UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI "M.FANNO"

CORSO DI LAUREA IN ECONOMIA

PROVA FINALE

"SOCIO-ECONOMIC BACKGROUND AND EDUCATIONAL PERFORMANCE"

RELATORE:

CH.MO PROF. LORENZO ROCCO

LAUREANDO: ALBERTO ANTONELLO

MATRICOLA N. 1162447

ANNO ACCADEMICO 2019 – 2020

ABSTRACT (Italiano)

A partire dalla definizione di status socio-economico e da una breve introduzione sul tema oggetto della tesi, descrivo, in sintesi, il percorso che, a partire dalla prima metà del secolo scorso, la letteratura economica ha seguito per giungere, finalmente, in una delle sue propaggini moderne, allo studio della correlazione inter-generazionale tra background socio-economico e performance scolastica.

Espongo le principali complessità che caratterizzano il dataset impiegato per l'analisi empirica (PISA 2018), tra cui la scala di misura impiegata (Rasch model) e la pesatura delle osservazioni. Delineo, inoltre, la metodologia impiegata per ricavare l'indice socio-economico utilizzato nello studio empirico: trattasi di un'analisi delle componenti principali realizzata a partire da indicatori di educazione, occupazione e patrimonio familiari.

In una prima analisi in cui utilizzo tutte le osservazioni senza distinguere per paese di origine, trovo evidenza di una forte relazione positiva tra l'indice socio-economico e la performance nel test PISA. Essa può essere stimata in un incremento, nel punteggio nel test, di un terzo della sua deviazione standard per ogni incremento nell'indice socio-economico pari a una deviazione standard, a parità di altre condizioni e una volta posti in essere i controlli, opportunamente scelti.

In una seconda fase dell'analisi, trovo evidenza di un diverso impatto dell'indice socio-economico sulla performance scolastica a seconda del paese in cui questo viene stimato. Ciò è sintomo di un diverso grado di disuguaglianza nelle opportunità scolastiche. Raccolgo, quindi, dati a livello di paese che possano potenzialmente spiegare tale variazione. Soltanto il PIL pro-capite e il suo quadrato risultano (statisticamente) utili a spiegare differenze tra paesi nella disuguaglianza di opportunità scolastiche.

Infine, trovo evidenza di un trade-off tra performance scolastica media (a livello di paese) e disuguaglianza di opportunità scolastiche. È, quest'ultimo, un risultato particolarmente rilevante per i suoi risvolti di policy e richiede ulteriore ricerca.

ABSTRACT

I use the PISA 2018 student-level achievement database, recently made available, to estimate the impact of socio-economic background on students' school performance. Family background is found to exert a strong positive effect on educational achievement, which is estimated in an increase of one third of a standard deviation in PISA test performance for each one-standard deviation increase in student socio-economic index. I use the estimated cross-country variation in family background effect as a measure of inequality of educational opportunities and search for cross-country differences in institutions that may explain such variation. Further analysis provides evidence of a potential trade-off between equality of educational opportunities and efficiency in education production.

1. INTRODUCTION

Political philosophy has studied the drivers of equalization of opportunities since at least the second half of the twentieth century (Rawls (1971), Dworkin (1981a; 1981b), Roemer (1998)). A correlation between socio-economic status (SES) and educational performance among children - which is at the core of this work - is an early signal for inequality of opportunities: indeed, an extensive economic literature finds a strong correlation between school performance during early childhood and adult income (cf. Currie and Thomas (1999)). That is, the economic literature on intergenerational earnings mobility (cf. Solon (1999)) is deeply linked to the literature on intergenerational mobility in education.

In the framework set by Becker and Tomes (1986), parents make choices (invest in their children's human capital, spend time in helping them doing homework, teach them values, etc.) that arguably affect the outcomes of all the members of the family (the timing of these choices has also been found to be crucial). Parental choices often depend on parents' education and socio-economic situation. Moreover, children share genes with their parents and, thus, they can inherit at least part of their skills. Empirical evidence is unanimous in backing this theoretical frame of reference. Björklund and Salvanes (2011) state:

> *In every society for which we have data, people's educational achievement is positively correlated with their parents' education or with other indicators of their parents' socio-economic status*

Nonetheless, there is no consensus on the magnitude of this phenomenon.

As stressed by Haveman and Wolfe (1995), sociology and psychology provide a complementary perspective, emphasizing parental role models, the impact of household shocks such as divorce and the life-span influence of family on each individual's development. In this framework, children will later make their own decisions regarding education and entrance in the job-market. Still, parents' education remains the most fundamental determinant of children's success in school.

This topic has significant policy implications: the extent to which family and neighborhood environment impact school performance - which is a measure of (in)equality of opportunities, as previously stated - is not a given. Many characteristics of the school system (e.g. a different organizational structure) can increase or reduce the impact of background factors on performance. An example will make the point clearer: early tracking (cf. Hanushek and Wößmann (2006); Brunello and Checchi (2007)), i.e. an early channeling of children into different schooling paths (e.g. one oriented to blue-collar jobs and one to white-collar jobs), could increase intergenerational correlation in earnings as children from poorer households will be probably channeled into a blue-collar schooling path, as white-collar paths are usually longer lasting and more expensive (these children usually face financial constraints, resulting in underinvestment in education). Another implication is that policies targeting parents' education (e.g. a reform extending compulsory schooling) may have additional spillover effects on children which should be taken into account in a cost-benefit analysis.

As highlighted by Black and Devereux (2010), reaching zero intergenerational correlation is not necessarily the optimal goal: because richer households invest more in their children's education, no intergenerational correlation would imply no return to human capital investment. Still, as long as intergenerational education correlation is driven by a higher level of investment and not by genetic differences in ability, public financing can have a role in equalizing opportunities.

It is fundamental to stress that every empirical research studying the impact of socio-economic status on whatever measure of educational attainment has to deal with an issue related to the measurement of SES. The definition of socio-economic status (SES) is unquestioned and dates back to Chapin (1928, p. 99), who defined SES as:

> *The position that an individual or family occupies with reference to the prevailing average of standards of cultural possessions, effective income, material possessions, and participation in group activity in the community (p. 99)*

The Michigan State Department of Education (1971, p. 5) gave another widely accepted definition:

> *Student socioeconomic status is often thought to be a function of three major factors: 1) family income; 2) parents' educational level; and 3) parents' occupation*

Nonetheless, as boldly stressed by White (1982), the literature is anything but unanimous on the variable(s) that should be employed to measure socio-economic status. Along with the *Index of Status Characteristics* and Hollingshead's *Two-Factor Index of Social Position* - probably the best known measures of SES -, parents' income, occupation, years of schooling, books at home and even ethnicity are all widely and alternately used in the literature to proxy SES. This inconsistency may in part explain the variation in the magnitude of the correlation between SES and academic achievement that is found by the literature on the topic (cf. White (1982) for a thorough examination of previous literature).

The choice of a valid SES measure is particularly delicate when using PISA dataset as there is no measure of household income available for a sufficient amount of countries in the dataset. Section 4 thoroughly describes the solution I resorted to in order to address the problem.

The remainder of the study is structured as follows. Section 2 summarizes previous literature on the topic. Section 3 describes the complexities of the dataset. Section 4 sets the empirical setup and section 5 presents the main results. Section 6 concludes.

2. REVIEW OF THE LITERATURE

Social sciences began to study the determinants of economic success in 1920s, focusing on the correlation between parents and children's occupation (see Haveman and Wolfe (1995)).

A branch of that literature evolved into the modern analysis of the impact of children's socio-economic background on educational attainment; this literature is manifold and articulates along many directions.

A first distinction to be made is that between studies focusing on quantitative outcomes (e.g. years of schooling) and studies focusing on qualitative outcomes (e.g. performance in tests). Even though years of schooling have been found to exert a substantial effect on adult earnings, Wößmann (2004) states that qualitative measures are better indicators for future economic opportunities than quantitative measures, on the basis of a review of previous literature (for example, while returns to years of formal schooling were found to decrease with an individual's time spent in the labor market, returns to educational quality and skills - as measured by performance in tests - were found to increase).

