Do Schools Reinforce or Reduce Learning Gaps between Advantaged and Disadvantaged Students? Evidence from Vietnam and Peru

December, 2014

Paul Glewwe
Department of Applied Economics
University of Minnesota

Sofya Krutikova
Institute for Fiscal Studies
London

Caine Rolleston
Institute of Education
University of London

Abstract
This paper examines – for two developing countries, Vietnam and Peru – whether disadvantaged children learn less than advantaged children when both types of children are enrolled in the same school. This is done by estimating education production functions that contain two school fixed effects for each school, one for advantaged children and one for disadvantaged children. The paper examines six different definitions of advantage, based on household wealth, cognitive skills at age 5, gender, ethnicity (Peru only), maternal education, and nutritional status. The results show no sign of discrimination against disadvantaged groups in Vietnam; indeed if anything one traditionally advantaged group, males, seems to do worse in school than the corresponding disadvantaged group, females. In contrast, in Peru ethnic minority students and students who enter primary school with low cognitive skills appear to learn less in school than ethnic majority students and students with relatively high cognitive skills, respectively, who are enrolled in the same school.
I. Introduction

Investing in education increases economic growth and, at the individual level, raises incomes and living standards (Krueger and Lindahl, 2001). Developing countries have achieved substantial increases in primary and secondary enrolment rates in the last 25 years (Glewwe et al., 2013), yet students in many developing countries learn much less than students in developed countries (Glewwe and Kremer, 2006). Equally disturbing is that, within developing countries, some children learn much more than others, and in some countries schools appear to widen learning gaps between advantaged and disadvantaged students (Banerjee and Duflo, 2011). While eliminating inequality in student learning is probably impossible, there may be scope for reducing within-country learning gaps by implementing education policies that disproportionately benefit disadvantaged students.

The existing literature on learning gaps focuses on whether disadvantaged students attend lower quality schools (McEwan, 2004; McEwan and Trowbridge, 2007). Yet recent evidence (reviewed below) finds large learning differences between children of different backgrounds even within the same schools. This paper uses panel data from Vietnam and Peru to investigate *intra*-school differences in learning. It uses six different advantage indicators to assess whether advantaged students learn more than disadvantaged students when they attend the same school. Comparison of Vietnam and Peru is of particular interest given their starkly different scores on the 2012 PISA assessments (OECD, 2013); of the 65 countries assessed, Vietnam ranked 17th in mathematics and 19th in reading (ahead of the U.S. and the U.K.), while Peru ranked last (65th) in both subjects. This difference is particularly striking given that Peru’s income per capita in 2012 ($6060) was four times larger than Vietnam’s ($1550) (World Bank, 2013).

More specifically, this paper estimates the combined impacts of all school and teacher characteristics on student learning by estimating separate fixed effects for advantaged and dis-
advantaged students in each school. For each country, the relative sizes of the two fixed effects are compared for six types of advantage (being wealthy, having higher skills at age 5, being male, being from an ethnic majority group, higher maternal education, and better nutrition in early childhood) to examine whether advantaged children learn more than otherwise comparable disadvantaged children in the same schools.

The paper advances knowledge on student learning in developing countries in three ways. First, the rich data set, which links detailed individual level panel data from early childhood to very detailed school surveys, reduces omitted variable bias. Second, separate school fixed effects for advantaged and disadvantaged children allow for within-school heterogeneity in the impact of school and teacher variables on learning, which is usually ignored in the literature, but may be an important driver of inequality. Third, using very similar data for two very different countries allows for cross-county heterogeneity in whether schools raise or lower student learning gaps.

The paper is organized as follows. Section II reviews the relevant literature. Section III describes the Vietnam and Peru school systems, and Section IV presents the methodology. The next three sections describe the data, discuss estimation issues and present results, respectively. Section VIII provides robustness checks and an extension to the analysis. Section IX concludes.

II. Literature Review

Gaps in education outcomes between children from more and less socio-economically advantaged backgrounds are widely documented. Filmer and Pritchett (1999) use data from 35 developing countries to show that most short-falls in basic education indicators (e.g. enrollment and drop-out) are due to children in the bottom 40% of the wealth distribution. Indeed, the link between socio-economic background and schooling has been found in both countries examined in this paper: Vietnam (Glewwe and Jacoby, 2004) and Peru (Jacoby, 1994). 
A large literature considers why children from advantaged backgrounds have better educational outcomes. This literature focuses mainly on differences at home and in school, which can affect learning through many mechanisms, including the availability of good quality schools, household demand for and attitudes toward education, and children’s learning capacity, which may be reduced by their exposure to adversity (Duncan and Murnane, 2011).

Home factors that seem to matter include income, parental education and early childhood nutrition (Glewwe and Miguel, 2008; Behrman, 2010; Alderman and Bleakley, 2013). There is evidence that school and teacher quality affect education outcomes in many contexts (Aaronson, 2007; Altonji and Mansfield, 2011; Lai et al., 2011), yet the available evidence is inconclusive on exactly what constitutes school quality. Glewwe et al. (2013), reviewing 20 years of studies on developing countries, conclude that much of what matters is unlikely to be easy to observe.

Whatever school quality is, the literature finds that access to it varies by children’s socio-economic background, and this explains much of the differences in their educational outcomes. For example, U.K. and U.S. studies find that learning gaps between more and less advantaged children widen during their time in school, and that differences in school quality explain part of this trend (Currie and Thomas, 2001; Fryer and Levitt, 2004). There is also developing country evidence: two recent studies find that a significant portion (e.g. 50-69% in Guatemala) of the learning gap between indigenous and nonindigenous children is due to differences in school quality in Latin American countries (McEwan, 2004; McEwan and Trowbridge, 2007).

Much of the literature focuses on differences between schools attended by advantaged and disadvantaged children. Our paper adds to a smaller literature on within-school differences in learning between children of different backgrounds. Banerjee and Duflo (2011) suggest that this leads to inequality in educational outcomes in many developing countries. We know of no studies that directly examine the effect of within-school differences in the educational experiences of
children from different backgrounds on learning gaps, yet there is evidence of within-school differences in treatment of children of different backgrounds that is consistent with such an effect.

For example, Hanna and Linden (2012) randomly assigned child characteristics (gender, age and caste) to cover sheets of pupil exams in India. Teachers recruited to grade those exams, who could see the characteristics from the sheets, gave low-caste children lower scores. Botelho et al. (2010) found similar discrimination by race in Brazil, and in the U.S. Dee (2005) showed that racial and ethnic distance between teachers and pupils affects teachers’ perceptions of pupil performance, especially of disadvantaged students. Other studies find that discrimination may lower pupils’ performance via “stereotype” threat. Hoff and Pandey (2006) show that a caste gap emerges in India when school-age task participants’ caste is revealed. This supports the emphasis on identity by Akerlof and Kranton (2010), who argue that within a school the dominant group’s norms can reduce the performance of “outsiders”. These findings motivate our focus on within-school differences in learning among children from advantaged and disadvantaged backgrounds.

III. Overview of the School Systems in Vietnam and Peru

The education systems of Vietnam and Peru provide an interesting contrast. Enrollment has increased in both countries, but Vietnam’s expansion has been much more equitable.

Primary school has long been compulsory in both countries. In Vietnam, the 1991 Law on Universal Primary Education made primary education (grades 1–5) compulsory for 6-14 year old children. Universal primary education was essentially achieved by 2010, with a net enrollment rate of 98% (World Bank, 2012). Peru’s 2003 General Education Law also stipulated compulsory basic education, and its net enrollment rate was 95% in 2010 (World Bank, 2012).

Education quality is a focus in both counties. Vietnam has invested heavily in primary school facilities, teacher training, and curricular and textbook reform, and raised hours per day in
primary school, which were relatively low. Vietnam’s high rank in the 2012 PISA tests suggests that its efforts succeeded in raising school quality; Peru’s PISA results indicate less success.

Evidence also suggests large inequities in educational provision in both countries. Vietnam has introduced several new programs to ensure “minimum standards” of quality to all, especially in disadvantaged areas (World Bank, 2004; UNESCO, 2011). Although concern remains in Vietnam about growing inequity (World Bank, 2011), comparisons with other developing countries suggest that Vietnam’s educational outcomes are quite equitably distributed (Holsinger, 2005). In contrast, in Peru there are large gaps in pupils’ educational outcomes by socio-economic group, ethnicity, and location (Murray 2012). Indeed, Peru had the largest ratio of performance at the 95th percentile over performance at the 5th percentile of all 43 countries participating in the 2000 PISA assessment (Crouch et al., 2009).

These differences between Vietnam and Peru suggest that their education systems have different effects on learning gaps between advantaged and disadvantaged students. We use comparable data to examine these effects for a relatively high-performing, and equitable, education system (Vietnam) and a more unequal, less well performing system (Peru).

IV. Methodology

This paper examines whether the impact of school characteristics on learning differs within schools between advantaged and disadvantaged pupils. This section presents the equations to be estimated. Data and estimation strategy are then discussed in Sections V and VI.

To begin, consider a linear approximation of a general production function for cognitive skills (C) of primary school students:

\[
C = \beta_0 + \beta_1 C_p + \beta_2 N_p + \beta_3 N + \beta_4 PE + \beta_5 PT + \beta_6 EI_3 + \beta_7 IA + \gamma_1 \text{SC} + \gamma_2 \text{TC} + u
\] (1)
In this equation, \( C_p \) is “preschool” cognitive skills (skills when starting primary school at age 5), \( N_p \) is “preschool” nutritional status (nutrition from birth to age 5), \( N \) is current nutritional status, \( PE \) is parental education, \( PT \) is parental time spent with the child during primary school on activities that raise cognitive skills, \( EI \) is purchased educational inputs such as textbooks, school supplies and tutoring, \( IA \) is the child’s innate ability, and \( SC \) and \( TC \) are vectors of school and teacher characteristics. The residual, \( u \), accounts for errors due to the linear approximation and measurement errors in \( S \) and is assumed to be uncorrelated with the explanatory variables.

Equation (1) assumes linear separability between household/child characteristics and teacher/school characteristics; interactions between household and child variables and school and teacher variables are assumed to be negligible. This is a strong assumption, yet studies of student learning in developed and developing countries often invoke it (e.g. Rockoff, 2004; Rivkin et al., 2005; Chay et al., 2005; and Rothstein, 2010). We will relax this assumption by allowing the effects of teacher/school variables to differ across advantaged and disadvantaged students.

Glewwe et al. (2013) explain how difficult it is to estimate the impacts of school and teacher characteristics on student learning. Fortunately, for the main objective of this paper – to assess whether schools reinforce or reduce learning gaps between advantaged and disadvantaged students – one can simply use a summary measure of school quality that captures both easy and difficult (if not impossible) to measure school and teacher characteristics. More specifically, one can combine the impacts of school and teacher variables in equation (1) into school fixed effects:

\[
C = \beta_1 C_p + \beta_2 N_p + \beta_3 N + \beta_4 PE + \beta_5 PT + \beta_6 EI + \beta_7 IA + \sum_{s=1}^{S} \delta_s D_s + u \quad (2)
\]

where \( \delta_s = \beta_0 + \gamma_1 SC_s + \gamma_2 TC_s \), and \( D_s \) is a dummy variable for school \( s \). Note that each \( \delta_s \) includes \( \beta_0 \); one can estimate only the relative, not the absolute, impact of each school on student learning. Two notable features of equation (2) are: 1. No restrictions were imposed on (1) to
obtain (2), the latter simply converts each school’s $\gamma_1^{\text{SC}} + \gamma_2^{\text{TC}}$ term into a school fixed effect; and 2. These fixed effects include all possible interactions between the SC and TC variables.

The disadvantage of estimating (2) instead of (1) is that (2) does not reveal which school characteristics increase student learning. Yet the advantage is that the fixed effect captures both observed and unobserved school variables that affect learning and so can be used to assess the role of school (and teacher) characteristics in explaining that outcome. This method has been used to study both school and teacher quality (Aaronson et al. 2007; Rivkin et al. 2005; Rockoff 2004).

To date, studies that use school fixed effects assume there are no interactions between (observed or unobserved) school variables and student and household variables. Thus this method assumes that within schools all students benefit equally from what schools offer regardless of student characteristics. To our knowledge, this assumption has never been relaxed. Yet the evidence discussed in Section II of differential performance among students in the same schools due to caste or race discrimination, or the effects of stereo-type threat, belie this assumption.

One way to allow the impacts of school and teacher characteristics to differ among students in the same school is to divide students in each school into two or more groups and allow each group to have a distinct school fixed effect. In particular, if children can be classified as “advantaged” and “disadvantaged” one can rewrite (2) as:

$$C = \beta'X + \sum_{s=1}^{S} \kappa_s D_s + \sum_{s=1}^{S} \theta_s D_s A + u_3 \quad (3)$$

where $A$ is a dummy variable indicating “advantaged” students and, to reduce clutter, $\beta'X$ denotes $\beta_1 C_p + \beta_2 N_p + \beta_3 N + \beta_4 P + \beta_5 PT + \beta_6 EI + \beta_7 IA$. The $\kappa_s$ term measures the impact of school $s$ on disadvantaged students, and $\kappa_s + \theta_s$ measures its impact on advantaged students. If schools affect the learning of both types of students equally, then $\theta_s = 0$. This is what we test in this paper.
We use six different definitions of advantage: being wealthy; having high skills at school starting age; being male;\(^1\) belonging to the dominant ethnic group;\(^2\) having an educated mother; and adequate nutrition in early childhood. How they are measured is discussed in the next section.

