

WPS3537

Families, Schools, and Primary-School Learning: Evidence for Argentina and Colombia in an International Perspective

Ludger Wößmann

with the help of Thomas Fuchs

Ifo Institute for Economic Research
at the University of Munich
and CESifo
Poschingerstr. 5
81679 Munich, Germany
Phone: (+49) 89-9224 1699
E-mail: woessmann@ifo.de
Internet: www.ifo.de/link/hs_e.htm

JEL Classification: I21, J24, O54

Keywords: Education production function, primary schools, international student achievement test, PIRLS, Argentina, Colombia, Latin America

World Bank Policy Research Working Paper 3537, March 2005

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent. Policy Research Working Papers are available online at <http://econ.worldbank.org>.

* We gratefully acknowledge financial support from the World Bank for the project “Using International Achievement Results to Improve Public Sector Management of the Education System”. We also appreciate the constructive comments and discussions of Harry Patrinos, Emiliana Vegas and the other participants at the Argentina Skills Study workshop of the World Bank in Washington, D.C.

1. Introduction

The quality of education is of prime importance for the economic fates of individuals and nations alike. Recent research has emphasized that the returns to early learning are particularly large, both because young people's time horizons to recoup their educational investments are longer and because skills acquired early on facilitate later learning in a dynamic synergistic process of skill formation (cf. Heckman 2000; Carneiro and Heckman 2003). A particular element in this synergistic process is the ability to read, because basic reading abilities are a crucial means to acquire further knowledge in all other areas. A general pedagogical assessment is that by the end of fourth grade, children should have learned how to read and start to read to learn (cf. Martin et al. 2003). Lacking reading literacy by the end of primary school thus diminishes the future possibilities of students to accumulate human capital in other subjects and harms their career opportunities in the society and the economy. It therefore does not come as a surprise that empirical evidence consistently shows that rates of return to education are highest at the primary level (Psacharopoulos and Patrinos 2004).¹ This gives vital relevance to policy questions of how to improve primary-school learning.

This paper estimates empirically the importance of family features and school policies for learning achievement in primary school in two Latin American countries, Argentina and Colombia. It has long been realized that education is one of the key challenges for government policy in Latin America in the era of government withdrawal due to economic liberalization and privatization (cf., e.g., Edwards and Baer 1993). Furthermore, the supply of skilled workers through widespread high-quality education, rather than the liberalization reforms, can be expected to affect the long-run income distribution in Latin America (Baer and Maloney 1997). Therefore, it is a cause for worry that the educational performance of Latin America over recent decades has been dismal (cf. Mizala and Romaguera 2002; Pritchett 2004). A focus on the basic foundations of learning in Latin American societies thus seems particularly warranted.

While a few empirical studies of the determinants of educational quality in individual Latin American countries exist, a drawback often consists in the lack of a benchmark for estimated parameters. For example, finding a statistically significant relationship between parental education and student performance does not yet answer the question of whether such

¹ It has also be argued that the basic literacy skills generated by primary schooling may carry particularly widespread positive externalities to the population at large (cf. Psacharopoulos 1996), e.g. in terms of the basic functioning and stability of societies.

a relationship is particularly strong or rather weak. In order to address this issue, this paper uses data from an internationally comparable student achievement test which provides data for Argentina, Colombia and selected comparison countries, thereby allowing for a benchmarking of estimated parameters relative to other countries.

Specifically, we use data from the most recent international achievement study, the Progress in International Reading Literacy Study (PIRLS), which tested the reading performance of fourth-grade students in 2001, with data released in 2003. A uniform data generation process for representative samples of students in all participating countries ensures comparability across countries. Latin America's poor educational performance is well documented in PIRLS: Argentina and Colombia ranked 31st and 30th, respectively, out of a total of 35 participating countries,² while at the same time featuring relatively large inequality among their students.³

The PIRLS study not only provides comparable information on the reading performance of students, but also extensive information on student and school background to an extent not generally available in student achievement studies. This allows for an analysis of the determinants of learning both in the realm of the family and of the school. The available information stems from background questionnaires answered, respectively, by students, teachers, school principals, and – as a particular advantage over other recent international achievement tests – a home background questionnaire completed by the parents of tested students. The latter questionnaire provides particularly relevant information which primary-school children themselves could not have provided. The available student background features include information on the students' pre-school reading performance, their kindergarten attendance, school starting age, language spoken at home, immigration status, parental education, occupation and income and rural versus urban location, among several others. The school features include class size, instruction time, proxies for the material endowment of the classroom, teacher characteristics including gender, education, recent training participation and experience, testing policies, curriculum determination, class formation rules and length of students' staying with the same teacher, among others.

² The only countries that performed lower were Iran, Kuwait, Morocco and Belize.

³ At the secondary level, Colombia participated in the 1995 TIMSS study, where it ranked 40th of 41 participating countries both in math and in science, with only South Africa performing lower. Similarly, Argentina participated in the 2002 PISA study, ranking 35th of 41 in reading, 34th in math and 37th in science, with in each case at least half of the few lower-performing countries also being Latin American.

Estimating education production functions based on such observational data certainly has its limitations (cf., e.g., Glewwe and Kremer 2005). While the extensive PIRLS data does allow for several important features, including the estimation of a limited value-added specification, it cannot solve all the problems imminent to retrospective studies using observational data. What the presented results provide are informative knowledge on the relationship between students' reading performance and a host of student, family and school characteristics, conditional on all the other characteristics considered. These are useful descriptive results which can inform policymaking when properly – and thus cautiously – interpreted. It should be clear from the start that they do *not* necessarily always reflect a causal relationship in the sense that the student or school characteristics causally influenced reading performance by the amount of the estimated parameters. Given that observational data always emanate from individual choices, unbiased causal empirical estimates – if ever possible – can ultimately only come from experimental setups of one form or the other. It is hoped that the descriptive results presented here can serve as a first step that leads towards a better-informed quest for causal identifications in the future.

So far, little research has been done on the determinants of educational quality in Argentina and Colombia, and actually in Latin America more widely. The existing empirical research on educational production in Latin America is ably reviewed in Mizala and Romaguera (2002). For Argentina, the only work they cite is Llach et al. (1999), who analyze a national achievement test. Galiani and Schargrotsky (2002) provide an analysis of secondary school decentralization on educational test performance in Argentina. Research on Colombian educational quality focuses mainly on the relative performance of private and public schools (Cox and Jimenez 1991; Piñeros and Rodríguez 1998) and the Colombian voucher system (King et al. 1999; Angrist et al. 2002). With respect to other Latin American countries, somewhat more research has been done on Chile (cf. Sapelli 2002 and the contributions introduced therein), again with special emphasis on its voucher system (e.g., Mizala and Romaguera 2000; Sapelli and Vial 2002; Vegas 2002; Hsieh and Urquiola 2003) and the relative performance of private and public schools (Mizala et al. 2002). Further examples of the little recent work on educational quality in Latin America include Jimenez and Sawada (1999) on the effect of community involvement in El Salvador, Urquiola (2000) on class-size effects in Bolivia and Lopez-Acevedo (2004) on the effects of teacher development policies in Mexico. Schultz (2004), Coady and Parker (2004) and Bourguignon et al. (2003) analyze the effects on the enrollment of students of cash transfer schemes that are

contingent on school participation in Mexico and Brazil, but they do not have educational performance data to estimate effects on educational quality.⁴

Research on the determination of educational quality in Latin America would be particularly welcome since Latin America is the region featuring the highest rates of return to education among the studies surveyed in Psacharopoulos and Patrinos (2004). Also, the returns to education typically increased in Latin America (including Argentina and Colombia) during the 1990s (cf. Pritchett 2004), which has been attributed to the economy-wide reforms that affected technological progress and created higher returns to education in the adaptation to the generated disequilibria (Behrman et al. 2000). While economy-wide reforms such as capital-market and trade liberalization, privatization and tax and labor-market reforms influence the profitability of educational investments, particularly during transitions to new equilibria, the same can be expected of the education system itself, which will lower the long-run returns to education if it does not succeed in producing basic skills to a satisfactory degree.

The remainder of the paper is structured as follows. Section 2 describes the database of the PIRLS 2001 international student achievement study and presents descriptive statistics on educational achievement and student and school characteristics in the selected Latin American and comparison countries. Section 3 derives the empirical model and its econometric implementation. Section 4 presents results on the relationship between student background and educational performance in the different countries. Section 5 estimates the relationship between school characteristics, both material and institutional, and educational performance. Section 6 summarizes the main findings and concludes.

2. The Database

2.1 The PIRLS 2001 International Student Achievement Test

The dataset used in the analysis is the PIRLS 2001 International Database. The PIRLS assessment was conducted in 2001 by the International Association for the Evaluation of

⁴ Cf. Hanushek (1995), Glewwe (2002), Pritchett (2004) and Glewwe and Kremer (2005) for reviews of research on the determinants of educational quality in developing countries. Velez et al. (1993) review the results of 18 early research reports on primary education in Latin American countries, mostly not published in refereed journals or books. While the evidence from other countries is informative, specific national evidence is indispensable because to some extent, the success of any service delivery depends on specific institutional circumstances that cannot be replicated elsewhere (cf. Pritchett and Woolcock 2004).

Educational Achievement (IEA) in 35 developed and emerging countries.⁵ The aim was to obtain an internationally comparable database for the reading literacy of students in what is the fourth grade of primary school in most education systems. The assessment was designed to capture the skills and features that are necessary “to understand and use those written language forms required by society and/or valued by the individual” (Campbell et al. 2001, p. 3). This definition leads to a real-life focus of the study. Because most national curricula require that students are able to use reading to learn, the definition also ensures that the international comparability of the measured reading performance should not be severely affected by differing national curricula (cf. Gonzalez and Kennedy 2003).

The Latin American countries that participated in PIRLS are Argentina and Colombia. To be able to compare and benchmark the Latin American results, we also analyze an additional six countries. These are Turkey and Macedonia, which show average performance levels similar to Argentina and Colombia. Together, these four countries form our group of “low performing countries” (LPC). We also look at benchmark countries from the industrialized world. Germany and England have been chosen as two large developed countries representing a Continental-European and Anglo-Saxon background, respectively.⁶ In addition, two further Western European countries that might constitute interesting benchmarks for the Latin American countries are Greece and Italy. The two countries are presumably the European countries participating in PIRLS that are most comparable to Argentina and Colombia. Many Argentines are of Italian origin, while Greece is perhaps closest among the European countries to the Latin American countries in terms of its level of economic development. Germany, Greece, Italy and England together form our group of “high performing countries” (HPC).

The specific target population of the PIRLS study was defined as the upper of the two grades with the highest share of nine-year-olds of a country at the time of testing. This led to a student population which was in the fourth grade in the countries considered in this paper, with the exception of England, where it was fifth grade. The mean age of the students in the

⁵ For detailed information on the study and the database, cf. Mullis et al. (2003), Martin et al. (2003), Gonzalez and Kennedy (2003), Campbell et al. (2001) and the PIRLS homepage at <http://timss.bc.edu/pirls2001.html>.

⁶ The United States could not be used for the analysis because two crucial student background variables, the number of books at home and parental education, were not administered in the US. This would result in a lack of comparability to the other countries, as well as to considerable bias in the estimation of the remaining student background and school effects in the US. Instead of the US, England is included in the analysis as a representative of the Anglo-American country group.

considered countries is 10 years and 3 months, with a variation between countries of at most 5 months.

While virtually all children in developed countries are enrolled in fourth grade, this is not necessarily the case in developing countries. If a substantial part of the potential student population of a country is not enrolled in school in fourth grade, this might introduce sample selection bias relative to full-enrollment countries. The most detailed data on this are Filmer's (2003) enrollment profiles based on national survey data, which include Colombian data for 2000. The data show that while 98.5% of Colombian children have been enrolled in first grade, this declines to 91.6% in fourth grade. While the rich and middle class, as well as the urban population in general, come relatively close to universal enrollment at fourth grade, this is not true for the poorest 40% of the Colombian population, or for the rural population. Even though more than 97% of the poor and of the rural population in Colombia enroll in first grade, early drop-out rates among them are very large. At the age of 15-19, only 83.4% of the poorest 40% of the Colombian student population, and only 81.4% of the rural Colombian population, report that they were enrolled in fourth grade (Filmer 2003; cf. Pritchett 2004). This selective non-attendance at the primary-school level may somewhat distort the Latin American samples of students tested by PIRLS, although the enrollment data show that the effect can be expected to be limited.

The PIRLS sampling procedure ensured that a representative sample of the target population was tested in each country. PIRLS used a two-stage sampling method, sampling in the first stage schools and in the second stage intact class(es) of the school. This led to a sample size of 3,300 students in Argentina and 5,131 students in Colombia. In the comparison countries we analyze, the sample size ranges from 2,494 to 7,633 students.