A second point of fracture in the literature refers to the explanatory variables employed. Björklund and Salvanes (2011) sum up the factors which may be responsible for the intergenerational correlation between parental SES and children's educational, which were explored by different branches of the literature: parents' education; parents' cognitive abilities transmitted through genes; parenting skills and risk preferences; economic endowments; public resources and investments.

Part of the literature aimed at deepening the knowledge on the *aggregate* impact of these variables on educational attainment. Early seminal papers by James Coleman (Coleman et al. (1966), the renowned *Coleman report*) and Gary Becker (Becker (1964)) found a strong relation between measures of academic achievement and nearly every variable describing household's socio-economic status. Subsequent literature curbed the unconditional reliance on the existence of a massive impact of SES on educational outcomes: White (1982) carries out a meta-analysis of 101 previous studies, to find that results changed dramatically based on unit of analysis (school, class, individual, etc.), dependent variable (IQ, GPA, years of schooling, etc.), age of the students, measures of SES (parents' income, occupation, education, etc.) and many other variables. For instance, he concludes that the socio-economic background is much more relevant in determining educational achievement at the school level than at the individual level, that a bigger impact of family background is found when we use IQ as the dependent variable rather than GPA, and, finally, that income is much more related to schooling outcomes than parents' occupation and education (although the literature provides counter-examples on this point).

At the beginning of the new millennium, international organizations such as the OECD developed new tests to assess students' knowledge and skills in mathematics, science and reading. Most of these programs (e.g., PISA, TIMSS, PIRLS, etc.) come with a student questionnaire that allows providers to gather valuable information on socio-economic background. The new data made available triggered the rise of a new literature employing these fresh qualitative outcomes as the dependent variable. These studies - from which my work draws deeply - stress the role of the correlation between socio-economic background variables and educational performance as a *proxy* for equality in educational opportunities: the higher the value of the regression coefficient estimated in a given country, the more unequal the educational opportunities in that educational system.

Thanks to the availability of many diverse variables describing the environment in which the child grows up and acquires education, these studies were able to separate the impact of family (and neighborhood) background from that of institutions and resource endowments (even though results on resource endowments are not conclusive due to endogeneity problems, which cannot be overcome using the available data).

For example, Fertig and Schmidt (2002) and Fertig (2003) carry out an in-depth analysis of PISA 2000 dataset and employ background factors (whose effect on children's performance is found to be higher than in previous studies) as controls to isolate the impact of the educational system across countries. Findings are that tangible differences in school systems (school conditions) exert an important role in determining differences in performance among OECD countries.

Along with international tests, many countries developed national tests that helped researchers study within-country variation in schooling performance and educational system quality. For instance, since the academic year 2005/2006, INVALSI administers mathematics and reading/grammar tests to the entire population of Italian students enrolled in 2nd, 5th, 8th, 10th and 13th grades (background questionnaires are available only for 5th and 10th grade). While this data helps explore country-level phenomena more in detail (the universal nature of most of national tests, as opposed to the sample nature of international tests, makes estimates drawn from the former even more precise), it cannot shed light on the impact of institutions on students' performance, as institutions (at least formal ones) are usually the same in the entire country. Thus, international tests are useful in that they allow researchers to study cross-country differences in institutions. Schütz et al. (2008) employ this feature of TIMSS dataset to estimate the impact of cross-country institutional differences on equality of educational opportunities. In particular, they look at the role of school tracking, full/half day school, pre-school and public/private investments.

This branch of the literature gave birth to niches which aimed at further develop the knowledge on the correlation between SES and children outcomes. For example, Lindahl et al. (2015) studied the long-term intergenerational persistence of human capital estimating the correlation between grandparents and grandchildren's SES. Björklund and Salvanes (2011) summarizes the parallel literature that employs sibling correlation as a broad measure for family/neighborhood factors, allowing to isolate the fraction of the total variance in educational performance that is attributable to factors shared by siblings. In line with this literature, recent studies have also employed data on MZ and DZ twins in order to separate the impact of "nature" (i.e. genetic factors) from that of "nurture" (i.e. how children are raised by their parents).

Another branch in the literature seeks to estimate the *causal* effect of single factors that were previously used in joint with other factors as a *proxy* for SES; in other words, moving from simple correlation to causality is at the core of these recent studies.

On the one hand, a rich literature developed around the causal influence of household's income and financial endowments on children outcomes, as masterfully summarized in Brooks-Gun and Duncan (1997). Dahl and Lochner (2012) provides a glimpse of the approach used by these studies to search for causal relations: they employed EITC (a U.S. federal program aimed at providing financial support to low- and middle-income households) as an instrument for income and find a stronger effect of income on educational performance variables than that found by OLS and FE approaches (probably due to attenuation bias from measurement errors).

On the other hand, many studies tried to estimate the causal effect of parental education on children's education. There are three main approaches employed by this branch of literature to find a causal relationship: twin studies, adoption studies and instrumental variables studies. Holmlund et al. (2011) provide a thorough summary of previous literature and applies all three approaches to the same dataset (based on Swedish data) to find that, while results are consistent with previous studies when we take into consideration each approach separately, results are not consistent with each other when we compare different approaches (and also, when we compare different datasets, though this variation may be explained by cross-country variation in institutions); in particular, while twin and adoptees studies find that, once controlled for assortative mating, the impact of father's education on children education is higher than that of mother's education, the reverse holds for IV studies. The authors call for differences in remaining biases between different identification strategies as the most likely explanation for those inconsistencies.

3. DATA

I use data from the 2018 wave of the OECD Programme for International Student Assessment (PISA). This data was made available in December 2019. The target population is 15-year-old students. At this age in most OECD countries, students are approaching the end of their compulsory schooling. Furthermore, part of the target population is attending lower secondary school, while the other part is attending upper secondary school. Students are tested on three domains: mathematics, science and reading.

The specific PISA target population in the 2018 wave is defined as all students between 15 years and 1 months old and 16 years and 4 months old at the beginning of the testing period. This causes the students to be enrolled in different grades.

The full dataset contains information on 612,004 students from 80 countries (between 2,000 and 36,000 students per country). The core of this dataset is represented by OECD countries.

A two-stage sampling procedure is used in PISA. First, a sample of schools is selected from a list of all the schools in which 15-year olds are enrolled. Then, a simple random sample of students is drawn from within the selected schools. In the second stage, 35 students per school are drawn. PISA requires a minimal student participation rate of 80% in order to limit the size of the bias due to non-response.

Differences in school size is a relevant phenomenon to consider; for example, schools in urban settings tend to enroll more students than schools in rural settings do. Though, in theory, all schools have the same probability of being drawn, the probability of drawing a certain student differ among schools due to differences in school size. To overcome this problem, schools are actually not drawn with equal probability; on the contrary, they are selected with probabilities proportional to their size (larger schools have higher selection probability than smaller ones). This procedure should guarantee that each student has the same selection probability; however, students' data still has to be weighted due to (1) missampling of some strata of the population,

(2) lack of accuracy in the measurement of school size and (3) adjustments for student non-response.

Also, because students cannot be considered as independent observations (due to the two-stage sampling), a replication method is suggested for calculating unbiased variances. In particular, each student is assigned 80 replicate weights (calculated generating 80 replicate samples) according to a Balanced Repeated Replication (BRR) method, in its Fay's variant (with a deflating factor K of 0.5). The statistic of interest will, thus, be computed on the whole sample and then again on each replicate. The sampling variance will be computed as:

$$\sigma^2_{(\hat{\theta})} = \frac{1}{G(1-K)^2}\sum_{i=1}^{G}(\hat{\theta}_{(i)} - \hat{\theta})^2 = \frac{1}{80(1-0.5)^2}\sum_{i=1}^{G}(\hat{\theta}_{(i)} - \hat{\theta})^2 = \frac{1}{20}\sum_{i=1}^{G}(\hat{\theta}_{(i)} - \hat{\theta})^2$$

where $\hat{\theta}$ is the statistic computed on the whole sample and $\hat{\theta}_{(i)}$ is the statistic computed on the replicate i.

A description of how performance in PISA test is computed is also needed. Performance is not simply computed as the percentage of correct answers: PISA applies the Rasch model, which estimates student's ability based on both correct answers and items' difficulty. This means that final scores are represented by weighted averages of the correct responses to all questions, with the difficulty of the item used as weight. Items' difficulty is calibrated through a complex process that generates a relative scale of difficulties (a continuum of difficulties): in other words, the (relative) difficulty of an item results from the comparison with all the other items, where the share of students who manage to get the item right is considered.