To clarify what we intend to estimate, consider the ideal “experiment” that would yield such an estimate. Assume that students are randomly “labeled” as advantaged or disadvantaged conditional on their values of \(X\). These labels may be correlated with some \(X\) variables – e.g. wealthy parents may help their children more with schoolwork (an \(X\) variable) – but conditional on \(X\) the labels are randomly assigned and so do not predict student performance unless teachers or schools, intentionally or unintentionally, favor students with a particular label. If no favoring exists, estimated school fixed effects conditional on \(X\) will not differ between students randomly labeled as advantaged or disadvantaged: \(\theta_s\) should be 0. Yet if favoritism exists, estimated school fixed effects conditional on \(X\) will differ across advantaged and disadvantaged students: \(\theta_s \neq 0\).

This “experiment” closely resembles that of Hanna and Linden (2012), yet, our data are not from a randomized control trial (RCT). Yet they were able to randomize only favoritism when grading exams, not favoritism in daily classroom interactions with students, and it would be very difficult to randomize labeling to detect this more general favoritism. While we cannot use an RCT to investigate such general favoritism, an advantage we do have is unusually detailed data that allow us to control for a wide variety of student and household variables, including cognitive skills at age 5 (before starting school), which could account for many difficult to measure factors such as innate ability (IA) and many types of parental time (PT) and educational inputs (EI). This completes the discussion of our methodology. We now turn to the data we use to implement it.

\(^1\) While some evidence suggests that girls outperform boys in reading (e.g. Guiso et al, 2008), other studies (e.g. Fryer and Levitt, 2010) find a significant, persistent gender gap favoring boys in math in both developed and developing countries. However, there is little evidence, at least in the U.S., of a gender gap before starting school.
\(^2\) We use this definition only for Peru as only 8% of the Vietnam school survey sub-sample are ethnic minorities.
V. Data

This section presents the data we use, and discusses the definitions of advantage.

A. Survey and Sampling. This paper uses data from Young Lives, a multi-country longitudinal study of child poverty in four developing countries: Ethiopia, India, Peru and Vietnam. Each country collects data on two cohorts of children: 1,000 children born in 1994-95 and 2,000 born in 2000-01. This study uses data only from the latter cohort, known as the “younger cohort”. In each country, the 2,000 younger cohort children were sampled by randomly selecting 100 children who were 6-18 months old in 2001-02 from each of 20 sites. The data are statistically representative of the site level populations, but are not nationally representative. However, the 20 sites in each country were purposively selected to represent each country’s socio-economic and geographic diversity (except for the wealthiest areas).

To date, three rounds of household level data, collected in 2002, 2006-07 and 2009, are available. School-level data were then collected in 2011 for sub-samples of the younger cohort children. The 2011 school surveys differ somewhat by country, reflecting differences in school systems and policies. This study utilizes both the household and school survey data, focusing on Peru and Vietnam due to the availability of recently collected school survey data.

The Vietnam school survey was conducted in October, 2011, in all 20 sites. The school survey sub-sample, the focus of this paper, consists of all younger cohort children enrolled in grade 5 (the appropriate grade for their age) in the 2011-12 school year; all schools attended by these children were surveyed. This sub-sample consists of 1,138 children in 92 primary schools.

The Peru school survey was conducted from October to December, 2011. The sampling was stratified, with four school types as strata so that each type – private urban, public urban,
public rural bilingual and public rural non-bilingual (Spanish only) – is represented. A random sub-sample was drawn of younger cohort children in each of these four school types; the school sample consists of these children’s schools; all younger cohort children in them were surveyed. This sub-sample has 572 children in 132 schools. Most are in Grade 4 (59%), while 32% are in Grade 5 and 9% in Grade 3. After dropping a few observations with missing data, the final samples are 547 children in 132 schools in Peru and 1,129 children in 90 schools in Vietnam.

**B. Testing.** Child learning is measured using scores on two curriculum-based tests, a math test and a reading comprehension test, administered as part of the school surveys. In Vietnam, both tests had 30 multiple-choice items that reflected the Grade 5 curriculum; they are similar to the Ministry of Education’s Grade 5 Assessment tests (World Bank 2011). In Peru, grade-specific math and reading tests with 30-35 items were developed by *GRADE* to reflect the curricula in grades 3-5. The tests were designed to reflect each country’s curriculum and so are not comparable across countries. The Peru tests have common items across grades, so one can use Item Response Theory (IRT) to create scores comparable across grades (as in Das and Zajonc, 2010).

**C. Advantage.** The first of the six advantage categories used in this paper is being wealthy, defined as living in a household in the top two wealth quintiles, based on a wealth index that combines measures of housing quality, ownership of consumer durables and access to key services. Having high skills at school starting age, the second advantage category, is defined as performing in the top 40 per cent on the Cognitive Development Assessment (CDA) test, which measures children’s basic quantitative skills at age five (the school starting age in both countries.

---

5 The English versions of the complete set of tests and questionnaires used in the Vietnam and Peru school surveys are available from the Young Lives website: http://www.younglives.org.uk/what-we-do/school-survey.
6 The curriculum itself can induce gaps between advantaged and disadvantaged students; we assess whether the entire school system generates such gaps and leave to future research which aspects of the systems produce any gaps we find.
7 The wealth index is a simple average of three individual indices ranging from 0 to 1, including housing quality (rooms per person, floor and roof quality), consumer durables (mean of ownership of radio, bicycle, TV, car/truck, motorbike/scooter, landline telephone, bed/table) and access to services (electricity; piped water into dwelling/yard; access to own pit latrine or flush toilet; cooking fuel is electricity, gas or kerosene).
Robustness checks in Section VIII show that the main findings hold if other cut-off points are used for the wealth and skill advantage classifications.

Being male is the third advantage type. The fourth, ethnicity, is used only for Peru; only 8% of the sample in the Vietnam school survey are ethnic minorities. Ethnicity is a sensitive issue in Peru. Thus in the Young Lives sample the great majority of mothers identify their children as *Mestizo*. For the purpose of identifying disadvantage, a more accurate definition of ethnicity is the mother’s first language; it is an ethnicity marker and so is associated with exclusion in Peru (Escobal et al., 2012). Thus children are classed as being an ethnic minority if their mother’s first language is not Spanish (but rather Quechua, Aymara, or another indigenous language).

The last two definitions of advantage are having an educated mother and adequate nutrition in infancy. The former is defined as having a mother who completed at least primary school. The latter is defined as not being stunted at age one, i.e. having a height-for-age Z-score greater than -2.

While these types of advantage are correlated, they overlap only partially (see Appendix Tables 2 and 3). Table 1 shows that they represent advantage in terms of skill acquisition; for all six types, the advantaged group almost always has significantly higher scores on cognitive tests (the CDA, described above, and the Peabody Picture Vocabulary Test – PPVT) taken at age 5, before the children started school, suggesting that advantaged children enjoy favorable conditions for acquiring skills before starting primary school.

Equation (3) can be estimated only for schools with at least one advantaged child and at least one disadvantaged child in our sample. Thus the estimation samples vary by the definition of advantage; Appendix Table 4 shows the number of observations and schools in each of these.

---

8 The CDA is a 15 item instrument developed by the International Evaluation Association to assess the cognitive development of 4-year-old children. The sub-scale administered to the Young Lives children links to perceptions of quantity by testing children’s understanding of concepts such as *few, most, half, many, equal, and pair*.

9 Ethnicity is not used as an advantage category for Vietnam, yet a variable indicating that a child an ethnic majority group is used as a control variable in the estimations. It is defined as being *Kinh* (i.e. not *Hmong, Dao, Tay* or other).

10 Meaning a person of combined European and indigenous descent.
estimation samples. In the Vietnam sample, of the 90 schools, 82 are attended by two or more sample children. Of these 82, 48 have sample children from both the poorer and richer groups, 61 include both more able and less able children, 77 have both boys and girls, 61 have children with both more and less educated mothers, and 63 have children who were and were not malnourished at age 1. Of the 132 schools in Peru, 80 have two or more sample children. Of these 80, 36 have both richer and poorer children, 47 have both more able and less able children, 63 have both boys and girls, 25 have both ethnic minority and ethnic majority children, 48 have children with both more educated and less educated mothers, and 54 have children who were and were not stunted at age 1. Limiting the analysis to schools with both advantaged and disadvantaged sample children raises issues of selection bias in the estimates of $\theta_s$; we return to this issue in Section VIII.

VI. Estimation

Equation (3) is estimated using data on child, parent and household characteristics from the Young Lives data. Skills acquired before entering primary school ($C_p$) are measured by scores on the CDA and PPVT tests in Round 2, when the children were five years old. The CDA measures basic quantitative skills, and the PPVT measures receptive vocabulary. Both tests have been validated in many contexts and are highly correlated with broad-based IQ measures, which suggests that they can also serve as controls for “innate ability” (IA). Nutritional status before entering school ($N_p$) is measured by maternal assessments of the child’s size at birth (5-point scale, very small to very large) and of the child’s health at age 1 relative to that of other children, as well as the height-for-age Z-score measured at age five (in Round 2). Current nutritional status

---

11 Perhaps the precision of estimates for the non-school variables would be higher, and bias in the school effects lower, if we used the whole sample (not just the “overlap” sub-sample of schools that have at least one advantaged child and one disadvantaged child), still estimating the school effects only for the “overlap” sub-sample. Yet the estimates of differences in school effects change little when the whole sample is used (results available upon request).

12 For details on the tests, their psychometric properties, and construction of the scores, see Cueto et al. (2009).
(N) could, in principle, be measured by the height-for-age Z-score measured in Round 3, but it is highly correlated with Round 2 height-for-age (0.80 for Vietnam, 0.79 for Peru), so it is excluded.

Parental education (PE) in equation (3) is measured by dummy variables for different levels of education, separately for mothers and fathers, with no education as the base category. Parental time with the child (PT) is measured by a dummy variable indicating whether someone outside the household cared for the child (including attending a crèche) in Round 1 (when the child was about one year old), a measure of the mother’s mental health in Round 1,\(^{13}\) a dummy variable for whether the child is an only child, and household size. Educational inputs (EI) are measured by time in pre-school and several asset variables (the wealth index, land owned, and per capita consumption) that should be highly correlated with unobserved educational inputs.\(^{14}\)

A few control variables are also added to equation (3): age when the test was taken; and dummy variables for being in grade 3 and 4 (Peru only), for ethnic minority and for being male. Finally, school identifiers from the school survey are used to generate school fixed effects; these are interacted with the relevant advantage dummy variable to estimate \(\theta_s\).

Table 2 presents summary statistics for all these variables. The children were around 10 when they took the math and verbal tests used as outcome measures. Test scores were transformed using a three parameter IRT model, with means and standard deviations standardized to 500 and 100, respectively.\(^{15}\) Most of these variables are self-explanatory, except that an IRT model was also used to generate CDA and PPVT scores that (for the full Young Lives sample) were standardized to have a mean of 300 and standard deviation of 50. Finally, child nutrition is measured by height-for-age Z-scores using WHO standards.

---

\(^{13}\) Maternal mental health is measured using the Self-Reporting Questionnaire; see the Table 2 notes for details.

\(^{14}\) This means that (2) and (3) are no longer “pure” production functions as other explanatory variables, such as child ability, may also affect educational input purchases. But this will have little effect on estimated school fixed effects.

\(^{15}\) The scores and standard deviations in Table 2 are not exactly 500 and 100, respectively, as the standardization was done for the whole school survey sample, including school peers.
A number of estimation issues require further discussion. First, recall from Section IV the “thought experiment” underlying our analysis. It is straightforward when disadvantage is defined as ethnicity or gender: conditional on $X$, ethnicity and gender should be considered as randomly assigned labels that have no predictive power for educational outcomes unless teachers or schools intentionally or unintentionally favor students with those labels. Even if girls tend to have better language skills, and boys better math skills, after controlling for skills at age 5 there should be little difference in the learning of boys and girls.\textsuperscript{16} Things are a bit more complicated when advantage is defined by household wealth; it should not directly affect learning after conditioning on purchased educational inputs (EI),\textsuperscript{17} but a complication is that it is a continuous variable. Teachers or schools may favor wealthier students in two distinct ways: gradual favoring that steadily rises with wealth, or “sharp” favoring that “turns on” when wealth crosses a specific threshold. In the former case, the school fixed effects will vary continuously with wealth, so estimates of their difference when the sample is divided roughly in half will estimate the difference in favoring between an average advantaged student and an average disadvantaged student. In the latter case, school fixed effects take only two values in any given school, and if the “jump” in favoritism occurs near the middle of that variable’s distribution one can estimate that jump. However, if the “jump” is in one of the distribution’s tails it is difficult to estimate because for most schools the data will not have both advantaged and disadvantaged children.