The performance tests were paper and pencil tests, with the assessment lasting a total of 80 minutes for each student. The PIRLS tests were constructed to test a range of relevant skills and competencies that reflected how well the reading literacy of students was developed. Each subject was tested using a broad sample of tasks with differing levels of difficulty in order to represent a coherent and comprehensive indicator of the continuum of students' abilities. The test items included multiple-choice items and questions requiring the students to construct their own responses. Using Item Response Theory (IRT), PIRLS mapped performance on a scale with an international mean of 500 test-score points and an international standard deviation of 100 test-score points.

In addition to the performance tests, PIRLS provides a rich set of background information. This background information was gathered through a student questionnaire, a home questionnaire (completed by parents), a teacher questionnaire and a school questionnaire (completed by school principals). By merging these data with the student achievement data, the database used in this paper contains extensive information on the family background of students, on the resource endowment and teacher characteristics of the class and school and on schools' institutional features.

A problem immanent in any empirical work based on survey data is non-response and thus missing data. Table A1 in the appendix reports the share of missing values for each variable used in this paper. While there are only very few missing values on such variables as student age and gender, kindergarten attendance and teacher gender and education, the amount of missing data on some variables in some countries is quite substantial, reaching more than 30 percent in Argentina in the case of books at home and parental education, working status and occupation. In such a situation, dropping any student that misses data on at least one variable in multivariate analyses such as the ones conducted in this paper means that one is left with relatively small samples. Furthermore, since data points are unlikely to be missing at random, such a procedure could introduce severe biases and reduce the representativeness of the results. In effect, when sticking to the original data, there is an inescapable trade-off between controlling for many variables and keeping many observations. Even more, dropping all observations with some missing values always means dropping relevant information, because there is abundant information on other variables for each student that misses data on one or a few of the numerous considered variables.

Thus, to ensure the usage and exploitation of all the information available in the dataset, we imputed missing values in the data using a simple median imputation method. This method assigns the median of the available observations at the lowest possible level, i.e. either the class, school or country median, for any missing observation. In this procedure, we used sampling weights to take account of the survey structure of the PIRLS dataset (see Section 3 below). In the estimated regressions, we include imputation dummies which ensure that none of the substantive results are driven by imputed rather than original data, and we do additional robustness tests of the imputation data (again, see Section 3 below).

2.2 *Descriptive Statistics on Student Achievement*

The mean reading performance of the eight considered countries in PIRLS, reported in Table 1, clearly reveals the division into the two groups. The group of relatively “low performing countries” (LPC), Argentina, Colombia, Turkey and Macedonia, shows a mean performance between 418 and 449 achievement points (AP), while the group of relatively “high performing countries” (HPC), Germany, Greece, Italy and England, performs between 525 and 553 AP.⁷ As the lowest performer among the selected countries, Argentina scores 4 AP lower than Colombia, 24 AP lower than Macedonia, 31 AP lower than Turkey and 122 AP lower than the HPC group. The size of these performance gaps can be illustrated by the unconditional difference of 42 AP between third graders and fourth graders in Sweden, which was the only country in PIRLS that tested third graders in addition to fourth graders. This difference can be interpreted as the additional value-added of one school year in Sweden, since both grades are tested in representative samples. Thus, the Argentine and Colombian performance is the equivalent of one-half to two-thirds of a (Swedish) school year behind the Macedonian and Turkish performance, and roughly three years behind the HPC group.

The mean performance hides substantial differences in the performance distribution between countries, though. The performance variation is much larger in Argentina than in all the countries of the HPC group. This is independent of the measure of performance variation, which include the standard deviation in national test scores (either absolute or relative to the national mean) and the performance difference between percentiles such as the 90th and 10th or the 75th and 25th. Only Macedonia shows a spread in performance that is similarly large as in Argentina, while depending on the measure, Colombia and Turkey show a variation that is slightly larger than in the HPC group, sometimes surpassed by England, but substantially lower than in Argentina and Macedonia. This pattern is also evident in the graphical depiction of Figure 1. The descriptive pattern shows that the Argentine education system has a shortcoming with respect to ensuring educational equity and providing support for the relatively weak students.

Another feature emerging from the descriptive data is that the performance of the 75th and 90th percentile students in Argentina is actually similar to (or even slightly better than in) Colombia. The lower mean performance and larger variation in Argentina arise from the

⁷ The fact that England is the highest performing among the considered countries may be related to the fact that it is the only country where most students start school at age five, so that England tested fifth-grade rather than fourth-grade students.

particularly poor performance of the lower half of the Argentine student population. Still, even the 75th percentile of students in Argentina and Colombia only roughly reaches the performance level of the 25th percentile in the HPC group, and only the 90th percentile in Argentina and Colombia comes close to the median performance in the HPC group.

In sum, both Argentina and Colombia show a relatively low reading performance among their students. But Argentina has a relatively large variation among its student body, with the lower part of its distribution performing particularly weak.

2.3 Definition of Variables and Descriptive Statistics on Student Background

Table 2 reports descriptive statistics on the input variables which will be used to estimate education production functions. These variables combine two main categories, student background and school characteristics.⁸

In terms of students' background, we have a measure of pre-school reading performance based on retrospective parental evaluation of how well their children could read when they began primary school. Specifically, parents reported reading performance in five areas, namely recognizing most of the letters of the alphabet, reading some words, reading sentences, writing letters of the alphabet and writing some words. Each of the areas was evaluated on a 4-point scale, with 1 = not at all, 2 = not very well, 3 = moderately well and 4 = very well. We use the mean of these five performance evaluations to control for pre-school reading performance in our value-added specification. It is interesting to note that the mean values of this measure do not vary systematically between the LPC and HPC group, always ranging between not very well and moderately well pre-school reading performance. This suggests that this measure is a relative one, formed by the parents under the perception of other students in their country. This does not undermine its value for a value-added specification within each country, though, and as will be seen in the results below, the measure is very strongly correlated with fourth-grade reading performance in every single country.

We also have data on kindergarten attendance, school commencement age and current age for each tested student. At 94%, kindergarten attendance in Argentina exceeds the rates of other LPCs and is comparable to the kindergarten attendance in HPCs. In Colombia, it is 81%. At 50% and 42%, respectively, most Argentine and Colombian children attend

⁸ Cf. Mullis et al. (2002) for additional background information on the education systems and reading instruction in the countries participating in PIRLS.

kindergarten for one to less than two years. The fraction of children entering school at age five or younger is larger in the two Latin American countries than in the other LPCs, at 20% and 39%. A large fraction of 93% of Argentine students attends school before they are seven years old. By contrast, late enrollment after age six is a larger problem in Colombia, at 18%.

At 84%, the share of students who always speak the test language at home is relatively low in Argentina and Colombia, albeit not substantially smaller than in the other LPCs. Both countries also feature relatively large fractions of immigrated students, at 38% and 30%.

A proxy for the socio-economic status of the students' family is the number of books in the household, which reflects the educational, social and economic background of the home. More of the Argentine and Colombian students grow up in a household possessing very few books relative to the comparison countries. As many as 45% and 47% of students in the two countries have at most 10 books in their home (not counting magazines, newspapers or children's books), and only 4% and 3% have more than 200 books at home. This compares to a mean among the HPCs of 9% with at most 10 books (31% in the other two LPCs) and 22% with more than 200 books (6% in the other two LPCs).

Another indicator of students' family background, the educational level of the parents (measured as the highest level reached by either father or mother, whichever is higher), also suggests a relatively weak educational background in Argentina and Colombia. However, the share of parents holding a university degree is larger than in the other two LPCs, indicating a larger educational inequality already at the parental level. The strong difference in the two family background indicators between the LPCs and the HPCs is evident in their graphical depictions in Figures 2 and 3.

In the LPCs, about one third of the parents report that neither of them has work, which is substantially larger than in the HPCs, where the fraction is around 10%. While the fraction of blue-collar workers in Latin America is surprisingly similar to the HPCs, the share of white-collar workers in the LPCs is 16 percentage points lower than in the HPCs.⁹ Parents were also asked to report their before-tax annual household income. However, this question is not

⁹ White-collar workers were defined as major group 1-3 of the International Standard Classification of Occupations (ISCO), encompassing legislators, senior officials and managers; professionals; and technicians and associate professionals. Blue-collar workers were defined as ISCO 8-9, encompassing plant and machine operators and assemblers; and sales and services elementary occupations. The residual category between the two, ranging from ISCO 4-7, encompasses clerks; services workers and shop and market sales workers; skilled agricultural and fishery workers; and craft and related trades workers. The variable was set to white-collar if at least one parent was in ISCO 1-3, and to blue-collar if no parent was in ISCO 1-7.

available in Argentina, Turkey and Italy. Furthermore, the categories are scaled to each national currency, so that they are not directly comparable across countries.¹⁰

Finally, at 11%, less tested schools in Argentina were located in villages or rural areas than in the other LPCs, while the share of schools located in cities with more than half a million inhabitants is largest in Turkey (35%), Colombia and Macedonia (both 24%).

2.4 Definition of Variables and Descriptive Statistics on School Characteristics

The data on school characteristics entail information both on the resource endowment of classrooms and schools and on their institutional features. In terms of resource endowments, the average class size in Argentina is 27 students per class, which is smaller than in Colombia (30), Turkey (35) and England (29), but larger than in the other three HPCs.

Total instructional time (measured as the product of hours per day and days per year) in Argentina is among the lowest observed, 15% lower even than in Germany's half-day schooling system.¹¹ The time that students report spending on reading homework per week in Argentina is the lowest among the LPCs, but not lower on average than in the HPCs.

The equipment of schools with instructional material, instructional staff and qualified teachers is measured by principals' assessment to what extent the shortage or inadequacy of the respective resource affects the school's capacity to provide instruction, where "lack" means "a lot", "no lack" means "not at all" and the residual category is in-between ("a little" or "some"). In general, the thus measured equipment is better in Argentina than in Colombia, which again is better than in the other LPCs.¹²

93% of the Argentine and 82% of the Colombian teaching force are female, which is larger than in most other, particularly LPC, countries. Remarkably few teachers (2%) in the Argentine school system hold a university degree, far less than in the other countries. Most of the teachers in all considered countries also hold a specific teaching certificate. 28% of the

¹⁰ The national categories are: Colombia (in Minimum Monthly Household Income, where 1 MMHI = 286000 pesos): less than 1, 1-2, 3-4, 5-6, more than 7, do not know; Macedonia (in Denars): less than 3900, 3900-6199, 6200-9299, 9300-12399, 12400-15499, 15500 or more; Germany (in 1000 DM): less than 40, 40-59, 60-79, 80-99, 100-119, 120 or more; Greece (in 1,000 drachmas): below 2400, 2400-3599, 3600-4799, 4800-7199, 7200-9599, 9600 or more; England (apparently in 1000 \$): less than 20, 20-29, 30-39, 40-49, 50-59, 60 or more.

¹¹ This measure of instructional time is reported by the school principal. We also experimented with an alternative measure, the specific time for reading instruction and activities per week reported by the teacher, with similar qualitative results in our estimations.

¹² As alternative measures of material endowment, we also experimented with information from teacher questionnaires on whether they use or never use textbooks, workbooks, computer software and other materials. These measures did not yield statistically significant results in our estimations.

Argentine teaching force, and 17% of the Colombian teaching force, reported having participated in a professional development program (workshop or seminar) on the teaching of reading for more than one week (of 35 hours) in the preceding two years, which is larger than in all comparison countries. The Argentine teaching force has a relatively low mean professional experience, at slightly less than 14 years on average.

In terms of more institutional features, the teacher background questionnaire gives information on student assessment practices. 59% of Argentine teachers and 65% of Colombian teachers report that they place major emphasis on classroom tests (for example, teacher-made or textbook tests) to monitor students' progress in reading. Only in England do teachers of more than a third of the student population place major emphasis to monitor student progress on national or regional examinations.

In all countries, a huge majority of principals reports that the national or regional curriculum has a lot of influence on their school's fourth-grade curriculum. About two thirds of students in Argentina and Colombia are taught in schools where fourth-grade classes are formed on the basis of students' ability, so that all students in a class are about the same ability. This prevalence of student tracking is considerably larger than in the other countries, particularly the HPCs. Staying with the same classroom teacher for four or more years is highly uncommon in Argentine primary schools, but much more so in most other countries. 44% of Argentine students stay with the same teacher for at most one year, which is only surpassed by England.

3. The Empirical Model

This section derives the empirical model and discusses its econometric implementation, including discussions of the employed value-added specification, the implemented estimation equation, imputation controls, the covariance structure and sampling weights.

