ECONOMIC POLICY

LIETUVOS BANKAS

Economic Policy
Fifty-eighth Panel Meeting
Supported by the Bank of Lithuania
Vilnius, 25-26 October 2013

School and Family Effects on Educational Outcomes across Countries

Richard B. Freeman (Harvard University)
Martina Viarengo (The Graduate Institute, Geneva)

The organisers would like to thank the Bank of Lithuania for their support.
The views expressed in this paper are those of the author(s) and not those of the funding organization(s).

CEPR Centre for Economic Policy Research

CES

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

School and Family Effects on Educational Outcomes across Countries [1]

Richard B. Freeman
Harvard University
Department of Economics and NBER
1050 Massachusetts Avenue
Cambridge, MA 02138

Martina Viarengo
The Graduate Institute
Department of Economics
Rue de Lausanne 132
1202 Geneva

ABSTRACT

This study assesses the link between student test scores and the school students attend, the policies and practices of the schools, students' family background and their parents' involvement in their education using data from the 2009 wave of the Program for International Student Assessment. We find that 1) a substantial proportion of the variation of test scores within countries is associated with the school students attend; 2) national tracking policies which affect sorting of students among schools explain part of the fixed school effects but most of the effects are associated with what schools do; 3) school policies and teaching practices reported by students explain a sizable proportion of school fixed effects but still leave unexplained a substantial part of school effects; 4) school fixed effects are a major pathway for the association between family background and test scores. The implication is that what schools do is important in the level and dispersion of test scores, suggesting the value of more attention to what goes in schools and pinning down causal links between policies and practices and test score outcomes.

Keywords: education, public policy, development
JEL subject codes: I2, H5, O12

[1] Corresponding author is Martina Viarengo (martina.viarengo@graduateinstitute.ch). We are grateful to the Editor and three anonymous Referees for helpful comments and suggestions. All errors remain our own.

This study uses the Organisation for Economic Cooperation and Development's (OECD) Program for International Student Assessment data set (PISA) to assess the link between student test scores and the school students attend, school policies and practices, students' family background and parental involvement in education. PISA provides data on test scores, schools, and family background for hundreds of thousands of students around the world, which makes it the largest and arguably best cross-country data set in the world for analyzing the relation between test scores and their potential determinants. Our focus on schools and school level policies/practices and parental involvement in education provides a more granular picture of the factors associated with student achievement than is given by studies that analyze country-level policies such as national standards and external exit examinations.[2]

The major finding of our analysis is that schools are the most important factor associated with student test scores. Within countries, school fixed effects explain a substantial proportion of the variation of test scores and much of the relation between family background and scores. To paraphrase the famous campaign slogan of President Clinton, the data tell us "it's the schools, stupid".[3] Going a step further, we identify school level teaching practices as reported by students as major correlates of the relation between the schools students attend and test scores.

There are two pathways by which test score outcomes are likely to be associated with the school students attend. The first pathway is through national policies or private behavior by parents or schools that sort students with similar ability together. Sorting creates differences in

[2] Checchi et al. 2011, Woessmann 2003a, Schuetz et al. 2008, find that standards and exit exams affect national test scores. Studies using other data also show that standards and exit examinations are associated with better student performance (Bishop 1997, 2006; Woessmann 2003b for the TIMSS 1995; Woessmann 2005a; Jurges et al. 2005; Greene and Winters 2005). Studies also find that having an effective link between assessment and curriculum (Bishop 2006) and keeping high standards (Peterson and Hess 2006, 2008) improve outcomes, and that increased school autonomy positively affects student performance when combined with an accountability system and external exams (Woessmann 2005a). An increasing number of countries and states in the US require following a specified curriculum and passing an exit examination for high school graduation. (Ou 2010, Peterson and Hess 2008).

[3] http://en.wikipedia.org/wiki/It's_the_economy,_stupid.

outcomes across schools without necessarily affecting the performance of students. The second pathway is through educational policies or practices that differ across schools in ways that affect student achievement, with "good schools" raising the test scores of students more than "poor schools". To differentiate between sorting and educational effects, we contrast school effects between countries with policies that assign students at early ages to schools beyond their first school and countries that assign them later. School effects are larger in countries that assign students to schools beyond their first school earlier through early tracking systems than in other schools, which we interpret as reflecting tracking. But school effects are still large in countries without such practices, which implies that the school effect on test performance goes beyond sorting. PISA-based measures of school policies and practices account for a substantial part of the school fixed effect on test scores, with the largest factor being the mode of teaching as reported by students, but these measures still leave a sizable proportion of the fixed school effect unexplained and do not pin down the line of causality between policies and test scores.

Our analysis also shows that addition of school dummy variables to equations for test scores substantially reduces the estimated impact of family background on scores in countries with tracking policies and in countries without tracking policies. This implies that schools are major pathway by which parental background relates to student test scores irrespective of the way students are sorted among schools. Extant PISA measures of parental involvement in student education, by contrast, do not account for much of the family background effect on test scores.

The paper has four sections. Section one describes the PISA data, the equation we use to relate test scores to school and family background variables, and the way we draw inferences

about school effects from the data.[4] Section two presents estimates of school fixed effects and the impact of school policies and practices on test scores. Section three gives estimates of the relation between family background and test scores and the role of schools and parental practices in the link between background and test scores. The concluding section considers the implications for policy research of the evidence that schools have a sizable effect on test scores.

1. PISA Dataset and Empirical Analysis

1.A) Data

Every three years since 2000 the PISA study has tested 15 year-old students on their skills, knowledge, and ability to use this knowledge to solve problems in real-life situations.[5] The tests cover mathematics, science, and reading. Participating countries randomly select schools to administer the test and randomly select students within the schools to provide a nationally representative sample of students. PISA standardizes test scores to a mean of 500 and standard deviation of 100 on the basis of the 2003 PISA. Differences in test scores can thus be interpreted in terms of percentage points of an international standard deviation. An increasing number of countries have participated in each succeeding wave, including developing countries outside the OECD.

We use data on students and schools from the 2009 wave. This data set provides information on about 520,000 students in 18,641 schools from 74 economies[6], of whom 34 are

[4] Woessmann (2005b) compared the importance of family background of five Asian countries (Hong Kong, Japan, Singapore, Korea and Thailand) to three Western countries (France, Spain and the U.S.) in the TIMSS 1995 study.

[5] Students' age varies between 15 years 3 months and 16 years 2 months and to be eligible to participate must have completed at least six years of formal schooling (OECD 2010a).

[6] Around 470,000 students in 65 countries participated to PISA 2009. In addition to this, about 50,000 students were tested in 2010 in the framework of PISA 2009+ (from 7 countries and 2 Indian states (Himachal Pradesh and Tamil Nadu).This was the same assessment as PISA 2009 but carried out one year later. The complete list of countries in our sample is presented in Table A2. More information on PISA 2009 and PISA 2009+ is available at

OECD countries and 31 are "partner countries" outside the OECD. In addition to the PISA tests, students fill a questionnaire on their characteristics, family structure, and background. Principals of each participating school report on school characteristics, policies, and practices. Fourteen countries administered a questionnaire to parents regarding their involvement at home with the child and in school-related activities,[7] including involvement when children were in primary schooling (ISCED 1).[8]

We analyze students in the full set of countries in the PISA and then group the countries into five subsets: OECD economies, European Union economies, Asian economies, other high income economies, and middle income economies.[9] Countries can appear in more than one subset in our groups: for instance the EU countries are also part of the OECD country subset. We differentiate the Asian countries because five of the top ten scoring countries in PISA 2009 were Asian, which raises questions about how they differ from other countries in educational practices.[10] We give the results of analyzing the mathematics test scores in PISA. We analyzed test outcomes in reading and science, and some non-cognitive measures of performance as well,[11] and obtained similar results to those for mathematics. These results are available on

---

https://mypisa.acer.edu.au/.

[7] The data set is publicly available from the PISA website: http://pisa2009.acer.edu.au/.

[8] Throughout, we use the sampling weights in the database as described in PISA Technical Report 2009 (OECD 2012a). To deal with missing observations, we kept all students with complete test score data and imputed the values of missing explanatory variables by using mean values and including in the regression an indicator for missing data. (See Hanushek and Woessmann 2011).

[9] Table A2 provides the list of countries in our analysis. Korea and Japan are Asian countries members of the OECD. Mexico is a middle-income country member of the OECD. For the purpose of the analysis that follows Korea and Japan are included in the Asian group and Mexico in the middle-income group.

[10] The Asian countries have large shares of students at the proficient level 5 (OECD 2010b). PISA classifies students into five proficiency levels, "moderate" performers achieve proficiency levels of 2 or 3 whereas "top" performers achieve proficiency levels of 5 or 6. OECD (2010a, p.8; Peterson et al. 2011, p.8). More than 20 percent students in Korea, Japan and Hong Kong perform above proficiency level 5 (OECD 2010c, p.63)) and above the advanced international benchmark of 625 or better. Cross-national surveys of learning achievement show that Korea, Macao, Shanghai combine top scores with levels of inequality lower than the OECD average (e.g., (OECD 2010b, p.53)).

[11] Our analysis of non-cognitive learning outcomes considered indicators related to disruptive behavior and two

request from the corresponding author.

1.B) Empirical Framework

At the heart of our analysis are estimates of the following equation:

$$T_{isc} = \alpha_1 F_{isc} + \alpha_2 FP_{isc} + \beta_1 X_{isc} + \beta_2 S_{sc} + \beta_3 SD_{sc} + \mu_c + \varepsilon_{isc} \quad (1)$$

where:

T is the test score of student *i* in school *s* of country *c*,

F is a measure of family background characteristics for student *i* in school *s* of country *c*,

FP is a measure of the parenting practices for student *i* in school *s* of country *c*,

X is a vector of the student's characteristics, for student *i* in school *s* of country *c*,

S is a vector of the policies and practices of school *s* of country *c*.

SD is a vector of dummy variables for each school in the data set[12]

μ_c is a vector of country dummy variables

ε is an error term[13]

$\alpha_1, \alpha_2, \beta_1, \beta_2, \beta_3$ are vectors of coefficients to be estimated.

Because the PISA gathers data from principals and teachers on school policies and practices and from about 35 students in each school, we have two ways to examine the relation between the

psychological traits central to the five factors model (Costa and McCrae 1992): openness to learning (attitude towards school) and locus of control (self-confidence in ICT high-level tasks). . Non–cognitive skills have been largely neglected in the cross-country studies due to the paucity of information available and difficulties of measurement (Woessmann et al. 2007; Brunello and Schlotter 2011).

[12] In the estimation of the school fixed effect models country fixed effects are absorbed by school fixed effects.

[13] Clustering-robust linear regression (CRLR) has been used in the estimation to estimate standard errors given the fact that the error term has both a school-level and a student-level components.

schools students attend and test scores. The first way is a "sibling" type estimate based on estimating the school level dummies SD. To obtain the maximum fixed school effect we omit the policy and practice vector in equation (1) (that is, set β_2 at 0). Just as studies of family background effects on sibling outcomes use the fact that siblings grow up in the same family to estimate the variance in outcomes associated with family background without measuring that background,[14] the PISA data allow us to estimate school effects through the similarity of outcomes among students in the same school absent measures of school policies and practices. The second way to examine the effects of schools on outcomes is to estimate the coefficients in the S vector of policies and practices in equation (1). To do this we omit the school dummy variables (that is, set β_3 at 0).

If students were randomly assigned to schools, the difference between the contribution to the variance of test scores associated with students attending the same school (the estimated school fixed effect) and the contribution to the variance in scores associated with school practices and policies among schools would reflect the effect of schooling on outcomes through mechanisms rather than observable policies and practices. But governments do not assign students randomly among schools. Countries assign students after a specified age to the next level of schooling. Some do so at early ages while others do so at later ages. At the same time, some parents try to place their children into schools with stronger academic standing, for instance by moving to areas with schools that have good academic reputations, while some schools have selective admissions policies. To the extent that parents of more able students successfully enroll their students in better schools and that selective admission schools attract students of greater ability, students of similar ability will be sorted among schools, producing

[14] Bjorklund and Salvanes (2010) report that factors shared by siblings explain more than 50% of the overall variation in years of schooling for children in a sample of developed countries.

school fixed effects independent of what schools actually do for students.[15]

To help differentiate what goes on in schools from sorting we contrast differences in the school fixed effects on test outcomes between countries with early sorting policies and countries with late sorting policies. We expect greater variation in the average scores among schools in countries with early tracking policies than among schools in countries without those policies. Large school effects in countries without tracking policies would indicate that something other than tracking underlies school differences. Evidence that policies and practices affect outcomes similarly among countries with and without tracking would also indicate that the huge school fixed effects in our data are associated in part with those policies or practices. We also contrast variance in test scores among schools with different admission policies and openness to residential and parental choice to see whether those measures of sorting account for the bulk of school fixed effects. They do not. The implication of these diverse calculations is that a sizable proportion of school fixed-effects is associated with what schools do as opposed to sorting similar students into the same schools.

2. The Role of Schools and Other Sources of Sorting

2.A) The Role of Schools

Table 1 gives some statistical properties of the math test score on which we focus. Column 1 records the mean and (in brackets below the mean) standard deviation of the test score in all the countries in the data set and countries in the five country groupings. The mean scores for the poorest countries in our sample, "middle income countries" by the World Bank definition, fall

[15] If different schools had no discernible effect on student performance, parental efforts to get their students into better schools would make little sense in terms of their education, so it is likely that sorting practices by parents are related to actual educational effects. Demonstrably better school effects on students should produce greater efforts to get able students into higher quality schools. We do not explore this interaction in this paper.

below those of the other and higher income countries. The mean scores for the Asian countries are higher than the scores for other countries. The coefficient of variation of the test scores in column 2 show modestly higher variation in scores among the middle income countries than others and modestly lower variation in test scores for the Asian countries than for the others. Variation of scores is, however, large among all groups, which reflects the wide distribution of scores among students within countries.

The columns under the heading "Percentage of Variance in Scores" record the proportion of the variance in test scores associated with three factors: country; measures of background; and schools[16]. The proportions of variance are calculated separately for the characteristics. The column labeled country shows the percentage of variance from regressions of student test scores on country dummies. The column labeled background shows the percentage of variance from regressions of test scores on the family background measures given in the table note. The column labeled schools shows the percentage of variance from regressions of test scores on school dummies.

Country factors are substantial in the all country group because of the sizable difference in scores between the middle income countries and the other countries. But within each of the groups country is associated with only a small proportion of the variance in scores. Background factors have a substantial impact on the variance in the overall sample and within every category. Background factors contribute least to the variance among the Asian countries.

The largest source of variation in test scores in the table is the school the student attends.[17] The regression of test scores on the school dummies explains a substantial proportion

[16] Standard errors in the empirical analysis here and in the remaining of the study are corrected for clustering.

[17] This can be explained either by the effect that schools have on determining student's learning outcomes or by some sort of self-selection that affects the student composition into the different schools. In the analysis presented in

of the variation of scores in all groups. Estimates of the contribution of schooling to the variance of test scores separately by country confirm this result for all countries, albeit with variation in the magnitudes (see Appendix table A2). The evidence that the students' school is the single biggest measurable contributor to the variance of test scores among the approximately 520,000 students in the PISA 2009 data motivates ensuing analysis.