A final complication is when the advantage variable is both continuous and directly affects learning, which occurs for three of our variables:\textsuperscript{18} skills at age 5, parental education, and nutritional status. Such variables must be included in $X$. If favoring of advantaged students

\textsuperscript{16} In fact, differences in cognitive skills by gender at age 5 are small in Vietnam and insignificant in Peru; see Table 1.

\textsuperscript{17} Our data contain few purchased educational inputs (e.g. preschool), so wealth could have a direct effect as proxy for unmeasured educational inputs. Variation in skills at age 5 should control for most variation in the tendency to purchase educational inputs, yet we also add consumption expenditure as a regressor; it is only marginally significant.

\textsuperscript{18} If advantage were a dummy variable that directly affected learning, it would be impossible to estimate favoring of advantaged students because that $X$ variable would be a linear combination of those students’ school fixed effects.
gradually increases as the advantage variable rises, it would be impossible to estimate such favoring because it cannot be disentangled from that variable’s direct effect on learning; the estimated school fixed effects would not differ across the two types of students. Yet if favoring consists of a discrete “jump” when a threshold is crossed, it can be identified if that threshold is roughly in the middle (say, middle two quartiles) of that variable’s distribution and that variable’s direct impact on learning is smooth (does not “jump”). To define advantage in terms of nutritional status, we use the standard definition for not being stunted: a height for age Z-score above -2; for the other two variables we define disadvantage so that about 60% of the sample is advantaged.\textsuperscript{19}

A second estimation issue is omitted variable bias; any child and household variables in equations (2) and (3) not in the data become part of the error term, likely making it correlated with the observed variables, which can bias estimates of all the coefficients, including $\theta$s. More specifically, estimates of $\theta$s could include not only school fixed effects that affect only advantaged children but also general “non-school” advantages those children have after controlling for other $X$ variables. Fortunately, the Young Lives data have detailed information dating back to the students’ first year of life; those data capture many, if not most, of the “non-school” advantages of being an advantaged child. In particular, $X$ has a comprehensive array of child and household variables, including conditions in the first year of life and the following four years before starting school, and – most importantly – children’s cognitive skills before starting primary school, which should account for the effect of omitted education inputs (EI), and innate ability (IA), on $C$.\textsuperscript{20}

\textsuperscript{19} For Peru, 60.8% are “more able”, 61.8% have better educated mothers, and 68.7% are not stunted. The corresponding figures for Vietnam are 65.6%, 81.5% and 78.1%.

\textsuperscript{20} The estimated model is thus similar to the combined cumulative inputs value-added model favored by Todd and Wolpin (2007). A pure cumulative inputs model includes only the history of inputs and so assumes that observed inputs fully account for all inputs. In contrast, a value-added model includes only current inputs and last period’s test score; it assumes that the lagged score fully captures endowments and the effects of the complete history of inputs. In the mixed cumulative value-added model that we estimate identification relies on a weaker assumption than either of these, that combined lagged test scores and observed current and lagged inputs fully controls for all past inputs.
If students sort into schools based on unobserved traits, such as studiousness or parental tastes for education, that directly affect learning and are correlated with advantage indicators, such unobserved differences can be mistakenly interpreted as favoring of advantaged students even if schools show no favoritism. There is little school choice in rural Vietnam, where most of the Vietnam sample children live; children are assigned to schools based on their residence, and rural areas typically have only one primary school. Yet school choice is an option in Peru, where most of the sample children live in urban areas and can enroll in other public schools; also, and unlike Vietnam, many private schools are available. We cannot rule out selection bias completely, but we expect our extensive set of controls, including cognitive skills at age 5, will capture most unobserved factors correlated with school choice and test scores.

Omitted variable bias concerns are also reduced because this paper focuses on differences between groups rather than level effects. Omitted variables will bias estimates of $\theta$, only if they vary systematically across advantaged and disadvantaged children. Thus our analysis is more robust than it would if we estimated, say, the level effects of school inputs. We explore in Section VIII the sensitivity of our estimates to omitted variables by comparing them to estimates based on stronger assumptions (cf. footnote 20).

A third issue is that $\theta_i$ can be identified only for advantaged children who attend schools with at least one disadvantaged child from the school survey sample. This may lead to biased estimates of the school fixed effects due to unobservable selection into schools with both types of children. This possibility is also examined in Section VIII.

A fourth estimation concern is that while school fixed effects are allowed to vary across advantaged and disadvantaged children, coefficients on child and household variables are not. In fact, this appears to be plausible; tests of whether those variables’ impacts varied by advantaged status rejected the null of no difference in only 4 of 22 cases (see Appendix Table 5).
Finally, some of the explanatory variables could be measured with error, such as the \( C_p \), \( N \) and PT variables. In principle, we need instrumental variables for these variables, but finding valid instruments is a challenge. For this reason, and because we focus on the impact of school, not child or household, characteristics on student learning, IV methods are not used to estimate (2) and (3).

VII. Results

This section examines evidence on whether within-school differences in learning exist between advantaged and disadvantaged children. It begins by briefly discussing estimates of the effect of the non-school determinants of math and reading skills, and then turns to estimates of the average value of \( \theta_s \), first for Vietnam and then for Peru, for each definition of advantage.

A. Non-school Determinants of Learning. Table 3 presents estimates of equation (2). It focuses on the estimates of non-school determinants of school attainment (X variables), separately for math and reading. The estimates use one school fixed effects for all pupils in each school.21

The first two columns of Table 3 the results for Vietnam. The 1129 students in 90 schools include all younger cohort children in Grade 5 in the 2011-12 school year. The coefficients on both cognitive skill (at age 5) variables are positive and highly significant. Several other variables have significant effects. First, other things equal, boys scored lower on both tests; the negative impact of being male on math skills is modest, -0.14 standard deviations, but the negative effect on reading, -0.32 standard deviations, is fairly large. Larger negative effects are seen for ethnic minority children; their math and reading scores are 0.38 and 0.47 standard deviations lower, respectively. Last, children of more educated parents had significantly higher scores on both tests.

Two of the five variables measured around age one have significant impacts.22 First, enrollment in childcare at age one significantly lowers math and Vietnamese scores (10%...
significance level). This may reflect reduced parental contact in early childhood. Second, there is a marginally significant negative impact of maternal stress or depression.

As expected (see Glewwe and Miguel, 2008, and Alderman and Bleakley, 2013), height-for-age at age 5 has positive impacts on both math and reading, though significantly so only for the latter. Household size at age 5 has a negative impact; this suggests that additional siblings reduces resources and parental attention for the child, but it is significant only for reading skills.

Turning to Peru (columns 3 and 4), the PPVT and CDA scores again have generally significantly positive effects on math and reading scores. Boys have higher math skills and lower the reading skills, yet unlike Vietnam neither effect is significant. Ethnic minority children have slightly lower math and reading scores, but not significantly so. As in Vietnam, parental education matters even after controlling for pre-school skills, but only mother’s education is significant.

Peruvian children enrolled in childcare at age one scored lower on both tests, and the large impact on Spanish skills (-0.25 standard deviations) is significant. Unlike Vietnam, the wealth index has a significantly positive but diminishing impact on Spanish, but not math, scores. As in Vietnam, nutritional status at age 5 has a significantly positive impact. Finally, land ownership has a significantly positive effect, but only on math skills.

**B. Differences in School Impacts by Advantaged Status.** The $\theta_s$ parameters in equation (3) indicate whether the impacts of school and teacher characteristics on learning are, on average, larger for advantaged children. The first two columns of Table 4 present estimates of the average $\theta_s$, the difference between the average school fixed effects of advantaged and disadvantaged children, for Vietnam for five definitions of advantage. The estimated means of $\theta_s$ reveal little

---

22 Some of these variables may be endogenous, which could bias the estimates of all variables, including school fixed effects. To check this we re-estimated the main tables after dropping the two most likely to be endogenous: someone outside the household cared for the child, and time in preschool. The results, available upon request, are very similar.

23 The average $\theta_s$ is weighted by the number of younger cohort children in each school. The standard errors for these averages are calculated using the standard formula for the variance of a sum of (possibly correlated) variables.
evidence that the average school impact in Vietnam varies between advantaged and disadvantaged children. For all but one definition, the estimated averages of $\theta_s$ are statistically insignificant. The exception is when advantage is defined as being male: Vietnamese schools convey a sizeable 

 disadvantage to being male. That is, the average impact of schooling on boys’ math scores is 12.7 points (0.13 standard deviations) lower than for girls, a difference significant at the 5% level, and 29.9 points (0.30 standard deviations) lower on the Vietnamese test, a difference significant at the 1% level even after controlling for age 5 test scores. This is surprising because, at age 5 boys outperform girls on both cognitive tests (Table 1).  

 Figure 1.A shows the distribution (over schools) of $\theta_s$ (on the left) as well as the (separate) distributions of boys’ and girls’ school fixed effects. The difference in the values of the fixed effects $\theta$ for math scores occurs mostly in the middle and upper range of those distributions, while for Vietnamese it is at all points in the distributions.

 Overall, we find no evidence that schools in Vietnam favor advantaged children; the means of $\theta_s$ are close to zero. Indeed, the only significant effect is that Vietnamese schools seem to favor girls, who are often considered disadvantaged, especially in math (e.g. Fryer and Levitt, 2010).

 In contrast, the results for Peru, in columns (3) and (4), show differences in school fixed effects between advantaged and disadvantaged children when advantaged is defined in terms of pre-school skills or in terms of ethnicity. When advantage is defined in terms of children’s pre-school skills (CDA scores at age 5), schools in Peru seem to favor more prepared children in their acquisition of Spanish skills, an impact of 38.3 (0.38 standard deviations).  

 The estimated average $\theta_s$ is also large (19.8), but not quite significant, for math. Figure 1.B shows that the difference in 

Note also that the joint significance of the schools fixed effects, separately for advantaged and disadvantaged children, is very high for all definitions of advantage, for both Vietnam and Peru (not shown in Table 4); this indicates that schools vary substantially in their contribution to children’s acquisition of cognitive skills.

24 We cannot directly control for any non-school advantages of boys, yet if such an advantage is not captured by age 5 test scores and other controls the estimated difference in boys’ and girls’ school effects would be biased toward zero.

25 This does not reflect that such children had better Spanish skills at age 5, as those skills are in the control variables.
estimated school fixed effects for math scores occurs mostly in the lower and middle of the
distributions of those fixed effects, while for Spanish the distributions differ at almost all points.

Unlike in Vietnam, schools in Peru do not favor girls over boys, or vice versa, in either
math or reading. There is also no evidence of non-school differences between boys and girls at
age 5 (see Table 1). Yet schools in Peru appear to contribute more to learning of ethnic majority
children. Even after controlling for age 5 skills, nutritional status and parental education (and all
other variables in Table 3), the impact of school and teacher characteristics on ethnic majority
children’s math skills is 24.3 points (0.24 standard deviations) higher than their impact on ethnic
minority children; the analogous differential impact for Spanish skills is 20.4 (0.20 standard
deviations). Both differences are statistically significant.

Figure 1.C shows that the difference in the values of $\theta$ for math is manifested in a long
lower tail for ethnic minority children, but not for ethnic majority children; for Spanish there is a
similar, though less clear, pattern. Ethnic majority children already had higher cognitive skills
than minority children at age 5 (before starting school), as seen in Table 1. One could argue that
unobserved child or household factors explain this difference – i.e. that these estimated differences
in school effects are biased upward – yet it is difficult to imagine mechanisms for such bias.

Finally, Table 4 also shows, as in Vietnam, no significant differences in school impacts
between children in Peru with less educated mothers and those with more educated mothers, as
well as no significant difference by children’s nutritional status in early childhood.

Overall, even after controlling for a large number of child and household variables, when
advantage is defined as having higher skills at age 5 and not being an ethnic minority, advantaged
children in Peru learn more than disadvantaged children who attend the same schools. This is in
stark contrast with Vietnam, where advantaged children never appear to learn more than dis-
advantaged children in the same school, and for one advantage type (male) appear to learn less.
VIII. Robustness Checks, and an Extension

This section checks the robustness of our results, and presents some evidence on the role of parental investments in explaining the gap between advantaged and disadvantaged children.

A. Omitted Variable Bias. As discussed in Section V, despite using an extensive set of controls for non-school inputs, including test scores at age 5, omitted variable bias remains a concern. To explore this, we assess the sensitivity of the main estimates to different assumptions about omitted variables. We compare three sets of estimates: (1) main model, i.e. combining current and lagged inputs with lagged skill scores; (2) using only lagged skill scores to capture all lagged inputs and innate ability; and (3) controlling only for current and lagged observed inputs. If the main results are driven by omitted variable bias, they are likely to be sensitive to different assumptions about omitted variables and so should vary noticeably across these three models.