As derived, e.g., in Hanushek (2002), under standard assumptions the education production function can be estimated in the value-added form:¹³

$$T_{is}^t = T\left(T_{is}^{t-1}, B_{is}^t, S_{is}^t\right), \quad (1)$$

¹³ Cf. Todd and Wolpin (2003) for a detailed discussion of the relative benefits of different models to estimate education production functions, and of the respective structural and behavioral assumptions underlying the different model specifications. Cf. also Glewwe (2002) and Glewwe and Kremer (2005) for discussions of underlying behavioral models and the limitations of estimating them using retrospective data.

where T_{is}^t is the test-score performance of student i in school s at time t , B are student background characteristics and S are school characteristics. This value-added specification is considered preferable to simple contemporaneous specifications because it accounts for the cumulative character of the education process by controlling for prior performance at time $t-1$ (cf. Hanushek 2002). Given the proxy of pre-school reading performance we derived from the PIRLS questionnaire, we can implement a version of this value-added specification using the PIRLS data. This is not a perfect solution because ideally, we would either have the students tested at the start of fourth grade or cumulative measures of school inputs S from first to fourth grade. However, given relative constancy of school factors over the course of the first four years of schooling, the approximations introduced by this specification should be limited.

Treating the input factors as additively separable and considering the imputed character of some observations, we implement the following microeconomic estimation equation of the education production function:

$$T_{is}^t = T_{is}^{t-1} \beta_1 + B_{is}^t \beta_2 + S_{is}^t \beta_3 + D_{is}^t \beta_4 + \varepsilon_{is} \quad , \quad (2)$$

where T_{is}^t is the PIRLS test score of student i in school s in fourth grade, T_{is}^{t-1} is the measure of pre-school reading performance, B is a vector of student background data, S is a vector of data on school characteristics and ε is a student-specific error term. The parameter vectors β_1 to β_4 will be estimated in the regressions.

The vector D contains one dummy for each variable in B and S that has missing data. The dummies equal 1 if a value was missing and thus imputed and 0 for original data (cf. Section 2.1 above on the data imputation). By including these imputation controls in the estimation, observations with missing data on each variable are allowed to have their own intercepts. This ensures that none of the results are driven by imputed data. While data imputation introduces measurement error in the explanatory variables, which should generally make it more difficult to observe statistically significant effects, the imputation controls ensure that the results are robust against possible bias arising from data imputation.

As an additional procedure to test for the robustness of our treatment of missing data, we also estimated our models under disregard of any missing data individually for each variable, re-running the regressions as many times as there are variables with missing data, for the two Latin American countries. We did not observe a single qualitative change of our results under these alternative specifications without each variable's imputed values, so that our presented results do not hinge on the data imputation.

Owing to the complex data structure produced by the PIRLS survey design and the multi-level nature of the explanatory variables, the error term ε of the regression has a non-trivial structure. Although we include a considerable amount of school-related variables, we cannot be sure that there are no omitted variables at the school level. Given the possible interdependence of students within the same school, the use of class- and school-level variables and the fact that schools were the primary sampling unit (PSU) in PIRLS, there may be unobservable correlation among the error terms ε_{is} at the school level (cf. Moulton 1986). We correct for potential correlations of the error terms by imposing an adequate structure on the covariance matrix. Thus, we suppose the error term to have the following structure:

$$\varepsilon_{is} = \eta_s + \nu_i \quad , \quad (3)$$

where η_s is the school-level element of the error term and ν_i is the student-specific element.

We use clustering-robust linear regressions (CRLR) to estimate standard errors that recognize this clustering of the student-level data within schools. The CRLR method relaxes the standard independence assumption and requires only that the observations be independent across the PSUs, i.e. across schools. By allowing any given amount of correlation within the PSUs, CRLR estimates appropriate standard errors when many observations share the same value on some but not all independent variables (cf. Deaton 1997). To avoid inefficiency due to heteroscedasticity, CRLR imposes a clustered covariance structure on the covariance matrix, allowing within-school correlations of the error term. Thus, the form of the covariance matrix V is

$$V = \begin{pmatrix} \Sigma_1 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \Sigma_l \end{pmatrix} \quad , \quad (4)$$

with Σ_i ($i=1, \dots, l$) representing the covariance matrices of the residuals of least-square regressions within each school cluster (PSU). Observations of two different PSUs are thus assumed to be independent and uncorrelated, leading to the block-diagonal matrix V with PSUs as diagonal elements. This method yields consistent and efficient estimates, enabling valid statistical inferences of the obtained estimation results (cf. White 1984).

Finally, the stratification of the PIRLS sampling design within each country produces varying sampling probabilities for different students. To obtain representative estimates from

the stratified survey data, we employ weighted least squares (WLS) estimation using the sampling probabilities as weights. WLS estimation ensures that the proportional contribution to the parameter estimates of each stratum in the sample is the same as would have been obtained in a complete census enumeration (cf. DuMouchel and Duncan 1983; Wooldridge 2001).

4. Student Background and Educational Achievement

This section reports the results of regressing students' educational performance on a host of student characteristics, which mainly measure their family background. Because this section is interested in the total impact of student background on educational performance, including any effect that might work through families' differential access to schools with different endowments and through their influence on the institutional features of schools, the estimation of the student-background effects does not control for school characteristics, such as resource endowment and institutional characteristics, as in equation (2):

$$T_{is}^t = T_{is}^{t-1} \beta_1 + B_{is}^t \beta_2 + D_{is}^t \beta_3 + \varepsilon_{is} \quad . \quad (5)$$

While this specification omitting school characteristics should be preferable for the research question considered in this section, we also tested the robustness of our student-background results to estimating equation (2), which includes the school characteristics. None of our qualitative results are sensitive to the alternative specification.

Tables 3a and 3b report the results of estimating equation (5) for each of the considered countries. Fourth-grade reading performance as tested in PIRLS is strongly positively related to reading performance at the start of primary school in each of the countries. The measure of pre-school reading performance has the highest t -statistic of all variables in Argentina and Colombia, and also one of the highest t -statistics in all other countries. This provides some confidence in the value-added specification employed.

Kindergarten attendance is not statistically significantly related to reading performance of Argentine students, and in Colombia, students who attended kindergarten for a considerable time even performed significantly worse. In the other considered countries, there is similarly very little evidence of positive relationships between kindergarten attendance and later student performance. In interpreting this result, it should be borne in mind that children are not randomly assigned to attending kindergarten, so that significant selection patterns may influence this result. However, assuming that students from superior backgrounds are more

likely to attend kindergarten, particularly in developing countries, the estimates might generally be expected to be positively biased. If this was the case, the results would suggest that kindergarten attendance alone does not improve primary-school learning. This might be the case where kindergartens are day-care centers rather than learning environments, so that the alternative of spending time at home would prepare students for later learning just as well.

There is no significant performance difference between Argentine children starting school at age five and age six, while in Colombia, children starting at age six perform slightly better. Because the effect of school commencement age controls for linear effects of the current age of the students in months, thereby holding effects of monthly age and possible grade repetition constant (as well as the grade level, which is the same for all students of a country), this result suggests that starting school early at age five is slightly detrimental for students in Colombia. However, later school commencement at age seven or older is related to lower performance in Argentina, but not in Colombia. Thus, starting school later than at age six is detrimental for students in Argentina.

Boys perform significantly worse than girls in all countries, with the sole exception of Colombia, where the performance difference between gender is small and statistically insignificant. In Argentina, the performance difference is 15.1 AP, which is equivalent to 15.1% of an international standard deviation in test scores and to 15.8% of a standard deviation in Argentine test scores.

Students who always speak the test language at home perform statistically significantly better than otherwise in all countries. Similarly, with the sole exception of Greece, students who were not born in the country perform statistically significantly worse than native students. The size of these effects is substantial and broadly similar between the Latin American and comparison countries. Holding the student's immigration status constant, belonging to the second generation of immigrants does not have a significant additional negative effect in Argentina, while there is an additional negative effect of the father being non-native in Colombia.

The first measure of general family background, the number of books in the household, shows a consistent and strong positive relation to student performance in all countries. This proxy for the social, economic and educational background of the students' home has a particularly large effect in Argentina, but a relatively small effect in Colombia. In Argentina, the achievement differential between the lowest and highest category of books at home is as large as 53.4 AP, surpassing all comparison countries. By contrast, it is only 14.4 AP in

Colombia, smaller than in all comparison countries. A small part of this difference can be attributed to the fact that data on household income are not available in Argentina, so that income is controlled for in Colombia but not in Argentina. However, excluding the income controls in Colombia and the other LDC with income data, Macedonia, does not change the picture significantly. In Macedonia, there is no substantive change at all in the results, and in Colombia, the currently significant estimates on books at home, parental education and working status get a bit larger and statistically significant at the 1% level. Still, the Colombian estimates remain smaller than in the comparison countries, e.g. at 21 AP in the case of the highest category of books at home.

The second measure of general family background, the education level of the parents, shows a very similar pattern. It is statistically significantly and strongly related to student performance in all countries, but the effect is relatively large in Argentina and particularly small in Colombia. The estimated effects of books at home and parental education in the different countries, depicted graphically in Figures 4 and 5, both suggest that the Argentine school system performs worse than other school systems, particularly among the LPCs, in terms of alleviating the effects of differential family backgrounds. By contrast, the Colombian school system achieves an equitable performance for children from different family backgrounds to a considerably larger extent. A tentative conclusion would be that these relatively large family background effects in Argentina might be one important factor underlying the relatively large performance variation detected in Argentina, particularly compared to Colombia.

For Argentina, these results on intergenerational effects on the quality of primary education correspond to findings on intergenerational mobility in the quantity of education between parents and their children. Behrman et al. (2001) and Dahan and Gaviria (2001) report that the intergenerational mobility in educational quantity between parents and their children in Latin America, including Argentina and Colombia, is relatively low compared to other countries (i.e., the effect of parental on children's educational quantity is relatively large). Our results suggest that while these quantitative effects go hand in hand with an equally differentiated educational performance in terms of the skills that children learn in Argentina, this does not seem to be the case in Colombia.

Both in Argentina and Colombia, students who have one parent working full-time perform statistically significantly higher than students whose parents do not work. However, in both countries there is no significant performance difference between students whose parents have

different occupations. This is exceptional, as such performance differences exist in all comparison countries. Household income is significantly and strongly related to student performance in Colombia.¹⁴ The same is true in all other countries with income data. However, due to differing national currencies and thus income categories, the estimates are not directly comparable across countries.

In contrast to all comparison countries, there is no statistically significant difference in student performance between any category of school location in Argentina. In the other LPCs, students in rural areas perform systematically worse than students in urban areas, and even than students in large cities of more than half a million inhabitants.

Not considering the imputation controls, the student background variables considered in the model together account for as much as 27% of the total variation in student performance both in Argentina and in Colombia. This is similar to Greece and England, slightly smaller than in Germany and Macedonia, but substantially larger than in Turkey and Italy.

5. School Characteristics and Educational Achievement

Tables 4a and 4b report the relationship between school characteristics and student achievement, estimated using equation (2) above, which controls for the effects of student background characteristics. The considered school characteristics fall into two broad categories, namely measures of the resource endowment of classrooms and schools and measures of the institutional features of schools.

5.1 Resource Endowment

In all considered countries, there is no consistent evidence of significant relationships between resource endowments and student performance. The few statistically significant estimates of a positive relationship tend to be matched by statistically significant estimates of negative relationships. This is true for Argentina and Colombia as much as for the comparison countries in the LPC and HPC group.

In no country are smaller classes related to better performance in a statistically significant way. However, estimates of class-size effects are particularly prone to endogeneity bias. For example, if class size is allocated in a compensatory way, so that weak students are placed in smaller classes, results from observational data will be biased against finding class-size effects (cf., e.g., Wößmann and West 2004). By contrast, if strong performers are placed into

¹⁴ Note that in Colombia, national category 5 refers to “I do not know”.

smaller classes, as may be the case in the more elitist education systems of many developing countries, results will be biased in favor of finding class-size effects. The extent to which either of this is the case in this study has to remain an open question.

Instructional time is also not statistically significantly related to better reading performance in any country at conventional significance levels. Likewise, time spent on homework is not significantly related to student performance in Argentina, and actually significantly negatively in Colombia.

The three groups of variables on lack of instructional material, instructional staff and qualified teachers are mostly not statistically significantly related to student performance in Argentina and Colombia. In Argentina, students in schools whose principals report that their instruction is not at all impeded by lack of instructional material actually perform significantly worse in Argentina. Students in Argentine schools whose principals report that their instruction is not at all impeded by lack of instructional staff perform significantly better, though. In Colombia, the only statistically significant relationship is that students in schools whose principals report that their instruction is strongly impeded by the lack of instructional material perform worse.

The only uniform result on teacher characteristics is that students of female teachers tend to perform better, with the relationship being statistically significantly in Argentina, Colombia, Macedonia and Italy. Teachers' holding a university degree is statistically significantly positively related to student achievement in Greece, but neither in Argentina nor in Colombia. Similarly, there is no statistically significant relationship between teachers' holding a teaching certificate and student performance in Argentina and Colombia. Students of teachers who participated in professional development seminars during the preceding two years actually perform statistically significantly lower in Argentina. Teacher's experience in the teaching profession is not significantly related to student performance in Argentina and Colombia.

