.01	-0.009		0.007	0.009
		(0.011)	(0.011)		(0.006)	(0.007)
Preprimary enrollment rate		-0.014	0.638		-0.087	0.51
		(0.137)	(0.474)		(0.143)	(0.583)
Preprimary enrollment rate squared			-0.016			-0.016
			(0.010)			(0.015)
N	162	141	141	162	141	141
R2	0.143	0.174	0.197	0.084	0.108	0.116
F	4.455	2.931	2.357	4.072	1.764	1.696

The dependent variable is IEO-OLS. Robust standard errors are in parenthesis superscript (*; **; ***) indicates significance at the (10, 5, 1) % level.

Table 5: The determinants of IEO at the middle of the test score distribution.

	Pooled-OLS			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP per capita)	1.263**	1.314*	1.028	3.198**	3.136**	3.464**
	(0.546)	(0.653)	(0.681)	(1.243)	(1.395)	(1.510)
GINI	0.123**	0.160***	0.129**	0.0969**	0.105**	0.106**
	(0.051)	(0.056)	(0.062)	(0.039)	(0.049)	(0.049)
ln(Expenditure)		1.444	1.438		1.191	1.117
		(1.498)	(1.478)		(1.628)	(1.633)
Class-Size		-0.007	-0.006		0.003	0.005
		(0.006)	(0.006)		(0.004)	(0.004)
Preprimary enrollment rate		-0.024**	0.331*		0.002*	0.343*
		(0.077)	(0.273)		(0.087)	(0.393)
Preprimary enrollment rate squared			0.009*			0.009*
			(0.006)			(0.009)
N	162	141	141	162	141	141
R2	0.126	0.162	0.184	0.079	0.101	0.108
F	3.976	2.898	2.49	3.454	2.288	2.06

The dependent variable is IEO-q50. Robust standard errors are in parenthesis superscript (*; **; ***) indicates significance at the (10, 5, 1) % level.

Table 6: The determinants of IOE at the bottom of the test score distribution.

	Pooled-OLS			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP per capita)	1.235**	1.424**	1.187*	2.703*	3.069**	3.129**
	(0.545)	(0.652)	(0.685)	(1.371)	(1.338)	(1.438)
GINI	0.101*	0.137**	0.111*	0.103***	0.111***	0.111***
	(0.052)	(0.059)	(0.063)	(0.037)	(0.039)	(0.040)
ln(Expenditure)		1.316	1.312		1.019	1.006
		(1.384)	(1.372)		(1.406)	(1.427)
Class-size		-0.010*	-0.010*		0.001	0.001
		(0.006)	(0.006)		(0.004)	(0.004)
Preprimary enrollment rate		-0.027	0.267*		0.053	0.114
		(0.070)	(0.244)		(0.050)	(0.298)
Preprimary enrollment rate squared			-0.007*			-0.002
			(0.005)			(0.007)
N	162	141	141	162	141	141
R2	0.110	0.164	0.179	0.088	0.123	0.123
F	4.826	3.919	3.119	5.827	3.309	2.866

The dependent variable is IEO-q20. Robust standard errors are in parenthesis superscript (*; **; ***) indicates significance at the (10, 5, 1) % level.

Table 7: The determinants of IOE at the Top of the Test score distribution

	Pooled-OLS				FE	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP per capita)	0.649*	0.535	0.164	1.048	(0.935)	(0.659)
	(0.381)	(0.476)	(0.506)	(1.046)	(1.224)	(1.320)
GINI	0.150***	0.162***	0.121***	0.060	0.045	0.045
	(0.036)	(0.040)	(0.042)	(0.038)	(0.045)	(0.045)
ln(Expenditure)		1.281	1.274		1.415	1.353
		(1.276)	(1.251)		(1.340)	(1.336)
Class-size		-0.003	-0.002		0.000	0.001
		(0.005)	(0.005)		(0.003)	(0.004)
Preprimary enrollment rate		0.023	0.483**		-0.125	0.162
		(0.067)	(0.218)		(0.082)	(0.280)
Preprimary enrollment rate squared			-0.011**			-0.008
			(0.005)			(0.007)
N	162	141	141	162	141	141
R2	0.181	0.192	0.242	0.023	0.062	0.068
F	6.481	4.028	3.361	1.12	1.386	2.278

The dependent variable is IEO-q80. Robust standard errors are in parenthesis superscript (*; **; ***) indicates significance at the (10, 5, 1) % level.