Table 5 shows the three sets of estimates for Vietnam and Peru for the types of disadvantage with significant differences: gender for Vietnam and pre-school cognitive skills and ethnicity in Peru. Estimates of $\theta_s$ for Vietnam when advantage is defined as being male are very stable across the three specifications for both math and Vietnamese. Turning to Peru, when advantage is defined as skills at age 5, estimates of $\theta_s$ across the three models are very stable for Spanish and reasonably stable for math. When advantage is defined as being from the ethnic majority group, estimates of $\theta_s$ are very stable for math skills, though there is more fluctuation for Spanish scores (estimates of $\theta_s$ range between 15.4 and 31.1). Overall, for all advantage categories where we find significant differences in school effects, the consistency of the direction, and in most cases size and statistical significance, of estimates of $\theta_s$ yield little evidence of omitted variable bias.

B. Sample Restrictions. Another plausible source of bias is selection of the analysis sub-samples. First, as explained in Section V, the school survey sub-sample for Vietnam includes
only those in grade 5 (the appropriate grade for their age) in the 2011-12 school year. In contrast, the Peru school survey sample is a sub-sample of the Young Lives children in grades 3, 4 and 5. Estimates using the Vietnam school survey sub-sample may be biased if the children excluded from that sub-sample are more likely to be disadvantaged (e.g. grade repeaters). Indeed, the data support this; unlike, Peru, the children included in the Vietnam school survey sub-sample are on average better off than those in the full Vietnam Young Lives sample (see Appendix Table 1).

To check whether our main findings are driven by this selection, we re-estimated them using the full sample of children and test scores from the last available round of the main Young Lives survey (Round 3). Table 6 reveals no systematic differences between the two sets of estimates. For math, the signs are the same for three of the five types of advantage, and in the two cases with different signs neither estimate is statistically significant. In the two cases where one estimate is significant and the other is not, the signs are the same, and in one case the significant result is for the school survey subsample and in the other case it is for the household sample. For reading, the signs are again the same for three of the five definitions of advantage, and in the two cases with different signs both estimates are insignificant. For the sole case when both samples find a significant differential, they agree: boys appear to be at a disadvantage relative to girls. Thus we conclude that excluding grade 3 and 4 students in the Vietnam school survey does not lead to systematic biases.

A second sample restriction is inclusion of only those schools with both advantaged and disadvantaged children in the school survey sub-sample. Selection into these schools may differ

---

26 The reason why the main analysis uses the sub-sample of children who were in the school survey is that the school survey tests were carefully designed to reflect the school curriculum and thus are the most appropriate outcome measures for exploring the impact of schooling on student learning. In contrast, the Round 3 tests were designed to reflect more general mathematics and verbal skills and so are not linked to specific school curricula.

27 Note that the Round 3 tests have a standard deviation of only 15, compared to a standard deviation of 100 for the school survey tests, so one would expect coefficients to be 6-7 times smaller for the Round 3 tests.
across these two types of children.\textsuperscript{28} For example, among children from richer households, those in the same schools as children from poorer families, rather than elite schools, may be the less able or less motivated among the richer children, while the reverse may hold for students in those schools from poorer households. While possible, the insensitivity of our estimates to different assumptions about omitted variable bias suggests that this is unlikely to drive the main results.

Third, the results may be sensitive to differences in the sample for different definitions of advantage, which occur because the estimates require at least one advantaged and one disadvantaged child in each school, so the analysis samples vary slightly for different types of advantage.\textsuperscript{29} To address this for Vietnam, we constructed a “master sample” of 38 schools with 822 children that are in the sample for all definitions of advantage; estimates of differences in school effects between more and less advantaged children for this sample are very similar to the main estimates, though less precise. For Peru, the smaller school sample rules out a general “master sample” (it would have only 96 children in 6 schools), so we focused on a somewhat larger sample for which each school had (at least) one child with relatively low skills at age five, one with relatively high age 5 skills, one with a relatively uneducated mother, and one with an educated mother. Again, we found similar, though very imprecise, results. Details for both countries are available on request.

Finally, sample restrictions may limit our findings’ external validity. If schools with both advantaged and disadvantaged children are atypical, our results may hold only for a small subset of schools in each country. To check this, Tables 7 and 8 show, by country, several school characteristics for the whole sample and the “overlap sub-samples” (schools with both types of children), with asterisks indicating significant differences between whole sample and sub-sample means. For Vietnam, the results show almost no significant differences for all definitions of

\textsuperscript{28} Recall that all unobservables affecting both groups are differenced out as $\theta$, is the difference in school fixed effects.

\textsuperscript{29} For example, the sub-sample of children in schools that have both more and less wealthy children from the school survey sample is different from the sub-sample in schools that have both ethnic majority and minority children.
advantage. Yet some differences emerge for Peru; this is unsurprising since, on average, the Peru sample has fewer children per school than the Vietnam sample, so fewer Peru schools have both advantaged and disadvantaged children. Indeed, schools in the overlap samples are larger than in the overall sample. The sub-sample most different from the main sample is the schools with both wealthier and poorer children. They are larger, with higher assets and fewer ethnic minority pupils and grade-repeaters. To the extent that these are “above average” schools, the external validity of the results for this type of advantage may be lower. Overall, however, other than school size there is little systematic selection of schools into the overlap samples, especially for the types of advantage for which we find significant within-school differences in school effectiveness.

C. Advantage Category Cut-offs. We also checked whether our results hold when different cutoffs (50th and 70th percentiles) are used to define advantage by wealth and preschool skills. The main results are unchanged (results available on request). In Vietnam, irrespective of the cut-off the results suggest that schools are equally effective at teaching children from more and less advantaged backgrounds. In Peru, the three different cut-offs (lowest 50%, 60% and 70%) yield the persistent finding that schools are more effective for children with higher pre-school skills.

D. Role of Parental Investments. This paper focuses on the impacts of teachers and schools on learning, and whether they differ for advantaged and disadvantaged children in the same school. Yet the data can also shed light on parental behavior. One approach is to remove from X all parental choice variables – child size at birth, child health (mother’s assessment), use of childcare, maternal stress/depression, household wealth, height for age, skills at age 5, time in preschool, only child status, household size, per capita consumption and land owned – keeping only child age, sex and ethnicity, and parental education. Estimates of \( \theta_s \) from this specification will be biased if parental behavior is correlated with a particular type of advantage. For example, parents of disadvantaged children may be discouraged and so not enroll their children in preschool,
which would lead to over-estimation of $\theta_s$; alternatively, such parents may do more to help their children, leading to underestimates of $\theta_s$. Of course, such bias could vary by the type of advantage.

Table 9 presents estimates of Table 4 after removing parental choices from $X$. For most types of advantage, the average estimated within-school differential in learning between advantaged and disadvantaged students barely changes. Yet there are two exceptions. First, for both countries the average difference in school fixed effects ($\theta_s$) increases sharply for advantage defined by nutrition in infancy. This implies that parents of well-nourished children take other actions to raise their children’s learning, so not controlling for parental behavior will overestimate the difference in school fixed effects between advantaged and disadvantaged students.

Second, defining advantage in terms of skills at age 5, the insignificant difference in fixed effects between advantaged and disadvantaged children in Vietnam becomes large and significant. This suggests that Vietnamese parents make additional educational investments in their “high ability” children, and not controlling for parental behavior may lead one to find “favoring” of advantaged students by Vietnamese teachers and/or schools.

**IX. Conclusions**

Students in many developing countries learn far less than their counterparts in developed countries, and within developing countries some children learn much more than others. This paper considers whether schools in Vietnam and Peru reinforce or reduce gaps in learning between advantaged and disadvantaged students for six different types of advantage by estimating separate school fixed effects for advantaged and disadvantaged students. This focuses on *intra-* rather than *inter-*school differences in schools’ contribution to learning among different types of students.

We estimate school fixed effects for advantaged and disadvantaged children, controlling for a variety of child and household characteristics, including cognitive skill tests at age 5 (before
starting primary school). We find no evidence that schools in Vietnam favor advantaged children. Indeed, the one significant effect is that girls, who some consider disadvantaged, pull ahead of boys between the ages 5 and 10. In contrast, for two definitions of disadvantage schools in Peru appear to favor advantaged students: those with higher skills at age five acquire more Spanish skills than do less prepared students, and ethnic majority students learn more math and Spanish than minority students, even after controlling for skills at age 5.

These findings conform with existing evidence on differences between the school systems of Vietnam and Peru. Vietnam’s expansion in primary education in the last two decades included effective investment in education quality and a focus on equity, emphasizing that all pupils attain “minimum standards”. In contrast, Peru’s schools suffer from low average quality and high inequality in student learning, with evidence of gaps in access and learning outcomes by, for example, income and ethnicity. Our results show that such gaps exist even among students attending the same school; more advantaged pupils learn more than less advantaged pupils, even after controlling for many pupil characteristics, including skills at age 5.

This paper does not explore the mechanisms behind these effects, yet the recent literature on within-school discrimination against disadvantaged groups and biases in school curriculum in favor of advantaged children in several contexts are potential explanations. Classroom observations to see whether teachers discriminate against disadvantaged children should be a high priority. Yet one conclusion is clear: estimates that assume that schools have the same impacts on different types of students may overlook a major source of inequality in student learning in developing countries.
References


Table 1: Difference in Pre-school Test Scores between More and Less Advantaged Children (advantaged-disadvantaged)

| Advantage category: | Vietnam | | Peru | |
|---------------------|---------|---------|---------|
|                     | (1) PPVT | (2) CDA | (3) PPVT | (4) CDA |
| Being richer (top 2 wealth quintiles) | 21.89*** | 13.29*** | 24.80*** | 30.94*** |
| Being more able (top 2 CDA quintiles) | 30.06*** | 77.35*** | 11.23*** | 76.39*** |
| Being male | 5.99** | 4.13* | -2.29 | -1.92 |
| Being an ethnic majority | NA | NA | 21.17*** | 15.19*** |
| Having a mother with at least completed primary schooling | 16.68*** | 6.33* | 22.81*** | 18.17*** |
| Having not been malnourished in infancy | 10.74*** | 2.82 | 13.91*** | -0.53 |

Notes: *** p<0.01, ** p<0.05, * p<0.1;
The Peabody Picture Vocabulary Test (PPVT) is a measure of receptive vocabulary.
The Cognitive Development Assessment (CDA) measures children’s perceptions of quantity - testing understanding of concepts such as few, most, half, many, equal, and pair.
Difference in CDA scores between more and less able children is there by construction since being more able is defined as having a CDA score in the top two quintiles of the distribution at age 5.

Advantage categories:
Being richer: Wealth quintiles are constructed using wealth index is a simple average of three individual indices ranging from 0 to 1, including housing quality (rooms per person, floor and roof quality), consumer durables (mean of ownership of radio, bicycle, TV, car/truck, motorbike/scooter, landline telephone, bed/table) and access to services (electricity; piped water into dwelling/yard; access to own pit latrine/flush toilet; cooking fuel is electricity, gas or kerosene).
Being an ethnic majority: Peru - mother’s first language is Spanish (i.e. not Quechua, Aymara, other).
Not being malnourished in infancy: defined using anthropometric indicator of adequate long-term growth: height-for-age z-scores > -2).
Table 2: Sample Descriptive Statistics