This general lack of resource effects resembles much previous research, also in Latin America (cf. Mizala and Romaguera 2002) and in other developing countries (cf., e.g., Hanushek 1995; Kremer 2003; Pritchett 2004). It should be borne in mind, though, that this correlational evidence does not necessarily depict causal effects, as the estimates might be biased by resource endogeneity, e.g. due to omitted variables and placement biases (cf. Glewwe 2002; Wößmann and West 2004; Glewwe and Kremer 2005). Thus, the presented results are suggestive at best, but clearly not definitive.

5.2 *Institutional Features*

Recent theoretical contributions have emphasized the important effects that institutional features of the school system can play in determining educational performance, because they determine the incentives that the involved actors have to further student learning (cf. Bishop and Wößmann 2004 and the references therein). Furthermore, evidence based on the cross-country variation in student performance shows that institutional effects matter empirically (e.g., Wößmann 2002, 2003; Fuchs and Wößmann 2004a, 2004b). However, most of the institutional features generally do not vary substantially within school systems, but only between countries. Empirically, institutional effects should therefore mainly be an issue in cross-country rather than within-country research. That is, altered institutional set-ups of their school systems could potentially shift countries onto production functions that are substantially different in terms of overall efficiency.

As is apparent from Table 2, the PIRLS background data reveal that a few institutional features do also vary within some of the considered school systems. Unfortunately, PIRLS does not provide data on such institutional features as the private or public operation or financing of individual schools, nor on decentralization and school autonomy on budgets, personnel and process decisions.¹⁵ But there is data on testing policies and curriculum determination, as well as on ability-based class formation and on how long primary students stay with the same teacher. While these institutional features seem much more important on a cross-country perspective (cf. Fuchs and Wößmann 2004b for the importance of institutional effects in cross-country PIRLS evidence), it seem worth including them also in the within-country analyses presented here.

Teachers' method to monitor students' progress, in terms of use of classroom tests or centralized examinations, are not statistically significant related to reading performance in Argentina and Colombia, as well as within most other considered countries.¹⁶

In Argentina, as well as in Turkey, students in schools whose curriculum is strongly influenced by the national or regional curriculum perform statistically significantly better than otherwise. This might suggest that external standards are conducive to student performance in these countries. At 36 AP, the effect size is substantial in Argentina.

¹⁵ Galiani and Schargrotsky (2002) present important results on the effects of decentralization on student performance in Argentina.

¹⁶ The positive relationship between use of national or regional examinations and student performance in Argentina is statistically significant only at the 20 percent level.

In Argentina, as well as in Macedonia, students in schools that have a policy of class formation based on student ability perform statistically significantly better than otherwise. This might indicate that tracking students into relatively homogenous groups helps their learning in Argentina. In Colombia, students in schools where they typically stayed with the same teacher for one year or less in primary school performed statistically significantly lower. However, the opposite is true in Argentina, suggesting that there is no consistent relationship of length of teacher assignments to student performance in the different countries.¹⁷

6. Summary and Conclusion

The general pattern of results on the relationship between families, schools and primary-school learning in Argentina and Colombia is similar to what has been found in other regions of the world and mostly at the secondary-school level (cf. Hanushek 2002 for the United States; Wößmann 2004 for Western Europe; Ammermüller et al. 2004 for Eastern Europe; Wößmann 2005 for East Asia). There is strong evidence of a relationship between students' educational performance and numerous measures of their family background, but results on school characteristics are very inconsistent. There is basically no consistent evidence of a relationship between educational performance and schools' resource endowment. Within each country, there is only some evidence that the variation in student performance is related to the institutional characteristics of schools which could be measured in this study. For example, a centralized curriculum and ability-based class formation within schools are positively related to student performance in Argentina. Using the same data source, evidence on institutional effects is much stronger between countries (cf. Fuchs and Wößmann 2004b), suggesting that institutional features of the school systems may exert strong systemic effects.

The evidence found in this paper pinpoints the crucial role that families play in determining the reading performance of students in primary school. Consistent with recent US evidence emphasizing basic family background factors in determining college readiness of students (cf. Carneiro and Heckman 2003), we find that learning readiness and family background have very strong effects on educational performance of Argentine and Colombian primary-school students. These effects do not seem to be easily amenable to simple policy interventions such as increased kindergarten attendance, as kindergarten attendance is not

¹⁷ The large negative coefficient in Argentina on staying with the same teacher for at least four years is due to only one principal reporting so. It may therefore reflect other specific characteristics of this school, for example a remotely located school which is hardly accessible and thus has only a single teacher.

found to have significant attenuating effects on performance differences. Given the importance of primary-school reading ability for later learning, these strong family-background effects may have important consequences for the longer-run intergenerational mobility of Latin American societies.

Comparing the evidence for Argentina and Colombia with the set of comparison countries, the general pattern of results in terms of effects of student background and school characteristics is rather consistent across countries. But benchmarking the different estimated parameters across the countries also reveals characteristic national patterns. In particular, the effects of family background, measured by the number of books at home and by parental education, are relatively strong in Argentina and relatively weak in Colombia. In this sense, the fact that the intergenerational transmission of skills is particularly strong in Argentina can be viewed as one important underlying factor behind the relatively large variation in student performance observed in Argentina, particularly in the comparison within Latin America to Colombia.

Another result specific to Argentina is that once the family-background factors are held constant, no significant performance difference is found between rural and urban areas. A specific result for Colombia is that it is the only country where boys do not score significantly worse than girls. Finally, a result that distinguishes the two Latin American countries from all comparison countries is that they do not show a significant performance difference between parental occupations.

The results presented here using the data from the PIRLS student achievement test provide useful descriptive features of the different primary-school systems. However, the limitations of such retrospective studies, particularly when it comes to causal inferences, are well known (cf., e.g., Glewwe 2002; Todd and Wolpin 2003; Glewwe and Kremer 2005). The need for further research on the underlying behavioral relationships, preferably based on quasi-experimental or truly experimental setups which are able to exploit exogenous variations in important family and school characteristics, is beyond doubt.

References

- Ammermüller, Andreas, Hans Heijke, Ludger Wößmann (2004). Schooling Quality in Eastern Europe: Educational Production During Transition. *Economics of Education Review*: forthcoming (revised version of: Research Memorandum ROA-RM-2003/2E. Maastricht: Research Centre for Education and the Labour Market, 2003).
- Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King, Michael Kremer (2002). Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment. *American Economic Review* 92 (5): 1535-1558.
- Baer, Werner, William Maloney (1997). Neoliberalism and Income Distribution in Latin America. *World Development* 25 (3): 311-327.
- Behrman, Jere R., Nancy Birdsall, Miguel Székely (2000). Economic Reform and Wage Differentials in Latin America. Inter-American Development Bank Research Department Working Paper 435. Washington, D.C.: Inter-American Development Bank.
- Behrman, Jere R., Alejandro Gaviria, Miguel Székely (2001). Intergenerational Mobility in Latin America. *Economia* 2 (1): 1-44.
- Bishop, John H., Ludger Wößmann (2004). Institutional Effects in a Simple Model of Educational Production. *Education Economics* 12 (1): 17-38
- Bourguignon, François, Francisco H.G. Ferreira, Phillippe G. Leite (2003). Conditional Cash Transfers, Schooling, and Child Labor: Micro-Simulating Brazil's Bolsa Escola Program. *World Bank Economic Review* 17 (2): 229-254.
- Campbell, Jay R., Dana L. Kelly, Ina V.S. Mullis, Michael O. Martin, Marian Sainsbury (2001). *Framework and Specifications for PIRLS Assessment 2001*. Second Edition. Chestnut Hill, MA: International Study Center, Boston College.
- Carneiro, Pedro, James Heckman (2003). Human Capital Policy. NBER Working Paper 9495. Cambridge, Mass.: National Bureau of Economic Research.
- Coady, David P., Susan W. Parker (2004). Cost-effectiveness Analysis of Demand- and Supply-side Education Interventions: the Case of PROGRESA in Mexico. *Review of Development Economics* 8 (3): 440-451.
- Cox, Donald, Emmanuel Jimenez (1991). The Relative Effectiveness of Private and Public Schools. *Journal of Development Economics* 34 (1): 99-121.
- Dahan, Momi, Alejandro Gaviria (2001). Sibling Correlations and Intergenerational Mobility in Latin America. *Economic Development and Cultural Change* 49 (3): 537-554.
- Deaton, Angus (1997). *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore: The Johns Hopkins University Press.
- DuMouchel, William H., Greg J. Duncan (1983). Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples. *Journal of the American Statistical Association* 78 (383): 535-543.
- Edwards, Jack K., Werner Baer (1993). The State and the Private Sector in Latin America: Reflections on the Past, the Present and the Future. *Quarterly Review of Economics and Finance* 33 (Special Issue): 9-19.

- Filmer, Deon (2003). Educational Attainment and Enrollment Profiles: A Resource Book based on an Analysis of Demographic and Health Survey Data. Mimeo. Development Research Group. Washington, D.C.: The World Bank.
- Fuchs, Thomas, Ludger Wößmann (2004a). What Accounts for International Differences in Student Performance? A Re-examination using PISA Data. CESifo Working Paper 1235. Munich: CESifo.
- Fuchs, Thomas, Ludger Wößmann (2004b). Governance and Primary-School Performance: International Evidence. Mimeo. Munich: Ifo Institute for Economic Research at the University of Munich.
- Galiani, Sebastian, Ernesto Schargrotsky (2002). Evaluating the Impact of School Decentralization on Educational Quality. *Economia* 2 (2): 275-314.
- Glewwe, Paul (2002). Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes. *Journal of Economic Literature* 40 (2): 436-482.
- Glewwe, Paul, Michael Kremer (2005). Schools, Teachers, and Education Outcomes in Developing Countries. Forthcoming in: Eric A. Hanushek, Finis Welch (eds.), *Handbook of the Economics of Education*. Amsterdam: North-Holland.
- Gonzalez, Eugenio J., Ann M. Kennedy (2003). *PIRLS 2001 User Guide for the International Database*. Chestnut Hill, MA: International Study Center, Boston College.
- Hanushek, Eric A. (1995). Interpreting Recent Research on Schooling in Developing Countries. *World Bank Research Observer* 10 (2): 227-246.
- Hanushek, Eric A. (2002). Publicly Provided Education. In: Alan J. Auerbach, Martin Feldstein (eds.), *Handbook of Public Economics, Volume 4*: 2045-2141. Amsterdam: North-Holland.
- Heckman, James J. (2000). Policies to Foster Human Capital. *Research in Economics* 54 (1): 3-56.
- Hsieh, Chang-Tai, Miguel Urquiola (2003). When Schools Compete, How Do They Compete? An Assessment of Chile's Nationwide School Voucher Program. NBER Working Paper 100008. Cambridge, Mass.: National Bureau of Economic Research.
- Jimenez, Emmanuel, Yasuyuki Sawada (1999). Do Community-Managed Schools Work? An Evaluation of El Salvador's EDUCO Program. *World Bank Economic Review* 13 (3): 415-441.
- King, Elizabeth M., Peter F. Orazem, Darin Wohlgemuth (1999). Central Mandates and Local Incentives: The Colombia Education Voucher Program. *World Bank Economic Review* 13 (3): 467-491.
- Kremer, Michael (2003). Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons. *American Economic Review* 93 (2): 102-106.
- Llach, Juan José, Silvia Montoya, Flavia Roldán (1999). *Educación para Todos*. Buenos Aires: IERAL.
- Lopez-Acevedo, Gladys (2004). Professional Development and Incentives for Teacher Performance in Schools in Mexico. World Bank Policy Research Working Paper 3236. Washington, D.C.: World Bank.
- Martin, Michael O., Ina V.S. Mullis, Ann M. Kennedy (2003). *PIRLS 2001 Technical Report*. Chestnut Hill, MA: International Study Center, Boston College.