To what extent does the huge fixed school effect found in table 1 result from national tracking policies or other sorting practices as opposed to potential impacts of schools on learning?

We begin our analysis of this question by estimating the effect on the variance of test scores of national policies regarding the placement or tracking of students into higher level schools. We gathered data on the first age at which students move from one level of schooling to the next level in different countries.[18] Researchers who study national tracking systems use these data to divide countries into two groups: countries which sort students at ages less than 16 are treated as having tracking systems while those with a first age of sorting above age 16 are treated as not having a tracking system.[19] Given that the students in the PISA are 15-16 years old, this definition works well with our data. In addition to this measure, we created a more refined measure of tracking by differentiating countries by the first age of movement to a different school. If tracking policies sort students by abilities, the variance of test scores *among schools* should be higher in countries that tracked students early in their lives than in countries that tracked them later in their lives while the *within-school* variance of test scores would be lower in

---

the following sections of the paper different possible explanations are going to be investigated.

[18] First age at which selection takes place in the education system (OECD 2005, Education at a Glance, Table D6.1; Hanushek and Woessmann 2006; Brunello and Checchi 2008, Woessmann 2005b; in addition to this, Prof. Ludger Woessmann kindly shared with us his international dataset on tracking acquired through detailed personal inquiries). This information refers to the policy set at the national level.

[19] For a review of the existing studies see Hanushek and Woessmann (2011, section 4.4.4).

the tracking countries than in the non-tracking countries.

We examined differences in the relation between test scores and schools at the level of countries and at schools within countries. At the country level, we calculated the coefficients on the vector of school dummies from equation (1), absent measures of school practices and policies. The estimated coefficients measure the fixed effects impact of each school. This is equivalent to taking the average performance of a school adjusted for the effects of student background factors. Then we related the estimated school effects in a country to how the school ranked in the country's distribution of school scores from highest to lowest, on the notion that tracking should produce a steeper relation between school effects and the ranking of schools in the country's distribution of schools by test scores.

Figure 1[20] graphs the scores of schools against the ranks of schools in the countries in the PISA data set. By construction the score-rank curve rises monotonically. The graphs show clear differences among countries in the gradients of the score-rank curves. We also we also checked the robustness of the country ordering by the gradient of Figure 1 using the other international studies. Specifically, we used both TIMSS and PIRLS. These surveys are not available for the year 2009; therefore we used the most recent wave available for both surveys. That is, TIMSS 2011, both 4th and 8th grade and PIRLS 2011. We would like to acknowledge the fact that, as widely documented in the existing literature, the international studies do not only differ with respect to the level of education but also with respect to the design and the structure of the survey. Overall, the ordering of the gradient seems to be consistent with the one we found using the PISA dataset. The greater similarity lays between the PISA study and TIMSS 8th grade, and the lower with the PIRLS study which focuses on the 4th grade and reading performance only.

[20] We also constructed the school ranks by relying on the decile of the distribution and checked the consistency of the two methods. These figures are available from the authors.

Among the international studies examined here, the PIRLS study is the one that presents the greater difference from the PISA one in terms of focus and design and the smaller number of countries to both the PISA and PIRLS surveys (15 as opposed to 26 common in PISA 2009 and TIMSS 2011, 8th grade and 37 common in PISA 2009 and TIMSS 2011, 8th grade).[21]

Regressing the estimated school fixed effects on the rank of the school for each country separately, we obtain the coefficient α_1 for country i that shows how much being higher in the distribution of schools affects the test score:

$$\beta'_{3sc} = \alpha_1 R_{sc} + \varepsilon_{sc,} \qquad (2)$$

where β'_{3sc} is the vector of estimated fixed school effects for school S in country C (ie the mean test score in the school adjusted for the other variables in regression (1)) and R_{sc} is the rank of school in the distribution of test scores for the schools in country C.

The estimated slopes for countries in appendix table A3 reflect the pictures in figure 1. They show high gradients for countries of Continental Europe (France, Germany, Hungary, the Netherlands), relatively high gradients for Asian countries and smaller gradients for Anglo-Saxon (Australia, Canada, United Kingdom) and Nordic European countries (Denmark, Finland, Norway).

Studies of the effects of sorting on test scores across countries indicate that early tracking increases subsequent inequality in educational outcomes without improving national scores (Hanushek and Woessmann 2006). They also find that sorting increases the importance of family background in educational achievement (Schuetz et al. 2008; Woessmann et al. 2009) and that it

[21] The complete analysis related to the comparison between the three international studies is available from the authors upon request.

raises the impact of family background on earnings in the labor market (Brunello and Checchi 2007). One implication of these analyses is that countries with greater sorting of students by school should not only have larger regression coefficients linking school fixed effects to the rank of schools in the distribution of scores than countries but should also show a larger contribution of school effects to the variance in test scores among students.

The regressions in Panel A in Table 2 test these expectations using two measures of country tracking policies. The first measure is the dummy variable for whether the first age at which students move school levels is below 16, which is the standard measure in analysis of tracking systems in the cross-country literature. The second measure is the "tracking index" based on the actual age given in appendix table A3. The index varies from 10 to 18 to reflect the age of tracking, where we reverse coded ages so that higher numbers reflect earlier ages and thus stronger tracking (i.e., the index is 18 for schools that send students to another school level at age 10, 17 for those that do so at age 11, and so on.).

The first two columns give the results of regressing the estimated gradients of school test scores on school rank (from the appendix table) on the specified measures of tracking. The first regression excludes covariates for student characteristics. The second regression includes country averages of the covariates listed in the table note[22]. The positive coefficients on the tracking dummy variable and on the tracking index show that earlier sorting of students into another level of schooling is associated with a larger gradient of test scores on school rank. A high ranking school has a relatively higher test score compared to others in the same country when the country tracks earlier.

The next two columns change the dependent variable to the share of country test score

[22] Descriptive statistics of these covariates are presented in Tables A1 and 6.

variance due to schools. They give the coefficients for the regression of the ratio of the variance in test scores attributed to the school dummies over total variance in each country on the tracking policies with and without covariates for country characteristics of students. These regressions show that the proportion of variance between schools is larger among countries with lower ages of sorting.

Panel B of the table shifts from country level regressions to school level regressions. Rather than using the summary statistic of the gradient of test scores on school rank to test the effect of country-wide tracking on school scores, we regress the estimated school fixed effects on the rank of each school in their own country and on country dummy variables *separately* for the countries with early tracking policies and those with later tracking policies. This specification forces a single coefficient on school rank for all of the countries in each of the groups. Because the regression is a school-based analysis it gives greater weight to countries with more schools in the survey. The results confirm the regression findings in the country data. The slope of the test score on the rank of schools is 0.44 in countries with an early tracking policy and 0.29 for countries without such a policy when the tracking measure is the sole explanatory variable; while the slope is 0.25 for countries with an early tracking policy compared to 0.20 for countries without the policy with inclusion of the covariates listed in the table note. We also regressed the test score on the rank of school, the age-based measure of the tracking policy and the interaction of the rank with the policy measure. If tracking produces a steeper gradient, the estimated coefficient on the interaction term should be positive. As the last line of table 2 shows, this is the case. Countries with stronger tracking policies sort students among schools in ways that produce greater differentiation in average test scores among schools.

The variance in test scores among schools in countries that do not track students before

age 15 cannot by construction be associated with tracking policies. While not all of the variance in test scores among schools in these countries will be due to true school impacts on students, the school fixed effects in these countries will be more reflective of true school impacts than the school fixed effects in countries with early tracking policies. Applying the maximum estimate of the effect of tracking on the variance of test scores in table 2 – the 14 point impact in the single variate regression – to the estimated impact of school fixed effects on the variance of test scores from table 1, would reduce the school effect from 62% of the variance in test scores among students to 48%, which still leave schools as the major factor associated with student test scores.

2.B) School Variance Due to Other Sources of Sorting

The age at which students first move to a higher school is not, however, the only way students may sort among schools and thus contribute to the variance of test scores among schools. Some countries allow parents greater choice in their enrollment. Schools may have selective admission policies, limiting enrollment to high performing students. Parents may affect student placement by residing in areas with reputedly higher quality schools.

The PISA country-level publication (OECD 2010d) and principals' questionnaire contain two questions which allow us to assess the potential impact of these forms of selection on the estimated school fixed effects. The first question relates to national practice. It asks whether in the country "*Families are given a general right to enroll in any traditional public school they wish."* Assuming that families with more academically able students seek and succeed in getting their children into schools with stronger academic reputations, the greater the rights of families to enroll in any public school should increase the variance in test scores among schools. The second question, extracted from the school principals' questionnaire, is *"How often is a student's*

record of academic performance (including placement tests) considered when students are admitted to your school?." Schools that take account of performance in admissions policy should have a more homogeneous student body and correspondingly lower variance of scores within the school than other schools.

The PISA has a third question regarding school policies that has a more ambiguous relation to the variance of school test scores. It asked principals *"How often is residence in a particular area considered when students are admitted to your school?"* If residence[23] is a strong factor in admission, a school will likely admit nearly all students in its catchment area – the less able as well as the more able, and thus be less selective on academic performance, producing greater within school variance in test scores and smaller variance in school fixed effects. On the other side, if wealthier and more educated families live in different catchments than poorer and less educated families, such a policy would increase the variance in school effects. If families with more academically able students move to areas with better schools, the dependence of residence on schools would add yet another way that sorting contributes to higher variance of scores among schools.[24]

Table 3 examines the relation between these potential contributors to the variance of test

[23] In order to further examine the sorting mechanism through residential location the empirical analysis has been carried out for a finer set of countries. These four countries do not have national tracking policies but two are more egalitarian (Finland and Norway) whereas in the other two residential sorting appears to play a more important role (United Kingdom and United States). Results related to this analysis are available from the authors upon request. We also acknowledge the fact that other sorting mechanisms may play an important role (e.g., parental choices related to public vs. private schools).

[24] We are aware of the fact that that given the structure of the database we are using, we are not been able to explore further sorting mechanisms but we acknowledge their potential role. Moreover, we acknowledge the multi-dimensionality of sorting and the complexity in defining and measuring tracking across countries (Betts 2011). Specifically, grouping of students by ability or other criteria may take place within schools in countries like Canada or the United States as documented by country-specific studies (Oakes 2005). We also acknowledge the fact that other mechanisms of sorting may play an important role. That is, sorting may take place in terms of parental background leading to residential segregation or parental choices related to public vs. private schools. An additional source of sorting may be related to the differentiation of schools in the public sector. Nechyba (2011) provides an insightful summary of these sorting mechanisms by relying on a broad and recent literature. Most studies reviewed are country-specific and often nationally representative.

scores among schools in two steps. The column labeled "share of schools with practice" shows that almost 2/3rds of the schools in the sample are in countries with tracking systems that move students to a higher level of school before the age 16 while 41% of the schools are in countries where families have a general right to enroll students in a preferred school. On the school side, 31% report selective admissions and 36% report giving heavy weight to enrolling students in the relevant geographic area.

The regression coefficients in the table relate the variance of test scores to policies likely to affect sorting. Column 1 examines the relation between the tracking policy and the three additional policies on the *variance of test scores among students within the school*. It is based on regressing the natural logarithm of the variance among students in schools on the four practice variables, the average test score in the school, and number of schools in the country. If a practice reflects sorting, it should be associated with lower within-school variance of test scores. The coefficients for the tracking and parents' choice measures (both country-level policy variables) and for the selective admissions variable (a school level policy variable) obtain negative coefficients. The selective residence measure, by contrast, is associated with greater variance in test scores, indicating that it most likely measures the greater strength of local catchment areas in enrolling all children in the area than sorting of students due to parents with more able students moving residences near good schools.

The regression in column (2) give coefficients relating the *variance in test scores among schools* in a country to the policy variables. It is based on country level regressions in which the selective admissions and selective residence variables are measured at the country level in terms of the percentage of schools in a country that report those policies. These regressions show that tracking, parental choice, and selective admissions add to the variance in test scores among

schools while selective residence reduces the variance. The main contributor to the variance in test scores beyond national tracking policies are the percentage of schools with selective admissions policies, whose effect in raising the variance is offset in part by the selective residence policy, which reduces the variance.

The average variance in test scores among schools is 41.76 Using the equation in column (2) we estimate that in a country with none of the four policies, the variance in test scores across schools would be substantially lower at 33.08. We acknowledge three possible complementary sources of a differential school effects on student tests: self-sorting of students, self-sorting of teachers and selective allocation of resources. We are aware of the fact that given the nature of our dataset we have not been able to examine these additional mechanisms of sorting.

With respect to the self-sorting of students, Nechyba (2011) provides an insightful summary of the following sorting mechanisms: parental background leading to residential segregation, parental choices related to public vs. private schools, as well as sorting related to the differentiation of schools in the public sector. The focus of the review is on the US but "while sorting mechanisms may differ across countries, there is little reason to doubt that sorting itself is a critical component of most primary and secondary school systems" (p.1029). The main focus is on sorting related to income and household peer quality. We discuss here some possible mechanisms through which sorting may take place. Many studies have relied on country-specific datasets and looked at the sorting especially in countries like the United Kingdom and the United States. Studies on sorting related to the differentiation in the public school sector are relatively more recent and limited to countries where such a differentiation exists (e.g., United States). The source of sorting in this case may be related to ability or other characteristics. Finally, sorting between public and private schools has been explored both with observational and experimental

methods. The source of sorting in this case may be related to ability, income or other characteristics. Specifically, the choice of residential areas correlated with the quality of schools may be related with family economic resources (Bjorklund and Salvanes 2011). That is, where admissions depend on the residential area then housing markets become a mechanism of sorting. Housing markets also play a similar role in case where admission do not depend on the residential area but transportation is costly (Nechyba 2011). Another sorting mechanism can occur through through parental education. More educated parents may have better information on the different quality of schools (Bjorklund and Salvanes (2011, p.215)). Sorting related to education may be closely related to sorting through economic resources or may also be independent as suggested by Bjorklund and Salvanes (2011). Sorting may also occurr through peer effects and student ability. That is, "cream skimming" in the private school sector may work as a sorting mechanism in case where private schools set standards to ensure given peer quality levels.

In the context of selective allocation of resources, a limited pool of resources may lead schools to introduce mechanisms of academic selectivity. That is, more and better resources are available to high performing students. On the other hand, the allocation of resources may be of a compensatory nature (e.g. Holmlund, McNally and Viarengo 2010; West and Woessmann 2003), In this case, more resources are allocated to more disadvantaged and low-performing students.

