IDB WORKING PAPER SERIES No. IDB-WP-435

Intercontinental Evidence on Socioeconomic Status And Early Childhood

Cognitive Skills: Is Latin America Different?

Florencia Lopez Boo

August 2013

Inter-American Development Bank
Social Protection and Health Division

Intercontinental Evidence on Socioeconomic Status And Early Childhood

Cognitive Skills: Is Latin America Different?

Florencia Lopez Boo

Inter-American Development Bank

2013

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Lopez Boo, Florencia, 1975-
Intercontinental evidence on socioeconomic status and early childhood cognitive skills: is Latin America different? / Florencia Lopez Boo.
p. cm. (IDB working paper series ; 435)
Includes bibliographical references.
1. Cognition in children—Developing countries. 2. Children—Developing countries—Economic conditions. 3. Children—Developing countries—Social conditions. I. Inter-American Development Bank. Social Protection and Health Division. II. Title. III. Series.
IDB-WP-435

http://www.iadb.org

Intercontinental Evidence on Socioeconomic Status And Early Childhood Cognitive Skills: Is Latin America Different?

Florencia Lopez Boo*

Abstract

This paper documents disparities in cognitive development— as measured by a receptive vocabulary test—between children from households with high and low socioeconomic status (SES) in two different phases of childhood (before and after early school years) in four developing countries: Peru, Ethiopia, India, and Vietnam. Intercontinental evidence on the timing, shape, pattern, and persistence of these disparities is provided. The nonparametric analysis suggests that disparities found at age 5 persist into the early school years across all four countries, and the conditional analysis shows that SES disparities seem to fall over time. However, both the magnitude of the gap and the degree of persistence vary. The main result is that Peru stands out, not only as the country with the largest cross-section disparity between rich and poor (of around 1.30–1.40 standard deviations), but also as the country with the highest persistence in cognitive development, as shown by the value-added specification. The latter suggests fewer opportunities for convergence in cognitive development between rich and poor over time in this Latin American country. Some channels behind these trends are discussed, but overall, the SES gradient persists even when controlling for a large number of important mediators, such as preschool, early nutrition, and schooling. Past performance on the Peabody Picture Vocabulary Test (PPVT) is the most important mediator of the SES gradient at age 8.

JEL classification: I14, I24, J13, J24, O53, O54, O55.

Keywords: cognitive skills, inequality, Latin America, Peru, Vietnam, Ethiopia, India

* Social Protection and Health Division, Inter-American Development Bank, and Institute for Study of Labor (IZA). Corresponding author: florencial@iadb.org, Inter-American Development Bank, 1300 New York Ave., NW (Room E0605), Washington, DC, 20577, USA. I am grateful for the Round 3 data provided by the Young Lives project. I am also grateful to Marta Favara for her early work on this piece, and to Fiorella Benedetti and Romina Tome for excellent research assistance.

1. Introduction

Household poverty is one of the main risk factors for a child to have poor nutrition and stimulation (Walker, Vera-Hernández, and Grantham-McGregor, 2011). Moreover, research from a number of developed (Currie and Thomas, 2000; Feinstein, 2003; Case and Paxson, 2008; Heckman et al., 2010) and developing countries (Hoddinott et al., 2008; Maluccio et al., 2009; Behrman et al., 2009; Walker, Vera-Hernández, and Grantham-McGregor, 2011) find that poor nutritional status and low levels of cognitive development in early childhood are important determinants of fewer years of schooling, less learning while in school, lower cognitive skills and worse health in adolescence (including mental health) and adulthood, lower probabilities of employment, lower earnings, lower wage rates, and more criminal activity. This evidence points to the fact that disadvantages found at early ages will result in the intergenerational transmission of poverty and inequalities.

Understanding socioeconomic status (SES) "gradients" in cognitive skills early in a child's life is therefore a crucial step toward understanding the intergenerational transmission of poverty and inequality. The persistence between parents and children's outcomes is well documented (Black and Devereux, 2010), but the magnitude of the intergenerational correlation varies dramatically across countries and regions. In particular, Hertz et al. (2008) provide a survey of parent-child schooling correlations for a sample of 42 countries using comparable data and find that the correlation is 0.66 in Peru, 0.40 in Vietnam, and 0.10 in Ethiopia, which are three of the four countries of interest in this paper.[1] Parent-child wages correlations are similar. Moreover, decompositions have shown that part of the observed intergenerational correlation can be explained by the fact that parental SES strongly predicts cognitive and noncognitive skills and health (Black and Devereux, 2010),[2]—a finding that motivates my interest in the SES-cognitive skills link in the first place.

Until very recently, there were no comparable data on cognitive development of young children (0–6 years of age) for most developing countries (Harpham, 2002). Therefore, not much research exists on disparities in cognitive development by SES, when they arise,

[1] Unfortunately, Hertz et al. (2008) do not have data available for India, the fourth country of interest in this paper.