Table 8: The determinants of IOE on average in OECD countries.

	Pooled-OLS			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP per capita)	1.679	2.523	3.531	4.522	7.57	17.42
	(1.423)	(2.608)	(2.772)	(13.560)	(13.300)	(13.450)
GINI	0.055	0.176*	0.109	0.623*	0.334	0.343
	(0.092)	(0.100)	(0.096)	(0.347)	(0.338)	(0.322)
ln(Expenditure)		2.483	3.584		18.560	12.690
		(5.290)	(5.014)		(22.870)	(19.740)
Class-Size		0.022	0.026		-0.006	-0.025
		(0.019)	(0.020)		(0.021)	(0.021)
Preprimary enrollment rate		-0.250	1.539*		-0.542	-4.650**
		(0.162)	(0.858)		(0.352)	(1.971)
Preprimary enrollment rate square			-0.0461**			0.0943**
			(0.021)			(0.045)
N	62	54	54	62	54	54
R2	0.173	0.213	0.308	0.305	0.374	0.439
F	12.33	12.04	14.87	7.173	9.389	5.565

The dependent variable is IEO-OLS. Robust standard errors are in parenthesis superscript (*; **; ***) indicates significance at the (10, 5, 1) % level.

Table 9: The determinants of IOE on the middle of the test score distribution in OECD countries.

	Pooled-OLS				FE	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP per capita)	-0.483	-1.298	-1.906	3.559	4.721	11.07
	(0.785)	(1.481)	(1.548)	(7.896)	(7.828)	(7.459)
GINI	0.003	0.066	0.026	0.315	0.163	0.169
	(0.049)	(0.064)	(0.059)	(0.186)	(0.187)	(0.174)
ln(Expenditure)		2.477	3.142		11.070	7.292
		(2.947)	(2.765)		(13.190)	(11.230)
Class-size		0.010	0.012		-0.006	-0.019
		(0.012)	(0.012)		(0.013)	(0.014)
Preprimary enrollment rate		0.0940	0.985**		-0.252	-2.900**
		(0.098)	(0.458)		(0.210)	(1.117)
Preprimary enrollment rate Squared			0.028**			0.0608**
			(0.011)			(0.025)
N	62	54	54	62	54	54
R2	0.146	0.178	0.291	0.292	0.383	0.471
F	12.47	14.04	14.78	6.433	8.042	7.083

The dependent variable is IOE-q50. Robust standard errors are in parenthesis superscript (*; **; ***) indicates significance at the (10, 5, 1) % level.

Table 10: The determinants of IEO at the bottom of the test score distribution in OECD countries.

	Pooled-OLS				FE	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP per capita)	-0.271	-1.036	-1.695	7.59	5.124	11.14
	(0.763)	(1.317)	(1.298)	(7.048)	(7.804)	(7.526)
GINI	-0.017	0.049	0.005	0.222	0.083	0.089
	(0.049)	(0.068)	(0.065)	(0.180)	(0.176)	(0.167)
ln(Expenditure)		3.186	3.906		5.902	2.321
		(2.834)	(2.517)		(14.860)	(12.730)
Class-Size		0.005	0.007		-0.008	-0.02
		(0.011)	(0.012)		(0.011)	(0.013)
Preprimary enrollment		-0.079	1.089**		-0.124	-2.631**
		(0.106)	(0.428)		(0.233)	(1.106)
Preprimary enrollment rate squared			-0.0301***			0.0576**
			(0.011)			(0.025)
N	62	54	54	62	54	54
R2	0.118	0.143	0.271	0.263	0.266	0.35
F	7.155	3.232	3.726	5.371	3.652	2.795

The dependent variable is IOE-q20. Robust standard errors are in parenthesis
superscript (*; **; ***) indicates significance at the (10, 5, 1) % level.