Finally, we must take into consideration that PISA database reports student performance through *plausible values* (PVs). This means that (posterior) distributions of students' latent ability are computed around the reported values (i.e. the actual result in the test, calculated as a Rasch value); then, a series of random values are drawn from the posterior distribution and assigned to the observation. In PISA 2018, 10 plausible values are drawn for each student.

Population statistics are first estimated using each of the 10 PVs. Then, the reported population statistic is the average of the 10 previous estimates:

$$\theta = \frac{1}{M}\sum_{i=1}^{M}\theta_i$$

The uncertainty in the estimation of the latent variable (i.e. students' ability) is computed as:

$$U_M = \frac{1}{M-1}\sum_{i=1}^{M}(\theta_i - \theta)^2 = \frac{1}{9}\sum_{i=1}^{10}(\theta_i - \theta)^2$$

Though other methods (e.g. using only one of the plausible values, or averaging PVs at the student level) give unbiased estimates when computing means, the use of PVs as just described is necessary in that it provides estimates of variances closest to the population value. In particular, the final variance will be computed as:

$$V = \sigma^2_{(\hat{\theta})} + \left(1 + \frac{1}{M}\right) U_M = \sigma^2_{(\hat{\theta})} + 1.1 U_M$$

The PISA data analysis manuals contain a much more detailed description of all the relevant features of the PISA dataset, with examples.

When this procedure is followed in computing statistics, there is no need to generate a complex structure for the error term, as Fuchs and Wößmann (2008) did, where they decomposed the error term into a school-level and a student-level element.

All the background data is drawn either from questionnaires administered to students after the test or from school questionnaires administered to principals in each of the selected schools. The dataset containing student data has been merged with the one containing school data by school ID (CNTSCHID).

Table 1 contains descriptive statistics of all the variables employed in the first regression.

Table 1: Descriptive Statistics

	Mean	Std. Dev.	Min	Max	Observations
TEST SCORES					
Mathematics	445.078	106.771	14.846	895.299	606,627
Reading	446.153	110.886	9.208	888.468	570,684
Science	453.737	104.260	320.945	883.928	606,627
STUDENT CHARACTERISTICS					
ESCS	-0.468	1.544	-7.366	4.459	559,710
Female	0.501	0.500	0	1	612,002
Age	15.809	0.292	15.08	16.33	612,004
Grade	-0.224	0.746	-4	3	608,985
Student non-citizen	0.037	0.189	0	1	593,357
Father non-citizen	0.101	0.302	0	1	585,179
Mother non-citizen	0.01	0.299	0	1	589,395
Other language at home	0.205	0.404	0	1	597,177
SCHOOL CHARACTERISTICS					
Class size	32.221	10.903	13	53	547,272
Student-teacher ratio	17.885	11.639	1	100	510,163
Perceived teacher's interest	0.180	0.953	-2.271	1.825	551,305
Share of girls	0.499	0.139	0	1	505,530
Share of girls squared	0.268	0.145	0	1	505,530
Private school	0.199	0.399	0	1	552,772
Selective school	0.649	0.477	0	1	583,685
Isolated school	0.294	0.455	0	1	612,004
Poor conditions	0.380	0.485	0	1	585,393
Shortage of educational staff	0.089	1.074	-2.589	4.113	583,834

Missing values are a serious concern for researchers analyzing PISA dataset. Given the considerable number of controls used in the regressions, using simple listwise deletion for handling missing values - as in Fertig (2003) - would imply using only little more than 2,000 observations from the entire 612,004 available pool. This generates biases in the estimates of regression coefficients if values are not missing at random. Information on available variables should not be lost, considered that none of the variables used in this work misses more than half of the observations and most of them are recorded for more than 80% of the sample.

Most of missing values at the student-level are due to non-response, while part of missing values at the school-level are due to the fact that some of questions could not be administered in some countries.

Fuchs and Wößmann (2008) employ a multiple imputation approach using measures of gender, age, grade, books at home and GDP per capita as explanatory variables. Dummy variables to control for imputation bias are then introduced.

I employed a more conservative approach to deal with missing values: for those control variables that missed a significant number of observations, missing values were converted into a constant and a dummy variable was generated taking the value of 1 for observations where the control variable witnessed a missing value, and 0 otherwise. The dependent variable was then regressed on the variable with replaced missing values and on the dummy variable.

This approach avoids the loss of information, while the imputation bias is controlled for by the dummy variable. However, standard errors on controls' coefficients may be biased, and, thus, their estimates should be taken with caution.

Table 2 contains descriptive statistics for four additional country-level variables, taken from World Bank's World Development Indicator, which are used in section 5.2 for further analysis of the dataset.

Table 2: Descriptive Statistics

	Mean	Std. Dev.	Min	Max	Observations
GDP per capita	36,800.68	24,691.83	7,586.385	131,908.2	77
Net enrollment, pre-primary	77.917	19.210	18.022	99.842	56
Years of compulsory schooling	10.534	1.879	6	15	73
Expenditure in education as a % of tot. gov. expenditure	13.561	3.421	7.809	23.423	47

4. EMPIRICAL SETUP

To estimate the effect of family-background characteristics on students' performance, and, thus, to quantify inequality of opportunities, I employed the following form:

$$Y_{isc} = \alpha + B_{isc}\delta_1 + X_{isc}\delta_2 + W_{sc}\delta_3 + \epsilon_{isc}$$

where Y_{isc} is the performance in PISA test of student i in school s in country c, B_{isc} is a variable measuring family-background characteristics, X_{isc} is a vector of student-level controls and W_{sc} is a vector of school-level controls. The coefficient vectors α, δ_1, δ_2 and δ_3 are to be estimated.

The specification is run three times, once for each domain in which students are tested: mathematics, reading and science.

As a measure of socio-economic background, I used the index ESCS. Though already present in the dataset, I re-built the index using a Principal Component Analysis (PCA) based on indicators of parental education (PARED), parental occupation (HISEI) and parental home possessions (HOMEPOS), both cultural and not. Unfortunately, as previously anticipated, no measure of household income is available in the dataset for a sufficient amount of countries; household items and possessions have, thus, been used to proxy family wealth.

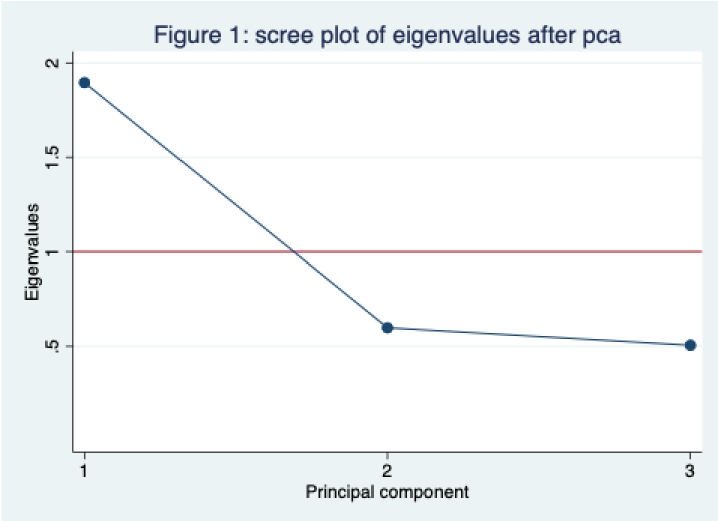
The use of these three heterogeneous measures reflects the manifold and arguably troubled definition of socio-economic background (see section 1). In particular, it is in line with the definition provided by the Michigan State Department of Education (1971).