Sorting of teachers may also be an underlying mechanism. For example, Hanushek and Rivkin (2006) review the existing studies on teacher labor markets. They acknowledge the importance of measuring teacher quality and the difficulty related to selection and unobserved characteristics: "Credible research into training versus selection issues as related to certification

policies, merit pay, and so forth clearly requires longitudinal observations that link teachers, programs, and student performance. Until recently, there has been little possibility of such work" (p.1073). A similar problem is found in the estimation of teachers earnings equations (e.g., Dolton (2006) ("Teacher Supply" Handbook of the Economics of Education, Volume 2, Ch. 19). On the other hand, Lankford et al. (2002) use a rich teacher dataset to examine the extent of teacher sorting in the state of New York. Their findings suggest that especially in urban areas disadvantaged students are in classrooms with many of the less skilled teachers. Among the possible explanations they discuss: 1) district may differ in terms of teacher preferences 2) districts may differ in terms of hiring efficiency 3) differences in teaching qualifications may be determined by differences with respect to the political power schools exert 4) Teacher preferences. They also acknowledge that "we know very little about sorting or the causal relationships that lead to sorting" (Lankford et al. 2002, p.39).

Sorting is important in the variation of test scores among schools[25], but however we adjust for sorting the data shows a substantial variance in scores among schools that could reflect within-school educational policies and practices.

2.C) School Policies and Practices

Given that measured tracking or sorting policies do not explain the bulk of school fixed effects, can we identify school-level policies and practices that account for the school fixed effects?

[25] We acknowledge the fact that some unobservable characteristics we are not able to control for may have an effect on the sorting of students into different schools. We previously referred to possible streaming of students with respect to the residential location that may be linked to school admission, as well as possible sorting with respect to school choice and school differentiation within the public sector. These less explicit mechanisms of sorting would imply a more limited effect of schools on learning outcomes.

PISA asked principals to report on a large number of school characteristics and policies.[26] It also asked students to report on the way they were taught in school. A substantial association between these policies and practices and our estimated school fixed effects would be at least suggestive evidence that within school practices contribute to student test score performance. If policies and practices are largely unrelated to the variance of student scores within schools, this would further indicate that the policies or practices affected students at different parts of the test score distribution similarly or not.

Table 4 presents our analysis of the effect of school level policies or practices on the school fixed effects test scores for the 18,340 schools in the PISA data set. The column labeled "Measures of School Policies or Practices" gives the specific policy variables and practices on which we focus. For ease of analysis, we grouped the school policies into three basic areas: school selectivity, school autonomy and accountability; and school resources, including the attributes of teachers. We group the measures of practices into two categories – approaches to teaching, which is reported by students; and staff attitude, reported by principals. Section A4.1 in the appendix provides the details on the questionnaire items we used.

The extensive educational literature on the relation of school policies and practices to student performance provides a valuable benchmark against which to assess our findings. Accordingly, we summarize below the current state of that literature regarding the policies and practices in our analysis.

On *selectivity*, as we saw in table 3, selective schools invariably have higher performing

[26] We acknowledge the fact that there may be some measurement error or limited variability in some policy variables (e.g., Dustmann et al. 2012). We also acknowledge the fact even though we tried to focus on those policies that are regarded as important determinants of educational attainment we are aware of the fact that we are not able to account for all the existing policies. We also looked closely at other sources (e.g., World Bank-Saber database, Eurydice reports) but given the nature of our database that includes both developing countries and advanced economies we could not find further consistent and comparable policy information.

students than less selective schools and lower variation in scores among students. Studies of whether sorting students by classrooms within the same school – the second policy measure in the selectivity group in table 4 – benefits students by providing more targeted instruction by ability or harms students by reducing the interaction between low performing and high-performing students give inconclusive results (Betts and Shkolnik 2000; Figlio and Page 2002).

Studies of *School Autonomy* and *Accountability*[27] suggest that the alignment of autonomy and accountability leads to higher outcomes (Fuchs and Woessmann 2007, Woessmann et al. 2009) but have an unclear impact on inequality of outcomes among students (Hanushek and Raymond 2003). Hanushek et al. (2011) shows that for 42 countries that participated to the PISA study over 2000-2009 school autonomy in academic content, hiring, salaries is associated with higher student learning outcomes in developed countries.

Studies of *resource* inputs such as spending per student, teacher-student ratio, class size, instructional time have produced some controversy. Reviewing extant studies through the early 2000s, Hanushek (2006) concludes that in advanced countries there is no compelling evidence that changes in the level of any of these resources, within the observable range of variation, have a substantive effect on educational outcomes. The Tennessee STAR study, however, gives a different picture – with a one-third reduction in class size associated with higher student achievement by the equivalent of 3 additional months of schooling. But other experiments find smaller or no impacts of class size (Chingos and Whitehurst 2011; Altinok and Kingdon (2009) and Woessmann (2010)). Studies relating to teachers' characteristics such experience, academic

[27] Research in this area using the international studies has examined various dimensions of school autonomy (e.g., autonomy over resources allocation, hiring and firing decisions), regulations of school governance and standardized exams (e.g., periodic measurement of teacher and principal performance, standardized exams).

and professional qualifications, have yielded a broad consensus: teachers matter but the observed characteristics are generally unrelated to teachers' performance (Hanushek and Rivkin (2006)).

With respect to school *practices* Bratti et al. (2008) find that cooperative behavior among students improves learning outcomes. Comparing student reports of teaching practices Bietenbeck (2011) reports that memorization and lecture-style teaching produce better outcomes whereas selected teaching practices such as working in groups, students' autonomy in problem-solving decisions have a small insignificant effect on learning outcomes. Schwerdt and Wuppermann (2011) compare lecture-style teaching to problem solving and find a positive association between the practice of lecture-style teaching and students' test scores. Examining learning practices in several waves of TIMSS, PIRLS and Algan, Cahuc and Shleifer (2012) find that "horizontal teaching practices" (i.e., students work in groups, ask questions) is related to pro-social beliefs while "vertical teaching practices" have the opposite effect.

There are considerable country differences in school policies and practices indicated in the Appendix (Table A5). Schools in Asian countries have greater autonomy over curriculum and assessment and a lower autonomy over resource allocation compared to schools in the OECD countries. But the Asian schools have higher levels of educational resources such as instructional material, laboratory equipment and availability of computers and software for instruction; and higher extra-curricular activities and much levels of selectivity in admissions. Nearly 72 percent of Asian schools compared to 33 percent of schools in the OECD countries report academic selectivity in admissions while just 8 percent of Asian schools report ability grouping between classes within schools compared to 11 percent of schools in OECD countries. Nearly all teachers in Asian countries are fully certified and have college degrees compared to around 84 percent in the advanced OECD countries. Schools in OECD countries use

achievement data to evaluate the principal schools in Asian countries (44 percent vs. 18 percent) while by contrast a larger share of OECD schools use achievement data to evaluate teachers (43 percent vs. 34 percent). With respect to school practices Asian students report less use of memorization than their OECD counterparts. The *index of elaboration*, which measures the way students relate new information to prior knowledge and knowledge acquired in other situations, is significantly higher in Asian countries.

Column 1 in table 4 gives the regression coefficients linking the school fixed effects to the policy variables for all schools and countries in the PISA sample. Schools with greater academic selectivity have higher test scores (as in table 3) while schools that group students by ability between classes have lower test scores. This could reflect that such grouping harms students or that schools with weaker student bodies are more likely to separate the more able from the pack. We find little relation between the measures of school autonomy or accountability and school performance, with most policies obtaining insignificant negative coefficients. The coefficients on the measures of school resources also show weak links to test scores with one notable exception: the student to teacher ratio has a significant negative coefficient, implying that scores are higher when there are more teachers per pupil, which is more consistent with the STAR results than the bulk of the literature on school resources.

The most striking results in table 4 relate to the "Approaches to Learning" variables, all of which obtain sizable and highly significant coefficients that tell a similar story. Memorization and use of elaboration are associated with low test scores while the use of comprehension checks are associated with higher test scores. What explains this strong relation? These data are based on student reports of actual teaching procedures, which may more closely reflect what goes on in classrooms than the school policy variables reported by the principals, whose link to actual

classroom activity is presumably minimal at best. Because the PISA files do not contain information on student achievements or abilities before being taught with these practices, however, the correlation could just as readily reflect schools choice of teaching techniques as the effect of the techniques on learning. Schools with weaker academic students may choose memorization and elaboration techniques while schools with stronger students favor comprehension checks. In any case, the approaches to learning variables are the only ones with a consistent and strong correlation to test scores.

The estimates in columns 2 and 3 are based on separate regressions for schools in countries without tracking systems and for schools in countries with such systems. There are some notable differences in the estimated relation between the effects of the measures of school autonomy and accountability on the school test scores between the schools in countries without tracking and those in countries with tracking. More policies have substantial and significant coefficients in column 2 than in column 3, which produces the noticeably higher R-square in the column 2 regression than in the column 3 regression. These results suggest the school level policies are more important in affecting test score outcomes in the non-tracking countries. Since these schools have a more diverse student body, there may be more scope for school level differences to affect outcomes.

By contrast, the estimated effects for the "approach to learning variables" are similar between schools in the two groupings. That the way students report they are being taught has similar effects regardless of national tracking policies makes this a robust result that suggests that the mode of teaching may be the key to understanding the strong school effects result in our calculations.

Finally, the last column in table 4 shows the result of regressing the *variation in test*

scores within schools on the policy and practice variables. Indicative of its role in sorting shown in table 3, academic selectivity reduces the coefficient of variation of scores. The school autonomy and accountability variables have no statistically discernible effect. The variables with the clearest impact are those relating to resources. The greater the availability of school resources, from the quality of resources to extra-curricular activities to the student-teacher ratio, the lower is the coefficient of variation in test scores. The approaches to learning variables have effects on the coefficient of variation that are the opposite of those on test scores: memorization and elaboration increase the variation while comprehension checks as a learning strategy reduce the variation. Overall, the policy and practice variables explain only a modest proportion of the variation of scores among students within a school, which implies that those factors do not have markedly different effects on students at different positions in the distribution of test scores within their skill.

Another way to examine the extent to which school fixed effects are associated with what schools do as opposed to the sorting of students among schools is to examine directly the link between policies and practices and the estimated fixed effects. To do this we run the following regression

$$\beta'_{3sc} = \alpha_1 S_{sc} + \varepsilon_{sc,} \qquad (3)$$

where β'_{3sc} is the vector of estimated fixed school effects for school S in country C and S_{sc} is the vector of school policies and practices in school S in country C.

Table 5 summarizes the results from estimating equation (3) with slightly different specifications. In columns 1 and 2 we exclude any covariates so that we focus solely on the relation between the fixed effects for schools and the policies and practices. Column 1 excludes

the country dummy variables effects while column 2 includes them. The estimated coefficients on specific policies and practices are necessarily similar to those in table 4 so we concentrate on the summary R-squared statistic which tells us the extent to which the whole set of policies and practices help explain the fixed effects. By itself, the vector of policies and practices accounts for 0.465 of the variance in fixed effects, and with the addition of the country dummy variables, the R-square increases to 0.655. Columns 3 and 4 give the coefficients and R-square for comparable regressions which include covariates for other school related factors, as noted in the table note.

In sum, the fixed school effects are sufficiently related to school practices and policies to suggest that they reflect more than sorting of students. While we lack experimental variation in the policies and practices to identify the path from the policies and practices to test scores separately from the potential endogeneity of policies and practices to the academic attributes of students, the association between the fixed effects and policies and practices directs attention at the fixed school effects as connected to school effects and decisions as opposed to sorting of students.

3. School Fixed Effects and Family Background

The analysis in Table 1 found that family background factors explained from 14% to 33% of the variance in test scores among students. Studies of the educational achievement of young persons invariably find that children from more advantaged backgrounds perform better on average in school in terms of test scores and other school outcomes, albeit with considerable heterogeneity in the magnitude of the impact of background factors among countries (Hanushek and Woessmann (2011a)).

Given our finding that much of the variance in test scores is associated with the school

which students attend, the natural question to ask next is to what extent, if at all, the school which the students attend mediates the relation between family background and student test scores.

The PISA contains three potential measures of family background with which to examine this question.[28] Descriptive statistics related to these variables are presented in Table 6. Students report on the education of their parent, [29] which is almost always positively associated with better school performance by children.[30] Among students taking the PISA, parental education differ substantially among countries as well as within countries. Students from Asian countries have a higher share of parents who have completed higher education (39%) compared to the students in the OECD countries (31%). Ammermuller et al. (2005) and Woessmann (2005b, 2008) use parental education to analyze background effects on children's education.

Students also report on the number of books in the household.[31] Some researchers favor the books at home measure as reflecting not only the parents' scholarly culture but also the socio-economic background of the household (Di Maggio 1982, Dronkers 1992). Ammermuller et al. (2005), Woessmann (2003a, 2005b, 2008) and Schuetz et al. (2008) use books at home in analyses of educational outcomes in different waves of TIMSS while Peterson and Woessmann (2007) analyze it in the PISA 2003. Here too there is considerable country variation as well as variation within countries. Asian students taking the PISA tests come from more literate families:

[28] See Schuetz et al. (2008) for a discussion of the advantages/disadvantages of these different measures.

[29] Answers are based on the ISCED classification (PISA 2009, Student Questionnaire, questions n. 10, 11, 14, 15). PISA derives an index of the highest level of education of either parent. It also converts this into the highest number of years of schooling.

[30] Some analysts argue that parents education increases the marginal productivity of children's education, by raising expectations about educational success (Davis-Kean 2005, Borjklund and Salvanes 2010) but it can also operate through the higher income associated with higher parental education.

[31] This variable comes from students' responses to a question about the number of books in their home exclusive of magazines, newspapers and schoolbooks. Students can choose one of the six categories: 0-10, 11-25, 26-100, 101-200, 201-500, more than 500 (PISA 2009, Student Questionnaire, question n. 22).

43% of Asian students live in families with more than 100 books compared to 34% of students in the OECD country sample.

The third measure is the "index of socio-economic and cultural status" (ESCS) developed by PISA. The ESCS is a composite[32] that includes the level of parental education converted into years of schooling,[33] an index of parental occupation, an index of family wealth, cultural possessions, home educational resources, and an index of home possessions, which includes the variable number of books at home. Jenkins et al. (2008) and Machin and McNally (2011) have found a significant relation between this index and learning outcomes.

We used all three measures to examine the mediating role of school fixed effects in the relation between background and test scores. We estimated variants of equation 1 with each of the three background factors introduced separately and no measures of school effects in the equation and then estimated the equation with the inclusion of fixed school effects.[34] Table 7 summarizes our findings for the ESCS variable, which has the virtue of encompassing parental schooling and books at home measures of background.[35] The results with the parental education and books at home are similar.[36]

Table 7 gives results for all countries, for particular country groupings, including with those with early tracking policies and the countries without them, the better to isolate effects

[32] The OECD Technical Report (OECD 2012a,) details the construction of this index. The index is standardized to have an OECD mean of 0 and a standard deviation equal 1

[33] The variable is derived from the student questionnaire, in which PISA asks students the highest level of schooling completed by their father/mother and whether their father/mother have completed undergraduate or postgraduate degrees

[34] Dustmann et al. (2012) used a similar approach to examine the effect of school resources, class size and accountability on the gap in learning outcomes among migrant and native students in PISA 2006.

[35] We estimate an international production function by using an approach commonly used in the existing literature (Hanushek and Woessmann 2011a). The main underlying assumptions are related to the fact that the production function is known and inputs are used to maximize student learning outcomes. We acknowledge possible limitations of this method as described in the existing literature (Hanushek (1979, 1986)), and in particular that country-specific results may be affected by the measurement of background factors (Schuetz et al. 2008).