<table>
<thead>
<tr>
<th>From School Survey</th>
<th>Vietnam</th>
<th>Peru</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics test score (IRT transformed; whole sample mean = 500, sd=100)</td>
<td>499.19</td>
<td>508.60</td>
</tr>
<tr>
<td>SD</td>
<td>96.48</td>
<td>102.65</td>
</tr>
<tr>
<td>Language test score (IRT transformed; whole sample mean = 500, sd=100)</td>
<td>497.44</td>
<td>505.77</td>
</tr>
<tr>
<td>SD</td>
<td>96.69</td>
<td>99.69</td>
</tr>
<tr>
<td>Age in months at the time of the test</td>
<td>123.24</td>
<td>122.16</td>
</tr>
<tr>
<td>SD</td>
<td>2.71</td>
<td>3.87</td>
</tr>
<tr>
<td>From Household Survey (time invariant variables)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (dummy)</td>
<td>0.51</td>
<td>0.48</td>
</tr>
<tr>
<td>SD</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>Ethnic minority (dummy)</td>
<td>0.08</td>
<td>0.36</td>
</tr>
<tr>
<td>SD</td>
<td>0.27</td>
<td>0.48</td>
</tr>
<tr>
<td>Father: Years of schooling</td>
<td>7.69</td>
<td>9.23</td>
</tr>
<tr>
<td>SD</td>
<td>3.53</td>
<td>3.77</td>
</tr>
<tr>
<td>Mother: Years of schooling</td>
<td>7.17</td>
<td>7.84</td>
</tr>
<tr>
<td>SD</td>
<td>3.30</td>
<td>3.11</td>
</tr>
<tr>
<td>Measured during Infancy (round 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth size (maternal assessment: 1=very large to 5=very small)</td>
<td>3.08</td>
<td>3.11</td>
</tr>
<tr>
<td>SD</td>
<td>0.66</td>
<td>0.96</td>
</tr>
<tr>
<td>Health better than other children (dummy, maternal assessment)</td>
<td>0.26</td>
<td>0.41</td>
</tr>
<tr>
<td>SD</td>
<td>0.44</td>
<td>0.49</td>
</tr>
<tr>
<td>Childcare: cared for by people outside the household /in crèche (dummy)</td>
<td>0.42</td>
<td>0.21</td>
</tr>
<tr>
<td>SD</td>
<td>0.49</td>
<td>0.41</td>
</tr>
<tr>
<td>Maternal mental health</td>
<td>4.4</td>
<td>5.78</td>
</tr>
<tr>
<td>SD</td>
<td>4.02</td>
<td>4.26</td>
</tr>
<tr>
<td>Wealth index</td>
<td>0.47</td>
<td>0.43</td>
</tr>
<tr>
<td>SD</td>
<td>0.20</td>
<td>0.19</td>
</tr>
<tr>
<td>Measured at Age 5 (round 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height for age z-score</td>
<td>-1.30</td>
<td>-1.53</td>
</tr>
<tr>
<td>SD</td>
<td>0.97</td>
<td>1.02</td>
</tr>
<tr>
<td>PPVT Score (IRT transformed; whole sample mean=300, sd=50)</td>
<td>305.33</td>
<td>301.34</td>
</tr>
<tr>
<td>SD</td>
<td>44.33</td>
<td>44.84</td>
</tr>
<tr>
<td>CDA Score (IRT transformed; whole sample mean=300, sd=50)</td>
<td>306.80</td>
<td>299.35</td>
</tr>
<tr>
<td>SD</td>
<td>46.56</td>
<td>46.39</td>
</tr>
<tr>
<td>Time spent in pre-school (hours per day in a “normal” week)</td>
<td>5.79</td>
<td>3.44</td>
</tr>
<tr>
<td>SD</td>
<td>2.45</td>
<td>1.83</td>
</tr>
<tr>
<td>Only child (dummy)</td>
<td>0.22</td>
<td>0.19</td>
</tr>
<tr>
<td>SD</td>
<td>0.42</td>
<td>0.40</td>
</tr>
<tr>
<td>Household size</td>
<td>4.57</td>
<td>5.52</td>
</tr>
<tr>
<td>SD</td>
<td>1.39</td>
<td>2.23</td>
</tr>
<tr>
<td>Log per capita real consumption (in local currency)</td>
<td>5.77</td>
<td>5.02</td>
</tr>
<tr>
<td>SD</td>
<td>0.54</td>
<td>0.67</td>
</tr>
<tr>
<td>Area of land owned (in hectares)</td>
<td>0.42</td>
<td>1.09</td>
</tr>
<tr>
<td>SD</td>
<td>1.07</td>
<td>3.34</td>
</tr>
</tbody>
</table>

Number of Observations 1,129 547

Notes: Ethnicity minority status: Peru - mother’s first language is not Spanish (i.e. is Quechua, Aymara, other); Vietnam – child reported by mother as not belonging to the Kinh group (i.e. are H’mong, Dao, Tay, other). Maternal mental health is measured using the Self-Reporting Questionnaire (SRQ20), a screening tool for risk of common mental disorders (rCMD) developed by the World Health Organization for developing countries (Beusenberg et al, 1994). It is a scale consisting of 20 statements with yes/no answer options. Score reflects the number of “yes” responses out of possible 20; a higher score reflects higher risk of rCMD. Wealth index is a simple average of three individual indices ranging from 0 to 1, including housing quality (rooms per person, floor and roof quality), consumer durables (mean of ownership of radio, bicycle, TV, car/truck, motorbike/scooter, landline telephone, bed/table) and access to services (electricity; piped water into dwelling/yard; access to own pit latrine/flush toilet; cooking fuel is electricity, gas or kerosene). Peabody Picture Vocabulary Test (PPVT) measures receptive vocabulary. The Cognitive Development Assessment (CDA) measures children’s perceptions of quantity - testing understanding of concepts such as few, most, half, many, equal, and pair.
**Table 3: Non-school Determinants of Cognitive Skills at School in Vietnam and Peru**

<table>
<thead>
<tr>
<th></th>
<th>Vietnam Math score</th>
<th>Vietnam Vietnamese score</th>
<th>Peru Math score</th>
<th>Peru Spanish score</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPVT score (IRT transformed; whole sample mean=300, sd=50)</td>
<td>0.292***</td>
<td>0.401***</td>
<td>0.184</td>
<td>0.220***</td>
</tr>
<tr>
<td></td>
<td>(0.077)</td>
<td>(0.089)</td>
<td>(0.116)</td>
<td>(0.082)</td>
</tr>
<tr>
<td>CDA score (IRT transformed; whole sample mean=300, sd=50)</td>
<td>0.196***</td>
<td>0.141**</td>
<td>0.208*</td>
<td>0.348***</td>
</tr>
<tr>
<td></td>
<td>(0.070)</td>
<td>(0.063)</td>
<td>(0.108)</td>
<td>(0.108)</td>
</tr>
<tr>
<td>Age at the time of the test (months after 9yrs old)</td>
<td>15.386</td>
<td>7.453</td>
<td>5.495</td>
<td>2.173</td>
</tr>
<tr>
<td></td>
<td>(9.291)</td>
<td>(10.100)</td>
<td>(6.667)</td>
<td>(7.494)</td>
</tr>
<tr>
<td>Age squared at the time of the test (months after 9yrs old)</td>
<td>-0.491</td>
<td>-0.299</td>
<td>-0.118</td>
<td>-0.090</td>
</tr>
<tr>
<td></td>
<td>(0.298)</td>
<td>(0.315)</td>
<td>(0.231)</td>
<td>(0.265)</td>
</tr>
<tr>
<td>In grade 3 at the time of the test</td>
<td>--</td>
<td>--</td>
<td>-66.474***</td>
<td>-50.490***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(20.071)</td>
<td>(20.914)</td>
</tr>
<tr>
<td>In grade 4 at the time of the test</td>
<td>--</td>
<td>--</td>
<td>-16.486</td>
<td>-12.951</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(12.568)</td>
<td>(12.870)</td>
</tr>
<tr>
<td>Ethnic minority (dummy)</td>
<td>-38.259***</td>
<td>-47.172***</td>
<td>-7.864</td>
<td>-12.297</td>
</tr>
<tr>
<td>Male (dummy)</td>
<td>-13.620**</td>
<td>-31.796***</td>
<td>4.930</td>
<td>-7.546</td>
</tr>
<tr>
<td></td>
<td>(5.994)</td>
<td>(4.434)</td>
<td>(8.167)</td>
<td>(8.271)</td>
</tr>
<tr>
<td>Father's schooling: 4-6 years</td>
<td>1.670</td>
<td>11.129</td>
<td>-3.306</td>
<td>2.899</td>
</tr>
<tr>
<td></td>
<td>(10.350)</td>
<td>(9.047)</td>
<td>(18.930)</td>
<td>(17.240)</td>
</tr>
<tr>
<td>Father's schooling: 7-9 years</td>
<td>6.978</td>
<td>17.339</td>
<td>14.177</td>
<td>-15.070</td>
</tr>
<tr>
<td>Father’s schooling: 10+ years</td>
<td>22.141**</td>
<td>28.035**</td>
<td>19.553</td>
<td>14.054</td>
</tr>
<tr>
<td></td>
<td>(10.322)</td>
<td>(12.147)</td>
<td>(21.344)</td>
<td>(17.694)</td>
</tr>
<tr>
<td>Mother's schooling: 4-6 years</td>
<td>12.039</td>
<td>16.818*</td>
<td>7.227</td>
<td>8.568</td>
</tr>
<tr>
<td></td>
<td>(7.869)</td>
<td>(9.376)</td>
<td>(12.964)</td>
<td>(11.387)</td>
</tr>
<tr>
<td>Mother's schooling: 7-9 years</td>
<td>13.426*</td>
<td>20.135*</td>
<td>14.272</td>
<td>25.028**</td>
</tr>
<tr>
<td></td>
<td>(7.695)</td>
<td>(10.641)</td>
<td>(13.082)</td>
<td>(10.196)</td>
</tr>
<tr>
<td>Mother’s schooling: 10+ years</td>
<td>28.050**</td>
<td>35.351**</td>
<td>35.150***</td>
<td>40.246***</td>
</tr>
<tr>
<td></td>
<td>(12.466)</td>
<td>(15.004)</td>
<td>(11.524)</td>
<td>(10.592)</td>
</tr>
<tr>
<td>Birth size (maternal assessment: 1=very large to 5=very small)</td>
<td>-1.608</td>
<td>1.639</td>
<td>0.776</td>
<td>0.949**</td>
</tr>
<tr>
<td></td>
<td>(3.210)</td>
<td>(3.855)</td>
<td>(0.677)</td>
<td>(0.394)</td>
</tr>
<tr>
<td>Health better than other children (dummy, maternal assessment)</td>
<td>-1.699</td>
<td>-4.625</td>
<td>4.426</td>
<td>6.900</td>
</tr>
<tr>
<td></td>
<td>(8.420)</td>
<td>(6.687)</td>
<td>(6.753)</td>
<td>(8.849)</td>
</tr>
<tr>
<td>Childcare: looked after by others / in crèche (dummy)</td>
<td>-9.501*</td>
<td>-10.996*</td>
<td>-11.386</td>
<td>-24.611***</td>
</tr>
<tr>
<td></td>
<td>(5.681)</td>
<td>(5.576)</td>
<td>(7.768)</td>
<td>(8.741)</td>
</tr>
<tr>
<td>Variable</td>
<td>Coefficient 1</td>
<td>Coefficient 2</td>
<td>Coefficient 3</td>
<td>Coefficient 4</td>
</tr>
<tr>
<td>----------------------------------------------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Maternal mental health</td>
<td>-1.276*</td>
<td>-1.065</td>
<td>-0.057*</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>(0.666)</td>
<td>(0.693)</td>
<td>(0.034)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Wealth index (at age 1)</td>
<td>-7.4291</td>
<td>-12.927</td>
<td>59.438</td>
<td>332.361***</td>
</tr>
<tr>
<td></td>
<td>(46.253)</td>
<td>(72.223)</td>
<td>(136.957)</td>
<td>(106.218)</td>
</tr>
<tr>
<td>Wealth index squared (infancy)</td>
<td>78.239</td>
<td>2.534</td>
<td>-40.820</td>
<td>-322.863***</td>
</tr>
<tr>
<td></td>
<td>(51.615)</td>
<td>(88.090)</td>
<td>(136.482)</td>
<td>(116.751)</td>
</tr>
<tr>
<td>Height for age z-score</td>
<td>1.725</td>
<td>7.007**</td>
<td>6.734**</td>
<td>9.039**</td>
</tr>
<tr>
<td></td>
<td>(3.480)</td>
<td>(3.117)</td>
<td>(3.386)</td>
<td>(3.703)</td>
</tr>
<tr>
<td>Time spent in pre-school (hours per day in a “normal” week)</td>
<td>-0.937</td>
<td>-0.807</td>
<td>0.565</td>
<td>-3.213</td>
</tr>
<tr>
<td></td>
<td>(2.124)</td>
<td>(2.019)</td>
<td>(2.467)</td>
<td>(2.015)</td>
</tr>
<tr>
<td>Only child (dummy)</td>
<td>-7.804</td>
<td>2.601</td>
<td>2.619</td>
<td>2.104</td>
</tr>
<tr>
<td></td>
<td>(6.765)</td>
<td>(4.683)</td>
<td>(8.021)</td>
<td>(9.079)</td>
</tr>
<tr>
<td>Household size</td>
<td>-1.119</td>
<td>4.472***</td>
<td>-0.484</td>
<td>1.542</td>
</tr>
<tr>
<td></td>
<td>(1.869)</td>
<td>(1.652)</td>
<td>(1.982)</td>
<td>(1.937)</td>
</tr>
<tr>
<td>Log per capital real consumption (in local currency)</td>
<td>2.237</td>
<td>3.103</td>
<td>4.678</td>
<td>6.547</td>
</tr>
<tr>
<td></td>
<td>(6.725)</td>
<td>(6.754)</td>
<td>(8.311)</td>
<td>(6.816)</td>
</tr>
<tr>
<td>Area of land owned (hectares)</td>
<td>-1.729</td>
<td>0.126</td>
<td>1.998**</td>
<td>-0.807</td>
</tr>
<tr>
<td></td>
<td>(2.500)</td>
<td>(2.360)</td>
<td>(0.922)</td>
<td>(0.972)</td>
</tr>
<tr>
<td>Constant</td>
<td>253.384***</td>
<td>301.134***</td>
<td>294.146***</td>
<td>228.450***</td>
</tr>
<tr>
<td></td>
<td>(76.393)</td>
<td>(87.763)</td>
<td>(104.182)</td>
<td>(80.886)</td>
</tr>
</tbody>
</table>

**School Fixed Effects**: Yes

**Number of observations**: 1,129

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors shown in parentheses; they are not clustered at the site level because there are only 20 sentinel sites and so not enough degrees of freedom, and because school fixed effects will capture most of the correlation of the error terms within the sentinel sites.