PCA aims at making a latent measure of socio-economic status emerge from the available measures of parental education, occupation and possessions. The measure of household possessions is itself a composite variable built from many items of the student questionnaire. Noticeably, it incorporates a measure of the number of books at home (*0 to 10*, *11 to 25*, *26 to 100*, *101 to 200*, *201 to 500* and *more than 500*). The relevance of this measure is stronger than it might seem: as emphasized by Wößmann (2004), not only it has the advantage of being more readily comparable across countries than parents' education variables (because educational tracks tend to differ across countries), but it simultaneously represents a specific measure of economic background in that books are goods, and they must be paid for. Books at home is, thus, a thorough measure of socio-economic background; indeed, the power of this measure as a descriptor of family background is witnessed in the literature. According to Schütz et al. (2008):

> *On average, books at home are the single most important predictor of student performance, considerably stronger than parental education*

Of the three principal components generated by PCA, I only kept the one with the highest eigenvalue (1.88) and discarded the other two (respectively, with an eigenvalue of .59 and .52). Figure 1 plots the eigenvalues for all three principal components.

Almost 63% of the variance is explained by the first principal component. Also, the first principal component has positive factor loadings of roughly equal size on all variables (.58 for PARED and HISEI, .56 for HOMEPOS). It can, thus, be interpreted as an overall adequate measure for SES, in which all three components play a similar role in explaining the latent phenomenon.



PCA using the three components just described is the same procedure employed by OECD to build their index, and in fact the correlation between the two indexes is close to .99. The only difference is that the OECD index reports a missing value when two or more of its components are missing, but when only one of the three components is missing, it is imputed using the other two. On the contrary, my index reports a missing value every time one of the three components is missing. Because the three components miss values only for a few observations, this choice only reduces the sample to 559,710, which represent over 90% of the original sample.

The endogeneity of the family socio-economic measure on student performance is guaranteed by the fact that background factors are for the most part determined prior to the test. This is fundamental in interpreting results.

Furthermore - as noted by Wößmann (2004) - a level estimation is required because family characteristics tend to be invariant over time (such that characteristics observed at the time the questionnaire is administered usually reflect those that could be observed in the past); thus, these characteristics impact not only the marginal educational growth just before the PISA test, but they affect children's performance throughout their life. In other words, a student's performance in the test is impacted by family background factors over time in an accumulative way.

Self-selection should not be a problem here because 15-year olds are usually enrolled in compulsory schooling. School tracking may bias results, but its impact should not be overestimated, given the young age of students.

PISA dataset contains a wide array of background data collected through student and parents questionnaires. School principals were also interviewed to collect school-level information.

As emphasized in Fertig and Schmidt (2002), some of the information collected in a set of individual questions - such as students' attitudes toward visiting school, study method, reading pleasure, etc. - can be considered as exogenously determined outcomes. In other words, they are likely determined by the same factors that determine test performance. Furthermore, there is arguably an inverse causality issue, as school performance may in turn impact these outcomes. For example, those who score high in reading may also perceive reading more as a pleasure. Similarly, those who score low in tests may perceive school as more stressful. For these reasons, this set of questions will be disregarded in the regression.

Controls employed in the regression can be divided into student-level and school-level controls. Some student characteristics may alter the impact of socio-economic background on student's inclination towards school and, in consequence, their performance. In particular, parents may change their attitude towards their children depending on their age, gender and school level. Moreover, parental influence may vary based on culture, and, thus, immigration status.

School-level controls are also fundamental, not only because they can be directly correlated to one or more of the variables from which the ESCS index was generated (parents' education, occupation and wealth), but also because they represent the "tangible aspects of the institutional arrangements" (Fertig 2003) and, as such, they will affect the extent to which background features impact students' performance, though the effect of the organization of the educational system cannot be clearly predicted in its extent and direction. Literature is not always consistent on this point.

The purpose of each single control will be discussed in detail in section 5.

It is fundamental to stress that the coefficients on the controls estimated in the regression must be interpreted cautiously, not only due to the particular approach employed to deal with missing values (see section 3), but also because school-level variables may be correlated with omitted variables measuring country cultural features.

The assumption I make is that, once individual- and school-level controls are introduced, no factor that enter the error term remains significantly correlated with the index of socio-economic background (ESCS) and ordinary least-squared approach yields an unbiased estimate of its regression coefficient.

The specification described so far is based on the assumption that the effect of background characteristics on students' performance is the same in all countries. Indeed, this assumption relies on the data-generating process being is the same in all countries: this feature would imply cross-country comparability in family background effect. However, because countries differ in their formal and informal institutions (i.e. school systems work differently), the impact of background factors on students' performance in test, and, thus, the inequality of educational opportunities, might differ across countries.

In the second specification, I test whether the impact of socio-economic background on test performance differs across countries. I put the hypothesis to the test by interacting ESCS (see above) with country dummies.

For each country in the dataset, the size of the impact of background characteristics on PISA test performance will be estimated by the sum of the coefficient on ESCS and the interaction coefficient. A country with a low interaction coefficient is a country where the impact of background characteristics on performance (a proxy for inequality of educational opportunities) is lower than the average of the countries in the dataset, and vice versa.

It is of interest to observe potential country-level institutional features that might explain these differences. In section 5, I explore the explanatory power of GDP per capita, enrollment rate in pre-primary school, length of compulsory schooling and governmental expenditure in education. First, I employ a graphical approach using scatter plots; then, I directly regress interaction coefficients on country institutional variables. Country-level data on these indicators are taken from the World Bank's World Development Indicators (WDI), which contain information on most of the countries in the PISA 2018 dataset.

As a technical note, in computing regression coefficients, I use the STATA package "Repest". This package was designed by the analysts of the OECD to compute regressions and descriptive statistics using replicate weights and plausible values of test scores, thus taking into account the complex design of PISA dataset described in the previous section (Avvisati and Keslair 2020).

5. MAIN RESULTS

5.1 Aggregate results

Table A1 presents the results of the regression using the whole sample of students. As previously stated, the specification relies on the assumption that the effect of background characteristics on students' performance is the same in all countries. In other words, the degree of inequality of educational opportunities is assumed to be the same for all students.

The results start with the index for socio-economic background (ESCS), built with PCA approach where parental education, occupation and wealth play a similar role (loadings are between .56 and .58 for all three components). ESCS is followed by a series of individual-level and school-level controls, whose specific purpose (anticipated in section 4) will be now explained in detail. The regression was repeated three times, changing the dependent variable: performance in all three domains where students are tested - mathematics, reading and science - is regressed on ESCS and controls.

Table A1: Socio-economic background and PISA test performance

	Mathematics	Reading	Science
Constant	592.8***	493.8***	535.1***
	(21.41)	(22.52)	(22.23)
STUDENT CHARACTERISTICS			
ESCS	24.00***	24.35***	22.87***
	(55.53)	(57.39)	(56.46)
Female	-10.15***	16.39***	-5.878***
	(-13.22)	(21.85)	(-8.44)
Age	-6.863***	-1.490	-3.350**
	(-4.00)	(-1.10)	(-2.26)
Grade	27.88***	28.61***	25.37***
	(18.58)	(25.10)	(22.33)
Student non-citizen	-8.676***	-12.71***	-10.29***
	(-3.28)	(-4.70)	(-3.76)
Father non-citizen	3.716**	10.75***	5.496***
	(2.49)	(6.08)	(3.09)
Mother non-citizen	17.91***	23.54***	17.09***
	(10.49)	(12.63)	(8.74)
Other language at home	-34.54***	-48.35***	-37.03***
	(-20.41)	(-32.03)	(-24.71)
SCHOOL CHARACTERISTICS			
Class size	0.244**	0.116	0.293***
	(2.39)	(1.37)	(3.35)
Student-teacher ratio	-1.166***	-0.792***	-1.044***
	(-10.78)	(-8.68)	(-10.15)
Perceived teacher's interest	2.736***	5.353***	3.856***
	(6.19)	(12.73)	(9.91)
Share of girls	46.77***	63.10***	57.72***
	(2.93)	(3.86)	(3.84)
Share of girls squared	-34.79**	-37.56***	-39.24***
	(-2.54)	(-2.73)	(-3.03)
Private school	-3.736	-3.971	-2.101
	(-1.30)	(-1.57)	(-0.84)
Selective school	7.113***	-1.235	4.864***
	(4.04)	(-0.73)	(2.93)
Isolated school	-19.10***	-21.74***	-15.83***
	(-8.69)	(-11.74)	(-8.05)
Poor conditions	-15.20***	-15.53***	-15.13***
	(-7.52)	(-8.27)	(-7.97)
Shortage of educational staff	1.273	0.226	2.063***
	(1.42)	(0.31)	(2.78)
Observations	551,718	517,869	551,718
R^2	0.327	0.354	0.309
Adjusted R^2	0.326	0.354	0.309
F	8,915.6	9,448.1	8,206.0

t statistics in parentheses
Source: OECD 2019
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

ESCS is strongly and positively correlated with performance in all three domains. Given the standard deviations of test performance and ESCS that can be found in Table 1, we can calculate that a one-standard deviation increase in ESCS is associated with one third of a standard deviation increase in performance, regardless of the domain. In other words, the expected gap in performance between students at the top of the distribution as regards socio-economic status and students at the bottom of it is estimated in 270 AP in science, 284 AP in mathematics and 288 AP in reading, once controls are in place. The *t*-test exceeds 50 for all domains.