[36] These results are available from the authors upon request.

within school unrelated to tracking. It records coefficient on the ESCS measure of background and the R^2 associated with the regression from regressions of test scores on the variables specified in the column and in the table note about the covariates. Column 1 records the estimated coefficients on family background with only one other covariate – country fixed effects so that the analysis focuses on within-country variation in test scores. Column 2 includes a set of student-level and a set of school-level controls that reflect differences in the urban location and size of schools across countries shown in appendix table A1 as well as country fixed effects. In both columns the coefficients of family background are large and significant.

Column 3 gives the results with the addition of the school dummy variables. The coefficients on the ESCS background measures drop greatly. For all countries, the coefficient drops from 24.20 to 10.28 – a 58 percent fall. While the extent of the decline differs among the groups, it is large in all cases. [37]

Do parental practices also account for the relation between family background and test scores?

Research on parental participation in home and school-related activities finds a positive association between parental involvement and students' learning outcomes (Bordieu 1986, Coleman 1988, Hoover-Dempsey and Sandler 1997). Desforges and Abouchaar's (2003) note the limitation of evaluations which do not satisfactorily address the endogeneity of parental involvement. Parents with greater education and resources are likely to invest more time and resources in their children's education than less educated parents or those with fewer resources. Parents are also more likely to invest in the education of children who are more receptive. Many parents will spend time and effort in helping children having problems in school – for instance

[37] Table 7 only shows the coefficients related to the family background variables. The complete regression results are available from the authors.

meeting with teachers or providing help with homework. Country-specific longitudinal studies (e.g., the UK National Child Development Study, the British Cohort Study, US National Educational Longitudinal Study) suggest that parental participation in learning activities at home (Schoon and Parsons 2002), aspirations for children's continuing in education (Singh et al. 1995), discussion of school experience and involvement in cultural and learning activities (George and Kaplan 1998) are positively associated with children's educational performance conditional on the socio-economic status of parents.

Fourteen countries[38] in the 2009 wave of PISA practices administered the parent questionnaire.[39] Caution is needed when analyzing the results given the high nonresponse rate in some countries (OECD 2012, p.328).[40] The OECD analysis of these questions (OECD 2012b) shows that parental activities are positively correlated with test scores but that specific activities with the child have different relations to outcomes. A parent discussing what their child is doing is positively related to test scores whereas a parent discussing their child's progress in school

[38] These are the countries: Hong-Kong, Korea, Macao-China, Chile, Denmark, Germany, Hungary, Italy, New Zealand, Poland and Portugal. The parent data file includes the variables derived from the parent context questionnaire (OECD (2012), "PISA 2009 Technical Report", Paris: OECD Publishing, p.328). The Parent Questionnaire of the PISA 2009 study was delivered to parents by students who had to return it to school once it had been filled in. At page 2 of the Parent Questionnaire it is explained that "This survey should be completed by a parent (or jointly by both parents) or other <primary caregiver> of the student".

[39] PISA 2006 also included some parent questions. Before the 2006 and 2009 waves, cross-country data sets of student performance did not provide much information on parents' participation in home and school-related activities.

[40] Some bias may also be generated by the design of the survey: better educated parents may be more likely to fill this optional questionnaire and more diligent students - with higher test scores – may be more likely to bring the questionnaire home, remember to ask their parents to fill it and after to return it to school (OECD, p.91). In this regard, descriptive statistics presented in the OECD study (2012b, p.163) [OECD (2012) "Parental Involvement in Selected PISA Countries and Economies" OECD Education Working Paper No. 73 – working paper prepared by F. Borgonovi and G. Montt] carried out for the entire PISA 2009 sample and for the sample of students who answered the parent questionnaire show a significant heterogeneity in terms of response rate across countries. Apart from Poland where the questions on parental involvement were not asked (OECD 2012, p.50), the response rate varies from about 100% in Macao-China to about 60% in Denmark. There are also some differences in terms of demographic characteristics. Students in the sample with answers to the parent questionnaire have on average better reading performance, a higher socio-economic status, are more likely to be a girl and less likely to be of immigrant status. Moreover, a difference between the information on parents' education as reported by students and parents themselves should be acknowledged, being information self-reported by parents on average than the one reported by their children (Jerrim and Micklewright 2012).

with a teacher is negatively related to test scores, presumably because the parent is responding to the student not doing that well in school. The OECD also finds that parental involvement is higher in more affluent households.

We selected from the parent questionnaire four areas of involvement for analysis: early parental investment in children such as how often they read books with the child at an early age; participation in school-related activities reflected in such variables as whether parents participate in the local school governance; resources directly available to the child such as subscription to a journal or magazine; and current activities with the child such as discussing political or social issues with the child.[41] Column (4) in table 7 shows the addition of the parental involvement indexes had a modest effect on the estimated impact of the ESCS background measure overall, decreasing it among OECD countries while increasing the background coefficient among Asian countries.[42] It does not have the consistent effect on the estimated background effect that the schooling fixed effects have.[43]

[41] We use indexes based on the parents' answers to the questions in each area available in the PISA database: *Early Parental Involvement Index* of activities with the child (when the child attended the first year of ISCED 1); *Current Parental Involvement Index* of current activities with the child; *Parent Participation in School Related Activities Index* of participation in school related activities; *Resources Directly Available to the Child Index* of resources available to the child. Section A4.2 in the appendix provides a detailed description of the indexes, items in the questionnaires on which they are based and how they have been constructed. The PISA 2009 Technical Report explains the computation, scaling methods and weights involved in the construction of the indexes (OECD 2010d, 2012a). The PISA variables represent a limited selection of the many aspects and activities that shape the relationship between parents and children and could impact on the students' learning progress. Descriptive statistics related to the parental involvement variables are available from the authors upon request.

[42] Columns 1-3 include all countries. Column 4 includes only the 14 countries where the parental questionnaire was administered. Regressions in columns 1-3 have also been estimated with the same sample as in column 4. Results show that parental involvement has only a limited impact in reducing the ESCS coefficient for all country groups. These additional results are available from the authors upon request.

[43] Our findings suggest that measures of parental practices do not account for much of the parental background factors. Instead the major factor that reduces the impact of background factors on outcomes are school fixed effects. It is important to notice that given the structure of the PISA database we are able to control for a limited number of family background characteristics and parental practices and we are not able to control for the unobserved parental characteristics. Therefore, given the existing asymmetry of our analysis we are not able to draw conclusions on the role that unobserved parental characteristics correlated with school fixed effects may play in affecting outcomes.

4. Discussion and Concluding Remarks

Our analysis of the 2009 wave of the Program for the International Student Assessment international data on test scores found considerable regularities in the association between the school students, school-level policies and practices on test scores and between family background and test scores. In all of the countries and groups of countries, we find that a substantial proportion of the variation of test scores within countries is associated with the school students attend. We also found that while national tracking policies and related modes sorting students among schools explain part of the fixed school effects, most of the effects are associated with what schools do and that school policies and teaching practices reported by students explain a sizable proportion of school fixed effects. More surprisingly, perhaps, we also found that school fixed effects are a major pathway for the association between family background and test scores, giving a more consistent and stronger pattern than indicators of parental involvement in student achievement. The implication is that what schools do is important in the level and dispersion of test scores, suggesting the value of more attention to what goes in schools and pinning down causal links between policies and practices and test score outcomes. The main weakness of our study is the inability to identify with the PISA data whether the association between school fixed effects and policies and practices reflects the impact of policies and practices on test scores, as we have modeled the relation, or the endogeneity of schools selecting policies and practices to the academic characteristics of students.

The finding that schools matter greatly in the variance of test scores and that the teaching practices reported by students show a stronger link to school fixed effects than other factors suggests the value of paying greater research attention to obtaining information about school experiences from students and to seeking experimental or pseudo-experimental ways to nail down the causal links from policies and practices to student outcomes.

References

Algan Y., P. Cahuc and A. Shleifer (2012) "Teaching Practices and Social Capital", mimeo Harvard University, unpublished manuscript, July 2012.

Altinok, N., Kingdon, G., (2009) "New evidence on class size effects: A pupil fixed effects approach". Centre for the Study of African Economies, University of Oxford, CSAE Working Paper WPS/2009-16.

Ammermueller, A., Heijke, H., Woessmann, L., (2005) "Schooling quality in Eastern Europe: Educational production during transition" *Economics of Education Review* 24 (5): pp. 579–99.

Betts J.R. (2011) "The Economics of Tracking in Education" in E.A. Hanushek, S. Machin and L. Woessmann (eds.), Handbook of the Economics of Education, Vol.3, The Netherlands: North-Holland, Chapter 7.

Betts J.R. and J.L. Shkolnik (2000), "The effects of ability grouping on student math achievement and resource allocation in secondary schools", *Economics of Education Review* 19, 1–15.

Bietenbeck J. C. (2011), "Teaching Practices and Student Achievement: Evidence from TIMSS", Master Thesis CEMFI No. 1104.

Bishop, J.H. (1997). "The Effect of National Standards and Curriculum-Based Exams on Achievement." American Economic Review, 87, 260–264

Bishop, John H. 2006. "Drinking from the Fountain of Knowledge: Student Incentive to Study and Learn – Externalities, Information Problems, and Peer Pressure." In Handbook of the Economics of Education, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland: 909-944.

Bjorklund A. and K.G. Salvanes (2010) "Education and Family Background: Mechanisms and Policies, IZA Discussion Paper No. 5002.

Bjorklund A. and K.G. Salvanes (2011) ("Education and Family Background: Mechanisms and Policies", Handbook of the Economics of Education, Chapter 3.

Bourdieu, P. (1986). "The forms of capital" (R. Nice, Trans.). In J. C. Richardson (Ed.), Handbook of theory and research for the sociology of education (pp. 241-258). New York: Greenwood Press.

Bratti, M., Checchi, D., Filippin, A., (2008), "Should you compete or cooperate with your schoolmates?" Institute for the Study of Labor, IZA Discussion Paper No. 3599.

Brunello, G., Checchi, D., (2007) "Does school tracking affect equality of opportunity? New

international evidence", *Economic Policy* 22 (52), pp. 781–861.

Brunello G. and M. Schlotter (2011), "Non Cognitive Skills and Personality Traits: Labour Market Relevance and their Development in Education and Training Systems", IZA Discussion Paper No. 5743.

Checchi D., M. Braga and E. Meschi (April 2012 Panel of Economic Policy), "Institutional Reforms and Educational Attainment in Europe: A Long Run Perspective", IZA DP No 6190.

Chingos M.M. and G.J. Whitehurst (2011), "Class Size: What Research Says and What it Means for State Policy" http://www.brookings.edu/research/papers/2011/05/11-class-size-whitehurst-chingos

Coleman, J. S. (1988). "Social capital in the creation of human capital". *American Journal of Sociology*, 94 (Supplement), pp. 95-120.

Dandy, J., and T. Nettelbeck (2002), "A cross-cultural study of parents' academic standards and educational aspirations for their children", *Educational Psychology*, 22: pp.621-627.

Davis-Kean PE. (2005), "The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment" *Journal of Family Psychology*, 19:294–304.

Desforges, C. and A. Abouchaar (2003), The Impact of Parental Involvement, Parental Support and Family Education on Pupil Achievements and Adjustment: A Literature Review, Department for Education and Skills, Nottingham.

DiMaggio, P. (1982), "Cultural capital and school success: The impact of status culture participation on the grades of U.S. High School Students", *American Sociological Review*, 47: pp. 189–201.

Dronkers, J. (1992), "Parents, love, and money: The relations between parental class, cognitive skill, educational attainment, occupation, marriage, and family income", *International Perspectives on Education and Society*", 2: pp. 277–293.

Dronkers J., R. van der Velden and A. Dunne (2012), "Why are migrant students better off in certain types of educational systems or schools than in others?" CREAM Discussion Paper No. 15/12.

Dunne, A. (2010), Dividing Lines: Examining the Relative Importance of between- and within-School Differentiation during Lower Secondary Education. Ph.D. European University Institute (Florence).

Dustmann C., T. Frattini and G. Lanzara (2012), "Educational achievement of second-generation immigrants: an international comparison" *Economic Policy*, CEPR & CES & MSH, vol. 27(69), pp. 143-185.

Figlio, David N., Marianne E. Page (2002), ''School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality?" *Journal of Urban Economics* 51(3): 497-514.

Fuchs, T., and L. Woessmann. (2007), "What accounts for international differences in student performance? A re-examination using PISA data." Empirical Economics 32, no. 2-3: pp. 433-462.

George, R., and Kaplan, D. (1998), "A structural model of parent and teacher influences on science attitudes of eighth graders: Evidence from NELS: 88", *Science Education*, 82, 93-109.

Greene, J. P. and M. Winters (2005) 'Public high school graduation and college readiness: 1991−2002", New York: Manhattan Institute for Policy Research.

Hanushek, E.A., (2006), "School Resources", In: Hanushek, E.A., Welch, F. (Eds.), Handbook of the Economics of Education. North Holland, pp. 865–908.

Hanushek E.A. and D.D. Kimko (2000), "Schooling, Labor Force Quality, and the Growth of Nations" *American Economic Review*, 90(5): pp. 1184-1208."

Hanushek, E.A. and Raymond, M.F (2003). "Lessons about the Design of State Accountability Systems", in: Paul E. Peterson and Martin R. West (eds). No Child Left Behind? The Politics and Practice of Accountability, Washington, DC: Brookings: pp. 127-51.

Hanushek, E.A. and Rivkin, S. G. (2006): "Teacher quality," in E. A. Hanushek, and F. Welch (eds) Handbook of the Economics of Education, vol. 1, pp. 1050–1078. North Holland, Amsterdam.

Hanushek, E.A., Woessmann, L. (2006), "Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries", *Economic Journal* 116 (510), C63–C76.

Hanushek E.A. and L. Woessmann (2011a), "The Economics of International Differences in Educational Achievement" in E.A. Hanushek, S. Machin and L. Woessmann (eds.), *Handbook in Economics*, Vol.3, The Netherlands: North-Holland, pp.89-200.

Hanushek E.A. and L. Woessmann (2011b), "Educational Outcomes" *Economic Policy*, July, pp. 427-91.

Hanushek E.A., S. Link and L. Woessmann (2011), "Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA" IZA Discussion Paper No. 6185.

Hoover-Dempsey, K. V., and Sandler, H. M. (1997). "Why do parents become involved in their children's education?" *Review of Educational Research*, 67(1), pp. 3-42.

Hunter J.E. and F.L. Schmidt (2004), "Methods of Meta-Analysis: Correcting Error and Bias in

Research Findings (Second Edition)", Newbury Park, CA: Sage.

Jenkins S.P., Micklewright, J., Schnepf, S.V. (2008) "Social segregation in secondary schools: How does England compare with other countries?" *Oxford Review of Education* 34(1): pp.21–38.

Jerrim J. and J. Micklewright (2012), "Socioeconomic gradients in children's cognitive skills: Are cross-country comparisons robust to who reports family background?" DoQSS – Institute of Education – Working Paper No. 12-06.