[2] Blanden, Gregg, and Macmillan (2007) examine the role of noncognitive skills and ability for intergenerational income persistence in Britain. Their work demonstrates that covariates can account for approximately half of the estimated intergenerational income elasticity (of 0.32), with a sizeable portion attributable to cognitive and noncognitive skills that work through educational attainment. Moreover, it has been fairly well established that better infant health has a positive causal effect on later adult outcomes (Currie and Moretti, 2003).

whom they affect, and how they evolve as children age. Gradients in cognitive development have indeed been found in Brazil (Victora et al, 2003) and Guatemala (Stein et al., 2008). More recently, Fernald et al. (2011) and Naudeau et al. (2011) use single cross-sections of data from low-income countries to study cognitive skills gradients, while related research from Ecuador (Paxson and Schady, 2007; Schady, 2011) showed substantial differences in cognitive development at young ages between children of higher and lower socioeconomic status, which increased between 3 and 5 year of age. The analysis from Ecuador included only households in the poorest half of the nationwide distribution of a composite measure of wealth, and it was limited to relatively young families in rural areas. That is why the study more related to this research is Schady et al. (2011). In their work, wealth gradients in five Latin-American countries (including Peru) are studied and substantial differences in children's performance on the Peabody Picture Vocabulary Test (PPVT) are found; these differences seem to persist as children age. However, I am not aware of any study providing intercontinental evidence on the magnitude of the disparities, as well as the persistence of PPVT scores using longitudinal data in developing countries.

Therefore, the contribution in this paper is the investigation of the relationship between SES and an important component of cognitive development—receptive language ability—for children 55 through 102 months of age in four developing countries (Ethiopia, India, Peru and Vietnam). In these four countries, the Young Lives (YL) study has collected extensive comparable data over three rounds of surveys.[3] Specifically, this paper provides intercontinental evidence on the magnitude of SES gradients as well as the persistence of PPVT scores using longitudinal data. It analyzes the shape, timing, and the pattern of these disparities in two different phases of childhood (before and after early school years). As far as I know, no analysis has been done in this respect.[4] An understanding of what children are most likely to show deficits in cognitive development, at what ages these deficits become apparent, and how these deficits evolve as children age is indispensable for the formulation of appropriate policies.

An additional motivation for this work is given by the fact that Latin America is the most unequal region in the world (Gasparini and Lustig, 2011; IDB, 2011). One simple and

[3] Young Lives, established in 2000, is an international research project on childhood poverty. The study traces the lives of children in Ethiopia, India, Peru, and Vietnam. See http://www.younglives.org.uk/ for more information.

[4] An additional, though more methodological, contribution is to see whether previous results from Latin America hold for a nationally representative sample (like the Peruvian one) and with a survey that covers a longer period; there are about four years between Round 2 and 3 of the YL data, while in previous studies there were only two years.

intuitive measure of inequality is the ratios of the income or consumption of households at the 90th percentile of the distribution to those at the 10th percentile. Using data from around 2000 and ranking countries by this ratio, 14 of the 15 countries in the world with the highest levels of inequality are found in Latin America. Moreover, much of Latin America's inequality is associated with inequality of opportunities, not just outcomes. That is, a substantial fraction of the inequality in incomes that is observed is determined by socioeconomic status (characteristics such as race, place of birth, wealth, or the education levels of one's parents).

In this context, I am most interested in Peru because it is the 21st most unequal country in the world, with a Gini of 0.48 (World Bank, 2005), while none of the other YL countries is anywhere close to that position in the ranking.[5] Not surprisingly, existing analysis of YL data shows that socioeconomic gradients between socioeconomic status (as measured by mother or father's education or wealth) and early childhood skills (as measured by the PPVT) are present at age 5 (Engle et al., 2011). Extending the analysis (with the availability of Round 3 data) to the children who are 8 years old will be useful for understanding how these differentials evolve, particularly once these children are in school. It might well be that differentials fade away due to the equalizing effect of schools, but this is the empirical question I will attempt to answer.

I will draw on the literature on SES and health (Currie, 2009; Case, Lubotsky, and Paxson, 2002) and SES and cognitive skills (Victora et al., 2003). In particular, I will build on the work of Paxson and Schady (2007) and Schady et al. (2011) for the descriptive section that uses nonparametric regressions. For the regression analysis, I build on the value-added production function approach in Todd and Wolpin (2007), which allows this research to go beyond all previous empirical studies on this topic.

Results show that, although differences in cognitive skills between SES are present in all countries, they arise more starkly in Peru, with a gap of around 1.30–1.40 standard deviations (SD) between the poorest and the richest quintiles. For Peru, which has norms provided by test developers, these differences imply developmental lags of up to one year at age 5. OLS regressions show that SES disparities seem to fall from age 5 to age 8; and that, overall, the SES gradient persists even when controlling for a large number of important mediators, such as preschool, early nutrition, and primary schooling. Using the value-added regressions that exploit the longitudinal data, I find that Peru shows the highest coefficient of

[5] Urban India is in the 84th place, Vietnam (urban-rural) in the 87th, rural India in the 119th, and Ethiopia (urban-rural) in the 121st place in this ranking.

past PPVT on current PPVT (0.42). This indicates that *ceteris paribus*, convergence between the rich and the poor will happen at a much slower speed in Peru than in the other three countries, unless accompanied by appropriate policies. These regressions also show that PPVT at age 5 is the most important mediator of the SES gradient at age 8.

The structure of the paper is as follows. Section 2 presents the analytical approach, economic framework and some brief descriptive analysis of the data. Section 3 presents the regressions, while Section 4 discusses the results and concludes.

2. Data, Economic Framework, and Descriptive Analysis

2.1 Data and Methodology

The data are from the longitudinal survey of the Young Lives (YL) project. Beginning in 2002, YL surveyed children in two cohorts (younger and older). I use data for the 8,000 children in the YL younger cohort dataset speaking the majority language of the region or country—Amarigna in Ethiopia, Telugu in India, Spanish in Peru,[6] and Vietnamese in Vietnam—following Cueto, Guerrero, and Munoz (2009). It is worth noting that the samples in Ethiopia, India, and Vietnam are not fully representative of each country, as the focus is on poor children; by contrast, the sample in Peru is representative for all but the richest 5 percent of districts. Moreover, in the case of India, the sample comes from one region: Andhra Pradesh. For all countries, the urban-rural divisions of the samples follow the same criteria given the identical sample designs of the YL study, but I pool the data to obtain more precise estimates.