Previous literature analyzing international tests performance is concerned with setting up an accurate educational production function reflecting theoretical models, using the data available through student and school questionnaires. The focus is on measuring the relative impact of as many of the determinants of student's performance as possible; therefore, not only most of the variables are treated as explanatory variables and not as controls, but the socio-economic index is often decomposed into its determinants, sometimes even distinguishing among paternal and maternal impacts. This makes the interpretation of coefficients not always immediate.

The choice made here is different: the focus is on accurately measuring the overall impact of family socio-economic status by building a reliable index of SES using the available measures of parental education, occupation and wealth and, then, regressing performance on this index and a set of individual- and school-level variables that control for confoundings, which may otherwise bias the estimated impact of SES on test performance. The assumption made here is that, once controls have been put into place and given the rich dataset employed (more than 550,000 students), the estimated impact of family background on student performance can arguably be taken as causal.

The role of all the controls employed in the regression will be now clarified and the regression coefficients on these controls will be briefly discussed, though it must be stressed that the impact of control variables on dependent variables cannot be considered as causal and their magnitude and sign must be taken with caution.

Most of the coefficients on controls are of the sign that can be expected. Girls score lower than boys in both mathematics (about a 10 achievement points difference) and science (about a 6 AP difference), but they score significantly higher than boys in reading test (about a 16 AP difference). This difference is strongly statistically significant in all three domains and it is in line with what previous literature found (Fertig and Schmidt 2002; Fertig 2003; Wößmann 2003; Woessmann 2004; Fuchs and Wößmann 2008). The relative size of these gaps (the gap in reading is positive and greater than the negative gap in mathematics, which is comparable to the gap in science) is also backed by literature (for example, see Fuchs and Wößmann (2008)). As a comparison, the gap in mathematics is close to one third of the expected gap between students in two adjacent grades, while the gap in reading is close to half of that gap.

Age and grade controls are introduced to check for difference in background impact due to different age group and educational levels. As previously noted, all students are aged between 15 years and 1 months and 16 years and 4 months. Once controlled for grade level, age is found to be negatively correlated with performance for all domains, though the difference is not statistically significant in reading. Previous literature is consistent with these results, which is unanimously ascribed to grade repetition. That said, the age gap is relatively small: the expected

difference in performance between the youngest and the oldest student is roughly equal to 8.5 AP, everything else being equal.
Grade is positively correlated with performance and the extent of its impact is relevant. However, the coefficient on grade variable should be considered as an upper bound of the impact of grade on performance, as the grade in which the student is enrolled is likely endogenous to school performance due to grade repetition. Still, I chose to keep grade level in the regression, as previous literature finds that it does not qualitatively affect results (see, for example, Fuchs and Wößmann (2008)).
As emphasized in section 4, immigration status is introduced in the regression as a control for culture, because parental attitude towards children and education differs across cultures. As regards regression coefficients, students born in their country of residence are expected to score roughly 10 AP better than students not born in the country, all other characteristics being equal. This is in line with what previous literature on PISA and TIMSS data found.
Results on parental immigration status differ from previous literature. I find that students whose parents were not born in the country score significantly higher than students whose parents are native. Both Fertig (2003) and Fuchs and Wößmann (2008) find a negative impact of non-native parents. A first explanation could be related to the different controls employed, where, on the one hand, Fertig (2003) introduces a control for second generation students (I do not), and, on the other hand, Fuchs and Wößmann (2003) does not introduce a control for other language at home (I do).
Another plausible explanation is that I use a more recent dataset, where new countries have participated in the test and administering of questionnaires. Indeed, further analysis (results are not shown for the sake of brevity) where I sorted students by country shows that the impact of non-native parents differs quantitatively and qualitatively across countries (in some countries it is positive, while it is negative in others). These findings could be explained through migrants' self-selection models. Further specific research is needed on this point.
As anticipated in the previous section, school-level variables are introduced in the regression to control both for direct correlations with the components of ESCS (parental education, occupation and wealth) and the characteristics of country educational institutions that are reflected in tangible school arrangements.
Class size has a positive and significant impact on performance in mathematics and science, but not in reading. However, the effect is only slightly relevant. Considering that class size varies between 13 and 53 students, the expected gap in science performance (largest regression coefficient) between students who are taught in the smallest and students who are taught in the

biggest class is a mere 11.72 AP in science, all else being equal. This reflects the findings of previous studies.

Similar considerations apply in relation to student-teacher ratio, which is found to have a significant and negative impact on performance in every domain; this is consistent in both sign and magnitude with Fertig (2003). *Ceteris paribus*, an increase of 10 units in the student-teacher ratio is expected to yield roughly a 10 AP decrease in PISA test performance.

A measure of perceived teacher interest is introduced as control for educational staff's attitude towards work and is predictably found to be significantly and positively correlated with performance in all domains, though the effect is small in magnitude.

The social environment where students grow up and are educated may shape children's role models and impact the way and extent to which parental lifestyle and values are transmitted to sons and daughters. A measure for the share of girls in the school where the student is taught is introduced; this measure is extracted from the questionnaire administered to school principals. Following Fertig (2003), who finds a quadratic relationship between test performance and the share of girls in school, I regress performance both on the share of girls and its square. Consistently with the literature, results show a positive impact of the share of girls in school on performance, though the slope is decreasing (the coefficient on the square of the share of girls is negative and significant). These results would imply that schools where both boys and girls are taught provide a better educational environment than schools where students are segregated by gender.

School ownership is notoriously correlated with socio-economic background. While Fertig (2003) does not find any significant impact of school ownership on performance - probably due to the small sample employed - Fertig and Schmidt (2002) and Fuchs and Wößmann (2008) find a significant and negative impact for publicly operated schools. In contrast with these findings, the coefficient on the dummy for private school is negative and not significant in my regression. This seems to be the result of controlling for the complex design of the dataset and, in particular, the choice of schools (instead of students) as unit of randomization: indeed, explorative regressions where plausible values are averaged at the student level and replicate weights are not taken into account yield similar results to those in previous studies. The coefficient turns not-significant and its sign is reverted when plausible values and replicate weights are used as suggested by the design of PISA dataset.

A dummy variable indicating whether schools select students at entry based on their records of academic performance (including placement tests) is introduced to control, at least weakly, for school tracking. The impact of selection is statistically significant and positive in mathematics and science and not-significant in reading. Differences with previous literature, where selection is found to have a statistically significant and positive impact on performance in reading, might be due to smaller sample size or a different sample of countries.

A dummy variable for school location is employed to control for differences between rural and urban environments: the dummy variable takes the value of 0 when the school is located in a town, city or large city and 1 when it is located in a village or small town. Consistently with what previous studies found, students enrolled in schools located in a rural environment are expected to score lower than students enrolled in schools located in an urban setting, all else being equal.

Two dummy variables - one taking the value 1 when school suffers from inadequate or poor quality physical infrastructure (e.g. building, grounds, heating/cooling, etc.) and 0 otherwise, the other taking the value 1 when school suffers from a lack of teaching staff and 0 otherwise - control for local investment in schooling and, consequently, how much the educational system is valued, relative to other expense items, when it comes to financial choices. Poor infrastructural conditions have a predictable negative impact on performance, while shortage of educational staff has the “wrong” sign and is statistically not-significant.

The specification has an explanatory power of between .31 and .35 of the variance in the dependent variable. The unexplained variation in student test performance is plausibly due to unobserved student-level ability differences and/or to unobserved variation in institutions (see section 5.2).