Jurges H., K. Schneider and F. Buchel (2005) "The effect of central exit examinations on student achievement: Quasi-experimental evidence from TIMSS Germany" Journal of the European Economic Association, Vol. 3, No. 5, pp.1134-1155

Lankford, Loeb and Wyckoff, J. (2002), "Teacher sorting and the plight of urban schools: A descriptive analysis". Educational Evaluation and Policy Analysis 24 (1), 37–62.

Machin S. and S. McNally, (2011) "The Evaluation of English Education Policies", CEE Discussion Paper No. 131.

Nechyba T.J. (2011) "Income and Peer Quality Sorting in Public and Private Schools" in E.A. Hanushek, S. Machin and L. Woessmann (eds.), Handbook of the Economics of Education, Vol.3, The Netherlands: North-Holland, Chapter 22.

Oakes J. (2005) "Keeping Track: How Schools Structure Inequality, second ed.", New Heaven: Yale University Press

OECD (2010a), "PISA 2009 at a Glance", Paris: OECD Publishing.

OECD (2010b), "PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science (Volume I)", Paris: OECD Publishing.

OECD (2010c), "PISA 2009 Results: Learning Trends: Changes in Student Performance Since 2000 (Volume V)", Paris: OECD Publishing.

OECD (2010d), "PISA 2009 Results: What Makes a School Successful? – Resources, Policies and Practices (Volume IV)", Paris: OECD Publishing.

OECD (2012a), "PISA 2009 Technical Report", Paris: OECD Publishing.

OECD (2012b) "Parental Involvement in Selected PISA Countries and Economies" OECD Education Working Paper No. 73 – working paper prepared by F. Borgonovi and G. Montt.

Okagaki, L., and P.A. Frensch (1998) "Parenting and children's school achievement: A multiethnic perspective", *American Educational Research Journal*, 35: pp.123-144.

Ou. D. (2010), "To Leave or Not to Leave? A Regression Discontinuity Analysis of the Impact of

Failing the High School Exit Exam" *Economics of Education Review*, 29(2):171-86.

Peterson, PE, and Hess, FM (2006) "Keeping an eye on state standards". Education Next 6(3), 28–29.

Peterson, PE, and Hess, FM (2008). Few states set world-class standards: In fact, most render the notion of proficiency meaningless. Education Next, 8(3), 70–73.

Peterson, P.E., Woessmann, L., (2007) "Introduction: Schools and the equal opportunity problem", In: Woessmann, L., Peterson, P.E. (Eds.), Schools and the Equal Opportunity Problem. MIT Press, pp. 3–27.

Peterson P.E, L. Woessmann, E.A. Hanushek and C.X. Lastra-Anadón (2011), "Globally Challenged: Are U. S. Students Ready to Compete? The latest on each state's international standing in math and reading", Harvard's PEPG Report No. 11-03.

Schoon, I., and Parsons, S. (2002), "Competence in the face of adversity: the influence of early family environment and long – term consequences", *Children and Society*, 16, 260-272.

Schuetz, G., Ursprung, H.W., Woessmann, L., (2008), "Education policy and equality of opportunity" Kyklos 61 (2): pp. 279–308.

Schwerdt G. and A.C. Wuppermann (2011): "Is Traditional Teaching Really All That Bad? A Within-Student Between-Subject Approach," *Economics of Education Review*, 30, pp. 365-379.

Singh, K., Bickley, P.G., Keith, T.Z., Keith, P.B., Trivette, P., and Anderson, E. (1995) "The effects of four components of parental involvement on eighth-grade student achievement: structural analysis of NELS-88 data", *School Psychology Review*, 24, 2, 299-317

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

Woessmann L. (2003b), "Central Exit Exams and Student Achievement: International Evidence". In: P.E. Peterson, M.R. West (eds.), No Child Left Behind? The Politics and Practice of School Accountability, Washington, D.C.: Brookings Institution Press, 292-323.

Woessmann L. (2005a). "The Effect Heterogeneity of Central Exams: Evidence from TIMSS, TIMSS-Repeat and PISA". Education Economics 13 (2): 143-169.

Woessmann L. (2005b) "Educational production in East Asia: The impact of family background and schooling policies on student performance", *German Economic Review*, 6 (3): pp.331–353; Kiel Working Paper No. 1152.

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

Woessmann L. and G. Schuetz (2006), "Efficiency and Equity in European Education and Training Systems", *Analytical Report Prepared for the European Commission*, Bruxelles: EU Commission.

Woessmann L., E. Ludemann, G. Schuetz and M. West (2007), "School Accountability, Autonomy, Choice, and the Level of Student Achievement: International Evidence from PISA 2003", OECD Education Working Papers No. 13.

Woessmann, L., E. Luedemann, G. Schuetz, and M.R. West. (2009) "School accountability, autonomy, and choice around the world". Cheltenham, UK: Edward Elgar.

TABLES AND FIGURES

Table 1: Test Scores and Analysis of Variance Decomposition of Test Scores: Country, Family Background and School Effects, PISA 2009

	Mean and Variation in Scores[1]		**Percentage of Variance in Scores due to Specified Factors[3]**		
Country Groups[7] (Number of countries and students in parenthesis**)**	**Average Test Score (Standard Deviation)**	**Coefficient of Variation[2]**	**Country Effects[4] (C)**	**Background Effects[5] (B)**	**School Effects[6] (S)**
All[8] (74; 515,956)	449.47 (104.74)	23.30	0.322	0.334	0.618
OECD (31; 249,125)	489.30 (94.54)	19.32	0.053	0.276	0.435
EU 15+[9] (17; 150,698)	498.24 (95.23)	19.11	0.024	0.291	0.475
Asian (7; 38,095)	540.28 (96.02)	17.77	0.026	0.143	0.502
Other High Income (8; 44,384)	465.71 (86.47)	18.57	0.021	0.191	0.348
Middle Income (27; 179,366)	388.28 (81.22)	20.92	0.076	0.258	0.503

Note: Sampling probability weights have been used in order to provide descriptive statistics representative for each country

[1] Test scores in mathematics here and in the remaining of the analysis unless otherwise specified.

[2] Coefficient of variation (CoV) is a normalized measure of dispersion calculated as the standard deviation divided by the mean score.

[3] Variance has been extracted as the share of test score variation explained by the model (ANOVA/R^2 In the underlying regression test score is a function of country dummies. The figure has been divided by test score variance.

[4] In the underlying regression test score is a function of country dummies. The figure has been divided by test score variance.

[5] In the underlying regression test score is a function of the family background measure, student-level and school-level controls. Here family background has been proxied by the ESCS index. The figure has been divided by test score variance. The ANOVA analysis has also been carried out by using the components of the ESCS unrestrictedly. Results are consistent with the analysis presented here and available from the authors.

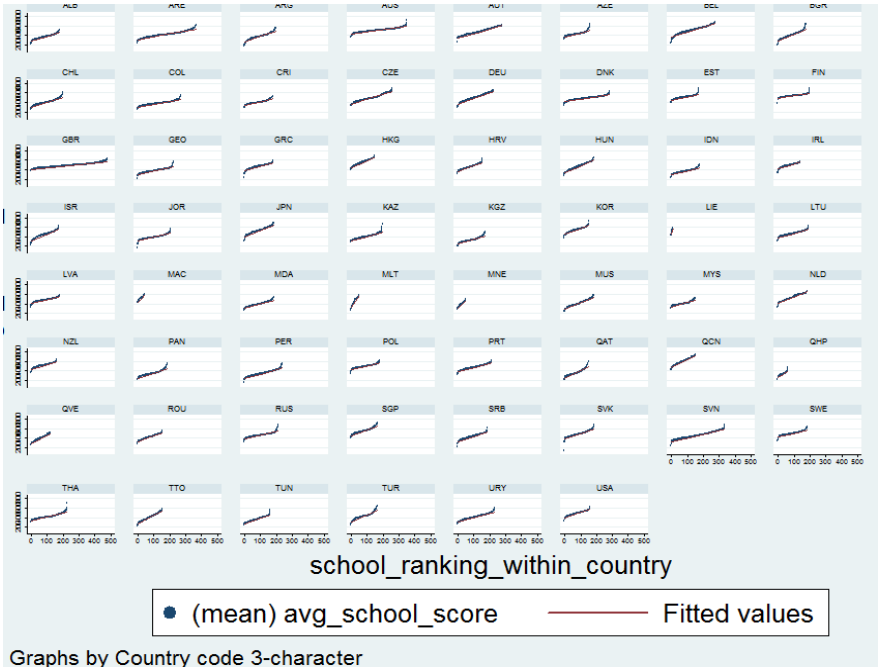
[6] In the underlying regression test score is a function of school dummies. The figure has been divided by test score variance

[7] Only a low-income country participated to the PISA 2009 study: Kyrgyzstan, so we did not include this group. The analysis has been carried out for this country and is available from the authors upon request.

[8] This group includes 72 countries and 2 Indian states.

[9] This category includes Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden and the United Kingdom, Norway and Switzerland.

Figure 1: School Rankings and Average Test Score, PISA 2009



Note**:** countries outliers in terms of number of schools (i.e., countries with more than 500 schools) have been omitted from this graph. These countries are: Brazil, Canada, Italy, Mexico and Spain. The graph related to these countries is available from the authors. Country codes are related to the following countries: ALB = Albania; ARE = United Arab Emirates; ARG = Argentina; AUS = Australia; AUT = Austria; AZE = Azerbaijan; BEL = Belgium; BGR = Bulgaria; CHE = Switzerland; CHL = Chile; COL = Colombia; CRI = Costa Rica; CZE = Czech Republic; DEU = Germany; DNK = Denmark; EST = Estonia; FIN = Finland; FRA = France; GBR = Great Britain; GEO = Georgia; GRC = Greece; HKG = Hong Kong; HRV = Croatia; HUN = Hungary; IDN = Indonesia; IRL = Ireland; ISL = Iceland; ISR = Israel; JOR = Jordan; JPN = Japan; KAZ = Kazakhstan; KGZ = Kyrgyzstan; KOR = Korea; LIE = Liechtenstein; LTU = Lithuania; LUX = Luxembourg; LVA = Latvia; MAC = Macao – China; MDA = Moldova; MLT = Malta; MNE = Montenegro; MUS = Mauritius; MYS = Malaysia; NLD = Netherlands; NOR = Norway; NZL = New Zealand; PAN = Panama; PER = Peru; POL = Poland; PRT = Portugal; QAT = Qatar; QCN = Shanghai – China; QHP = Himachal Pradesh - India; QTN = Tamil Nadu - India; QVE = Miranda – Venezuela; ROU = Romania; RUS = Russian Federation; SGP = Singapore; SRB = Serbia; SVK = Slovak Republic; SVN = Slovenia; SWE = Sweden; TAP = Chinese Taipei; THA = Thailand; TTO = Trinidad and Tobago; TUN = Tunisia; TUR = Turkey; URY = Uruguay; USA = United States

Table 2: Estimates on Effect of Country Tracking Practices on Gradient of Effect of Rank-School School-fixed-effects and Percentage of Country Variance in Test Scores, by Country Tracking Practices

Panel A: Country Level Regressions

	Coefficient from regression of gradient of school test score-school rank on tracking policies Country-level Regressions[2] (n=68)		**Coefficient of Percentage of country test score variance attributable to variance in scores across schools on tracking policies (n= 68)**	
Measure of Tracking in country[1]	**Single Variate Regression**	**Regression with Covariates[3]**	**Single Variate Regression**	**Regression with Covariates[3]**
School Tracking Policy Dummy=1 for Tracking	0.49 (0.37)	0.83** (0.36)	14.40*** (3.16)	7.07*** (2.58)
Tracking Index based on age	0.21** (0.09)	0.24** (0.01)	3.74*** (0.79)	2.54*** (0.78)

Panel B: School Level Regression

	Regression coefficient of school-fixed-effects on school rank, with country dummies[4] (number of schools =17,539)	
Country Group	**Single Variate regression**	**Regression with Covariates[5]**
School Tracking Policy = 0 [n. schools=6,288]	0.29*** (0.004)	0.20*** (0.004)
School Tracking Policy = 1[6] [n. schools=11,251]	0.44*** (0.006)	0.25*** (0.005)
Rank interacted with Tracking Index based on age * Rank[all schools]	0.015*** (0.002)	0.003** (0.001)

Note: [1] Analysis for 68 countries for which we have tracking information. Basic relations in the data similar to those with regression results for the entire sample

[2] The analysis is carried out at the country-level and the dependent variable is the country-specific coefficient on school rank presented in table A3

[3] These are country averages of: student's age, gender, grade, whether the language of test is the main language spoken at home, whether the student lives with two parents (natural or otherwise), immigrant status, parents' immigrant status, total school enrollment, dummies for community size where the school is located, index of socio-economic status (ESCS). Descriptive statistics related to these variables are presented in tables A1 and 6. Student sampling weights are included.

[4] The dependent variable is the average test score. The coefficient presented in the table is related to the school ranking. Test scores divided by 1,000 so coefficients are easier to read. Country dummies are included. School-level regressions are weighted by 1/number of school

[5] These are school averages of the control variables described in [3]

[6] School tracking is equal to 1 if age at first selection is below 16

Significance level (based on clustering-robust standard errors): * denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%

Table 3: Variance in Test Scores Within Different Types of Schools and Among Schools For Countries With different Proportions of Schools Having Specified Characteristics

VARIABLES	Share of Schools with Practice	**Variance in Test Scores among Students in Schools [school-level regression][1]** **(1)**	**Variance in Test Scores among Schools in Countries [country-level regression][2]** **(2)**
Tracking[3]	64.15	-0.17***	10.48***
		(0.009)	(2.658)
Parents' choice[4]	41.69	-0.11***	4.33
		(0.009)	(2.631)
Selective admissions[5]	31.43	-0.14***	13.88**
		(.009)	(6.303)
Selective residence[6]	35.87	0.12***	-17.23**
		(0.009)	(6.948)
Mean test score		0.001***	-0.06***
		(0.000)	(0.020)
Number of schools		-0.0002***	0.01
		(0.000)	(0.006)
Constant		7.94***	59.55***
		(0.029)	(9.554)
Observations		17,419	68
R-squared		0.117	0.489

[1] The natural logarithm (ln) of the variance among students in schools is regressed on the practice variable, average test score, number of schools in the country, standard errors clustered at the school level

[2] The between schools variance in test scores is regressed on the practice variable, average test score, number of schools in the country, standard errors clustered at the country level; and the percentages of country schools with specified practices

[3] This is related to *Age at First Selection*. Variable derived: 1=if age at first selection is below 16, 0=otherwise

[4] *Parents' Freedom to Choose a Public School for their Children*: question from the country-level data in PISA 2009 Results Volume 4, p.73 and related table IV.3.7, p.219:"Families are given a general right to enroll in any traditional public school they wish": scale: yes/no, Variable derived: 1= no, 0=yes

[5] This is related to *Selectivity by Student's Record of Academic Performance (including placement tests)*: question from the school questionnaire: "How often is student's record of academic performance (including placement tests) considered when students are admitted to your school? scale: never/sometimes/always. Variable derived: 1= always, 0=otherwise for school regression. Share of schools in the country that always use this form of selectivity in the country regression.