A major strength of this study is the use of a common measure of child cognitive development: performance on the widely used Peabody Picture Vocabulary Test. Raw scores on this test, given by the number of items answered correctly, can be used only to compare children of the same age. I therefore present results based on the following common standardization of PPVT scores separately by country. I constructed internally standardized, age-specific z-scores by subtracting the month-of-age-specific mean of the raw score and dividing by the month-of-age-specific standard deviation. These results can be used to make comparisons within countries, including comparisons between children of different ages.

[6] In Peru, the PPVT was translated into Quechua, an indigenous language spoken primarily in rural areas of the highlands, and children were given the option of taking the test in Spanish or Quechua. Twenty-two percent of children in rural areas, but only 0.1 percent of children in urban areas, chose to take the test in Quechua. In the robustness tests, I test whether our results are robust to including these children in the analysis.

However, internally standardized scores are not informative about differences across countries. I use this method to see SES gaps within countries, and then I compare those SES gaps (in standard deviations) across countries.

I use household expenditure as a proxy of SES.[7] Expenditures are calculated for all sample households in each country (about 2,000); they include expenditures on food, transport, security, telephone, electricity, water supply, housing, clothes, and footwear. Quintiles of expenditure are created separately for each country on the basis of the aggregation of all sampled households in that country. The distribution of expenditures in the country samples, as well as the distribution of wealth, are presented in Figures A1 and A2 in Appendix A. The wealth distribution in India and Ethiopia seems shifted to the left, while the Peruvian distribution of wealth looks bimodal. On the contrary, all four distributions of expenditures have the usual bell shape of a normal distribution, which emphasizes the need to use expenditure data instead. I then sort children into quintiles of the household expenditure distributions and compare outcomes for children in the top and bottom quintiles. The mean language scores for each expenditure quintile are presented for each country in terms of SDs.

Lastly, I build a balanced panel of children present in Round 2 (at age 5) and Round 3 (at age 8) to analyse the evolution of the gaps in PPVT by SES. Specifically, I inspect age patterns in SES gradients in child development using nonparametric regressions (Fan and Gijbels, 1996). I split the sample in two: households in the first quintile of household expenditure and those in the fifth quintile. These particular results are presented in Figure 2. The nonparametric regressions smooth out average PPVT scores by age and make patterns more apparent. Similar methods have been used extensively in the literature in economics (Deaton, 1997; Case, Lubotsky, and Paxson, 2002) and, more recently, in medicine (Park et al., 2005; Moscicki, et al., 2004; Fleming, et al. 2011). Standard errors and confidence intervals for the nonparametric regressions are constructed by bootstrapping (Efron and Tibshirani, 1993).

In Section 3, I run OLS regressions of the PPVT at age 5 and age 8, introducing potential mediators of the SES-PPVT relationship gradually. Finally, I run a value-added production functions, as explained in the next subsection, for the PPVT at age 8.

[7] The analysis of the associations between PPVT scores and wealth (with wealth indexes calculated separately for urban and rural areas) was done in the previous version of this paper. However, as noted by a reviewer, household expenditures (or consumption) are a better indicator of SES.

2.1.1 Motivation for the Valued-added Regression Analysis

The value-added regression analysis is inspired by the educational production function literature that commonly adopts value-added specifications when data on lagged inputs are missing or incomplete (Todd and Wolpin, 2007). In its most basic form, the value-added specification relates an achievement outcome measure to contemporaneous measures of school and family inputs and to a lagged (baseline) achievement measure. Then, if lagged cognitive skills (PPVT at age 5) is a sufficient statistic for input histories and unobserved ability, estimating equation (3) would result in consistent estimates of the production function of cognitive skills today (PPVT at age 8).

Following Todd and Wolpin (2007), the goal is to identify the relation between current (age 8) and past (age 5) cognition to explore the level of persistence of the performance on the PPVT. In equation (1), θ_{it} are cognitive skills of child i in time t, which depend on child's nutrition status at the beginning of the period (H_{it}), the family's characteristics (X_{it}), the child's ability (μ_{it}), and an error term (ε_{it}). α_{it} is a constant

$$\theta_{it} = \alpha_{it} + \beta H_{it} + \gamma X_{it} + \delta\, \mu_{it} + \varepsilon_{it} \qquad (1)$$

The assumptions for the estimation are that only contemporaneous inputs matter for the production of current cognitive skills (or inputs are unchanging over time) and contemporaneous inputs are uncorrelated with unobserved ability and unobserved inputs. Parents will then maximize lifetime utility derived from consumption (c), leisure (l), and their child's achievement (θ) intertemporally:

$$Max_{ht,it}\ \ U(c_t, l_t, \theta_t) \qquad (2)$$

subject to:

- a production function of cognitive skills, $\theta_t = f(\theta_{t\text{-}1}, i_t, H_t, X_t, \mu_t)$
- a production function for next period nutrition status, $H_{t+1} = g(H_t, i_t, X_t, \mu_t)$
- a time constraint, $l_t = 1 - h_t$
- a budget constraint,

where h is the total number of hours worked, l is leisure, and i_t are family inputs.