Notice that, when reading performance is the dependent variable, the regression is calculated on fewer observations: reading score contains more missing values than mathematics and science score.

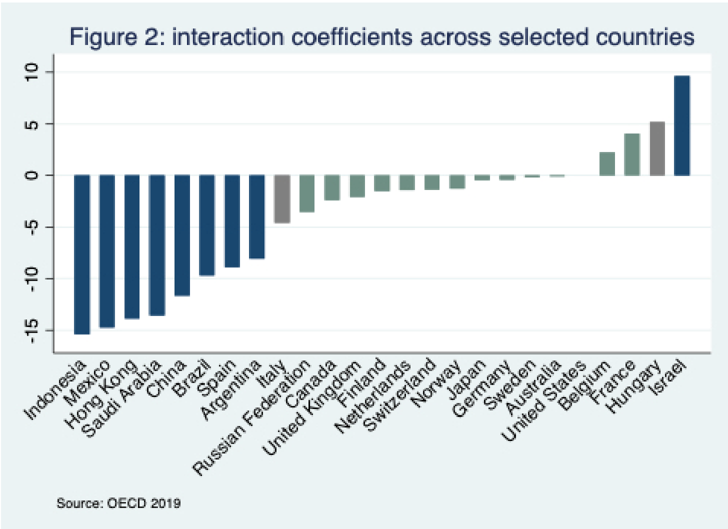
All considered, results tell us that, once controls have been put into place, ESCS - an overall measure of socio-economic background, given the factor loadings found in section 4 - has a significant positive impact on PISA test performance, whose Rasch model scale is purposefully designed to capture student’s school performance and ability. This impact can be estimated in one third of a standard deviation increase in test performance for every one-standard deviation increase in the socio-economic index. This means that, *ceteris paribus*, the expected gap in performance between a student at the top of the distribution as regards socio-economic status and one at the bottom of it is estimated in roughly 280 AP (regardless of the domain).

5.2 Equality of educational opportunities across countries

In this second part of the data analysis, I question the validity of the assumption that the impact of socio-economic background on children’s performance does not depend on the country in which it is estimated.

I interacted the socio-economic index (ESCS) with country dummies, keeping all the controls described in section 5.1 and choosing mathematics as the dependent variable. The size of the impact of background characteristics on PISA test performance for each individual country will be given by the sum of the coefficient on ESCS and the specific interaction coefficient. The result is considered as a cross-country measure of inequality of educational opportunities.

Results from regression are not shown due to difficulties in displaying it. However, Figure 2 pictures interaction coefficient for 25 selected countries through a bar graph. Interaction coefficients are displayed in dark blue when significant at the .01 level, gray when significant at the .05 level (but not at the .01 level) and jade-green when not significant at the .05 level.



The U.S. is taken as the base in the regression, and, thus, its interaction coefficient is null by construction. The coefficient on ESCS is equal to 25.54 (significant at the .01 level), meaning that family background impact on performance in the U.S. is close to the average impact for the entire dataset, as displayed in section 5.1 (the aggregate impact in mathematics is 24.00).

A first result is that 11 out of 79 countries (U.S. is null by assumption) display an interaction term which is statistically significant at the .05 level (but not at the .01 level) and 33 out of 79 countries display an interaction term which is statistically significant at the .01 level. This means that, in 44 out of 79 countries, the impact of SES on school performance is statistically different than in the U.S. (remember that the coefficient for the U.S. is close to the average found in section 5.1). In countries where the interaction coefficient is negative and significant, inequality of educational opportunities can be considered smaller than the average of countries in the dataset, and vice versa. Schütz et al. (2008) study cross-country differences in equality of educational opportunities employing a similar approach. They use TIMSS 1995 and TIMSS-Repeat 1999 instead of PISA 2018 and employ an educational production function-type of regression where the impact of socio-economic background on performance is measured by the regression coefficient on a variable measuring the number of books at home (see section 4).

Still, the correlation between the interaction coefficients I find and the country-specific coefficients on books variable they compute is close to .59. This result provides evidence that the phenomenon under investigation is the same and it subsists.

Variation in interaction coefficients is arguably due to differences in institutions across countries.

Country-level data that proxy for cross-country institutional variation could provide an insight on the determinants of inequality of educational opportunities. World Bank's World Development Indicators have the advantage of providing official data on all the countries where PISA test was administered in 2018; other datasets contain information only on a sub-sample of these countries.

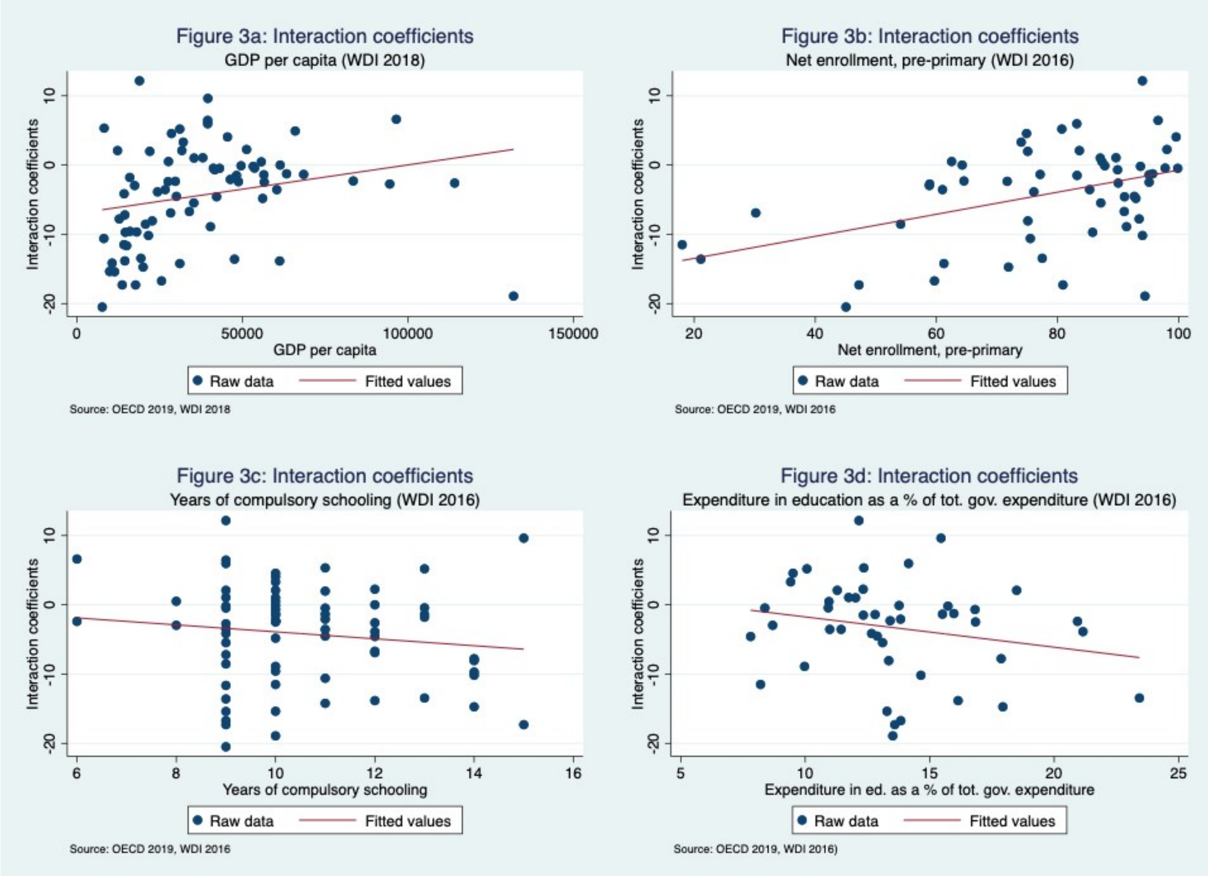
The indicators I used to proxy for institutions are GDP per capita, net enrollment rate in pre-primary school, years of compulsory schooling and expenditure in education as a percentage of total government expenditure. Section 3 contains descriptive statistics on all four indicators. Data on GDP per capita dates to 2018, while all other country-level data dates to 2016: this choice was made because more recent data contains too many missing values to provide relevant results.