Table 11: The determinants of IOE at the top of the test score distribution in OECD countries.

	Pooled-OLS			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP per capita)	-0.719	-0.920	-1.535	3.042	8.207	13.760
	(1.002)	(1.539)	(1.469)	(12.510)	(13.250)	(13.310)
GINI	0.113*	0.164***	0.123***	0.255	0.103	0.108
	(0.057)	(0.047)	(0.036)	(0.246)	(0.279)	(0.258)
ln(Expenditure)		-1.022	-0.350		9.644	6.340
		(2.940)	(2.606)		(10.550)	(9.875)
Class-size		0.012	0.014		-0.007	-0.017*
		(0.009)	(0.010)		(0.011)	(0.009)
Preprimary enrollment rate		-0.157	0.935**		-0.448*	-2.76**
		(0.092)	(0.355)		(0.230)	(1.249)
Preprimary enrollment rate squared			-0.028***			0.0531*
			(0.008)			(0.027)
N	62	54	54	62	54	54
R2	0.161	0.292	0.399	0.076	0.275	0.342
F	4.279	14.05	22.54	2.217	9.338	6.2

The dependent variable is IOE-q80. Robust standard errors are in parenthesis superscript (*; **; ***) indicates significance at the (10, 5, 1) % level.

Table 12: The determinants of IEO on average in non-OECD countries.

	Pooled-OLS				FE	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP per capita)	2.954	3.786	-3.531	4.522	7.57	17.42
	(1.973)	(2.547)	(2.772)	(13.560)	(13.300)	(13.450)
GINI	0.313***	0.334***	0.109	0.623*	0.334	0.343
	(0.090)	(0.111)	(0.096)	(0.347)	(0.338)	(0.322)
Ln(Expenditure)		5.326	3.584		18.560	12.690
		(3.818)	(5.014)		(22.870)	(19.740)
Class-size		-0.037	0.026		-0.006	-0.025
		(0.025)	(0.020)		(0.021)	(0.021)
Preprimary enrollment rate		0.233	1.539*		-0.542	-4.650**
		(0.223)	(0.858)		(0.352)	(1.971)
Preprimary enrollment rate squared			0.0461**			0.0943**
			(0.021)			(0.045)
N	64	56	54	62	54	54
R2	0.229	0.309	0.308	0.305	0.374	0.439
F	5.433	4.827	14.87	7.173	9.389	5.565

The dependent variable is IEO-OLS. Robust standard errors are in parenthesis superscript (*; **; ***) indicates significance at the (10, 5, 1) % level.

Table 13: The determinants of IEO on the middle of the test score distribution in OECD countries.

	Pooled-OLS				FE	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP per capita)	1.813	2.337	-1.906	3.559	4.721	11.07
	(1.175)	(1.545)	(1.548)	(7.896)	(7.828)	(7.459)
GINI	0.158***	0.174***	0.026	0.315	0.163	0.169
	(0.050)	(0.061)	(0.059)	(0.186)	(0.187)	(0.174)
ln(Expenditure)		2.483	3.142		11.070	7.292
		(2.207)	(2.765)		(13.190)	(11.230)
Class-size		-0.022	0.012		-0.006	-0.019
		(0.015)	(0.012)		(0.013)	(0.014)
Preprimary enrollment rate		0.107	0.985**		-0.252	-2.900**
		(0.131)	(0.458)		(0.210)	(1.117)
Preprimary enrollment rate squared			0.0278**			0.0608**
			(0.011)			(0.025)
N	64	56	54	62	54	54
R2	0.221	0.288	0.291	0.292	0.383	0.471
F	5.445	4.33	14.78	6.433	8.042	7.083

The dependent variable is IOE-q50. Robust standard errors are in parenthesis superscript (*; **; ***) indicates significance at the (10, 5, 1) % level.