- Mizala, Alejandra, Pilar Romaguera (2000). School Performance and Choice: The Chilean Experience. *Journal of Human Resources* 35 (2): 392-417.
- Mizala, Alejandra, Pilar Romaguera (2002). Equity and Educational Performance. *Economia* 2 (2): 219-273.
- Mizala, Alejandra, Pilar Romaguera, Darío Farren (2002). The Technical Efficiency of Schools in Chile. *Applied Economics* 34 (12): 1533-1552.
- Moulton, Brent R. (1986). Random Group Effects and the Precision of Regression Estimates. *Journal of Econometrics* 32 (3): 385-397.
- Mullis, Ina V.S., Michael O. Martin, Ann M. Kennedy, Cheryl L. Flaherty (2002). *PIRLS 2001 Encyclopedia: A Reference Guide to Reading Education in the Countries Participating in IEA's Progress in International Reading Literacy Study (PIRLS)*. Chestnut Hill, MA: International Study Center, Boston College.
- Mullis, Ina V.S., Michael O. Martin, Eugenio J. Gonzalez, Ann M. Kennedy (2003). *PIRLS 2001 International Report: IEA's Study of Reading Literacy Achievement in Primary School in 35 Countries*. Chestnut Hill, MA: International Study Center, Boston College.
- Piñeros, Luis, Alberto Rodríguez (1998). Los Insumos Escolares en la Educación Secundaria y su Efecto sobre el Rendimiento Académico de los Estudiantes. LCSHD Working Paper 36. Washington, D.C.: World Bank.
- Pritchett, Lant (2004). Access to Education. In: Bjørn Lomborg (ed.). *Global Crises, Global Solutions*: 175-234. Cambridge: Cambridge University Press.
- Pritchett, Lant, Michael Woolcock (2004). Solutions When *the* Solution Is the Problem: Arraying the Disarray in Development. *World Development* 32 (2): 191-212.
- Psacharopoulos, George (1996). Public Spending on Higher Education in Developing Countries: Too Much Rather Than Too Little. *Economics of Education Review* 15 (4): 421-422.
- Psacharopoulos, George, Harry A. Patrinos (2004). Returns to Investment in Education: A Further Update. *Education Economics* 12 (2): 111-134.
- Sapelli, Claudio (2002). Introducción: La Economía de la Educación y el Sistema Educativo Chileno. *Cuadernos de Economía – Latin American Journal of Economics* 39 (118): 281-296.
- Sapelli, Claudio, Bernardita Vial (2002). The Performance of Private and Public Schools in the Chilean Voucher System. *Cuadernos de Economía – Latin American Journal of Economics* 39 (118): 423-454.
- Schultz, T. Paul (2004). School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program. *Journal of Development Economics* 74 (1): 199-250.
- Todd, Petra E., Kenneth I. Wolpin (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *Economic Journal* 113 (485): F3-F33.
- Urquiola, Miguel (2000). Identifying Class Size Effects in Developing Countries: Evidence from Rural Schools in Bolivia. World Bank Policy Research Working Paper 2711. Washington, D.C.: World Bank.
- Vegas, Emiliana (2002). School Choice, Student Performance, and Teacher and School Characteristics: The Chilean Case. World Bank Policy Research Working Paper 2833. Washington, D.C.: World Bank.

- Velez, Eduardo, Ernesto Schiefbein, Jorge Valenzuela (1993). Factors Affecting Achievement in Primary Education: A Review of the Literature for Latin America and the Caribbean. World Bank, Human Resources Development and Operations Policy, Working Paper 2. Washington, D.C.: World Bank.
- White, Halbert (1984). *Asymptotic Theory for Econometricians*. Orlando: Academic Press.
- Wooldridge, Jeffrey M. (2001). Asymptotic Properties of Weighted M -Estimators for Standard Stratified Samples. *Econometric Theory* 17 (2): 451-470.
- Wößmann, Ludger (2002). *Schooling and the Quality of Human Capital*. Berlin: Springer.
- Wößmann, Ludger (2003). Schooling Resources, Educational Institutions, and Student Performance: The International Evidence. *Oxford Bulletin of Economics and Statistics*. 65 (2): 117-170.
- Wößmann, Ludger (2004). Educational Production in Europe. Paper presented at the 40th Panel Meeting of *Economic Policy* in Amsterdam. Mimeo. Munich: Ifo Institute for Economic Research at the University of Munich.
- Wößmann, Ludger (2005). Educational Production in East Asia: The Impact of Family Background and Schooling Policies on Student Performance. *German Economic Review*: forthcoming (revised version of: Kiel Working Paper 1152. Kiel: Institute for World Economics, 2003).
- Wößmann, Ludger, Martin R. West (2004). Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS. *European Economic Review*: forthcoming (revised version of: Harvard University, Program on Education Policy and Governance Research Paper PEPG/02-02, 2002).

Table 1: PIRLS Reading Achievement: Descriptive Statistics

	Argentina	Colombia	Turkey	Macedonia	Germany	Greece	Italy	England
Mean	417.8	422.2	449.0	441.4	538.9	524.5	541.0	553.4
Standard deviation	96.0	80.7	86.6	102.2	66.8	72.7	71.0	85.3
Standard deviation/mean	0.23	0.19	0.19	0.23	0.12	0.14	0.13	0.15
10 th percentile	286.4	319.6	330.9	297.2	448.1	428.8	446.8	437.9
25 th percentile	352.7	367.6	394.2	369.6	497.2	477.2	495.0	502.1
Median	423.1	423.6	450.9	450.1	543.5	529.2	546.3	559.0
75 th percentile	484.3	478.3	510.0	518.8	585.5	575.0	591.0	611.8
90 th percentile	534.7	524.2	558.4	568.9	619.0	615.0	627.5	657.8
Difference 90 th -10 th percentile	248.3	204.6	227.5	271.7	170.9	186.2	180.7	219.9
Difference 75 th -25 th percentile	131.6	110.7	115.8	149.2	88.3	97.8	96.0	109.7

Notes: Students weighted by sampling probabilities.

Table 2: Student and School Characteristics: Descriptive Statistics

	Source	ARG	COL	TUR	MKD	GER	GRE	ITA	ENG
Student background									
Pre-school reading performance ^a	H	2.6	2.6	2.5	3.0	2.3	2.9	2.5	2.9
Kindergarten attendance									
Not at all	H	5.7	19.3	76.6	19.2	3.4	4.4	2.4	8.3
Less than one year	H	3.4	2.3	4.1	5.2	1.0	1.2	0.8	3.4
One to less than two years	H	49.8	41.9	14.5	57.7	7.3	25.0	5.9	60.3
Two years or more	H	41.1	36.5	4.8	17.9	88.3	69.4	90.9	28.0
School commencement age									
Five years or younger	H	19.9	38.5	1.6	1.9	2.0	4.3	16.8	92.9
Six years	H	73.5	43.7	27.3	30.1	68.0	75.9	80.3	6.0
Seven years	H	5.8	13.7	64.3	66.0	29.5	19.2	2.8	0.8
Eight years or older	H	0.8	4.1	6.8	2.0	0.5	0.6	0.1	0.3
Age (in months)	St	122.2	126.4	122.8	127.9	126.5	119.3	118.2	122.5
Male	St	49.2	49.5	51.9	51.1	50.2	50.4	52.1	48.3
Speak test language at home									
Always	St	84.3	84.1	86.9	87.9	89.5	94.1	96.0	88.5
Sometimes	St	9.9	12.6	12.3	9.9	9.5	5.6	3.2	10.6
Never	St	5.8	3.3	0.8	2.2	1.0	0.3	0.8	0.9
Immigration status									
Student is immigrant	St	37.5	29.7	4.5	8.6	20.8	11.7	4.6	15.3
Father is immigrant	St	11.1	6.4	4.6	5.2	18.3	11.7	5.9	17.7
Mother is immigrant	St	11.2	6.0	4.8	6.5	17.5	15.5	7.4	17.1
Books at home									
0-10	H	45.0	47.0	39.7	22.4	6.4	10.3	11.9	6.7
10-25	H	24.0	25.1	27.7	28.7	12.5	16.4	21.2	10.9
26-100	H	21.8	20.4	22.7	33.8	36.0	43.9	35.8	33.6
100-200	H	4.7	4.3	5.2	7.1	17.9	13.1	13.9	20.7
More than 200	H	4.5	3.3	4.7	8.1	27.2	16.4	17.1	28.1
Parental education									
Not completed lower secondary	H	41.8	47.7	0.0	23.3	1.7	10.7	3.9	7.1
Lower secondary	H	21.9	17.3	70.9	30.4	11.9	18.3	32.9	42.1
Upper secondary	H	23.6	15.5	0.2	23.6	28.2	33.0	44.9	19.3
Post secondary (non tertiary)	H	0.0	4.5	20.2	12.9	38.4	18.6	3.4	4.5
University degree	H	12.8	15.0	8.6	9.8	19.8	19.3	14.9	27.1
Working status of parents									
None working	St	32.6	27.5	39.9	30.1	8.3	9.3	12.2	7.7
At least one works half-time	St	13.6	16.9	7.1	15.5	5.1	4.3	5.9	3.6
At least one works full-time	St	44.1	38.8	45.7	29.0	69.7	44.1	49.3	63.0
Both work full-time	St	9.7	16.8	7.3	25.4	16.9	42.3	32.6	25.7
Occupation of parents									
Blue-collar workers	H	7.2	9.7	7.1	12.2	6.2	6.2	6.6	6.9
White-collar workers	H	21.9	15.9	11.6	21.7	28.6	27.9	25.4	54.8
Household income									
Lowest national category	H	–	45.3	–	20.1	17.6	17.8	–	13.8
National category 1	H	–	25.3	–	17.8	20.7	16.0	–	19.8
National category 2	H	–	9.2	–	17.7	23.9	18.1	–	23.9
National category 3	H	–	3.1	–	15.0	16.1	21.9	–	17.4
National category 4	H	–	1.8	–	12.1	9.9	12.3	–	10.3
National category 5	H	–	15.4	–	17.4	11.7	13.9	–	14.8
Location of school									
Rural area	Sc	2.3	15.8	21.1	32.6	26.2	16.2	2.5	19.0
Village up to 3k people	Sc	8.7	10.2	1.5	1.6	1.5	4.7	3.8	2.1
Town 3k-100k people	Sc	60.1	25.1	20.7	37.2	46.0	51.0	77.6	42.1
City 100k-500k people	Sc	14.5	25.2	21.7	4.2	15.7	9.3	7.8	22.8
City over 500k people	Sc	14.4	23.7	35.0	24.4	10.6	18.8	8.3	14.0

Table 2 (continued)

	Source	ARG	COL	TUR	MKD	GER	GRE	ITA	ENG
School characteristics									
Class size in reading	T	26.6	30.4	35.4	26.3	22.5	19.9	20.5	29.0
Instructional time (in 1000 minutes per year)	Sc	41.3	59.1	47.9	40.5	48.3	47.4	62.3	57.4
Time spent on homework (in hours per week)	St	2.3	3.4	3.7	4.5	1.1	4.7	2.3	1.6
Lack of instructional material	Sc	18.2	29.6	54.2	26.4	0.1	33.2	1.0	–
No lack of instruct. material	Sc	30.7	25.8	6.5	13.4	71.7	23.4	45.6	–
Lack of instructional staff	Sc	4.6	14.2	67.2	31.0	2.9	38.9	1.8	–
No lack of instructional staff	Sc	78.6	58.8	8.5	17.7	40.5	26.5	40.9	–
Lack of qualified teachers	Sc	5.1	14.7	59.2	21.9	0.7	15.4	6.2	–
No lack of qualified teachers	Sc	58.8	39.8	7.4	27.4	82.8	28.3	24.7	–
Teacher characteristics									
Gender (female)	T	92.3	81.8	49.5	65.0	81.5	67.9	97.5	79.8
Education									
University degree	T	2.2	85.6	44.5	25.4	92.5	20.7	25.8	100.0
Teaching certificate	T	94.0	94.1	97.9	98.9	92.0	100.0	77.8	98.3
Professional development	T	27.8	17.3	12.0	13.6	1.3	8.4	17.0	7.9
Experience (in years)	T	13.7	17.1	14.8	19.4	23.4	15.6	21.8	14.3
Teacher emphasis on									
Classroom tests	T	58.6	64.8	79.5	70.1	33.4	83.8	82.3	24.3
National/regional examinations	T	15.5	7.0	9.6	25.6	3.8	–	4.3	39.7
National/regional curriculum	Sc	81.2	76.0	73.3	87.8	93.0	90.8	84.2	92.4
Ability-based class formation	Sc	61.9	66.4	13.2	55.7	5.4	10.9	12.1	15.1
Taught by same teacher									
One year or less	Sc	43.6	36.4	6.4	0.0	0.0	13.9	0.6	91.0
At least four years	Sc	0.7	18.2	58.7	99.4	58.7	1.9	84.5	0.6

Notes: Means, based on non-imputed data for each variable, weighted by sampling probabilities. – Source: Data source and thus level of observation, with St = student achievement test or student background questionnaire; Sc = school background questionnaire, H = home background questionnaire (completed by parents), T = teacher background questionnaire. – All variables are denoted in percentage points of students belonging to each category unless otherwise stated. – ^a Measured on a scale from 1 to 4.