Additional research is needed to gather further relevant data on institutional variation across all the countries in PISA dataset; in particular, this is true for information on school tracking and central examinations.

A first graphical analysis is carried out by plotting interaction coefficients and the variables that proxy for institutional features on a simple scatter-plot. Figures 3a to 3d show the results.

GDP per capita is a first obvious variable whose impact must be considered, though there is no obvious reason to predict in which direction the relationship will point and there could exist both direct and indirect channels (for instance, there could exist an indirect effect of increasing GDP per capita on income inequality).

Schütz et al. (2008) find no evidence of a statistically significant relationship between GNI per capita and family background impact on test performance. The relationship found in Figure 3a is weakly positive, but there seems to exist a quadratic relation between the two variables.



Because children's educational performance and attitude towards the learning process prior to school entering is for the most part determined by families, the earlier the school tracking, the more equal educational opportunities should be. The formal model developed in Schütz et al. (2005) emphasizes the equalizing effect of pre-primary education as it exposes children from both poor and rich environments to potentially the same formal education.

However, a second force, contrary to the first one, may drive results: it will be children from well-off socio-economic environment to enroll in pre-primary school first, as families must pay the cost of it. Therefore, as the enrollment rate raises, the average SES of children who remain excluded from pre-primary formal education will become lower and lower. It cannot be *a priori* ascertained which force is mainly driving results. Schütz et al. (2008) find a statistically significant inverted U-shaped pattern in the relationship between family-background effect and enrollment share in pre-primary school, possibly reflecting non-random sorting into early formal education.

WDI do not provide data on the age of first entry into pre-primary school. Still, information on net enrollment share in pre-primary education is available and is employed in Figure 3b. The figure shows only a positive but noisy relationship between the two variables. There is no evidence of an inverted U-shaped pattern. Results seem to be mainly driven by non-random sorting into pre-primary school, where the lowest strata of population remain confined out of early formal education as pre-primary enrollment rate raises.

The length of compulsory exposure to formal schooling should also exert an equalizing effect on educational opportunities in that it forces students from low socio-economic background to receive potentially the same formal education as children from higher SES for a number of years that is considered sufficient to build skills demanded by the labor market. Compulsory schooling is often subsidized by governments and its impact partly mitigates that of early tracking.

Schütz et al. (2008) use no data on length of compulsory schooling, but they interact books-at-home variable with a measure of length of pre-primary school. They find a statistically significant negative effect of a longer pre-school cycle on the measure of family background impact on performance.

WDI provide data on the length of compulsory schooling in each country. The scatter plot (Figure 3c) shows a negative relationship between compulsory schooling variable and interaction coefficient, but the relationship is too weak and noisy to draw any conclusions on this relation.

Finally, government expenditure in education may mitigate the impact of SES on student's educational path as it is often directed at ensuring common standards in the educational system and subsidize the poorest, who cannot afford formal education. Schütz et al. (2008) find no evidence of a relationship between educational expenditure per student and family-background effects. WDI provide information on the expenditure on education as a percentage of total government expenditure. Figure 3d shows a weak negative relationship between the two variables. Again, the scatter plot is very noisy and does not allow decisive conclusions.

As a direct extension of this first graphical analysis, in Table A2, I present the linear regression of interaction coefficients on each of the four country institutional variables.

As suggested by Figure 3a, I regress interaction coefficients both on GDP per capita and its square and I do the same for net enrollment rate in pre-primary school, following Schütz et al. (2008). GDP per capita is measured in thousands of U.S. dollars.

The signs of the relationships are the same highlighted in the scatter plots. However, only GDP per capita and its square are statistically significant. The relationship between interaction coefficients and GDP per capita is found to display an inverted U-shaped relation.

Notice that Schütz et al. (2008) use country-level institutional data directly in the educational production function regression by interacting them with their measure of family-background (books at home). On the contrary, I chose to regress the interaction coefficient extracted from the regression with interactions on country institutional indicators. Further research should put to the test the findings of Schütz et al. (2008) by interacting country variables with different (and, possibly, more accurate) family background indicators, such as ESCS.

Table A2: Interaction coefficients and country features

	(1)	(2)	(3)	(4)
Constant	-14.12***	-16.39**	1.120	2.637
	(-6.74)	(-2.13)	(0.23)	(0.63)
GDP per capita	0.441***			
	(4.93)			
Squared GDP per capita	-0.00336***			
	(-4.40)			
Net enrollment, pre-primary		1.150		
		(0.62)		
Squared net enrollment, pre-primary		0.0000694		
		(0.04)		
Years of compulsory schooling			-0.502	
			(-1.12)	
Expenditure in ed. as a % of tot. gov. expenditure				-0.438
				(-1.45)
Observations	76	56	72	46
R^2	0.2542	0.1875	0.0175	0.0454
Adjusted R^2	0.2337	0.1568	0.0035	0.0237
F	12.44	6.12	1.25	2.09

t statistics in parentheses
Source: OECD 2019, WDI 2016, WDI 2018
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

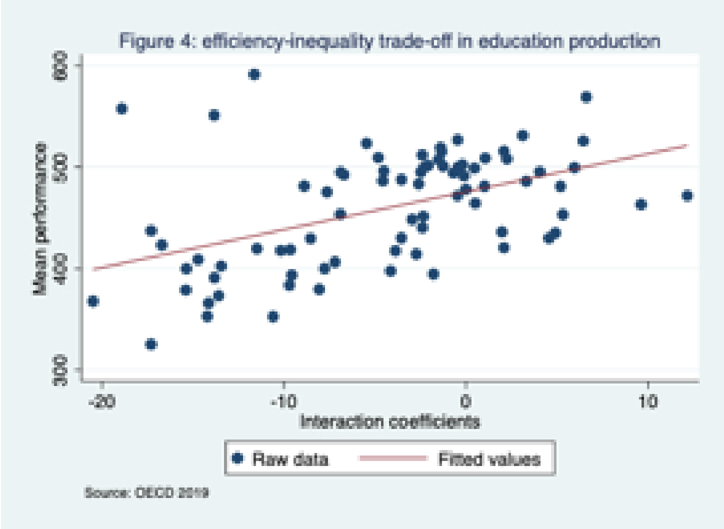
5.3 Equality of educational opportunities and efficiency of education production

Finally, I test whether there exists a trade-off between equality of educational opportunities and efficiency in the production of education; in other words, I test whether equalizing educational opportunities can only be bought at the price of a lower average level of performance. Wößmann (2004) uses the mean of the math achievement in each country as a measure of efficiency, and the performance difference between students with more than two bookcases at home and students with less than one shelf of books at home as a measure of (in)equality. The sample used contains 18 countries. The study finds no evidence of a trade-off between efficiency in education production and equality of education opportunities (the cross-country correlation coefficient between the two measures is equal to .002).

I use the same variable to proxy efficiency and performance as Wößmann (2004) does, but I employ interaction coefficients (see above) as a measure of (in)equality of educational opportunities. Moreover, my sample contains information on 80 countries. Figure 4 below shows results in a scatter plot.

Results are in stark contrast with Wößmann (2004) and show a clear direct positive relationship between mean performance and interaction coefficients. The cross-country correlation coefficient between the two measures is equal to .477. These results suggest the existence of a tradeoff between efficiency in education production and equality of educational opportunities. Still, three outliers - China, Hong Kong and Macao - witness the feasible coexistence of high performance and a relatively high level of equality of opportunities.

These results represent a simple illustrative impression and call for further in-depth research.



6. CONCLUSIONS

In this work, I provided an overview on the topic of the impact of socio-economic status on children's educational performance, giving the relevant definitions to understand the topic. I also provided a brief summary of the literature on the topic.

I described the complexities of the dataset (PISA 2018) employed to carry out an empirical analysis of the determinants of school attainment. Then, I set up the empirical framework.

The first part of the analysis focused on the impact of SES on school performance, regardless of the institutional setting (the entire dataset is employed without sorting observations by country of origin). Results are that, *ceteris paribus*, the expected gap in performance between a student at the top of the distribution as regards socio-economic status and one at the bottom of the distribution is estimated in roughly 280 achievement points (regardless of the domain), when 450 AP (roughly) is the average performance in PISA 2018 test.