Table 14: The determinants of IEO at the bottom of the test score distribution in non-OECD countries.

	Pooled-OLS				FE	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP per capita)	1.955	2.557	(1.695)	7.590	5.124	11.140
	(1.157)	(1.538)	(1.298)	(7.048)	(7.804)	(7.526)
GINI	0.127**	0.139**	0.005	0.222	0.083	0.089
	(0.049)	(0.057)	(0.065)	(0.180)	(0.176)	(0.167)
ln(Expenditure)		2.181	3.906		5.902	2.321
		(2.095)	(2.517)		(14.860)	(12.730)
Class-size		(0.020)	0.007		(0.008)	(0.020)
		(0.017)	(0.012)		(0.011)	(0.013)
Preprimary enrollment rate		0.084	1.089**		(0.124)	-2.631**
		(0.120)	(0.428)		(0.233)	(1.106)
Preprimary enrollment rate squared			-0.0301***			0.0576**
			(0.011)			(0.025)
N	64	56	54	62	54	54
R2	0.261	0.322	0.271	0.263	0.266	0.350
F	7.430	4.938	3.726	5.371	3.652	2.795

The dependent variable is IOE-q20. Robust standard errors are in parenthesis superscript (*; **; ***) indicates significance at the (10, 5, 1) % level.

Table 15: The determinants of IEO at the top of the test score distribution in non-OECD countries

	Pooled-OLS				FE	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP per capita)	0.483	0.275	(1.535)	3.042	8.207	13.760
	(0.761)	(0.863)	(1.469)	(12.510)	(13.250)	(13.310
GINI	0.177***	0.183***	0.123***	0.255	0.103	0.108
	(0.047)	(0.060)	(0.036)	(0.246)	(0.279)	(0.258)
ln(Expenditure)		2.158	(0.350)		9.644	6.340
		(1.676)	(2.606)		(10.550)	(9.875)
Class-size		-0.010	0.014		-0.007	-0.0174*
		(0.007)	(0.010)		(0.011)	(0.009)
Preprimary enrollment		0.103	0.935**		-0.448*	-2.761**
		(0.103)	(0.355)		(0.230)	(1.249)
Preprimary enrollment rate squared			-0.028***			0.0531*
			(0.008)			(0.027)
N	64	56	54	62	54	54
R2	0.207	0.276	0.399	0.076	0.275	0.342
F	4.593	5.464	22.540	2.217	9.338	6.200

The dependent variable is IOE-q80. Robust standard errors are in parenthesis superscript (*; **; ***) indicates significance at the (10, 5, 1) % level.

Table 16: The equity-efficiency tradeoff

	Pooled-OLS			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(GDP per capita)	56.00***	56.74***	51.37***	3.302	3.296	1.559
	(9.352)	(9.373)	(10.280)	(5.290)	(5.300)	(4.984)
IOE-OLS		-1.517*			(0.082)	
		(0.786)			(0.210)	
IOE-q20			4.295			2.778***
			(3.032)			(0.872)
IOE-q50			-1.656			-2.854***
			(3.052)			(0.999)
IOE-q80			-6.515***			0.429
			(2.307)			(0.644)
N	174	174	174	174	174	174
R2	0.516	0.541	0.593	0.004	0.005	0.119
F	35.860	18.790	51.800	0.390	0.255	3.842

The dependent variable is the mean mathematics test score. Robust standard errors are in parenthesis. The superscript (*; **; ***) indicates significance at the (10, 5, 1) % level.