Ethnicity minority status: Peru - mother’s first language is not Spanish (i.e. is Quechua, Aymara, other); Vietnam – child reported by mother as not belonging to the Kinh group (i.e. are H’mong, Dao, Tay, other). Grade at the time of the test: Vietnam - all children in grade 5 so no control; Peru – grade 5 is the omitted category. Father & Mother’s schooling: no schooling/incomplete primary is the omitted category. Maternal mental health is measured using the Self-Reporting Questionnaire (SRQ20), a screening tool for risk of common mental disorders (CMD) developed by the World Health Organization for developing countries (Beusenberg et al, 1994). It is a scale consisting of 20 statements with yes/no answer options. Score reflects the number of “yes” responses out of possible 20; a higher score reflects higher risk of CMD. Wealth index is a simple average of three individual indices ranging from 0 to 1, including housing quality (rooms per person, floor and roof quality), consumer durables (mean of ownership of park, bicycle, TV, car/truck, motorbike/scooter, landline telephone, bed/table) and access to services (electricity; piped water into dwelling/yard; access to own pit latrine/flush toilet; cooking fuel is electricity, gas or kerosene). Peabody Picture Vocabulary Test (PPVT) measures receptive vocabulary. The Cognitive Development Assessment (CDA) measures children’s perceptions of quantity - testing understanding of concepts such as few, most, half, many, equal, and pair.
<table>
<thead>
<tr>
<th>Advantage Category</th>
<th>Vietnam</th>
<th>Peru</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>Math score</td>
<td>Vietnamese score</td>
</tr>
<tr>
<td>Being richer (top 2 wealth quintiles)</td>
<td>-8.11 (7.42)</td>
<td>-3.55 (6.89)</td>
</tr>
<tr>
<td>Being more able (top 2 CDA quintiles)</td>
<td>9.12 (9.63)</td>
<td>11.10 (9.72)</td>
</tr>
<tr>
<td>Being male</td>
<td>-12.69** (5.09)</td>
<td>-29.88*** (4.88)</td>
</tr>
<tr>
<td>Being an ethnic majority</td>
<td></td>
<td>24.30** (9.67)</td>
</tr>
<tr>
<td>Having a mother with at least primary education</td>
<td>-8.98 (11.45)</td>
<td>-14.22 (10.96)</td>
</tr>
<tr>
<td>Having not been malnourished in infancy</td>
<td>4.81 (9.11)</td>
<td>-6.76 (7.72)</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1; Robust standard errors in parentheses.

$s_\bar{\theta}$ is the mean incremental effect of school on test scores of advantaged children compared to disadvantaged, as shown in Equation (3). It is the mean value of the coefficients for the interaction of school fixed effects with each of the advantage categories. All models control for full set of non-school child, parent and household characteristics shown in Table 3.

Outcome measures: Math and Verbal scores transformed using Item Response Theory (IRT); whole sample mean = 500, sd=100

Advantage categories:
- Being richer: Wealth quintiles are constructed using wealth index is a simple average of three individual indices ranging from 0 to 1, including housing quality (rooms per person, floor and roof quality), consumer durables (mean of ownership of radio, bicycle, TV, car/truck, motorbike/scooter, landline telephone, bed/table) and access to services (electricity; piped water into dwelling/yard; access to own pit latrine/flush toilet; cooking fuel is electricity, gas or kerosene).
- Being more able: The Cognitive Development Assessment (CDA) measures children’s perceptions of quantity - testing understanding of concepts such as few, most, half; many, equal, and pair.
- Being an ethnic majority: Peru - mother’s first language is Spanish (i.e. not Quechua, Aymara, other);
- Not being malnourished in infancy: defined using anthropometric indicator of adequate long-term growth: height-for-age z-scores >-2)
Table 5: Sensitivity of the Estimated Significant Differences in School Effectiveness by Advantage Group

<table>
<thead>
<tr>
<th></th>
<th>Vietnam</th>
<th>Peru</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Math score</td>
<td>Vietnamese score</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>**Mean incremental effect, ** $\bar{\theta}_s$ in eq (3) (se) **</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td><strong>Model 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Being male</td>
<td>-12.69** (5.09)</td>
<td>-29.88*** (4.88)</td>
</tr>
<tr>
<td>Being more able (top 2 CDA quintiles)</td>
<td>19.83 (14.80)</td>
<td>38.32** (17.12)</td>
</tr>
<tr>
<td>Being an ethnic majority</td>
<td>24.30** (9.67)</td>
<td>20.39* (11.76)</td>
</tr>
<tr>
<td><strong>Model 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Being male</td>
<td>-13.63*** (5.05)</td>
<td>-31.39*** (4.95)</td>
</tr>
<tr>
<td>Being more able (top 2 CDA quintiles)</td>
<td>29.10* (15.48)</td>
<td>41.37*** (15.90)</td>
</tr>
<tr>
<td>Being an ethnic majority</td>
<td>26.61*** (9.15)</td>
<td>31.07*** (8.74)</td>
</tr>
<tr>
<td><strong>Model 3</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Being male</td>
<td>-9.98** (5.09)</td>
<td>-26.61*** (4.95)</td>
</tr>
<tr>
<td>Being more able (top 2 CDA quintiles)</td>
<td>19.95** (8.40)</td>
<td>33.14*** (9.75)</td>
</tr>
<tr>
<td>Being an ethnic majority</td>
<td>22.03** (9.68)</td>
<td>15.39 (11.12)</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1; Robust standard errors in parentheses.

$\bar{\theta}_s$ is the mean incremental effect of school on test scores of advantaged children compared to disadvantaged, as seen in Equation (3).

Model 1: Main model (estimates reproduced for convenience) – controlling for all non-school child, parent and household characteristics shown in Table 3.

Model 2: Controlling for age 5 cognitive test scores only

Model 3: Controlling for age 1 and 5 child, parent and household characteristics, excluding age 5 cognitive test scores.

Outcome measures: Math and Verbal scores transformed using Item Response Theory (IRT); whole sample mean = 500, sd=100

Advantage categories:

Being more able: The Cognitive Development Assessment (CDA) measures children’s perceptions of quantity - testing understanding of concepts such as few, most, half, many, equal, and pair.

Being an ethnic majority: Peru - mother’s first language is Spanish (i.e. not Quechua, Aymara, other)
Table 6: Comparing Main Results to Estimates Using Household Survey Tests and Sample in Vietnam

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Math score (school survey)</td>
<td>Math score (household survey)</td>
<td>Vietnamese score (school survey)</td>
<td>Early Grade Reading Assessment score (household survey)</td>
</tr>
<tr>
<td>Mean incremental effect, $\bar{\theta}_s$ in eq (3) (se)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Being richer (top 2 wealth quintiles)</td>
<td>-8.11 (7.42)</td>
<td>-0.38 (0.77)</td>
<td>-3.55 (6.89)</td>
<td>-0.07 (0.84)</td>
</tr>
<tr>
<td>Being more able (top 2 CDA quintiles)</td>
<td>9.12 (9.63)</td>
<td>-0.80 (1.03)</td>
<td>11.10 (9.72)</td>
<td>-0.31 (1.11)</td>
</tr>
<tr>
<td>Being male</td>
<td>-12.69** (5.09)</td>
<td>-0.50 (0.58)</td>
<td>-29.88*** (4.88)</td>
<td>-2.30*** (0.61)</td>
</tr>
<tr>
<td>Having a mother with at least primary education</td>
<td>-8.98 (11.45)</td>
<td>0.86 (1.21)</td>
<td>-14.22 (10.96)</td>
<td>-2.56* (1.36)</td>
</tr>
<tr>
<td>Having not been malnourished in infancy</td>
<td>4.81 (9.11)</td>
<td>1.76** (0.82)</td>
<td>-6.76 (7.72)</td>
<td>-0.81 (0.89)</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1; Robust standard errors in parentheses

$\bar{\theta}_s$ is the mean incremental effect of school on test scores of advantaged children compared to disadvantaged, as shown in Equation (3). All models control for full set of non-school child, parent and household characteristics shown in Table 3.

Outcome measures: Col (1, 3) Math and Vietnamese scores on tests administered to subset of children in Grade 5 who were in the school survey (N=1,129); test scores transformed using Item Response Theory (IRT); whole sample mean = 500, sd=100. Col (2, 4) Math and Early Grade Reading Assessment scores on tests administered to all children in Round 3 of household survey (N=1,965; not the same tests as the school survey tests); raw test score standardized to have a mean of 300 and a standard deviation of 15.

Advantage categories:
Being richer: Wealth quintiles are constructed using wealth index is a simple average of three individual indices ranging from 0 to 1, including housing quality (rooms per person, floor and roof quality), consumer durables (mean of ownership of radio, bicycle, TV, car/truck, motorbike/scooter, landline telephone, bed/table) and access to services (electricity; piped water into dwelling/yard; access to own pit latrine/flush toilet; cooking fuel is electricity, gas or kerosene).

Being more able: The Cognitive Development Assessment (CDA) measures children’s perceptions of quantity - testing understanding of concepts such as few, most, half, many, equal, and pair.

Not being malnourished in infancy: defined using anthropometric indicator of adequate long-term growth: height-for-age z-scores >-2)
Table 7: Main Characteristics of Schools in Vietnam in the Whole Sample and Schools in the Analysis Overlap Sub-samples

<table>
<thead>
<tr>
<th></th>
<th>(1) All</th>
<th>(2) Wealth</th>
<th>(3) Ability</th>
<th>(4) Sex</th>
<th>(5) Maternal Education</th>
<th>(6) Malnutrition in infancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>School assets index</td>
<td>0.00 (1.64)</td>
<td>0.33 (1.36)</td>
<td>0.25 (1.59)</td>
<td>0.26 (1.57)</td>
<td>0.29 (1.46)</td>
<td>0.07 (1.45)</td>
</tr>
<tr>
<td>School size (number of students)</td>
<td>481.73 (347.04)</td>
<td>535.81 (387.28)</td>
<td>533.77 (387.38)</td>
<td>502.63 (367.53)</td>
<td>501.75 (375.76)</td>
<td>445.71 (196.72)</td>
</tr>
<tr>
<td>School offers free full-day schooling</td>
<td>0.24 (0.43)</td>
<td>0.17 (0.38)</td>
<td>0.13 (0.34)</td>
<td>0.25 (0.44)</td>
<td>0.23 (0.43)</td>
<td>0.26 (0.44)</td>
</tr>
<tr>
<td>School offers free lunch</td>
<td>0.88 (0.33)</td>
<td>0.91 (0.28)</td>
<td>0.83 (0.38)</td>
<td>0.87 (0.34)</td>
<td>0.88 (0.04)</td>
<td>0.90 (0.30)</td>
</tr>
<tr>
<td>Principal: Years of experience</td>
<td>10.78 (7.08)</td>
<td>10.91 (7.27)</td>
<td>10.1 (6.98)</td>
<td>10.76 (7.35)</td>
<td>10.15 (7.10)</td>
<td>10.48 (6.83)</td>
</tr>
<tr>
<td>Teacher: Years of experience</td>
<td>17.58 (6.37)</td>
<td>17.10 (6.26)</td>
<td>16.68 (5.39)</td>
<td>17.81 (6.33)</td>
<td>17.88 (6.45)</td>
<td>17.49 (6.72)</td>
</tr>
<tr>
<td>Teacher: Score on Vietnamese pedagogy test</td>
<td>70.27 (12.23)</td>
<td>69.97 (13.10)</td>
<td>71.17 (13.57)</td>
<td>70.68 (12.96)</td>
<td>70.13 (12.79)</td>
<td>71.57 (11.18)</td>
</tr>
<tr>
<td>Teacher: Score on Maths pedagogy test</td>
<td>68.84 (8.44)</td>
<td>68.30 (6.74)</td>
<td>70.61 (8.32)</td>
<td>68.93 (8.34)</td>
<td>67.90 (7.89)</td>
<td>68.38 (7.38)</td>
</tr>
<tr>
<td>Pupils: Mean number of days absent (per month)</td>
<td>1.94 (6.82)</td>
<td>1.57 (6.51)</td>
<td>2.52 (8.16)</td>
<td>2.18 (0.84)</td>
<td>2.46 (8.18)</td>
<td>1.94 (7.53)</td>
</tr>
<tr>
<td>Pupils: Proportion ethnic minority</td>
<td>0.26 (0.37)</td>
<td>0.09*** (0.19)</td>
<td>0.12** (0.24)</td>
<td>0.19 (0.31)</td>
<td>0.16* (0.28)</td>
<td>0.20 (0.32)</td>
</tr>
<tr>
<td>Pupils: Proportion of grade repeaters</td>
<td>0.05 (0.05)</td>
<td>0.04 (0.04)</td>
<td>0.04 (0.05)</td>
<td>0.05 (0.06)</td>
<td>0.05 (0.05)</td>
<td>0.05 (0.05)</td>
</tr>
<tr>
<td>Number of Schools</td>
<td>90</td>
<td>48</td>
<td>61</td>
<td>77</td>
<td>61</td>
<td>63</td>
</tr>
</tbody>
</table>

Notes: Col(1) shows mean for whole school survey sample; Overlap sub-samples include schools in the sample that have both advantaged and disadvantaged children when advantage is defined as: Col(2) top 2 wealth quintiles; Col(3) top 2 ability quintiles; Col(4) boys; Col(5) children of more educated mothers; Col(6) children who were not malnourished in infancy. See Notes in Table 4 for details on how each category is measured.