[6] This is related to *Selectivity by Residential Area*: question from the school questionnaire: How often is "residence in a particular area" considered when students are admitted to your school? scale: never/sometimes/always. Variable derived: 1= always, 0=otherwise in school regression. Share of schools in the country that always use this form of selectivity in country regression.

Significance level (based on clustering-robust standard errors): * denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%

Table 4 Regression Coefficients of Test Scores and Coefficient of variation of scores within schools on School Policies and Practices, PISA 2009[1]

"Measures of School Policies or Practices" [2]		all countries [n=68]	countries by tracking[3] no [n =21]	yes [n =47]	Coefficient of variation of scores within schools
		(1)	(2)	(3)	(4)
POLICIES					
SELECTIVITY	Academic selectivity --[index based on factors for admittance; higher = greater selectivity]	4.42***	2.81**	4.48***	-0.30**
		(0.662)	(1.192)	(0.914)	(0.145)
	Ability grouping between classes; [index derived from factor related to ability grouping	-3.07*	-2.98*	-3.18	0.24
		(1.582)	(1.632)	(1.974)	(0.145)
SCHOOL AUTONOMY AND ACCOUNTABILITY	Responsibility over curriculum and student assessment [index, mean=0, sd=1 for OECD]	-2.48	1.32	-3.54**	-0.06
		(1.650)	(2.378)	(1.747)	(0.075)
	Responsibility over school resource allocation; [index standardized: mean=0, sd=1 for OECD]	-2.10	-1.83	-2.34	0.15
		(1.300)	(2.795)	(1.519)	(0.119)
	Test info posted publicly [=1 if posted publicly]	3.38*	7.47***	2.05	0.04
		(1.951)	(1.773)	(2.406)	(0.309)
	Achievement data used to monitor teachers' performance [dummy variable=1 if use)	0.08	-5.70**	2.01	-0.11
		(2.242)	(2.238)	(2.923)	(0.311)
	Achievement data used to evaluate principal [dummy variable=1 if yes)	-3.30*	-0.23	-3.22	0.47
		(1.872)	(1.731)	(2.545)	(0.373)
	Teacher peer review[dummy variable=1 if teacher peer review for teaching/assessment	-0.80	-1.45	0.24	0.02
		(2.332)	(3.133)	(3.455)	(0.247)
	External observer [dummy variable=1 if external observers monitor teachers]	-2.47	-8.11*	-0.47	0.10
		(2.559)	(4.455)	(2.005)	(0.201)
RESOURCES, including teacher attributes	Quality of educational Resources – [index based on items that can affect instruction at school]	2.07	2.16**	2.21	-0.23**
		(1.247)	(0.769)	(1.722)	(0.091)
	Extra-curricular activities [index based on the number of extra-curricular activities offered]	1.28	2.86***	0.28	-0.25*
		(1.005)	(0.948)	(1.126)	(0.145)
	Student-teacher ratio=enrollment divided by the total number of teachers	-0.26**	-0.36**	-0.22*	0.02**
		(0.104)	(0.133)	(0.121)	(0.008)
	% of **certified** teachers [the number of certified teachers by total number teachers]	0.08	-10.70***	2.89	-0.24
		(3.200)	(3.360)	(3.102)	(0.216)
	% of **qualified** teachers [dividing teachers with ISCED 5A by total number teachers]	4.72	-8.80	8.88	0.02
		(5.928)	(9.938)	(6.501)	(0.518)
PRACTICES					
APPROACHES TO LEARNING	Memorization [index on the use of memorization as a learning strategy]	-32.31***	-26.82**	-33.59***	0.46
		(4.924)	(10.530)	(4.938)	(0.437)
	Elaboration [index on the use of elaboration as a learning strategy]	-18.57***	-16.65***	-18.72***	0.96**
		(2.889)	(4.364)	(4.547)	(0.472)
	Control of learning strategies[index on use of comprehension checks as a learning strategy]	55.44***	49.77***	59.34***	-1.41***
		(5.191)	(8.355)	(6.788)	(0.505)
STAFF ATTITUDE	Teachers' low expectations [dummy variable=1 if principal said detrimental to learning]	-1.70	-2.47	-0.04	-0.18
		(2.048)	(3.177)	(2.109)	(0.186)
	Staff resisting change [dummy variable=1 if principal said detrimental to learning]	2.44	11.64***	-1.49	-0.03
		(2.688)	(2.896)	(2.457)	(0.202)
	Constant	511.85***	328.23*	625.79***	49.18**
	Observations	18,340	6,247	11,005	18,340
	R-squared	0.757	0.869	0.718	0.168

Note: [1]All regressions at school-level. Covariates in each regression include country dummies. The following controls are included: school-level average student age, gender, grade, whether the language of test is main language spoken at home, whether student lives with two parents), immigrant status, parents' immigrant status, index of socio-economic cultural status, school enrollment, dummies for community size where the school is located. School sampling weights are used. France excluded as the school questionnaire was not administered. 134 schools have no school-level variance as they have only one student. [2] Full description of variables is in the appendix, section A4.1. Descriptive statistics are in Table A5.
[3] Robust standard errors clustered at the country-level in parentheses. Significance level: *** p<0.01, ** p<0.05, * p<0.1

Table 5: Estimated Coefficients (standard errors in parenthesis) for Regressions of School Fixed Effects on School Policies and Practices, with and without covariates and country fixed effects

VARIABLES[1]	(1)	(2)	(3)	(4)
	No covariates No country fe	No covariates Country fe	Covariates[2] No country fe	Covariates Country fe
Academic selectivity	12.01**	16.88***	7.67*	9.25***
	(5.558)	(2.944)	(4.120)	(1.690)
Ability grouping btw Classes	-5.26	-3.99*	-4.84*	-3.11*
	(3.226)	(2.134)	(2.572)	(1.569)
Responsibility over curriculum and assessment	4.66	-1.58	1.54	-2.38
	(2.837)	(2.508)	(2.464)	(1.649)
Responsibility over resource allocation	4.74*	5.00*	-2.51	-2.00
	(2.560)	(2.970)	(1.886)	(1.313)
Quality of educational Resources	16.80***	9.79***	4.83***	2.01
	(1.771)	(1.075)	(1.380)	(1.224)
Extra-curricular activities	7.88***	4.26***	2.94**	1.38
	(1.928)	(1.568)	(1.313)	(0.974)
Student-teacher-ratio	-0.90**	-0.28	-0.35**	-0.25**
	(0.351)	(0.198)	(0.154)	(0.100)
% of certified teachers	32.18***	-0.05	13.18**	-0.18
	(8.337)	(6.119)	(6.536)	(3.159)
% of qualified teachers	31.98***	16.19**	24.70**	4.31
	(11.378)	(6.894)	(10.025)	(5.945)
Test info posted publicly	18.91***	7.03**	7.74***	3.39*
	(4.330)	(2.980)	(2.573)	(2.019)
Achievement data used to evaluate teachers	-2.65	3.34	-1.96	0.29
	(4.346)	(2.839)	(3.266)	(2.260)
Achievement data used to evaluate principal	-1.73	-4.72*	-4.24	-3.41*
	(3.071)	(2.640)	(3.337)	(1.904)
Teacher-peer-review	-1.49	0.78	-0.98	-1.18
	(5.409)	(2.212)	(3.790)	(2.316)
External observer	-5.81	-7.65	-1.60	-2.47
	(5.844)	(5.184)	(3.224)	(2.554)
Memorization	-45.56***	-49.22***	-33.36***	-32.19***
	(8.295)	(6.978)	(6.122)	(4.890)
Elaboration	-54.93***	-26.23***	-25.70***	-18.52***
	(7.302)	(7.412)	(4.957)	(2.897)
Control of Learning Strategies	86.87***	88.89***	41.05***	54.93***
	(8.590)	(8.679)	(6.304)	(5.154)
Teachers' low expectations	-3.03	-4.23*	1.04	-1.71
	(4.135)	(2.288)	(3.338)	(2.118)
Staff resisting change	10.05***	5.10*	4.66	2.39
	(2.767)	(2.729)	(3.077)	(2.730)
Constant	-43.52***	-73.71***	336.17*	53.95
	(16.006)	(8.537)	(183.134)	(148.430)
Observations	18,340	18,340	18,340	18,340
R-squared	0.465	0.640	0.655	0.758

Note: [1] These are school-level regressions. The dependent variable is school fixed effects

[2] Covariates include these The following controls are included: school-level average student age, gender, grade, whether the language of test is main language spoken at home, whether student lives with two parents), immigrant status, parents' immigrant status, index of socio-economic cultural status, school enrollment, dummies for community size where the school is located. School sampling weights are used.

Significance level (based on clustering-robust standard errors): * denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%

Table 6: Mean of Family Background Measures, PISA 2009

Country Group[6]	**Years of Schooling**[1]	**Books at Home**[2]	**Index of Economic, Social and Cultural Status (ESCS)**[4]
	Mean	Mean Category[3]	Mean[5]
All	12.17 (3.63)	2.82 (1.42)	-0.54 (1.22)
OECD	13.11 (3.11)	3.16 (1.45)	-0.02 (1.02)
EU 15+[7]	13.45 (2.98)	3.32 (1.44)	0.04 (0.93)
Asian	13.56 (2.39)	3.45 (1.37)	-0.15 (0.82)
Other High Income	13.28 (1.77)	3.30 (1.36)	-0.19 (0.82)
Middle Income	10.77 (4.05)	2.26 (1.78)	-1.19 (1.22)

Note: Sampling probability weights have been used in order to provide descriptive statistics representative for each country

[1] Highest level of education of either parent according to the ISCED classification.

The first three columns related to parents' education show the share with completed primary education (ISCED 1); completed secondary education (ISCED 2 (lower secondary), ISCED Level 3B or 3C (vocational/pre-vocational upper secondary); ISCED 3A (upper secondary) and/or ISCED 4 (non-tertiary post-secondary)); completed tertiary education, that is theoretically oriented tertiary and post-graduate (ISCED 5A, 6). Parental education has been converted into years of schooling according to a country-specific scale provided in PISA 2009 (Vol. I, Table A1.1, p.168). The fourth column of this family background measure shows average years of schooling with standard deviation in parentheses

[2] Books at home shows the share of individuals according to the number of books they report having in their home.

[3] Mean of the books-at-home categories (1–6). These are the six categories: 1=1-10 books, 2=11-25 books, 3=26-100, 4=101-200, 5=201-500, 6=more than 500 books. Standard deviation in parentheses

[4] The ESCS index is constructed by using the following three sub-components: the higher parental occupation index, the higher parental education index and the index of home possessions

[5] Mean of the ECSC index. Standard deviation in parentheses

[6] Only a low-income country participated to the PISA 2009 study: Kyrgyzstan, so we did not include this group. The analysis has been carried out for this country and is available from the authors upon request

[7] This category includes Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden and the United Kingdom, Norway and Switzerland

Table 7: Regression coefficients of Student Test scores on ESCS Measure of Family Background PISA 2009[1]

Country Group [n. of observations[2]]	**Test Score on ESCS[3] (with *country fixed effects*)**	**R^2**	**Test Score on ESCS[3] *with other covariates[4], country fixed effects*)**	**R^2**	**Test Score on ESCS[3] with school fixed effects and other covariates[4],**	**R^2**	**Test Score on ESCS with Parental Involvement indicators[5] other covariates [4], *school fixed effects*)**	**R^2**
All [(1)-(3): 515,956;(4):101,370]	31.23*** (0.544)	0.42	24.20*** (0.497)	0.48	10.28*** (0.317)	0.66	9.39*** (0.51)	0.64
OECD [(1)-(3): 249,125; (4): 63,023]	39.52*** (0.80)	0.21	31.54*** (0.770)	0.32	16.05*** (0.571)	0.52	8.05*** (0.48)	0.66
EU15+ [(1)-(3): 150,696; (4): 48,106]	40.04*** (0.78)	0.18	27.89*** (0.757)	0.33	13.65*** (0.425)	0.56	7.58*** (0.56)	0.65
Asian [(1)-(3): 38,095; (4): 15,778]	38.19*** (1.68)	0.13	33.53*** (1.580)	0.17	8.21*** (0.898)	0.52	13.10*** (1.49)	0.45
Other High Income [(1)-(3): 44,384; (4): 18,600]	36.51*** (2.26)	0.14	27.31*** (2.206)	0.21	17.28*** (1.467)	0.44	12.48*** (1.03)	0.54
Middle Income [(1)-(3): 179,366; (4): 3,969]	24.78*** (0.77)	0.20	16.84*** (0.656)	0.32	5.32*** (0.334)	0.58	2.46*** (1.55)	0.63
Without school tracking[6] [(1)-(3): 171,885; (4): 21,356]	31.48*** (1.030)	0.41	25.62*** (0.899)	0.48	13.02*** (0.644)	0.62	15.80*** (1.07)	0.44
With school tracking [(1)-(3): 316,357; (4): 80,014]	31.09*** (0.609)	0.42	23.05*** (0.572)	0.49	8.53*** (0.309)	0.69	8.88*** (0.55)	0.65

Note: Regressions weighted by the students' sampling probabilities

[1] Estimates using parents' education and number of books at home available from the authors upon request

[2] Number of observations in (4) for the 14 countries that administered the parental involvement questionnaire. Regressions (1)-(3) estimated for the 14 country sample given similar results to those in the table. Results are available from the authors. Detailed descriptive statistics related to the parental involvement variables are available from the authors.