Given wages and prices, parents choose how much to work in the market and how much to invest in their child's production of cognitive skills and their child's nutrition. I assume that investment in nutrition affects the child's nutrition status only at the end of the period. I also assume that the production function has a value-added form, and it can be written as:

$$\theta_{it} = \alpha_{it} + \beta H_{it} + \gamma X_{it} + \psi \theta_{it-1} + \delta \mu_{it} + \varepsilon_{it} \quad (3)$$

The coefficient of interest in Section 3 is ψ.

Even if I estimate a production function for cognitive skills, it is worth clarifying that SES enters the production function not as an input directly affecting cognitive outcomes but as an indirect determinant affecting the productivity of inputs and the shape of the production function. Currie (2009) suggests potential mechanisms through which SES may affect child outcomes. She looks at the link between SES and child health. However, a very similar framework can be applied to the links between SES and child cognitive achievement, and that is the framework I follow here.

2.2 SES and Cognitive Skills

Table 1 reports differences in average z-scores of the PPVT between the fifth and first quintile in the first column in each country. The second column presents the p-value of the difference between the quintiles.

Table 1. PPVT z-score at Age 5 and 8, Gaps between 1st and 5th Quintiles and t-tests

	Ethiopia		India		Peru		Vietnam	
	(Q5-Q1)	*p-value*	*(Q5-Q1)*	*p-value*	*(Q5-Q1)*	*p-value*	*(Q5-Q1)*	*p-value*
PPVT age 5	**1**	0.000	0.3	0.000	**1.3**	0.000	1	0.000
PPVT age 8	**1**	0.000	0.5	0.000	**1.4**	0.000	0.7	0.000

Note: The calculation of the mean scores gives equal weight to each month of age, within a country.

Differences in language development between richer and poorer children within countries are large and statistically significant, both at age 5 and age 8. [8] Table 1 and Figure 1 show that these differences are biggest in Peru in both waves. At age 5 in Peru, there is a 1.30 SD gap between the first and the fifth quintile, followed by the gap in Ethiopia and Vietnam

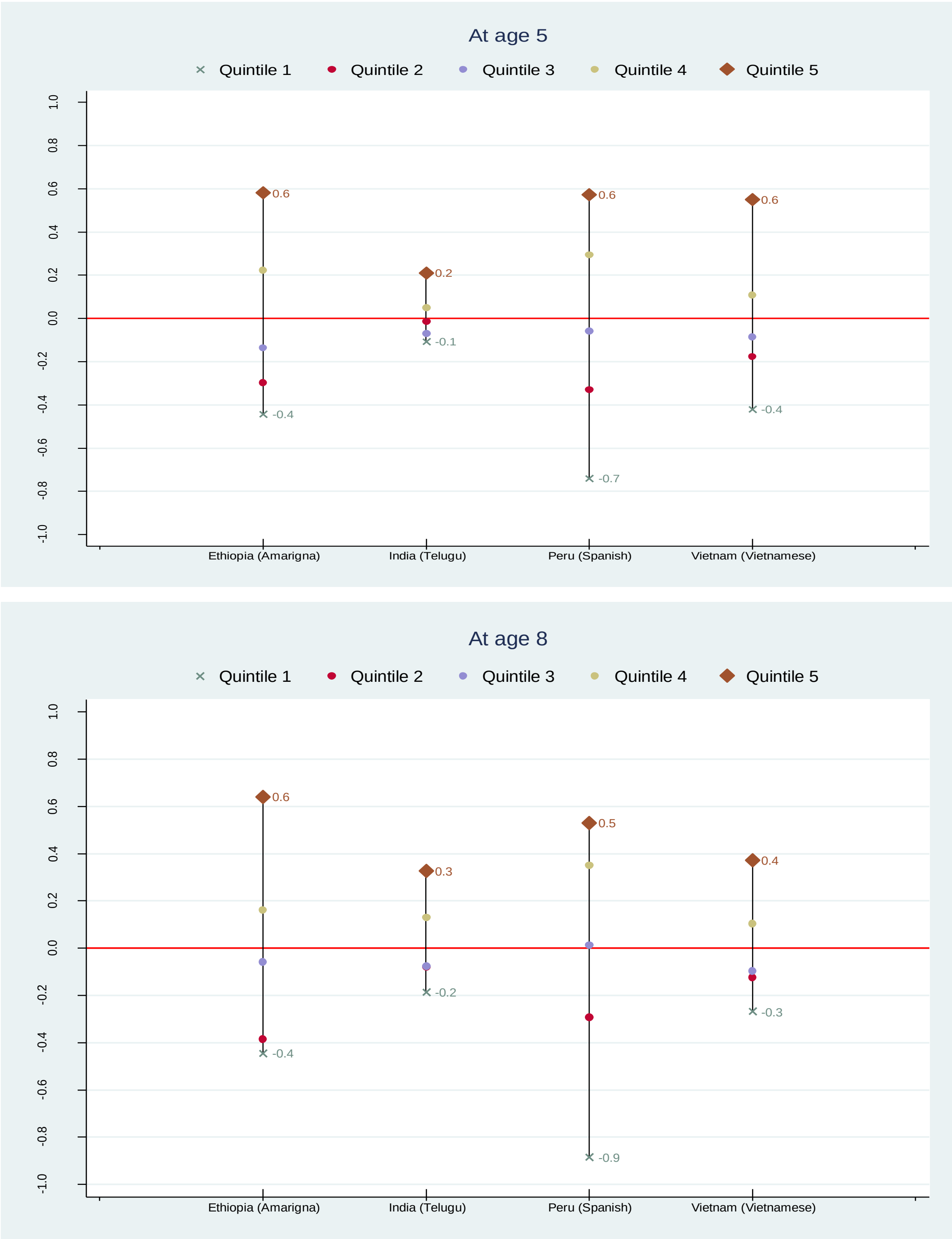
[8] t-tests available upon request.