Table 3a: Student Background and Reading Performance (LPC Group)

	Argentina		Colombia		Turkey		Macedonia	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Pre-school reading performance	17.69***	(2.85)	15.93***	(2.25)	14.34***	(2.00)	19.53***	(2.68)
Kindergarten attendance ^a								
Less than one year	8.05	(16.07)	-7.16	(8.36)	-5.14	(6.53)	-2.11	(7.37)
One to less than two years	-0.89	(9.93)	-13.03**	(5.61)	9.03*	(4.82)	-11.84*	(6.15)
Two years or more	7.31	(10.03)	-19.42***	(6.43)	10.78	(8.52)	-5.84	(6.55)
School commencement age ^b								
Six years	0.89	(4.76)	7.75**	(3.37)	16.17*	(8.79)	-1.37	(10.99)
Seven years	-44.22***	(7.99)	1.63	(5.74)	20.28**	(8.64)	-0.61	(11.87)
Eight years or older	-57.52*	(31.18)	-3.03	(8.23)	17.63*	(10.42)	-12.99	(19.02)
Age (in months)	-0.20	(0.20)	-0.44***	(0.14)	-0.22	(0.17)	-0.42	(0.41)
Male	-15.14***	(3.05)	-3.98	(2.83)	-16.20***	(2.23)	-19.54***	(2.76)
Speak test language at home ^c								
Sometimes	-25.72***	(5.80)	-28.48***	(4.93)	-27.06***	(5.38)	-23.65**	(9.81)
Never	-33.49***	(7.87)	-44.08***	(8.40)	-34.98***	(11.66)	-37.13**	(15.71)
Immigration status								
Student is immigrant	-26.40***	(4.85)	-23.82***	(4.04)	-25.66***	(8.30)	-19.29**	(8.34)
Father is immigrant	-5.97	(8.23)	-14.33***	(5.09)	-15.87	(10.14)	-9.29	(7.35)
Mother is immigrant	-7.11	(7.68)	2.40	(5.10)	-9.95	(7.82)	-24.05***	(8.26)
Books at home ^d								
10-25	10.16	(6.51)	-2.14	(4.71)	10.71***	(3.67)	-2.26	(6.79)
26-100	21.43***	(5.93)	7.47*	(4.09)	23.06***	(3.66)	20.90***	(7.10)
100-200	28.74***	(9.83)	12.95**	(6.42)	32.71***	(5.13)	18.66*	(10.45)
More than 200	53.43***	(10.88)	14.44**	(6.74)	39.45***	(5.78)	19.43**	(9.56)
Parental education ^e								
Lower secondary	-1.61	(7.17)	3.87	(5.56)	–	–	18.41**	(7.68)
Upper secondary	20.24***	(6.39)	6.02	(6.86)	6.00	(17.49)	17.00**	(8.01)
Post second. (non tertiary)	–	–	15.78*	(8.25)	18.72***	(4.22)	27.86***	(9.57)
University degree	39.27***	(8.97)	14.96*	(8.85)	31.85***	(5.45)	19.27**	(9.41)
Working status of parents ^f								
At least one works half-time	10.60	(8.74)	4.55	(5.26)	-7.08	(5.02)	13.58*	(7.72)
At least one works full-time	11.77*	(6.69)	10.03*	(5.24)	6.41	(4.09)	-1.71	(6.91)
Both work full-time	16.99*	(9.85)	3.34	(7.14)	10.09	(6.15)	5.48	(7.85)
Occupation of parents ^g								
Blue-collar workers	-0.69	(15.54)	2.89	(4.863)	12.72**	(5.57)	2.19	(5.33)
White-collar workers	4.58	(7.57)	1.01	(6.34)	17.95***	(5.26)	15.68**	(7.63)
Household income ^h								
National category 1	–	–	15.10***	(4.43)	–	–	3.05	(7.04)
National category 2	–	–	26.40***	(6.07)	–	–	3.51	(8.07)
National category 3	–	–	51.29***	(10.6)	–	–	16.52**	(6.81)
National category 4	–	–	66.85***	(12.83)	–	–	1.51	(8.02)
National category 5	–	–	-14.34**	(6.69)	–	–	18.45**	(8.27)
Location of school ⁱ								
Village up to 3k people	0.90	(22.97)	9.29	(14.43)	17.68**	(8.24)	-28.54***	(8.22)
Town 3k-100k people	1.08	(22.61)	26.40***	(9.74)	12.80	(9.86)	54.20***	(10.56)
City 100k-500k people	2.92	(23.82)	27.01**	(10.84)	24.26***	(9.11)	54.40***	(13.59)
City over 500k people	19.63	(23.48)	29.50***	(10.77)	16.86**	(8.49)	61.90***	(11.99)
Imputation controls + constant	incl.	–	incl.	–	incl.	–	incl.	–
Students (observations)	3,300	–	5,131	–	5,125	–	3,756	–
Schools (PSUs) [Classes]	138	[138]	147	[199]	154	[154]	138	[159]
R ² [without imputation controls]	0.35	[0.27]	0.29	[0.27]	0.22	[0.22]	0.37	[0.32]

Notes: Dependent variable: PIRLS international reading test score. – Regressions weighted by students' sampling probabilities. – Coefficient: Coefficient estimate. – S.E.: Clustering-robust standard error (taking account of correlated error terms within schools). – Reference categories: ^a Not at all. ^b Five years or younger. ^c Always. ^d 0-10. ^e Not completed lower secondary. ^f None working. ^g In-between ISCO groups. ^h Lowest national category. ⁱ Rural area.

Significance level (based on clustering-robust standard errors): *** 1 percent. – ** 5 percent. – * 10 percent.

Table 3b: Student Background and Reading Performance (HPC Group)

	Germany		Greece		Italy		England	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Pre-school reading performance	7.26***	(1.08)	11.21***	(1.89)	12.81***	(1.46)	26.61***	(2.50)
Kindergarten attendance ^a								
Less than one year	12.01	(10.34)	-15.82	(17.58)	3.85	(13.76)	-25.17**	(11.19)
One to less than two years	-8.85	(5.51)	-6.50	(11.89)	-0.08	(8.94)	-18.02***	(5.86)
Two years or more	-0.90	(4.13)	-3.54	(11.73)	9.32	(8.09)	-16.08***	(6.08)
School commencement age ^b								
Six years	8.08	(5.62)	1.57	(6.86)	15.61***	(3.58)	-13.37*	(6.78)
Seven years	17.19***	(5.83)	-8.27	(8.06)	12.47	(9.00)	-1.52	(23.58)
Eight years or older	-5.05	(12.09)	58.67**	(18.24)	17.81	(46.40)	-40.14	(26.00)
Age (in months)	-1.13***	(0.17)	0.34	(0.40)	0.70**	(0.31)	2.29***	(0.37)
Male	-8.24***	(1.75)	-17.90***	(2.57)	-5.48**	(2.23)	-17.20***	(2.98)
Speak test language at home ^c								
Sometimes	-27.78***	(2.99)	-11.21	(6.81)	-23.80***	(6.96)	-32.85***	(4.58)
Never	-10.65	(7.10)	8.70	(16.04)	-19.15	(16.27)	-60.24***	(11.21)
Immigration status								
Student is immigrant	-31.25***	(2.11)	-1.56	(6.58)	-21.16***	(7.90)	-46.94***	(4.46)
Father is immigrant	-2.15	(2.91)	-9.54	(5.86)	-8.25	(5.73)	5.62	(4.45)
Mother is immigrant	-2.63	(2.83)	-6.04	(4.81)	5.21	(5.41)	13.15***	(4.11)
Books at home ^d								
10-25	8.98**	(3.97)	-6.77	(7.19)	15.24***	(4.38)	0.08	(9.43)
26-100	18.61***	(3.56)	17.49***	(5.90)	23.10***	(4.32)	4.11	(9.51)
100-200	28.75***	(4.04)	21.12***	(7.35)	35.76***	(5.01)	16.08*	(9.57)
More than 200	38.13***	(3.65)	33.02***	(6.49)	47.79***	(4.76)	22.70**	(10.38)
Parental education ^e								
Lower secondary	1.15	(8.85)	5.99	(8.22)	22.47***	(6.53)	13.78	(9.15)
Upper secondary	21.61**	(8.38)	19.45**	(7.69)	33.27***	(6.43)	26.81***	(9.84)
Post second. (non tertiary)	22.91***	(8.37)	33.73***	(9.02)	19.00**	(9.34)	10.77	(11.02)
University degree	32.64***	(8.60)	44.62***	(9.29)	40.29***	(7.10)	27.69***	(9.88)
Working status of parents ^f								
At least one works half-time	6.66	(4.34)	-10.10	(10.12)	3.22	(7.46)	14.69	(10.03)
At least one works full-time	10.27***	(3.26)	-3.97	(7.87)	13.06***	(4.10)	12.66*	(7.03)
Both work full-time	5.05	(3.82)	-9.02	(7.36)	16.63***	(4.65)	10.55	(7.86)
Occupation of parents ^g								
Blue-collar workers	-3.31	(2.72)	-11.70*	(6.21)	0.36	(4.65)	5.11	(6.22)
White-collar workers	10.37***	(2.07)	-0.44	(3.93)	5.91*	(3.41)	7.39*	(4.29)
Household income ^h								
National category 1	2.87	(3.02)	4.98	(5.22)	–	–	16.60***	(6.23)
National category 2	8.84***	(3.33)	3.10	(5.93)	–	–	32.73***	(6.64)
National category 3	12.41	(3.64)	14.77***	(5.50)	–	–	32.74***	(7.24)
National category 4	20.72***	(4.05)	20.45***	(6.70)	–	–	35.43***	(8.10)
National category 5	17.74***	(3.93)	27.41***	(6.64)	–	–	27.58**	(11.08)
Location of school ⁱ								
Village up to 3k people	-10.69	(7.01)	13.58	(15.70)	-27.23***	(10.19)	-10.15	(7.80)
Town 3k-100k people	-3.41	(2.72)	13.19*	(6.80)	-2.09	(3.13)	-13.39**	(5.71)
City 100k-500k people	-6.11	(4.27)	4.75	(8.39)	-2.19	(5.54)	-4.12	(5.69)
City over 500k people	-13.71***	(4.53)	32.75***	(9.16)	-6.45	(7.55)	-6.75	(7.33)
Imputation controls + constant	incl.		incl.		incl.		incl.	
Students (observations)	7,633		2,494		3,502		3,156	
Schools (PSUs) [Classes]	211	[393]	145	[145]	184	[184]	131	[132]
R ² [without imputation controls]	0.35	[0.31]	0.28	[0.27]	0.19	[0.18]	0.30	[0.27]

Notes: Dependent variable: PIRLS international reading test score. – Regressions weighted by students' sampling probabilities. – Coefficient: Coefficient estimate. – S.E.: Clustering-robust standard error (taking account of correlated error terms within schools). – Reference categories: ^a Not at all. ^b Five years or younger. ^c Always. ^d 0-10. ^e Not completed lower secondary. ^f None working. ^g In-between ISCO groups. ^h Lowest national category. ⁱ Rural area.

Significance level (based on clustering-robust standard errors): *** 1 percent. – ** 5 percent. – * 10 percent.

Table 4a: School Characteristics and Reading Performance (LPC Group)

	Argentina		Colombia		Turkey		Macedonia	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Resource endowment								
Class size in reading	-0.10	(0.58)	0.25	(0.33)	-0.17	(0.28)	-0.37	(0.61)
Instructional time	0.42	(0.52)	-0.27	(0.39)	0.37	(0.24)	-0.53	(0.60)
Time spent on homework	-0.82	(0.81)	-3.04***	(0.64)	-1.25**	(0.53)	1.410*	(0.75)
Lack of instructional material	-14.63	(10.37)	-14.51*	(7.61)	-2.11	(5.47)	-31.63***	(10.84)
No lack of instruct. material	-15.32**	(7.27)	9.72	(8.46)	-2.62	(10.46)	-7.80	(9.40)
Lack of instructional staff	1.36	(16.23)	-9.62	(8.99)	-19.66***	(6.04)	8.51	(10.22)
No lack of instructional staff	28.87***	(9.62)	-10.58	(7.56)	-13.85	(10.40)	-1.67	(11.24)
Lack of qualified teachers	7.34	(21.08)	1.73	(8.26)	-4.51	(5.74)	1.55	(12.95)
No lack of qualified teachers	6.77	(7.33)	-4.72	(7.33)	3.15	(9.23)	-7.53	(8.32)
Teacher characteristics								
Gender (female)	19.02*	(11.15)	15.86*	(8.73)	5.58	(5.28)	22.10**	(8.71)
Education								
University degree	-7.91	(18.84)	-1.27	(9.86)	-10.82	(8.73)	11.70	(9.25)
Teaching certificate	7.69	(14.09)	4.43	(10.68)	39.81**	(17.85)	-66.85***	(19.34)
Professional development	-13.91*	(7.68)	-11.97	(8.28)	-13.05*	(7.69)	6.22	(9.93)
Experience	-0.47	(0.41)	-0.43	(0.30)	0.33	(0.47)	0.26	(0.39)
Institutions								
Teacher emphasis on								
Classroom tests	-7.15	(6.93)	4.03	(6.17)	9.51	(6.17)	-2.57	(9.01)
National/reg. examinations	11.03	(8.42)	-5.25	(12.50)	5.47	(10.16)	-12.31*	(7.23)
National/regional curriculum	35.62***	(8.61)	-4.34	(6.89)	15.14***	(5.57)	-17.53	(12.80)
Ability-based class formation	14.72**	(6.98)	5.63	(6.20)	-6.23	(11.11)	19.05**	(7.50)
Taught by same teacher								
One year or less	11.70*	(6.38)	-14.57**	(6.56)	5.44	(9.75)	–	
At least four years	-71.66***	(23.89)	-10.89	(8.48)	11.60**	(5.74)	30.23	(24.59)
Student background controls	incl.		incl.		incl.		incl.	
Imputation controls + constant	incl.		incl.		incl.		incl.	
Students (observations)	3,300		5,131		5,125		3,756	
Schools (PSUs) [Classes]	138	[138]	147	[196]	154	[154]	138	[159]
R ² [without imputation controls]	0.43	[0.35]	0.35	[0.31]	0.27	[0.24]	0.44	[0.34]

Notes: Dependent variable: PIRLS international reading test score. – Regressions weighted by students' sampling probabilities. – Regressions control for all the student background variables reported in Table 3. – Coefficient: Coefficient estimate. – S.E.: Clustering-robust standard error (taking account of correlated error terms within schools).