Means, sd’s in parentheses, asterisks indicating significance level of t-test of difference in means for all schools: *** p<0.01, ** p<0.05, * p<0.1 ; .

School asset index is the first factor from principal components analysis of school-level indicators for: need for repairs, separate room for each Grade 5 class, availability of library, computers, electricity and internet in school.
<table>
<thead>
<tr>
<th></th>
<th>(1) All</th>
<th>(2) Wealth</th>
<th>(3) Ability</th>
<th>(4) Sex</th>
<th>(5) Ethnicity</th>
<th>(6) Maternal Education</th>
<th>(7) Malnutrition in infancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>School assets index</td>
<td>0.00 (1.74)</td>
<td>1.33*** (1.47)</td>
<td>0.52 (1.83)</td>
<td>0.28 (1.88)</td>
<td>0.27 (1.76)</td>
<td>-0.27 (1.70)</td>
<td>-0.09 (1.89)</td>
</tr>
<tr>
<td>School size (number of students)</td>
<td>378.0 (398.6)</td>
<td>785.9*** (444.4)</td>
<td>553.7*** (449.2)</td>
<td>491.1 (456.9)</td>
<td>557.7*** (504.4)</td>
<td>417.3 (413.3)</td>
<td>410.3 (415.1)</td>
</tr>
<tr>
<td>School private</td>
<td>0.14 (0.35)</td>
<td>0.08 (0.28)</td>
<td>0.11 (0.31)</td>
<td>0.10 (0.30)</td>
<td>0.04 (0.20)</td>
<td>0*** (0)</td>
<td>0.06* (0.23)</td>
</tr>
<tr>
<td>School offers free lunch</td>
<td>0.66 (0.48)</td>
<td>0.58 (0.50)</td>
<td>0.62 (0.49)</td>
<td>0.67 (0.48)</td>
<td>0.80 (0.41)</td>
<td>0.71 (0.46)</td>
<td>0.72 (0.45)</td>
</tr>
<tr>
<td>Principal: Years of experience</td>
<td>12.14 (8.67)</td>
<td>13.6 (9.37)</td>
<td>11.66 (8.06)</td>
<td>11.76 (8.34)</td>
<td>9.92 (7.40)</td>
<td>11.57 (8.40)</td>
<td>11.81 (8.12)</td>
</tr>
<tr>
<td>Teacher: Years of experience</td>
<td>17.23 (7.69)</td>
<td>19.52 (6.40)</td>
<td>18.05 (7.22)</td>
<td>18.92 (7.05)</td>
<td>18.57 (6.32)</td>
<td>19.59* (6.34)</td>
<td>17.81 (7.38)</td>
</tr>
<tr>
<td>Teacher: Score on Maths pedagogy test</td>
<td>7.74 (2.02)</td>
<td>8.23 (1.44)</td>
<td>7.62 (2.15)</td>
<td>7.57 (2.08)</td>
<td>7.99 (1.42)</td>
<td>7.25 (2.35)</td>
<td>7.67 (2.38)</td>
</tr>
<tr>
<td>Pupils: Mean number of days absent (per week)</td>
<td>0.47 (1.06)</td>
<td>0.30 (0.47)</td>
<td>0.31 (0.47)</td>
<td>0.43 (0.76)</td>
<td>0.36 (0.50)</td>
<td>0.50 (0.86)</td>
<td>0.45 (0.79)</td>
</tr>
<tr>
<td>Pupils: Proportion ethnic minority</td>
<td>0.19 (0.32)</td>
<td>0.05** (0.11)</td>
<td>0.15 (0.29)</td>
<td>0.17 (0.30)</td>
<td>0.08 (0.170)</td>
<td>0.21 (0.32)</td>
<td>0.22 (0.34)</td>
</tr>
<tr>
<td>Pupils: Proportion of grade repeaters</td>
<td>0.26 (0.29)</td>
<td>0.14*** (0.13)</td>
<td>0.19 (0.18)</td>
<td>0.22 (0.19)</td>
<td>0.23 (0.19)</td>
<td>0.24 (0.16)</td>
<td>0.27 (0.21)</td>
</tr>
<tr>
<td>Number of Schools</td>
<td>132</td>
<td>36</td>
<td>47</td>
<td>63</td>
<td>25</td>
<td>48</td>
<td>54</td>
</tr>
</tbody>
</table>

Notes: Col(1) shows mean for whole school survey sample; Overlap sub-samples include schools in the sample that have both advantaged and disadvantaged children when advantage is defined as: Col(2) top 2 wealth quintiles; Col(3) top 2 ability quintiles; Col(4) boys; Col(5) ethnic majority children; Col(6) children of more educated mothers; Col(7) children who were not malnourished in infancy. See Notes in Table 4 for details on how each category is measured.

Means, sd’s in parentheses, asterisks indicating significance level of t-test of difference in means for all schools: *** p<0.01, ** p<0.05, * p<0.1 ; .

School asset index is the first factor from principal components analysis of school-level indicators for: need for repairs, separate room for each class, availability of library, computers, electricity and internet in school.
Table 9: Mean Incremental Effects with and without Controls for Parental Investments

<table>
<thead>
<tr>
<th></th>
<th>Vietnam</th>
<th>Peru</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>Math score</td>
<td>Vietnamese score</td>
</tr>
<tr>
<td>Mean incremental effect, $\bar{\theta}_s$ in eq (3) (se)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Model 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Being richer (top 2 wealth quintiles)</td>
<td>-8.11 (7.42)</td>
<td>-3.55 (6.89)</td>
</tr>
<tr>
<td>Being more able (top 2 CDA quintiles)</td>
<td>9.12 (9.63)</td>
<td>11.10 (9.72)</td>
</tr>
<tr>
<td>Being male</td>
<td>-12.69** (5.09)</td>
<td>-29.88*** (4.88)</td>
</tr>
<tr>
<td>Being an ethnic majority</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Having a mother with at least completed primary schooling</td>
<td>-8.98 (11.45)</td>
<td>-14.22 (10.96)</td>
</tr>
<tr>
<td>Having not been malnourished in infancy</td>
<td>4.81 (9.11)</td>
<td>-6.76 (7.72)</td>
</tr>
<tr>
<td><strong>Model 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Being richer (top 2 wealth quintiles)</td>
<td>-1.81 (6.62)</td>
<td>1.96 (6.23)</td>
</tr>
<tr>
<td>Being more able (top 2 CDA quintiles)</td>
<td>21.18*** (6.22)</td>
<td>21.79*** (5.92)</td>
</tr>
<tr>
<td>Being male</td>
<td>-9.94** (5.01)</td>
<td>-26.32*** (4.97)</td>
</tr>
<tr>
<td>Being an ethnic majority</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Having a mother with at least completed primary schooling</td>
<td>-18.01 (11.18)</td>
<td>-23.70*** (10.97)</td>
</tr>
<tr>
<td>Having not been malnourished in infancy</td>
<td>7.84 (7.79)</td>
<td>8.52 (6.32)</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01, ** p<0.05, * p<0.1; Robust standard errors in parentheses.

$\bar{\theta}_s$ is the mean incremental effect of school on test scores of advantaged children compared to disadvantaged, as shown in Equation (3).

Model 1: Main model (reproduced for convenience). Model 2: excluding controls for parental inputs. Including only controls for sex, age, ethnicity, parental education (Peru only: grade at the time of the test).

Outcome measures: Math and Verbal scores transformed using Item Response Theory (IRT); whole sample mean = 500, sd=100

Advantage categories:

- Being richer: Wealth quintiles are constructed using wealth index is a simple average of three individual indices ranging from 0 to 1, including housing quality (rooms per person, floor and roof quality), consumer durables (mean of ownership of radio, bicycle, TV, car/truck, motorbike/scooter, landline telephone, bed/table) and access to services (electricity; piped water into dwelling/yard; access to own pit latrine/flush toilet; cooking fuel is electricity, gas or kerosene).
- Being more able: The Cognitive Development Assessment (CDA) measures children’s perceptions of quantity - testing understanding of concepts such as few, most, half, many, equal, and pair.
- Being an ethnic majority: Peru - mother’s first language is Spanish (i.e. not Quechua, Aymara, other).
- Not being malnourished in infancy: defined using anthropometric indicator of adequate long-term growth: height-for-age z-scores >-2)
Figure 1: Distribution of Estimated $\theta$ Parameters for Selected Definitions of Advantage

A. Vietnam: By Gender

Math

B. Peru: By Being More Able

Math

C. Peru: By Ethnicity

Math
Appendix Table 1: Differences in Main Characteristics between Sub-sample of Children Included in the School Surveys (analysis sample in this paper) and Whole Young Lives Sample of Younger Cohort Children

<table>
<thead>
<tr>
<th>From Household Survey (time invariant variables)</th>
<th>Vietnam</th>
<th>Peru</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male (dummy)</td>
<td>0.51</td>
<td>0.50</td>
</tr>
<tr>
<td>Ethnic minority (dummy)</td>
<td>0.08</td>
<td>0.27</td>
</tr>
<tr>
<td>Father: Years of schooling</td>
<td>7.69</td>
<td>3.53</td>
</tr>
<tr>
<td>Mother: Years of schooling</td>
<td>7.17</td>
<td>3.30</td>
</tr>
<tr>
<td>Measured in Infancy (round 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth size (maternal assessment: 1=very large to 5 =very small)</td>
<td>3.08</td>
<td>0.66</td>
</tr>
<tr>
<td>Health better than other children (dummy, maternal assessment)</td>
<td>0.26</td>
<td>0.44</td>
</tr>
<tr>
<td>Age in months</td>
<td>12.82</td>
<td>2.61</td>
</tr>
<tr>
<td>Childcare: looked after by others / in crèche (dummy)</td>
<td>0.42</td>
<td>0.49</td>
</tr>
<tr>
<td>Maternal mental health</td>
<td>4.4</td>
<td>4.02</td>
</tr>
<tr>
<td>Wealth index</td>
<td>0.47</td>
<td>0.20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Diff (Std Err)</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Diff (Std Err)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.51</td>
<td>0.50</td>
<td>0.52</td>
<td>0.50</td>
<td>-0.00 (0.02)</td>
<td>0.48</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>-0.02 (0.02)</td>
</tr>
<tr>
<td>0.08</td>
<td>0.27</td>
<td>0.14</td>
<td>0.35</td>
<td>-0.06*** (0.01)</td>
<td>0.36</td>
<td>0.48</td>
<td>0.31</td>
<td>0.46</td>
<td>0.05** (0.02)</td>
</tr>
<tr>
<td>7.69</td>
<td>3.53</td>
<td>7.41</td>
<td>3.88</td>
<td>0.29** (0.14)</td>
<td>9.23</td>
<td>3.77</td>
<td>9.15</td>
<td>3.95</td>
<td>0.09 (0.19)</td>
</tr>
<tr>
<td>7.17</td>
<td>3.30</td>
<td>6.82</td>
<td>3.81</td>
<td>0.35*** (0.13)</td>
<td>7.84</td>
<td>3.11</td>
<td>7.81</td>
<td>4.54</td>
<td>0.02 (0.21)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Diff (Std Err)</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Diff (Std Err)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.02 (0.03)</td>
<td>3.11</td>
<td>0.96</td>
<td>3.12</td>
<td>0.99</td>
<td>-0.00 (0.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.02 (0.02)</td>
<td>0.41</td>
<td>0.49</td>
<td>0.40</td>
<td>0.49</td>
<td>0.02 (0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.18*** (0.11)</td>
<td>11.93</td>
<td>3.60</td>
<td>12.04</td>
<td>3.56</td>
<td>-0.11 (0.17)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04** (0.02)</td>
<td>0.21</td>
<td>0.41</td>
<td>0.24</td>
<td>0.43</td>
<td>-0.04* (0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04 (0.15)</td>
<td>5.78</td>
<td>4.26</td>
<td>5.68</td>
<td>4.29</td>
<td>0.10 (0.21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.02*** (0.01)</td>
<td>0.43</td>
<td>0.19</td>
<td>0.43</td>
<td>0.19</td>
<td>0.01 (0.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Measured at Age 5 (round 2)