(both with 1 SD gap), and then by the gap in India (0.30 SD).[9] In Round 3, Peru presents the widest gap (1.40 SD), followed by Ethiopia with 1 SD, Vietnam (0.70 SD), and India (0.50 SD). The last two are not significantly different from each other.

The gaps for Peru are larger than those found elsewhere—specifically, in Victora et al. (2003); Stein et al. (2008); Fernald et al. (2011) and Naudeau et al. (2011). Using the same dataset for Peru, Schady et al. (2011) found that the gap is 0.95 SD in urban areas and 0.77 SD in rural ones. Moreover, they find that differences in internally standardized scores between the first and fourth quartile in the distribution of wealth are biggest in urban Colombia (1.23 SD) and rural Ecuador (1.21 SD). Although these numbers are not completely comparable because they use quartiles of wealth (and not quintiles of expenditure), they are informative of the fact that what proxy is used for SES—wealth/mother schooling levels is used in Schady et al. (2011) versus household expenditures here—matters for the magnitudes found.

[9] In the case of India, the caste of the children might be the best indicator to portrait Indian social inequalities rather than the expenditure distinction made in this paper.

Figure 1. PPVT z-scores at Age 5 and 8 by Country and Quintile of Expenditure, Majority Language Sample



In Figure 2, I investigate age patterns in SES gradients in child development using nonparametric regressions on the full panel sample. I split the sample in two: households in the first quintile (solid blue line) of household expenditure, and households in the fifth quintile (dotted red line). The dotted redlines are the confidence interval for the fifth quintile, and the blue ones for the first quintile. These internally standardized scores suggest that the bulk of the difference between poorer and less poor children is apparent by age 5 (around 57 months) in all four countries. Looking at externally standardized scores, I observe that by the time children are 5½ to 6 years of age, the poorest children in Peru are 35 to 40 points (approximately 2½ SD, or 1 year) behind the reference population (see also Schady et al., 2011).[10] Unfortunately, it is not possible to perform this calculation for the other three countries, as there are no externally referenced tests in Vietnamese, Telugu, or Amarigna.

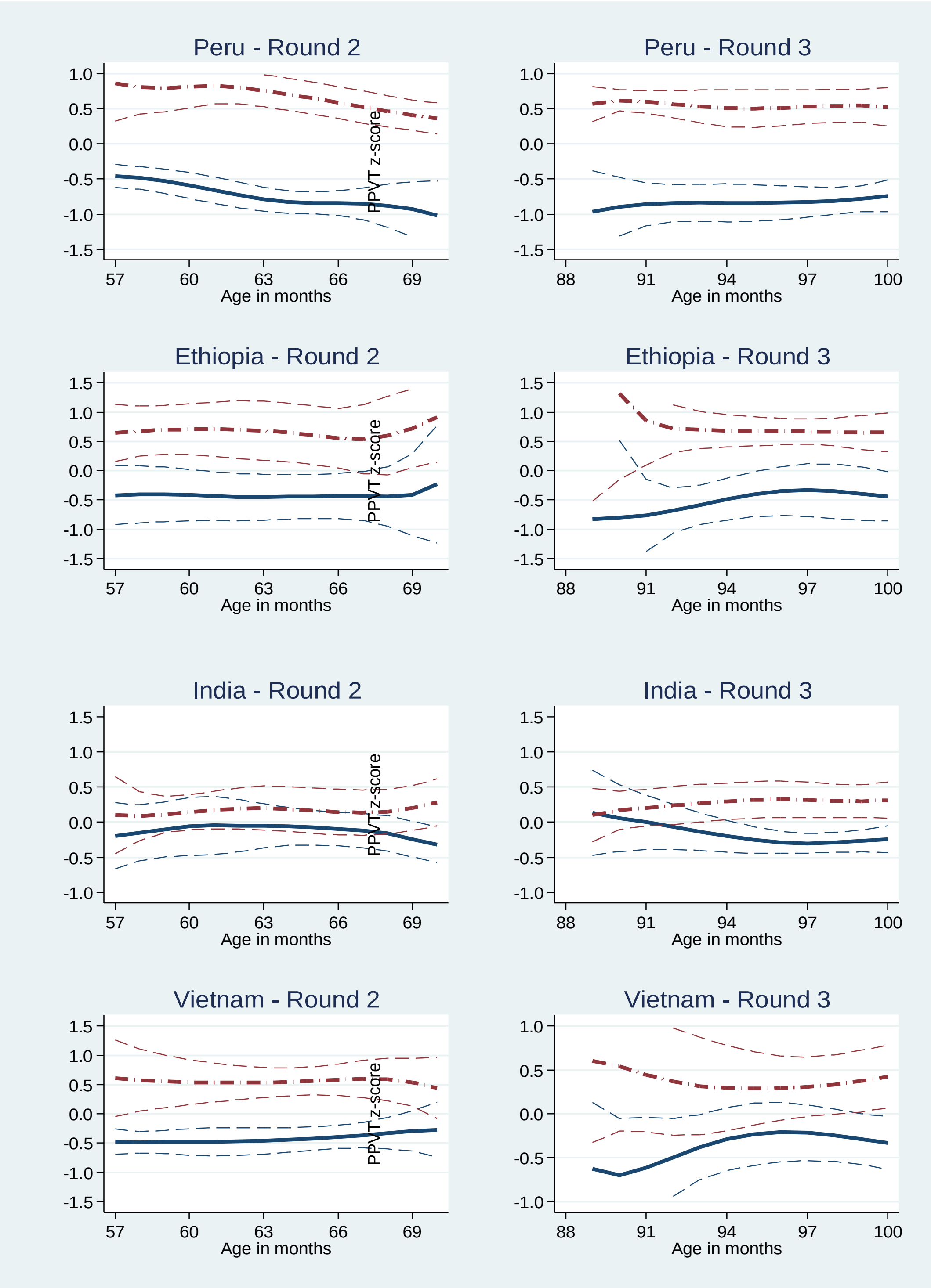
There is yet another peculiarity about Peru: Figure 2 shows that the gap immediately starts out somewhat ‘‘big” at age 55–58 months (1.4 SD) and stays at that level throughout that round. In all the other three countries, the gap is also stable, but smaller. In Peru, the scores of both the poorer and the richer children go down slightly after 63 months of age—a pattern worth exploring.