In the remainder of the analysis, I questioned the assumption that the impact of SES on children's performance does not depend on the country in which it is estimated, to find that, in 44 out of 79 countries, the assumption does not hold. Then, I tested the power of four country-level institutional variables in explaining the cross-country variation in SES impact on performance. GDP per capita and its square are the only variables that are found to be statistically relevant in explaining the variation.

Finally, I provided descriptive evidence of a potential trade-off between efficiency in the production of education and inequality of educational opportunities, though China represents a noticeable outlier. Further in-depth analysis is required on this point.

REFERENCES

Alivernini, Fabio, and Sara Manganelli. 2015. 'A Multilevel Structural Equation Model Testing the Influences of Socio-Economic Status and Pre-Primary Education on Reading Literacy in Italy'. *Procedia-Social and Behavioral Sciences* 205: 168–172.

Avvisati, Francesco, and François Keslair. 2020. 'REPEST: Stata Module to Run Estimations with Weighted Replicate Samples and Plausible Values'.

Barabanti, Paolo, and Valeria F. Tortora. n.d. 'Uno Studio Sugli Studenti Top Performer Attraverso Le Prove INVALSI 2015'.

Becker, Gary. 1964. 'S. Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education'. *Chicago: University of Press* 19 (6): 4.

Becker, Gary S., and Nigel Tomes. 1986. 'Human Capital and the Rise and Fall of Families'. *Journal of Labor Economics* 4 (3, Part 2): S1–S39.

Björklund, Anders, and Kjell G. Salvanes. 2011. 'Education and Family Background: Mechanisms and Policies'. In *Handbook of the Economics of Education*, 3:201–247. Elsevier.

Black, Sandra E., and Paul J. Devereux. 2010. 'Recent Developments in Intergenerational Mobility'. National Bureau of Economic Research.

Breen, Richard, and John H. Goldthorpe. 1997. 'Explaining Educational Differentials: Towards a Formal Rational Action Theory'. *Rationality and Society* 9 (3): 275–305.

Brooks-Gunn, Jeanne, and Greg J. Duncan. 1997. 'The Effects of Poverty on Children'. *The Future of Children*, 55–71.

Brunello, Giorgio, and Daniele Checchi. 2007. 'Does School Tracking Affect Equality of Opportunity? New International Evidence'. *Economic Policy* 22 (52): 782–861.

Brunello, Giorgio, and Lorenzo Rocco. 2013. 'The Effect of Immigration on the School Performance of Natives: Cross Country Evidence Using PISA Test Scores'. *Economics of Education Review* 32: 234–246.

Chapin, F. Stuart. 1928. 'A Quantitative Scale for Rating the Home and Social Environment of Middle Class Families in an Urban Community: A First Approximation to the Measurement of Socio-Economic Status.' *Journal of Educational Psychology* 19 (2): 99.

Checchi, Daniele, Carlo Fiorio, and Marco Leonardi. 2007. 'Sessanta Anni Di Istruzione Scolastica in Italia'.

Coleman, James S. 1966. *Equality of Educational Opportunity [Summary Report*. Vol. 2. US Department of Health, Education, and Welfare, Office of Education.

Considine, Gillian, and Gianni Zappalà. 2002. 'Factors Influencing the Educational Performance of Students from Disadvantaged Backgrounds'. In *Competing Visions: Refereed Proceedings of the National Social Policy Conference*, 2001:91–107.

Currie, Janet, and Duncan Thomas. 1999. 'Early Test Scores, Socioeconomic Status and Future Outcomes'. National bureau of economic research.

Dahl, Gordon B., and Lance Lochner. 2012. 'The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit'. *American Economic Review* 102 (5): 1927–56.

Dworkin, Ronald. 1981a. 'What Is Equality? Part 1: Equality of Welfare'. *Philosophy & Public Affairs*, 185–246.

———. 1981b. 'What Is Equality? Part 2: Equality of Resources'. *Philosophy & Public Affairs*, 283–345.

Fertig, Michael. 2003. 'Who's to Blame? The Determinants of German Students' Achievement in the PISA 2000 Study'.

Fertig, Michael, and Christoph M. Schmidt. 2002. 'The Role of Background Factors for Reading Literacy: Straight National Scores in the PISA 2000 Study'.

Fuchs, Thomas, and Ludger Wößmann. 2008. ‘What Accounts for International Differences in Student Prformance? A Re-Examination Using PISA Data’. In *The Economics of Education and Training*, 209–240. Springer.

Hanushek, Eric A., and Ludger W ößmann. 2006. ‘Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries’. *The Economic Journal* 116 (510): C63–C76.

Haveman, Robert, and Barbara Wolfe. 1995. ‘The Determinants of Children’s Attainments: A Review of Methods and Findings’. *Journal of Economic Literature* 33 (4): 1829–1878.

Hertz, Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith, and Alina Verashchagina. 2008. ‘The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends’. *The BE Journal of Economic Analysis & Policy* 7 (2).

Holmlund, Helena, Mikael Lindahl, and Erik Plug. 2011. ‘The Causal Effect of Parents’ Schooling on Children’s Schooling: A Comparison of Estimation Methods’. *Journal of Economic Literature* 49 (3): 615–51.

Lindahl, Mikael, M\aarten Palme, Sofia Sandgren Massih, and Anna Sjögren. 2015. ‘Long-Term Intergenerational Persistence of Human Capital an Empirical Analysis of Four Generations’. *Journal of Human Resources* 50 (1): 1–33.

Lundborg, Petter, Anton Nilsson, and Dan-Olof Rooth. 2014. ‘Parental Education and Offspring Outcomes: Evidence from the Swedish Compulsory School Reform’. *American Economic Journal: Applied Economics* 6 (1): 253–78.

Marco Serino, Michele Cardone di, and Emiliano Campodifiori. n.d. ‘7. Spazio Sociale, Capitale Culturale e Performance Scolastica. Un’analisi Multidimensionale Dei Dati Del Questionario Studente INVALSI’. *UNO SGUARDO SULLA SCUOLA*.

Michigan State Department of Education. 1971. ‘Distribution of Educational Performance and Related Factors in Michigan: A Supplement’. ERIC Document Reproduction Service No. ED 059254.

Monseur, Christian. 2009. ‘PISA Data Analysis Manual: Spss Second Edition’.

Montanaro, Pasqualino, and Paolo Sestito. 2014. ‘La Qualità Dell’Istruzione Italiana: Un Confronto Tra PISA E Le Rilevazioni Nazionali Dell’Invalsi (The Quality of Italian Education: A Comparison between the International and the National Assessments)’. *Bank of Italy Occasional Paper*, no. 218.

Rawls, John. 1971. ‘Http://Www. Bookstore. Org/Id1971,’A Theory of Justice’’.

Roemer, John E. 1998. *Theories of Distributive Justice*. Harvard University Press.

Russo, Felice. 2010. ‘Risorse, Fattori Di Contesto Socio-Economici e Esiti Scolastici: Una Applicazione Su Dati Invalsi 2005-06’. In . Atti XXII Conferenza Siep, Pavia.

Schütz, G., H. W. Ursprung, and L. Wößmann. 2005. ‘Education Policy and Equality of Opportunity (CESifo Working Paper No. 1518)’. *Munich, Germany: CESifo*.

Schütz, Gabriela, Heinrich W. Ursprung, and Ludger Wößmann. 2008. ‘Education Policy and Equality of Opportunity’. *Kyklos* 61 (2): 279–308.

Solon, Gary. 1999. ‘Intergenerational Mobility in the Labor Market’. In *Handbook of Labor Economics*, 3:1761–1800. Elsevier.

White, Karl R. 1982. ‘The Relation between Socioeconomic Status and Academic Achievement.’ *Psychological Bulletin* 91 (3): 461.

Woessmann, Ludger. 2004. ‘How Equal Are Educational Opportunities? Family Background and Student Achievement in Europe and the US’.

Wößmann, Ludger. 2003. ‘Schooling Resources, Educational Institutions and Student Performance: The International Evidence’. *Oxford Bulletin of Economics and Statistics* 65 (2): 117–170.