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Height for age z-score</strong></td>
<td>-1.30</td>
<td>0.97</td>
<td>-1.34</td>
<td>1.04</td>
<td>0.05 (0.04)</td>
<td>-1.53</td>
<td>1.02</td>
<td>-1.54</td>
</tr>
<tr>
<td><strong>PPVT Score (IRT transformed; whole sample mean=300, sd=50)</strong></td>
<td>305.33</td>
<td>44.33</td>
<td>300.50</td>
<td>48.09</td>
<td>4.84*** (1.75)</td>
<td>301.34</td>
<td>44.84</td>
<td>300.10</td>
</tr>
<tr>
<td><strong>CDA Score (IRT transformed; whole sample mean=300, sd=50)</strong></td>
<td>306.80</td>
<td>46.56</td>
<td>300.17</td>
<td>49.42</td>
<td>6.63*** (1.79)</td>
<td>299.35</td>
<td>46.39</td>
<td>300.00</td>
</tr>
<tr>
<td><strong>Time spent in pre-school (hours per day)</strong></td>
<td>5.79</td>
<td>2.45</td>
<td>5.05</td>
<td>3.07</td>
<td>0.74*** (0.10)</td>
<td>3.44</td>
<td>1.83</td>
<td>3.60</td>
</tr>
<tr>
<td><strong>Only child (dummy)</strong></td>
<td>0.22</td>
<td>0.42</td>
<td>0.24</td>
<td>0.43</td>
<td>-0.02 (0.02)</td>
<td>0.19</td>
<td>0.40</td>
<td>0.19</td>
</tr>
<tr>
<td><strong>Household size</strong></td>
<td>4.57</td>
<td>1.39</td>
<td>4.67</td>
<td>1.51</td>
<td>-0.09* (0.05)</td>
<td>5.52</td>
<td>2.23</td>
<td>5.51</td>
</tr>
<tr>
<td><strong>Log per capita real consumption (in local currency)</strong></td>
<td>5.77</td>
<td>0.54</td>
<td>5.76</td>
<td>0.61</td>
<td>0.01 (0.02)</td>
<td>5.02</td>
<td>0.67</td>
<td>5.01</td>
</tr>
<tr>
<td><strong>Area of land owned (hectares)</strong></td>
<td>0.42</td>
<td>1.07</td>
<td>0.48</td>
<td>1.03</td>
<td>-0.06 (0.04)</td>
<td>1.09</td>
<td>3.34</td>
<td>2.54</td>
</tr>
</tbody>
</table>

**Number of Observations**

|       | 1,129 | 1,965 | 547  | 1,962 |

**Notes:** Asterisks indicating significance level of t-test of difference in means (analysis sub-sample – whole sample): ***p<0.01, **p<0.05, *p<0.1. Ethnicity minority status: Peru - mother’s first language is not Spanish (i.e. Quechua, Aymara, other); Vietnam – child reported by mother as not belonging to the Kinh group (i.e. are H’mong, Dao, Tay, other). Maternal mental health is measured using the Self-Reporting Questionnaire (SRQ20), a screening tool for risk of common mental disorders (rCMD) developed by the World Health Organization for developing countries (Beusenberg et al, 1994). It is a scale consisting of 20 statements with yes/no answer options. Score reflects the number of “yes” responses out of possible 20; a higher score reflects higher risk of rCMD. Wealth index is a simple average of three individual indices ranging from 0 to 1, including housing quality (rooms per person, floor and roof quality), consumer durables (mean of ownership of radio, bicycle, TV, car/truck, motorbike/scooter, landline telephone, bed/table) and access to services (electricity; piped water into dwelling/yard; access to own pit latrine/flush toilet; cooking fuel is electricity, gas or kerosene). Peabody Picture Vocabulary Test (PPVT) measures receptive vocabulary. The Cognitive Development Assessment (CDA) measures children’s perceptions of quantity -testing understanding of concepts such as few, most, half, many, equal, and pair.
### Appendix Table 2: Overlap in Advantage Definitions (Vietnam)

<table>
<thead>
<tr>
<th>Wealth</th>
<th>Ability</th>
<th>Gender</th>
<th>Maternal education</th>
<th>Nutrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rich</td>
<td>Poor</td>
<td>More able</td>
<td>Less able</td>
<td>Male</td>
</tr>
<tr>
<td>Richer</td>
<td>100</td>
<td>0</td>
<td>38.0</td>
<td>62.0</td>
</tr>
<tr>
<td>Poorer</td>
<td>0</td>
<td>100</td>
<td>26.9</td>
<td>73.1</td>
</tr>
<tr>
<td>More able</td>
<td>50.7</td>
<td>49.3</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Less able</td>
<td>38.0</td>
<td>62.0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Male</td>
<td>41.8</td>
<td>58.2</td>
<td>32.8</td>
<td>67.2</td>
</tr>
<tr>
<td>Female</td>
<td>42.2</td>
<td>57.8</td>
<td>30.3</td>
<td>69.7</td>
</tr>
<tr>
<td>Mother more educated</td>
<td>48.7</td>
<td>51.3</td>
<td>32.3</td>
<td>67.7</td>
</tr>
<tr>
<td>Mother less educated</td>
<td>12.5</td>
<td>87.5</td>
<td>28.4</td>
<td>71.6</td>
</tr>
<tr>
<td>Not malnourished in infancy</td>
<td>45.7</td>
<td>54.3</td>
<td>32.0</td>
<td>68.0</td>
</tr>
<tr>
<td>Malnourished in infancy</td>
<td>27.5</td>
<td>72.5</td>
<td>29.7</td>
<td>70.3</td>
</tr>
</tbody>
</table>

Notes: Table shows the proportion of children in each of the row categories who are in each of the column categories. Proportions add up to 100 for each of the advantage/disadvantage column pairs. For example, Row 1, columns 3 and 4 show that among the children in the “rich” category, 38% are in the “more able” category and 62% are in the “less able” category. See Notes in Table 4 for details on how each category is measured.
### Appendix Table 3: Overlap in Advantage Definitions (Peru)

<table>
<thead>
<tr>
<th>Wealth</th>
<th>Rich</th>
<th>Poor</th>
<th>Ability</th>
<th>More able</th>
<th>Less able</th>
<th>Gender</th>
<th>Male</th>
<th>Female</th>
<th>Ethnicity</th>
<th>Ethnic Majority</th>
<th>Ethnic Minority</th>
<th>Maternal education</th>
<th>Mother more educated</th>
<th>Mother less educated</th>
<th>Nutrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rich</td>
<td>100</td>
<td>0</td>
<td>More able</td>
<td>49.8</td>
<td>50.2</td>
<td>50.2</td>
<td>49.8</td>
<td>88.6</td>
<td>11.4</td>
<td>94.1</td>
<td>5.9</td>
<td>83.6</td>
<td>16.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poorer</td>
<td>0</td>
<td>100</td>
<td>Less able</td>
<td>21.0</td>
<td>79.0</td>
<td>46.7</td>
<td>53.4</td>
<td>47.3</td>
<td>52.7</td>
<td>58.5</td>
<td>41.5</td>
<td>69.2</td>
<td>30.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>More able</td>
<td>61.2</td>
<td>38.8</td>
<td>Male</td>
<td>100</td>
<td>0</td>
<td>46.1</td>
<td>53.9</td>
<td>70.2</td>
<td>29.8</td>
<td>85.4</td>
<td>14.6</td>
<td>72.5</td>
<td>27.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less able</td>
<td>29.8</td>
<td>70.2</td>
<td>Female</td>
<td>31.2</td>
<td>68.8</td>
<td>49.1</td>
<td>50.9</td>
<td>60.7</td>
<td>39.3</td>
<td>66.7</td>
<td>33.3</td>
<td>76.2</td>
<td>23.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>41.8</td>
<td>58.2</td>
<td>Ethnic majority</td>
<td>100</td>
<td>35.8</td>
<td>64.2</td>
<td>49.6</td>
<td>50.4</td>
<td>100</td>
<td>88.5</td>
<td>11.5</td>
<td>81.1</td>
<td>18.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>38.4</td>
<td>61.6</td>
<td>Ethnic minority</td>
<td>26.8</td>
<td>35.8</td>
<td>64.2</td>
<td>49.6</td>
<td>50.4</td>
<td>100</td>
<td>88.5</td>
<td>11.5</td>
<td>81.1</td>
<td>18.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother more educated</td>
<td>51.8</td>
<td>48.2</td>
<td>Mother more educated</td>
<td>38.2</td>
<td>35.8</td>
<td>64.2</td>
<td>49.6</td>
<td>50.4</td>
<td>100</td>
<td>88.5</td>
<td>11.5</td>
<td>81.1</td>
<td>18.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother less educated</td>
<td>08.7</td>
<td>91.3</td>
<td>Mother less educated</td>
<td>17.4</td>
<td>35.8</td>
<td>64.2</td>
<td>49.6</td>
<td>50.4</td>
<td>100</td>
<td>88.5</td>
<td>11.5</td>
<td>81.1</td>
<td>18.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not stunted in infancy</td>
<td>44.6</td>
<td>55.4</td>
<td>Not stunted in infancy</td>
<td>31.5</td>
<td>35.8</td>
<td>64.2</td>
<td>49.6</td>
<td>50.4</td>
<td>100</td>
<td>88.5</td>
<td>11.5</td>
<td>81.1</td>
<td>18.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stunted in infancy</td>
<td>26.3</td>
<td>73.7</td>
<td>Stunted in infancy</td>
<td>17.4</td>
<td>35.8</td>
<td>64.2</td>
<td>49.6</td>
<td>50.4</td>
<td>100</td>
<td>88.5</td>
<td>11.5</td>
<td>81.1</td>
<td>18.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Table shows the proportion of children in each of the row categories who are in each of the column categories. Proportions add up to 100 for each of the advantage/disadvantage column pairs. For example, Row 1, columns 3 and 4 show that among the children in the “richer” category, half are in the “more able” category and half are in the “less able” category. See Notes in Table 4 for details on how each category is measured.
## Appendix Table 4: Individual and school sample sizes in “overlap group” in each of the advantage-disadvantage categories

<table>
<thead>
<tr>
<th>Vietnam</th>
<th>Children</th>
<th>Schools</th>
<th>Peru</th>
<th>Children</th>
<th>Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richer</td>
<td>420 (474)</td>
<td>48</td>
<td>175</td>
<td>(219)</td>
<td>36</td>
</tr>
<tr>
<td>Poorer</td>
<td>510 (655)</td>
<td>127 (328)</td>
<td>47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>More able</td>
<td>346 (357)</td>
<td>61</td>
<td>144</td>
<td>(178)</td>
<td>47</td>
</tr>
<tr>
<td>Less able</td>
<td>660 (772)</td>
<td>223 (369)</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>568 (578)</td>
<td>212 (263)</td>
<td>232</td>
<td>(284)</td>
<td>63</td>
</tr>
<tr>
<td>Female</td>
<td>540 (551)</td>
<td>232 (284)</td>
<td>232</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnic majority</td>
<td>128 (198)</td>
<td>25</td>
<td>90</td>
<td>(349)</td>
<td></td>
</tr>
<tr>
<td>Ethnic minority</td>
<td>90 (349)</td>
<td>25</td>
<td>90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>More educated mother</td>
<td>816 (921)</td>
<td>61</td>
<td>235</td>
<td>(309)</td>
<td>48</td>
</tr>
<tr>
<td>Less educated mother</td>
<td>185 (208)</td>
<td>145 (238)</td>
<td>145</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not malnourished in infancy</td>
<td>775 (900)</td>
<td>264 (410)</td>
<td>264</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malnourished in infancy</td>
<td>217 (229)</td>
<td>120 (137)</td>
<td>120</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Overlap group at the child level (Col 1, 3) includes children who are in schools that have both advantaged and disadvantaged children for each definition of advantage. Overlap group at the school level (Col 2, 4) includes all the schools in the sample that have both advantaged and disadvantaged children for each definition of advantage. The total number of children in each of the advantage and disadvantage groups is included in brackets. See Notes in Table 4 for details on how each category is measured.
## Appendix Table 5: Tests for equality of non-school variable coefficients in Equation (3) across advantaged and disadvantaged children

<table>
<thead>
<tr>
<th>Advantage Indicator</th>
<th>Vietnamese Math Score</th>
<th>Vietnamese Score</th>
<th>Spanish Math Score</th>
<th>Spanish Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Being richer (top 2 wealth quintiles)</td>
<td>1.22 (0.21)</td>
<td>1.49* (0.053)</td>
<td>1.18 (0.25)</td>
<td>0.57 (0.96)</td>
</tr>
<tr>
<td>Being more able (top 2 CDA quintiles)</td>
<td>0.87 (0.65)</td>
<td>1.28 (0.16)</td>
<td>1.11 (0.33)</td>
<td>1.28 (0.16)</td>
</tr>
<tr>
<td>Being male</td>
<td>0.74 (0.83)</td>
<td>0.87 (0.65)</td>
<td>1.39 (0.102)</td>
<td>1.34 (0.12)</td>
</tr>
<tr>
<td>Being an ethnic majority</td>
<td></td>
<td></td>
<td>1.30 (0.17)</td>
<td>2.18*** (0.00)</td>
</tr>
<tr>
<td>Having a mother with at least completed primary schooling</td>
<td>1.30 (0.15)</td>
<td>1.36 (0.11)</td>
<td>1.51* (0.06)</td>
<td>1.33 (0.14)</td>
</tr>
<tr>
<td>Having not been malnourished in infancy</td>
<td>1.55** (0.04)</td>
<td>1.28 (0.15)</td>
<td>3.14*** (0.00)</td>
<td>1.66** (0.02)</td>
</tr>
</tbody>
</table>

Notes: Table shows F-test statistics and p-values in brackets for tests of joint significance of interactions of non-school variables with each of the advantage indicators used in the main analysis: *** p<0.01, ** p<0.05, * p<0.1. See Table 3 for complete set of non-school variables and details of how they are measured. See Notes in Table 4 for details on how each advantage category is measured.