Round 3 data on the same children show that the household expenditure gradients that are apparent among 4–5 year old children continue to be apparent as these children enter the first years of primary school in all countries. On the other hand, the poorest children do not appear to fall further behind, either. For all four countries, the gap between the PPVT score of children living in households in the first quintile of the household expenditure distribution and those in the fifth quintile seems pretty stable over time once children are in school. Overall, there is no clear pattern of a systematically widening gap, which challenges the evidence in Paxson and Schady (2007) for the Ecuadorian data (the only longitudinal data on PPVT I am aware of besides the YL data in a developing country).

As a robustness test, I have pooled the two rounds of data for each country and regress PPVT on expenditure quintiles and the interactions of expenditure quintiles with month dummies. The sum of the coefficient of the fifth quintile and its interaction with the coefficient of any given age month is the estimate of the disparity in PPVT between fifth and first quintile (that is, the first quintile is the excluded dummy) at the given age. These coefficients were all statistically significant, positive, and showed the same pattern as the ones described in Figure 2.

[10] The calculation of the relative delays follows the next steps. First, I calculate the average raw score by month of age, separately for children in urban and rural areas and separately for children in the poorest and richest quintile. Second, I use the tables provided by the developers of the PPVT (TVIP in Spanish) test to identify the age in months at which this raw score corresponds to a score of 100 in the reference population. The difference in the age at which children in my sample and children in the reference population attain the same vocabulary level is my estimate of the months delayed. Third, I report the difference in months delayed between children in the first and fifth expenditure quintiles. The calculation of the relative delays gives equal weight to each month of age, within a country and by place of residence (urban or rural).

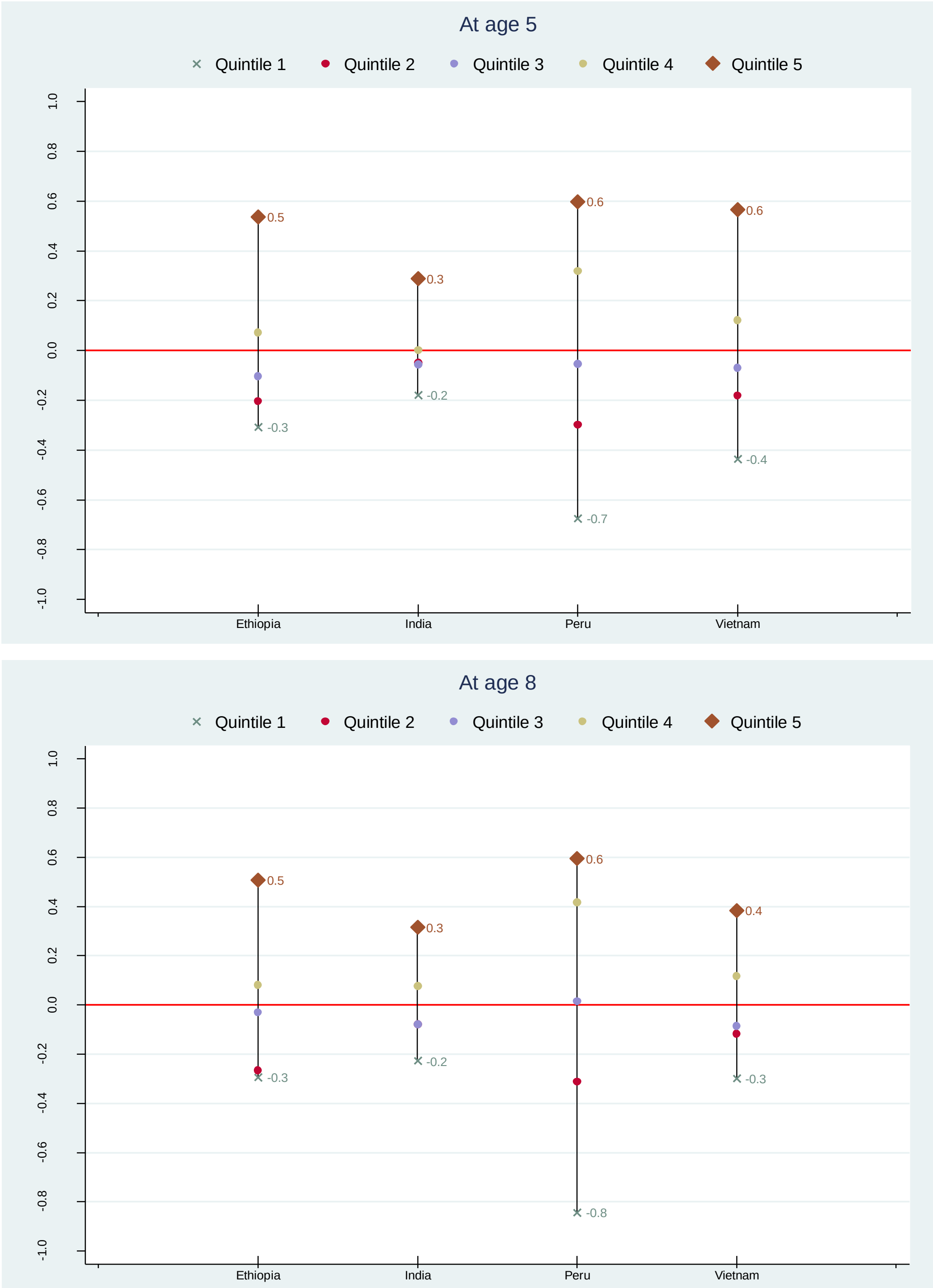
Figure 2. Panel Data Analysis of PPVT Age Patterns, Four Young Lives Countries, Majority Language Sample