[3] The ESCS index is constructed from the following three sub-components: the higher parental occupation index, the higher parental education index and the index of home possessions

[4] Covariates: student age, gender, grade, whether the language of test is the main language spoken at home, whether the student lives with two parents, immigrant status, parents' immigrant status, school enrollment, dummies for community size where the school is located

[5] Parental involvement is measured with four indices extracted from the parental questionnaire described in the appendix: "parental involvement" "resources" "early support at ISCED1", "current support"

[6] The country-specific information on school tracking is available in the appendix, Table A3. The "without school tracking" group includes countries where age at first selection of students into different tracks is equal to 16 or higher. The group "with school tracking" includes countries where age at first selection is below Information on school tracking. The country-specific information is presented in Table A3

Significance level (based on clustering-robust standard errors): * denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%

APPENDIX

Table A1: Descriptive Statistics of the Control Variables, PISA 2009

	Students								Schools			
Country Group	**Female (%)**	**Age**	**Grade[1]**	**Immigrant Background (%)[2]**	**Mother Immigrant Background (%)**	**Father Immigrant Background (%)**	**Language of Test Spoken at Home (%)[3]**	**Live with Two Parents (%)[4]**	**School Size[5]**	**Village (%)[6]**	**Town (%)[7]**	**City (%)[8]**
All	50.39	15.77 (0.29)	-0.12 (0.59)	3.80	9.50	9.58	85.02	75.46	967.82 (812.80)	12.69	46.80	40.50
OECD	49.34	15.79 (0.29)	-0.07 (0.60)	8.22	17.09	17.32	90.18	78.45	995.22 (719.58)	9.19	53.60	37.21
EU15+	49.90	15.79 (0.29)	-0.16 (0.63)	6.85	16.20	16.45	90.04	81.88	761.56 (464.77)	5.63	66.53	27.83
Asian	48.29	15.75 (0.29)	-0.10 (0.34)	1.54	3.35	2.61	95.49	84.40	1159.59 (922.74)	0.26	23.77	75.97
Other High Income	50.29	15.80 (0.29)	0.15 (0.65)	7.29	17.53	17.97	90.17	71.78	559.60 (490.22)	19.47	35.6	44.93
Middle Income	51.90	15.77 (0.29)	-0.07 (0.87)	1.22	2.26	2.36	76.92	70.77	959.62 (880.66)	17.41	47.6	35.00

Note: mean value weighted by sampling probabilities; standard deviation (only for discrete variables)
[1] Grade is compared to modal grade in country (classified as 0)
[2] Students foreign-born, that is born in a country different from country of test
[3] Language of test spoken at home most of the time
[4] Live with two (natural or otherwise) parents
[5] Total school enrolment
[6] Community size where the school is located: a village, hamlet or rural area (fewer than 3,000 people)
[7] Community size where the school is located: either a small town (from 3,000 to about 15,000 people) or a town (from 15,000 to about 100,000 people)
[8] Community size where the school is located: either a city (from 100,000 to 1 million people) or a large city (with over 1 million people)

Table A2: Analytical Sample: Summary Statistics for Students and Schools, PISA 2009

Country	Students			Schools		
	No. Students	Average Test Score[1]	Coefficient of Variation	No. Schools	Variance Across Schools (in %)[2]	Variance Within Schools (in %)
Albania	4,596	377	23.94	181	39	61
Argentina	4,774	388	24.03	199	58	42
Australia	14,251	515	18.18	353	27	73
Austria	6,590	495	19.13	282	56	44
Azerbaijan	4,691	431	14.93	162	46	54
Belgium	8,501	516	20.18	278	60	40
Brazil	20,127	386	20.96	947	59	41
Bulgaria	4,507	428	23.3	178	49	51
Canada	23,207	526	16.56	978	27	73
Chile	5,669	421	19.05	200	53	47
Chinese Taipei	5,831	543	19.29	158	38	62
Colombia	7,921	381	20.01	275	39	61
Costa Rica	4,578	410	17.46	181	50	50
Croatia	4,994	461	19.15	158	41	59
Czech Republic	11,812	493	18.83	261	51	49
Denmark	5,924	503	17.2	285	20	80
Estonia	4,727	512	15.73	175	24	76
Finland	5,810	540	15.17	203	14	86
France	4,298	497	20.41	168	50	50
Georgia	4,646	380	22.64	226	30	70
Germany	4,979	512	19.23	226	62	38
Great Britain	12,179	493	17.61	482	33	66
Greece	4,969	465	19.22	184	39	61
Himal Pradesh – India	1,616	339	21.08	66	38	62
Hong Kong	4,837	555	17.13	151	43	57
Hungary	4,605	490	18.85	187	66	33
Iceland	3,646	507	17.85	131	17	83
Indonesia	5,136	371	18.94	183	46	54
Ireland	3,937	487	17.71	144	26	74
Israel	5,761	447	23.39	176	45	55
Italy	30,905	483	19.23	1,097	55	45
Japan	6,088	529	17.79	186	57	43
Jordan	6,486	387	21.65	210	44	56
Kazakhstan	5,412	406	20.63	199	46	54
Korea	4,989	546	16.24	157	43	57
Kyrgyzstan	4,986	331	24.35	173	42	58
Latvia	4,502	481	16.56	184	27	73
Liechtenstein	329	534	16.33	12	37	63
Lithuania	4,528	476	18.47	196	32	68
Luxembourg	4,622	488	20.16	39	34	66
Macao-China	5,952	525	16.26	45	23	77

Table A2 (cont'd): Analytical Sample: Summary Statistics for Students and Schools, PISA 2009

Country	Students			Schools		
	No. Students	Average Test Score[1]	Coefficient of Variation	No. Schools	Variance Across Schools (in %)[2]	Variance Within Schools (in %)
Malaysia	4,999	404	18.2	152	30	70
Malta	3,453	463	22.54	53	38	62
Mauritius	4,654	420	21.09	185	62	38
Mexico	38,250	419	18.88	1,535	47	53
Miranda-Venezuela	2,901	397	21.3	121	52	48
Moldova	5,194	397	21.61	186	38	62
Montenegro	4,825	403	20.82	52	31	69
Netherlands	4,760	526	16.97	186	68	32
New Zealand	4,643	520	18.62	163	31	69
Norway	4,660	498	17.18	197	16	84
Panama	3,969	360	22.21	188	60	40
Peru	5,985	365	24.47	240	60	40
Poland	4,917	494	17.78	185	21	79
Portugal	6,298	487	18.73	214	32	68
Qatar	9,078	368	26.46	153	54	46
Romania	4,776	426	18.58	159	51	49
Russian Federation	5,308	468	18.17	213	31	69
Serbia	5,523	443	20.63	190	45	55
Shanghai-China	5,115	600	17.1	152	42	58
Singapore	5,283	563	18.61	171	36	64
Slovak Republic	4,555	497	19.25	189	44	66
Slovenia	6,155	501	19.04	341	60	40
Spain	25,887	484	18.68	889	22	78
Sweden	4,567	494	19.01	189	24	76
Switzerland	11,812	535	18.62	426	35	65
Tamil Nadu – India	3,210	350	19.54	147	49	51
Thailand	6,225	419	18.96	230	41	59
Trinidad and Tobago	4,778	414	23.75	158	67	33
Tunisia	4,955	371	20.92	165	47	53
Turkey	4,996	446	20.94	170	71	29
United Arab Emirates	10,867	421	21.94	369	49	51
United States	5,233	487	18.53	165	39	61
Uruguay	5,957	427	21.24	232	46	54

Note:
[1] Descriptive statistics are shown for mathematics. Comparable statistics for science and reading are available from the authors upon request. The mathematics literacy scale was introduced in PISA 2003 (with a mean of 500 and a standard deviation of 100). Scores in PISA 2009 are compared to this benchmark.
[2] Variance in test score between and within schools has been calculated by regressing test score on school dummies using the school weight and by double-checking the results using the ANOVA method

Table A3: Schools, School Ranking, Family Background and Test Scores, PISA 2009

Country	**No. Schools[1]**	**Coefficient on School Rank[2]**	**Coefficient on Family Background (with controls) [3]**	**Coefficient on Family Background (with controls, school fixed effects)[4]**	**Tracking (first age)[5]**
Albania	181	0.994*** (0.023)	18.768*** (2.195)	12.001*** (1.589)	14
Argentina	199	1.105*** (0.028)	26.780*** (2.428)	9.443*** (1.207)	15
Australia	353	0.413*** (0.011)	40.447*** (1.712)	26.012*** (1.332)	16
Austria	282	0.887*** (0.008)	33.389*** (2.648)	7.967*** (1.458)	10
Azerbaijan	162	0.870*** (0.060)	7.637*** (1.632)	5.387*** (1.153)	15
Belgium	278	1.005*** (0.032)	27.108*** (1.484)	9.986*** (1.091)	12
Brazil	947	0.172*** (0.003)	17.760*** (1.715)	3.158*** (0.702)	17
Bulgaria	178	1.487*** (0.046)	31.682*** (2.571)	16.736*** (1.704)	13
Canada	978	0.154*** (0.002)	27.509*** (1.239)	20.951*** (0.992)	16
Chile	200	0.981*** (0.027)	25.848*** (2.193)	6.782*** (1.050)	13
Chinese Taipei	158	1.408*** (0.037)	35.838*** (2.790)	16.611*** (1.843)	15
Colombia	275	0.505*** (0.013)	17.926*** (1.701)	8.701*** (1.005)	15
Costa Rica	181	0.698*** (0.027)	14.302*** (1.336)	4.795*** (0.748)	.
Croatia	158	1.149*** (0.018)	26.981*** (2.214)	10.004*** (1.775)	14-15
Czech Republic	261	0.911*** (0.018)	42.620*** (2.760)	16.579*** (1.531)	11
Denmark	285	0.484*** (0.010)	29.743*** (1.634)	24.427*** (1.529)	16
Estonia	175	0.679*** (0.026)	21.443*** (2.345)	13.745*** (1.647)	15
Finland	203	0.475*** (0.012)	24.055*** (1.664)	21.852*** (1.716)	16
France	168	1.489*** (0.038)	36.200*** (3.373)	16.636*** (1.579)	15
Georgia	226	0.722*** (0.025)	23.615*** (1.896)	17.903*** (1.402)	.
Germany	226	1.184*** (0.021)	24.666*** (1.963)	7.612*** (1.022)	10
Great Britain	482	0.313*** (0.004)	40.696*** (1.761)	25.220*** (1.453)	16
Greece	184	0.905*** (0.034)	28.696*** (2.066	17.116*** (1.374)	15
Hong Kong	151	1.383*** (0.030)	15.522*** (2.083)	1.863 (1.390)	16

Table A3 (cont'd): Schools, School Ranking, Family Background and Test Scores, PISA 2009

Country	**No. Schools**[1]	**Coefficient on School Rank**[2]	**Coefficient on Family Background (with controls)** [3]	**Coefficient on Family Background (with controls, school fixed effects)**[4]	**Tracking (first age)**[5]
Himal Pradesh – India	66	2.268*** (0.116)	8.962*** (2.542)	4.422*** (1.439)	16
Hungary	187	1.355*** (0.025)	38.355*** (2.641)	7.560*** (1.310)	11
Iceland	131	0.960*** (0.050)	25.108*** (1.742)	23.920*** (1.573)	16
Indonesia	183	0.870*** (0.028)	11.264*** (1.737)	3.440*** (1.110)	15
Ireland	144	0.943*** (0.054)	33.224*** (1.830)	26.313*** (1.695)	15
Israel	176	1.294*** (0.033)	36.325*** (2.387)	16.847*** (1.632)	15
Italy	1,097	0.205*** (0.002)	19.382*** (1.481)	3.346*** (0.660)	14
Japan	186	1.327*** (0.032)	28.140*** (3.335)	0.843 (1.468)	15
Jordan	210	0.765*** (0.023)	21.344*** (2.083)	14.373*** (1.150)	16
Kazakhstan	199	0.931*** (0.037)	27.282*** (2.591)	14.144*** (1.538)	.
Korea	157	1.065*** (0.035)	32.017*** (3.416)	15.342*** (1.654)	14
Kyrgyzstan	173	0.936*** (0.026)	19.795*** (1.871)	11.957*** (1.164)	14-15
Latvia	184	0.717*** (0.019)	24.393*** (1.942)	17.936*** (1.743)	16
Liechtenstein	12		9.710 (6.063)	7.189 (5.416)	11
Lithuania	196	0.858*** (0.020)	30.518*** (1.816)	20.476*** (1.381)	14-15
Luxembourg	39	4.571*** (0.350)	23.311*** (2.287)	12.797*** (2.295)	13
Macao-China	45	3.316*** (0.252)	6.099** (2.611)	0.771 (1.255)	16
Malaysia	152	0.840*** (0.036)	24.587*** (1.820)	14.951*** (1.297)	18
Malta	53	4.300*** (0.276)	32.687*** (3.332)	15.465*** (2.650)	11
Mauritius	185	1.396*** (0.028)	22.067*** (1.652)	5.616*** (1.027)	.
Mexico	1,535	0.105*** (0.001)	13.599*** (0.821)	4.823*** (0.482)	12
Miranda – Venezuela	121	1.707*** (0.037)	30.887*** (2.381)	9.127*** (1.612)	.

Table A3 (cont'd): Schools, School Ranking, Family Background and Test Scores, PISA 2009

Country	No. Schools[1]	Coefficient on School Rank[2]	Coefficient on Family Background (with controls) [3]	Coefficient on Family Background (with controls, school fixed effects)[4]	Tracking (first age)[5]
Moldova	186	0.959*** (0.023)	20.741*** (1.593)	14.964*** (1.187)	17
Montenegro	52	3.026*** (0.144)	24.842*** (2.477)	9.928*** (1.442)	15
Netherlands	186	1.219*** (0.028)	23.490*** (2.191)	5.383*** (1.166)	12
New Zealand	163	0.895*** (0.030)	47.646*** (2.132)	37.599*** (1.629)	16
Norway	197	0.535*** (0.017)	33.496*** (1.849)	31.143*** (1.812)	16
Panama	188	0.923*** (0.035)	14.517*** (2.459)	2.800 (1.470)	15-16
Peru	240	0.847*** (0.023)	20.103*** (1.737)	8.424*** (0.898)	14-15
Poland	185	0.627*** (0.021)	32.158*** (1.683)	28.193*** (1.457)	15
Portugal	214	0.834*** (0.023)	18.695*** (1.218)	15.009*** (0.901)	15
Qatar	153	1.306*** (0.061)	11.872*** (2.416)	4.148*** (1.081)	15
Romania	159	1.157*** (0.019)	28.060*** (2.382)	12.840*** (1.511)	14
Russian Federation	213	0.703*** (0.023)	27.441*** (2.681)	18.034*** (1.741)	14.5
Serbia	190	1.012*** (0.022)	27.760*** (2.279)	9.437*** (1.313)	.
Shanghai-China	152	1.606*** (0.024)	29.956*** (2.896)	2.407 (1.552)	14
Singapore	171	1.067*** (0.040)	36.465*** (2.340)	22.775*** (1.774)	10
Slovak Republic	189	1.006*** (0.041)	36.567*** (2.653)	17.960*** (1.595)	11
Slovenia	341	0.705*** (0.008)	27.765*** (2.414)	4.431*** (1.243)	11
Spain	889	0.157*** (0.003)	14.645*** (1.322)	10.893*** (0.847)	16
Sweden	189	0.737*** (0.026)	37.783*** (1.924)	30.040*** (1.741)	16
Switzerland	426	0.417*** (0.008)	28.538*** (1.600)	17.715*** (1.399)	15
Tamil Nadu – India	147	0.967*** (0.034)	6.698*** (2.891)	-1.120 (1.397)	16
Thailand	230	0.735*** (0.021)	17.013*** (1.876)	5.760*** (1.081)	15
Trinidad and Tobago	158	1.685*** (0.035)	20.229*** (2.604)	-1.657 (1.143)	11

Table A3 (cont'd): Schools, School Ranking, Family Background and Test Scores, PISA 2009

Country	No. Schools[1]	Coefficient on School Rank[2]	Coefficient on Family Background (with controls) [3]	Coefficient on Family Background (with controls, school fixed effects)[4]	Tracking (first age)[5]
Tunisia	165	1.129*** (0.036)	10.100*** (1.608)	4.275*** (0.835)	16
Turkey	170	1.255*** (0.067)	27.206*** (2.443)	7.925*** (0.819)	11
United Arab Emirates	369	0.566*** (0.010)	24.709*** (2.058)	12.582*** (1.207)	15
United States	165	0.953*** (0.032)	33.493*** (2.022)	19.867*** (1.629)	16
Uruguay	232	0.871*** (0.022)	17.840*** (1.305)	9.415*** (0.882)	12

Note: Regressions weighted by sampling probabilities

[1] Total number of schools in the country

[2] The coefficients have been extracted from the country-specific regressions run at school-level. The dependent variable is average test score; the independent variables include school ranking (i.e., national ranking: ranking of the school with respect to the other schools in the country), family background, student-level and school-level controls, standard errors clustered at school level

[3]Family background is proxied in this and the next two columns by the index of socio-economic and cultural status (escs). The same analysis has been carried out for the other two family background variables: parents' education and books at home and results are consistent; The dependent variable is student test score; the independent variables include family background, standard errors clustered at school level

[4]The dependent variable is student test score; the independent variables include family background, student-level and school-level controls, standard errors clustered at school level

[5] First age at which selection takes place in the education system. This information refers to the policy set at the national level. Several sources have been used to extract this information including Prof. Woessmann who kindly shared his dataset with us that was constructed through detailed country-specific inquiries (OECD 2005, Education at a Glance, Table D6.1; OECD 2010d Table IV 3.2a; Hanushek and Woessmann 2006; Brunello and Checchi 2008, Schuetz et al. 2008; Woessmann 2005b; country-specific sources for Malta and India). "." denotes information not available.