Significance level (based on clustering-robust standard errors): *** 1 percent. – ** 5 percent. – * 10 percent.

Table 4b: School Characteristics and Reading Performance (HPC Group)

	Germany		Greece		Italy		England	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Resource endowment								
Class size in reading	0.16	(0.34)	0.65	(0.61)	0.01	(0.54)	1.21**	(0.61)
Instructional time	0.26	(0.20)	-0.70*	(0.40)	0.32	(0.21)	-0.31	(0.63)
Time spent on homework	-4.53***	(0.70)	3.06***	(0.66)	-3.33***	(0.75)	0.44	(1.28)
Lack of instructional material	4.12	(11.12)	-1.42	(5.95)	52.04***	(7.40)	–	
No lack of instruct. material	-2.24	(2.54)	1.45	(6.42)	6.48	(4.27)	–	
Lack of instructional staff	-2.50	(6.13)	6.74	(5.95)	4.31	(15.86)	–	
No lack of instructional staff	4.38*	(2.52)	11.68*	(6.71)	-12.24***	(4.00)	–	
Lack of qualified teachers	-14.62***	(4.69)	-6.27	(6.72)	-13.94*	(7.79)	–	
No lack of qualified teachers	-0.42	(3.37)	5.18	(5.59)	-8.11	(4.95)	–	
Teacher characteristics								
Gender (female)	3.94	(2.53)	3.88	(5.26)	14.44**	(6.65)	-4.29	(6.23)
Education								
University degree	-3.79	(6.48)	39.94**	(19.52)	3.22	(4.00)	–	
Teaching certificate	6.13	(6.29)	–		-3.30	(3.97)	8.43	(10.60)
Professional development	-8.02	(12.87)	-11.88	(8.84)	-5.94	(4.59)	20.82***	(5.35)
Experience	0.10	(0.09)	-0.06	(0.24)	0.50**	(0.20)	-0.09	(0.30)
Institutions								
Teacher emphasis on								
Classroom tests	1.09	(2.51)	3.94	(5.44)	-0.86	(4.77)	-0.93	(5.36)
National/reg. examinations	3.78	(5.35)	–		-16.22	(13.80)	10.43*	(5.66)
National/regional curriculum	0.32	(3.75)	-6.66	(7.49)	6.11	(5.25)	-12.07	(13.08)
Ability-based class formation	5.46	(4.47)	-3.62	(6.52)	4.25	(6.29)	0.46	(7.39)
Taught by same teacher								
One year or less	–		16.02**	(7.94)	-6.53	(10.73)	-7.59	(6.75)
At least four years	-3.71*	(2.17)	16.64	(10.55)	-5.06	(4.76)	-12.33	(13.95)
Student background controls	incl.		incl.		incl.		incl.	
Imputation controls + constant	incl.		incl.		incl.		incl.	
Students (observations)	7,633		2,494		3,502		3,156	
Schools (PSUs) [Classes]	211	[393]	145	[145]	184	[184]	131	[132]
R ² [without imputation controls]	0.37	[0.31]	0.34	[0.28]	0.25	[0.20]	0.31	[0.26]

Notes: Dependent variable: PIRLS international reading test score. – Regressions weighted by students' sampling probabilities. – Regressions control for all the student background variables reported in Table 3. – Coefficient: Coefficient estimate. – S.E.: Clustering-robust standard error (taking account of correlated error terms within schools).

Significance level (based on clustering-robust standard errors): *** 1 percent. – ** 5 percent. – * 10 percent.

Figure 1: PIRLS Student Achievement by Percentiles

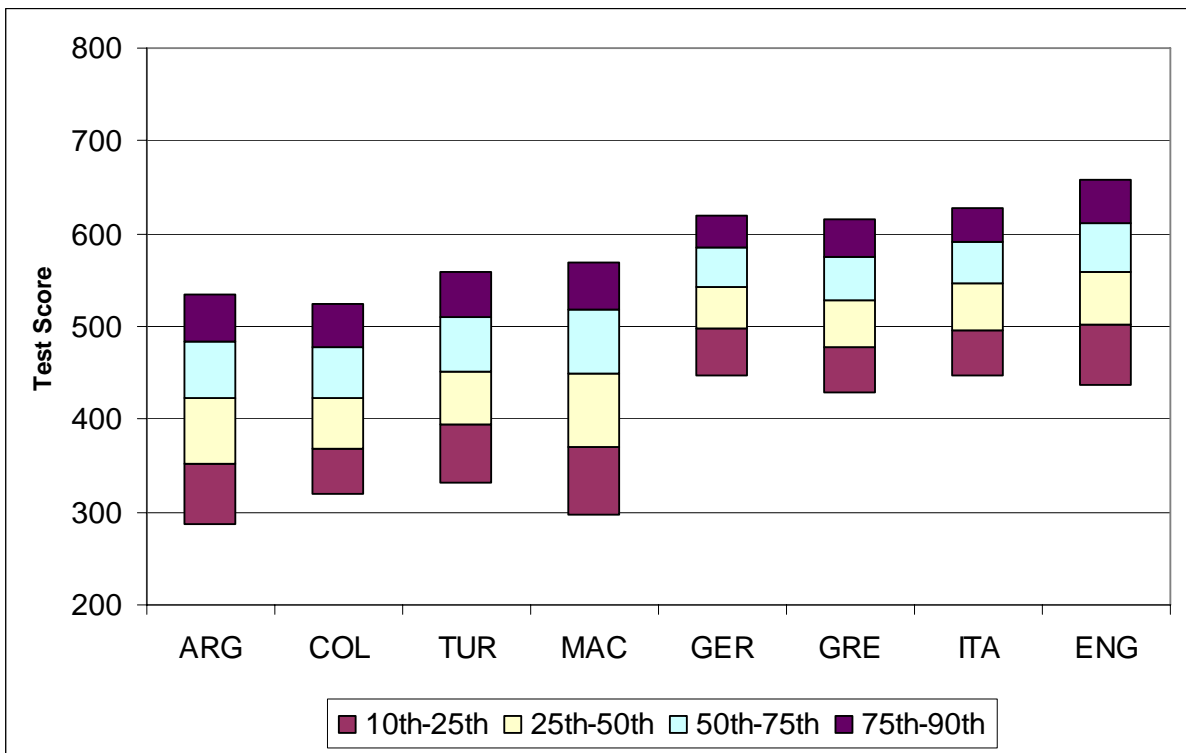


Figure 2: Number of Books in Students' Home

In percentage points, weighted by sampling weights.

■ 0-10 books. ■ 11-25. ■ 26-100. ■ 101-200. ■ More than 200.

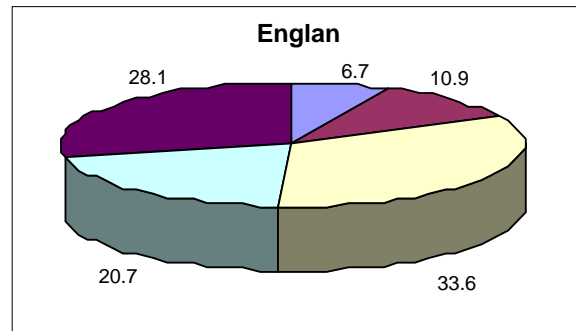
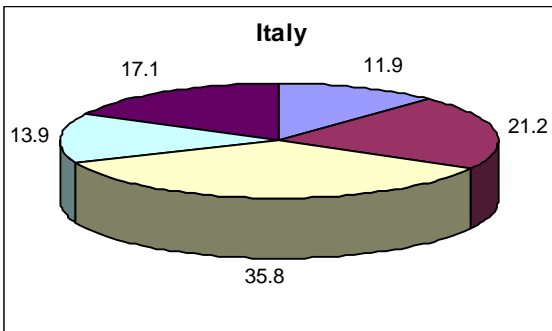
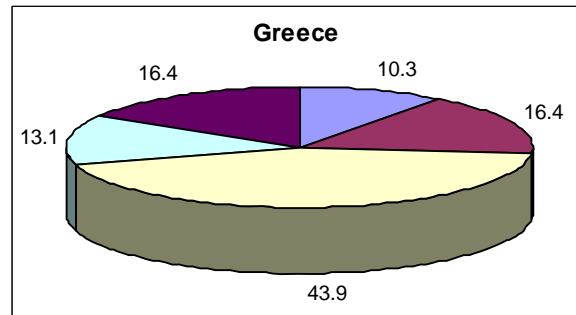
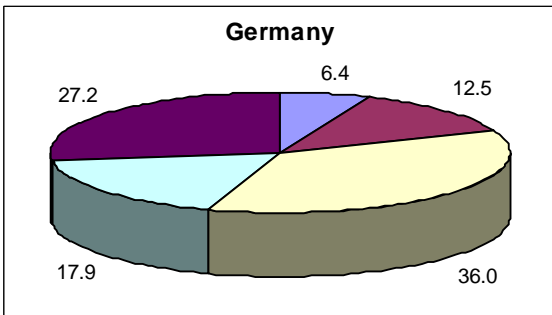
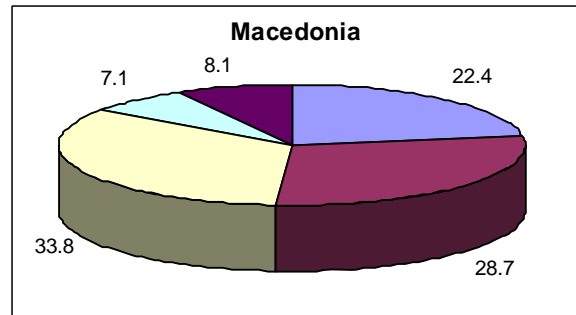
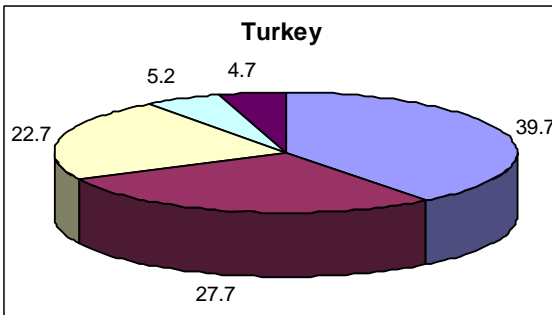
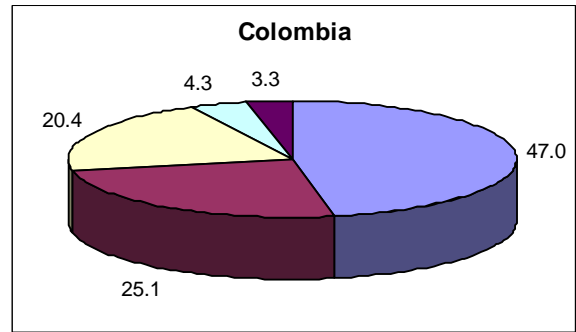
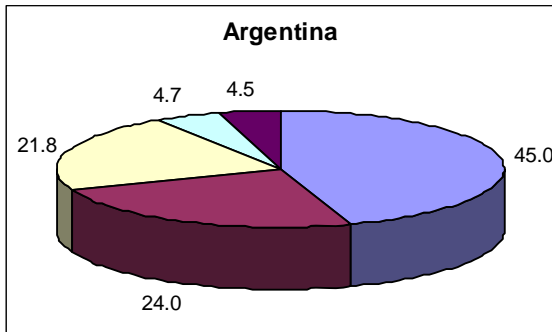


Figure 3: Educational Level of Students' Parents

In percentage points, weighted by sampling weights.