Note: As in all calculations in this paper, regressions are restricted to the panel of children observed in both Rounds 2 and 3. The nonparametric regressions of PPVT score (internally standardized for main language group) on age in months, by household expenditure quintile (Q1 vs. Q5). The bandwidth of the regressions is 5. Data for children 88–91 months of age in Round 3 in Ethiopia have confidence intervals that are too wide because of small sample size at those ages (when the data are restricted to children whose first language is Amarigna, the sample from Ethiopia is reduced by half of the total size). The same applies for India for children 57–59 and 88–91months of age.

I conduct an important robustness checks on this main result. Mainly, it is likely that the poorest children in all countries are being dropped from the sample above as a consequence of the strategy of considering only those children who speak the main language of the country. Because of this consideration, I present Figure 3, which replicates Figure 1 except for the full sample of children in each country sample. The ordering of countries in terms of the gaps remains exactly the same. There are only very small changes in magnitudes relation to Figure 1 in the case of Ethiopia (both waves) and India in Round 2. Replicating Figure 2 for the full sample also gives the same results (available upon request). Regressions analysis will also consider both samples.

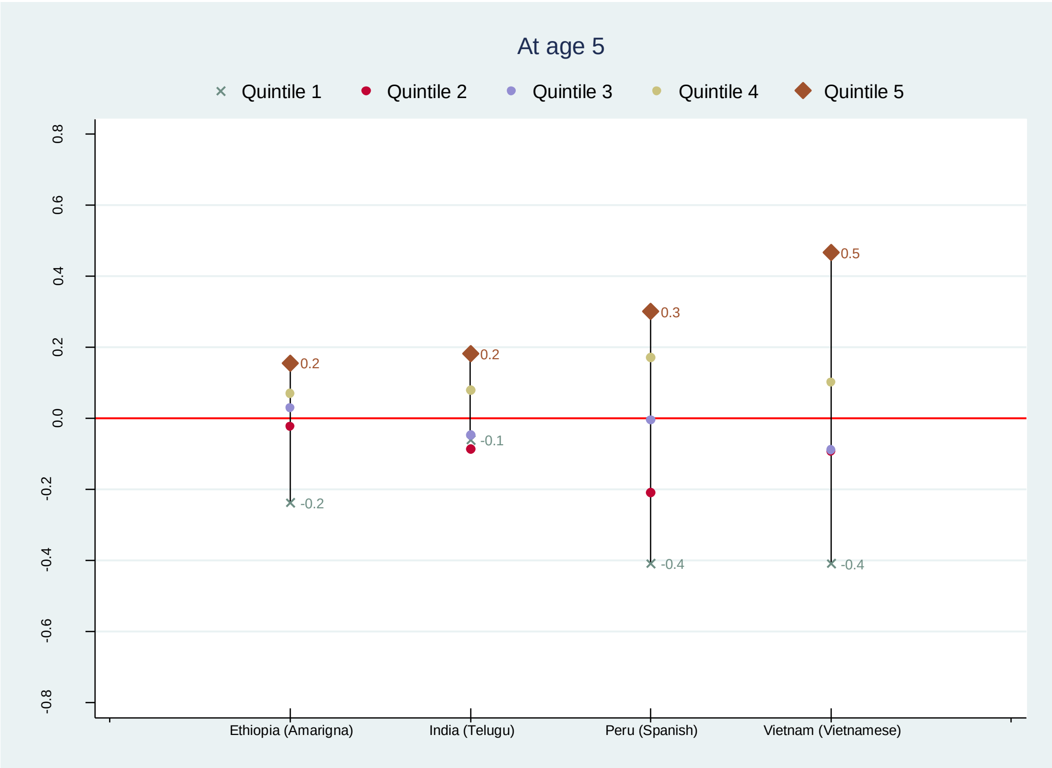
Figure 3. PPVT z-scores at Age 5 and 8 by Country and Quintile of Expenditure, Full Sample



2.3 SES and Nutrition

Looking for important covariates behind the trends found in Subsection 2.2, I look at one variable that has been shown to be the most important determinant of cognitive development: the World Health Organization (WHO) standardized score of height-for-age at age 5 (HAZ at age 5), a proxy of chronic malnutrition, known to be strongly related with mental functioning, particularly during the first five years of life (Alderman, 2000; WHO, 2000; among many others). The HAZ distribution in Round 2 (Figure 4) shows that the biggest gap in HAZ between the richest and the poorest at age 5 is found among Vietnamese children (0.90 SD). The Peruvian children follow closely, with a gap of 0.70 SD.[11] This shows that this variable must be closely examined in the regression analyses.