Significance level (based on clustering-robust standard errors): * denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%

A4.1: School Policies and Practices, PISA 2009

The list below describes the policies and practices used in this study that have been extracted from the PISA database (http://pisa2009.acer.edu.au). Complete information can be found in the School, Student Questionnaires, the related Codebooks and PISA 2009 Technical Report (OECD 2010d, 2012a).

POLICIES

School Autonomy

The **Index of the Relative Level of Responsibility of School Staff Relating to Curriculum and Assessment** was computed from four items measuring the school principal's report concerning who had responsibility for curriculum and assessment ("Establishing student assessment policies", "Choosing which textbooks are used", "Determining course content", and "Deciding which courses are offered"). The index was calculated on the basis of the ratio of "yes" responses for principal or teachers to "yes" responses for regional/local education authority or national educational authority. Higher values indicate relatively higher

levels of school responsibility in this area. The index was standardized to having an OECD mean of zero and a standard deviation of one

The **Index of the Relative Level of Responsibility of School Staff in Allocating Resources** was derived from six items measuring the school principals' report on who has considerable responsibility for tasks regarding school management of resource allocation ("Selecting teachers for hire", "Firing teachers", "Establishing teachers' starting salaries", "Determining teachers' salaries increases", "Formulating the school budget", "Deciding on budget allocations within the school"). The index was calculated on the basis of the ratio of "yes" responses for principal or teachers to "yes" responses for regional/local education authority or national educational authority. Higher values on the scale indicate relatively higher levels of school responsibility in this area. The index was standardized to having an OECD mean of 0 and a standard deviation of 1

Resources

The **Index on the School's Educational Resources** was computed on the basis of seven items measuring the school principal's perceptions of potential factors hindering instruction at school. All items were inverted for IRT scaling and positive WLE scores indicate better quality of educational resources. The following items are included:

SC11Q07 Shortage or inadequacy of science laboratory equipment
SC11Q08 Shortage or inadequacy of instructional materials (e.g. textbooks)
SC11Q09 Shortage or inadequacy of computers for instruction
SC11Q10 Lack or inadequacy of Internet connectivity
SC11Q11 Shortage or inadequacy of computer software for instruction
SC11Q12 Shortage or inadequacy of library materials
SC11Q13 Shortage or inadequacy of audio-visual resources

The **Index of Extra-Curricular Activities Offered by School** is based school principals reports what extra-curricular activities occur at their school. Responses to the items were coded such that positive WLE scores indicate higher levels of extra-curricular school activities. The following items are included:

SC13Q01 Band, orchestra or choir
SC13Q02 School play or school musical
SC13Q03 School yearbook, newspaper or magazine
SC13Q04 Volunteering or service activities, e.g. <national examples>
SC13Q05 Book club
SC13Q06 Debating club or debating activities
SC13Q07 School club or school competition for foreign language, math or science
SC13Q08 <Academic club>
SC13Q09 Art club or art activities
SC13Q10 Sporting team or sporting activities
SC13Q11 Lectures and/or seminars (e.g. guest speakers such as writers or journalists)
SC13Q12 Collaboration with local libraries
SC13Q13 Collaboration with local newspapers

The **Student-Teacher Ratio** is enrollment by the total number of teachers. The number of part-time teachers is weighted by 0.5 and the number of full-time teachers is weighted by 1.0

Teacher characteristics

The **Proportion of Fully Certified Teachers** was computed by dividing the number of fully certified teachers by the total number of teachers

The **Proportion of Teachers with a University Educatio**n (ISCED 5A qualification) is the number of these kinds of teachers by the total number of teachers

Selectivity

The **Index of Ability Grouping between Classes** was derived from the two items by assigning schools to three categories: (1) schools with no ability grouping for any subjects, (2) schools with one of these forms of ability grouping between classes for some subjects and (3) schools with one of these forms of ability grouping for all subjects. [One item asked about the occurrence of ability grouping into different classes and the other

regarding ability grouping within classes (with the response categories "For all subjects", "For some subjects" and "Not for any subject")]

The **Index of Academic School Selectivity** was computed by assigning schools to three different categories: (1) schools where neither of these two factors is considered for student admittance, (2) schools considering at least one of these two factors, (3) schools where at least one of these two factors is a prerequisite for student admittance. [Principals were asked how much consideration was given to the following factors when students are admitted to the school, based on a scale with the categories "not considered", "considered", "high priority", and "pre-requisite": students' academic record (including placement tests) and the recommendation of feeder schools]

Accountability

Achievement data are posted publicly (e.g. in the media) is a dummy variable =1 if school principals said achievement data was posted publicly (SC22Q01) during the past year.

Achievement data are used in evaluation of the principal's performance is a dummy variable =1 if school principals said achievement data (SC22Q02) was used to monitor the principal's performance during the past year.

Achievement data are used in evaluation of teachers' performance is a dummy variable =1 if school principals said achievement data (SC22Q03) was used to monitor teachers' performance during the past year.

Teacher peer-review is a dummy variable =1 if school principals said teacher peer review (of lesson plans, assessment instruments, lessons) (SC23Q02) was used to monitor teachers during the past year.

External observers is a dummy variable =1 if school principals said external observers (SC23Q04) were used to monitor teachers during the past year.

PRACTICES

Approaches to learning

I**ndex of Memorization.** Positive WLE scores on these indices indicate higher importance attached to the given reading strategy. Four items measuring the construct of memorization were included in the PISA 2009 main study. There are four response categories varying from "almost never", "sometimes", "often" to "almost always". Positive WLE scores on a given learning strategy index indicate greater use of that learning strategy. The following items are included:

ST27Q01 When I study, I try to memorize everything that is covered in the text
ST27Q03 When I study, I try to memorize as many details as possible
ST27Q05 When I study, I read the text so many times that I can recite it
ST27Q07 When I study, I read the text over and over again

Index of Elaboration. Positive WLE scores on these indices indicate higher importance attached to the given reading strategy. Four items measuring the construct of elaboration were included in the PISA 2009 main study. There are four response categories varying from "almost never", "sometimes", "often" to "almost always". Positive WLE scores on a given learning strategy index indicate greater use of that learning strategy. The following items are included:

ST27Q04 When I study, I try to relate new information to prior knowledge acquired in other subjects
ST27Q08 When I study, I figure out how the information might be useful outside school
ST27Q10 When I study, I try to understand the material better by relating it to my own experiences
ST27Q12 When I study, I figure out how the text information fits in with what happens in real life

Index of Control Learning Strategies. Positive WLE scores on these indices indicate higher importance attached to the given reading strategy. Five items measuring the construct of control learning strategies were included in the PISA 2009 main study. There are four response categories varying from "almost never", "sometimes", "often" to "almost always". Positive WLE scores on a given learning strategy index indicate greater use of that learning strategy. The following items are included:

ST27Q02 When I study, I start by figuring out what exactly I need to learn
ST27Q06 When I study, I check if I understand what I have read
ST27Q09 When I study, I try to figure out which concepts I still haven't really understood

ST27Q11 When I study, I make sure that I remember the most important points in the text
ST27Q13 When I study and I don't understand something, I look for additional information to clarify this

Staff Attitude

Teachers' Low Expectations as detrimental to the learning of students is a dummy variable =1 if school principals said it was to some extent or a lot detrimental based on: "not at all", "very little", "to some extent" "a lot".

Staff Resisting Change as detrimental to the learning of students is a dummy variable =1 if school principals said to some extent or a lot based on: "not at all", "very little", "to some extent" "a lot".

A4.2: Parental Involvement, PISA 2009

The list below provides a description of the variables related to parental involvement used in this study that have been extracted from the PISA database (http://pisa2009.acer.edu.au). Complete information can be found in the Parent Questionnaire, the related Codebook and PISA 2009 Technical Report (OECD 2010d, 2012a).

PARENTAL INVOLVEMENT

The **Index of the Early Activities with the Child when the Child attended the First Year of ISCED 1** (PRESUPP) is related to parental support of child's reading literacy. The index is based on the information provided by parents and constructed using an IRT scaling method. Parents are asked when their child attended the first year of ISCED 1, how often they (or someone else at home) undertook specific activities with child. Eight items (each one is related to an activity) are included: "read books", "tell stories", "tell songs", "play with alphabet toys", "talk about what you had read", "play word games", "write letter or words", "read aloud signs and labels". Possible answers include: never or hardly ever; once or twice a month; once or twice a week; every day or almost every day. Positive WLE scores on this index indicate greater parental support of child's reading literacy at the beginning of ISCED 1.

The **Index of the Current Activities with the Child** (CURSUPP) is related to parental support in home activities. The index is based on the information provided by parents and constructed using an IRT scaling method. Parents are asked how often they (or someone else in your home) undertake specific activities with their child. Six items (each one is related to an activity) are included: "discuss political or social issues", "discuss books", "films or television programs", "discuss how well your child is doing at school", "go to a bookstore or library with your child", "talk with your child about what he/she is reading on his/her own", "help your child with his/her homework". Possible answers include: never or hardly ever; once or twice a month; once or twice a week; every day or almost every day. Positive WLE scores on this index indicate greater parental support of child's reading literacy.

The **Index of Parental Involvement** (PARINVOL) is related to parents' participation in school-related activities. The index is based on the information provided by parents and constructed using an IRT scaling method. Parents are asked whether the last academic year they participated in any of the specified school-related activities. Possible answers include: yes; no. Eight items (each one is related to an activity) are included: "discuss your child's behavior or progress with a teacher on your own initiative", "discuss your child's behavior or progress on the initiative of one of your child's teachers", "volunteer in physical activities, e.g. building maintenance, carpentry, gardening or yard work", "volunteer in extra-curricular activities, e.g. book club, school play, sports, field trip", "volunteer in the school library or media center", "assist a teacher in the school", "appear as a guest speaker", "participate in local school government (e.g. parent counsel or school management committee)". Positive WLE scores on this index indicate greater parental involvement in their child's school.

The **Index of Resources Available to the Child** (READRES) is related to resources available to the child at home. The index is based on the information provided by parents and constructed using an IRT

scaling method. Parents are asked whether the specified items are available to their child in their home. Possible answers include yes; no. Six items are included: “email”, “chat on line/MSN”, “internet connection”, “daily newspaper”, “a subscription to a journal or magazine”, “books of his/her very own (do not count school books)”. Positive WLE scores on this index indicate greater availability of reading resources at home.

Table A5: School Policies and Practices, PISA 2009

A] POLICIES

Country	School Autonomy		Resources			Selectivity	
	Responsibility over curriculum and assessment[1]	Responsibility over resource allocation[1]	Teacher-student ratio[1]	Quality of educational resources[1]	Extra-curricular activities[1]	Academic Selectivity[2] (%)	Ability Grouping between Classes[3] (%)
All	-0.18 (0.98)	-0.12 (0.93)	17.77 (11.92)	-0.28 (1.20)	0.36 (1.03)	41.45	18.97
OECD	-0.14 (0.96)	0.05 (0.98)	14.82 (5.98)	0.14 (1.08)	0.54 (0.92)	33.46	11.36
EU15+	0.04 (0.96)	-0.10 (1.00)	12.81 (4.79)	0.10 (0.93)	0.21 (0.89)	36.43	12.16
Asian	0.81 (0.83)	-0.17 (0.98)	14.62 (5.22)	0.36 (1.02)	0.52 (0.92)	71.82	8.42
Other High Income	-0.37 (0.80)	-0.09 (0.70)	11.86 (4.51)	-0.57 (0.96)	0.70 (0.88)	26.52	36.23
Middle Income	-0.41 (0.92)	-0.27 (0.88)	22.24 (15.94)	-0.76 (1.17)	0.10 (1.12)	43.83	25.97

Country	Teachers		Accountability				
	Share of qualified teachers	Share of certified teachers	Achievement data posted publicly (% yes)	Achievement data used to evaluate principal (% yes)	Achievement data used to evaluate teachers (% yes)	Teacher peer-review (% yes)	External observers (% yes)
All	82.26	75.48	42.59	50.33	60.49	66.32	44.97
OECD	83.67	82.81	53.03	44.29	42.92	50.00	34.47
EU15+	69.22	87.68	26.44	26.65	33.06	42.11	24.38
Asian	96.05	97.16	16.38	18.07	34.45	60.21	36.41
Other High Income	86.90	94.79	70.42	76.79	93.84	93.19	68.86
Middle Income	77.07	60.63	33.99	59.47	78.59	79.68	53.40

Note: Sampling probability weights have been used in order to provide descriptive statistics representative for each country. Policy and Practice variables are described in the appendix, section A4.1.

[1]Mean of the index and standard deviation (in parentheses) [2]At least one factor is considered for school admittance (see variable description) [3] for all subjects.

Table A5 (cont'd): School Policies and Practices, PISA 2009

B] PRACTICES

Country	Approaches to Learning			Staff Attitude	
	Memorization[1]	Elaboration[1]	Control of learning strategies[1]	Teachers' low expectations[3] (% yes)	Staff resisting change[3] (% yes)
All	0.09 (0.97)	0.10 (0.98)	-0.04 (0.97)	27.98	29.14
OECD	0.02 (1.00)	-0.01 (1.02)	0.05 (1.02)	23.76	31.01
EU15+	0.02 (0.94)	-0.04 (0.96)	0.07 (0.96)	17.06	24.41
Asian	-0.34 (1.05)	-0.31 (1.09)	-0.41 (1.02)	31.17	36.18
Other High Income	0.22 (0.85)	0.09 (0.93)	-0.11 (0.86)	37.68	33.24
Middle Income	0.24 (0.89)	0.30 (0.88)	-0.03 (0.90)	29.88	25.06

Note: Sampling probability weights have been used in order to provide descriptive statistics representative for each country. Policy and Practice variables are described in the appendix, section A4.1.
[1]Mean of the index and standard deviation (in parentheses).
[3] Perceived by the school principal as detrimental to the learning of students.