■ Incomplete lower secondary. – ■ Lower secondary. – ■ Upper secondary. – ■ Post secondary (non tertiary). – ■ University degree.

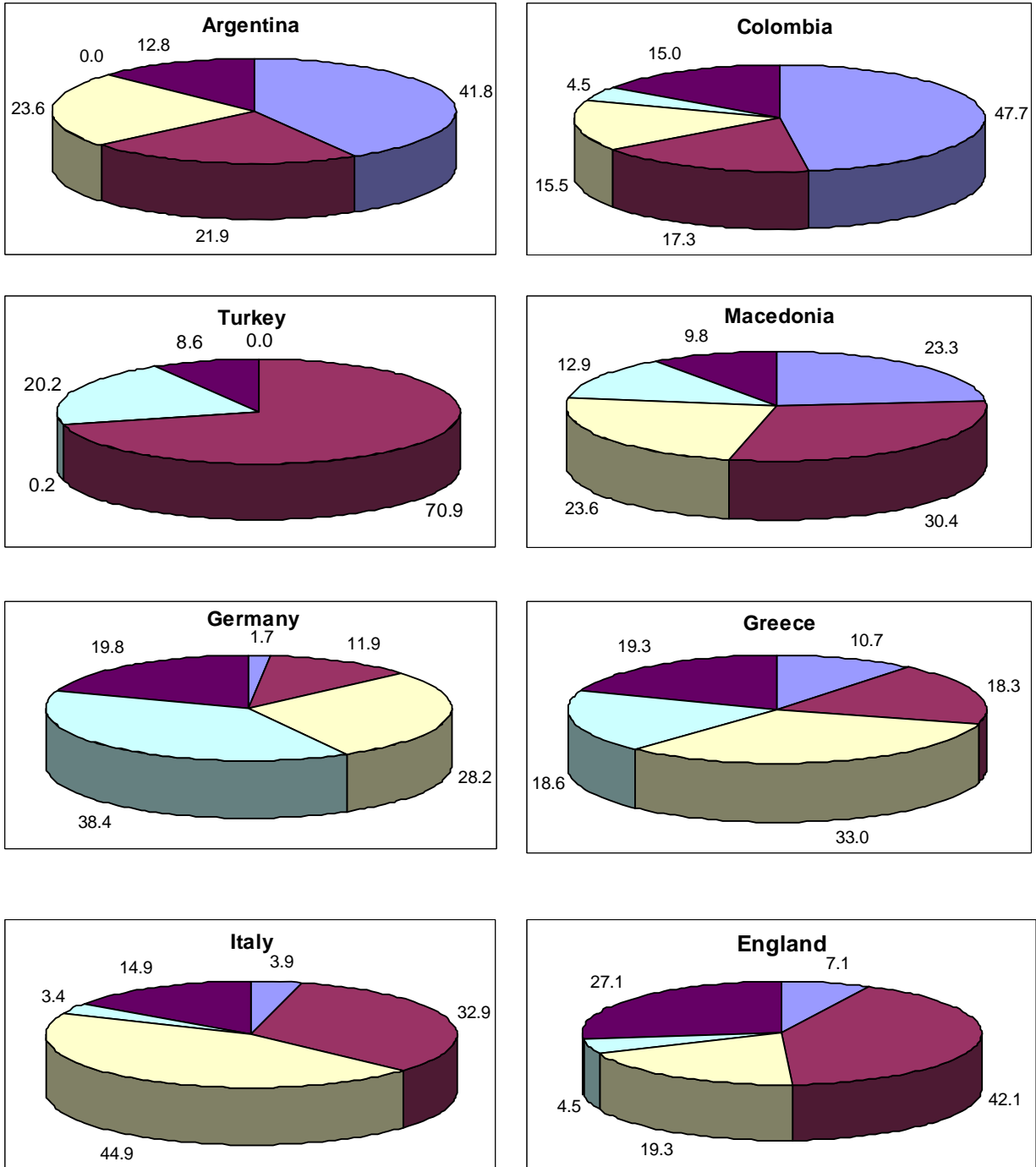


Figure 4: Effects of Books at Home on Student Achievement

Reference group: 10 books or less at home. Estimated using equation (5).

■ 11-25 books. – ■ 26-100. – ■ 101-200. – ■ More than 200.

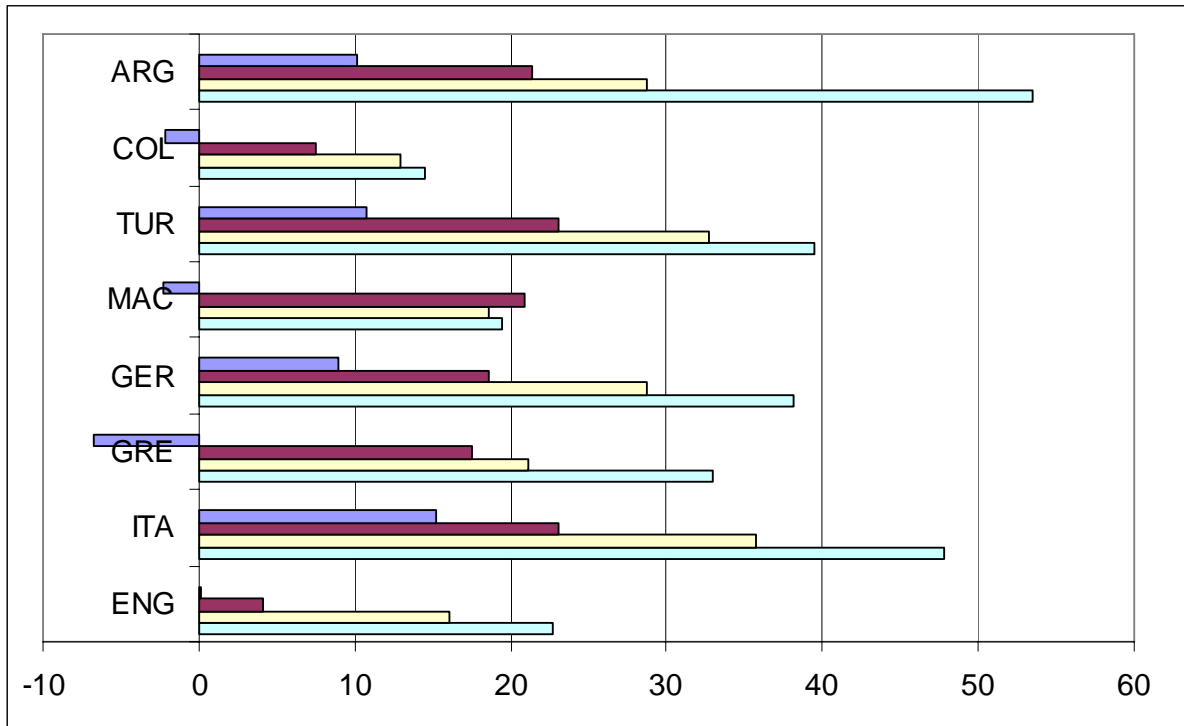


Figure 5: Effects of Parental Education on Student Achievement

Reference group: not completed lower secondary education. Estimated using equation (5).

■ Lower secondary. – ■ Upper secondary. – ■ Post secondary (non tertiary). – ■ University degree.

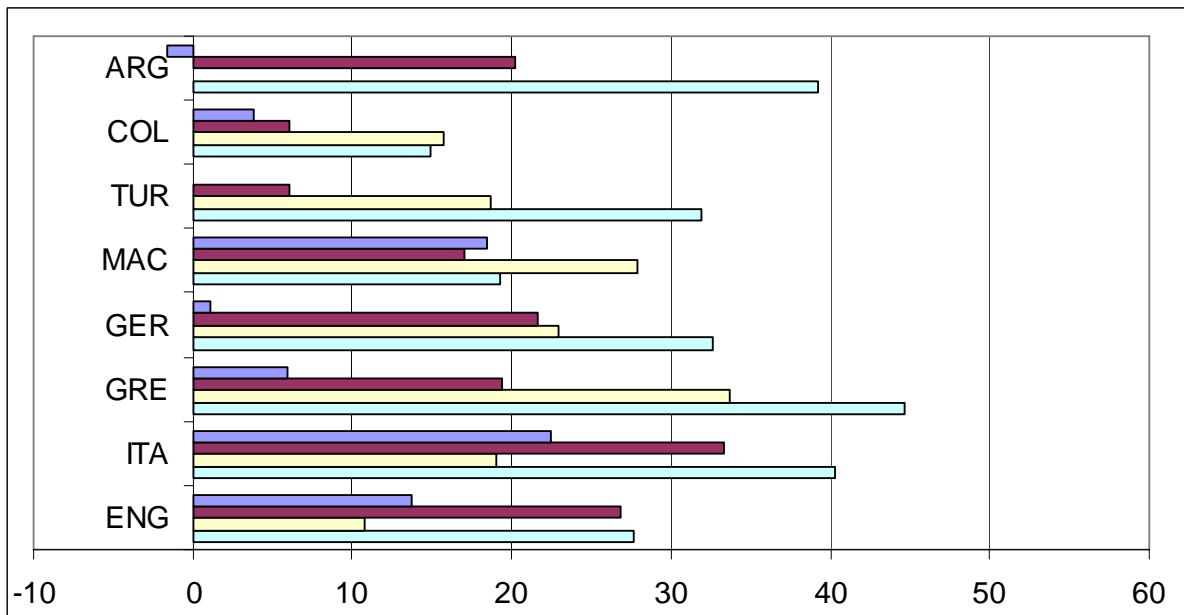


Table A1: Missing Data

	Source	ARG	COL	TUR	MKD	GER	GRE	ITA	ENG
Student background									
Pre-school reading performance	H	29.9	11.0	2.2	24.3	12.8	12.0	2.9	44.9
Kindergarten attendance	H	1.0	0.8	2.6	1.4	1.2	0.4	0.0	0.2
School commencement age	H	31.1	12.4	2.3	24.6	12.8	13.4	3.8	45.1
Age	St	0.5	1.9	0.0	0.0	0.0	0.2	0.0	0.0
Gender	St	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Immigration status									
Student is immigrant	St	22.1	3.3	2.1	11.9	5.7	4.3	0.0	1.4
Father is immigrant	St	21.4	3.5	2.7	10.4	6.4	2.4	1.9	0.8
Mother is immigrant	St	21.4	2.6	2.5	9.9	6.3	2.5	1.8	0.6
Speak test language at home	St	20.8	3.7	2.2	7.3	2.9	1.8	0.3	0.7
Books at home	H	30.9	12.3	3.0	25.5	13.4	12.5	3.7	44.8
Parental education	H	39.4	22.4	21.7	40.0	35.1	15.8	4.3	48.5
Working status of parents	St	35.7	17.0	11.8	36.2	16.0	16.6	7.3	45.5
Occupation of parents	H	39.6	19.9	11.5	37.4	16.8	16.4	5.9	46.9
Household income	H	–	12.7	–	30.5	26.2	19.1	–	51.2
Location of school	Sc	15.3	8.9	1.3	4.0	4.6	8.8	0.0	9.2
School characteristics									
Class size in reading	T	6.6	10.1	1.3	3.0	4.8	5.2	0.0	6.1
Instructional time	Sc	20.3	27.6	6.9	3.5	25.8	17.6	0.0	17.4
Time spent on homework	St	5.5	7.5	0.8	6.2	4.9	5.2	0.0	9.3
Lack of instructional material	Sc	12.8	2.0	1.3	5.1	4.9	6.0	1.8	–
Lack of instructional staff	Sc	11.1	3.8	2.2	7.0	4.6	6.6	3.3	–
Lack of qualified teachers	Sc	12.9	2.9	0.7	8.7	7.4	6.0	3.3	–
Teacher characteristics									
Gender	T	4.4	2.6	0.0	4.5	6.5	5.2	0.0	7.7
Education									
University degree	T	4.4	3.8	0.6	4.5	7.8	5.2	0.5	7.7
Teaching certificate	T	5.6	4.4	0.0	5.8	8.8	21.0	3.2	7.6
Professional development	T	5.3	3.1	0.0	5.8	10.2	5.2	1.1	7.6
Experience	T	7.2	7.9	1.1	15.1	6.6	7.4	2.2	6.9
Teacher emphasis on									
Classroom tests	T	8.1	3.8	0.6	6.5	12.7	6.2	3.5	6.8
National/reg. examinations	T	13.3	8.2	4.3	13.7	17.8	–	11.8	6.8
National/regional curriculum	Sc	4.0	2.2	2.2	1.6	5.2	6.0	1.9	3.8
Ability-based class formation	Sc	4.3	1.9	0.0	1.6	6.4	7.8	0.5	3.8
Taught by same teacher	Sc	6.3	14.0	2.9	3.3	8.7	35.8	6.7	6.5

Notes: Imputed data in percent of all observations, weighted by sampling probabilities.