Figure 4. Height-for-Age z Scores at Age 5, by Country and Quintile of Expenditure



[11] Given what is stated in the medical literature (WHO, 2000) in the regressions analysis, I include HAZ at age 5, as it at this stage when nutritional status builds up.

3. Regression Analysis

3.1 The SES Gradient in PPVT Scores

To further investigate the relationship between SES and PPVT across ages and countries, Table 2 shows an OLS regression of PPVT at age 5 (in the first four columns) and age 8 (in the last five columns). Both types of regressions control for standard predetermined variables at the child level (such as sex, age, and birth-order) and household level (region, urban, size, expenditure quintile, education of the mother and father, and mother's height) measured at the relevant age. Standard errors are clustered at the sentinel site level. All these controls have the expected sign, with the urban dummy, and parental education co-variates showing singular relations with the PPVT (see the complete version of Table 2, which reports all controls, in Appendix Table A1 of Appendix A).

To explore the factors mediating the link between SES and cognitive development, I will sequentially include a set of relevant controls. In the first specification (first and fifth columns), only predetermined controls at the child and household level are included. In a second specification (the second and sixth column of each country in Table 2), I include the preschool attendance dummy, as this is one important variable that could explain the gaps found, if richer children go to better private schools and poorer children do not attend school at all or go to (bad) public preschools. In the third specification (the third and seventh column of each country), HAZ at age 5 is included, as nutrition during the first five years of life is an important determinant of cognition (Walker et al., 2011). In the fourth specification (only presented for PPVT at age 8 regressions), I include an indicator of whether the child is attending schooling or not (this indicator is not present at age 5, as children are too young to be in formal schooling).

The coefficient of interest is that of the fifth household expenditure quintile (with the first quintile as the omitted category). This coefficient shows that *ceteris paribus*, at age 5 richer children in Peru will have a PPVT score that is 0.63 SD higher than poorer children. This figure is 0.51 SD in Ethiopia, 0.34 SD in India, and 0.53 SD in Vietnam. At age 8, these figures are 0.54, 0.41, 0.45, and 0.40, respectively. This shows that Peru exhibits the largest "gradient" both at age 5 and age 8; see columns (1) and (5).

When controlling for preschool attendance, results remain unaltered for the age 5 regressions (with the exception of the coefficient in Ethiopia, which decreases by 4 percentage points (pp)—although that does not alter their ranking position). Preschool attendance is significant only for Peru and Ethiopia, and stands at 0.25 SD and 0.34 SD, respectively. At age 8, however, the inclusion of preschool alters the ranking, with Ethiopia falling to last place and Vietnam to third place. This is due to Ethiopia's coefficient decreasing by 4 percentage points. The coefficient on preschool attendance is still significant only for Peru and Ethiopia, and stands at 0.21 SD and 0.34 SD, respectively.

When adding HAZ at age 5, results for the PPVT age 5 regressions remain unaltered in relation to the benchmark in column (1) for Peru and India.[12] HAZ is significant for Peru, India, and Vietnam, and ranges from 0.12 SD (India and Peru) to 0.18 SD (Vietnam). Vietnam, whose coefficient on the fifth household expenditure quintile has decreased by 12 pp, becomes the third in the ranking, while Ethiopia becomes the second. At age 8, the inclusion of HAZ does not alter the ranking either, although the coefficient of Peru decreases to 0.51 SD. The coefficient on HAZ is significant in all countries, and of a similar magnitude to four years previously, at the exception of Vietnam. Adding whether the child is at school at age 8—in column (8)—does not change the ranking results, and this coefficient is significant (and large) in all countries.

Probably the most important specification is the one in columns (4) and (9), in which all the potential mediators are included. The ranking shows that Peru is followed by Ethiopia, Vietnam, and India at age 5; while at age 8, it is followed by India, Ethiopia, and Vietnam. Moreover, even if the individual inclusion of each of these controls does not change the magnitude of the coefficient of interest, the inclusion of all of them does indeed decrease the magnitude of the SES coefficient.

Overall, the SES gradient persists even when controlling for a large number of important mediators. Peru has the largest coefficients on the fifth quintile of expenditure in all specifications across ages. Moreover, these finding suggests that preschool might be one important channel that mediates the SES gradients in Ethiopia, but less so in the other countries; while early nutrition seems to be an important channel in Vietnam.

Comparisons of the coefficient of the fifth quintile of the consumption expenditure distribution between these regressions—using the preferred specifications in columns (4) and (9)—would suggest that SES disparities in cognitive achievement decrease over time (with

[12] I have added the interaction of expenditure quintiles with HAZ to see if that pattern provides some hints on the channel nutrition-cognition; however, although these interactions are all significant, they are all very close to zero.

the exception of India, where it increases by 11 percentage points). The decline is 9 percentage points in Peru and Vietnam, and 10 percentage points in Ethiopia.

Table 3 replicates Table 2, but includes all observations in the country samples (that is, not only the majority language) and a dummy that indicates whether the language used by child during administration was different from the majority language. Main results remain unchanged, although in the regression of the PPVT at age 5, the coefficient on the fifth quintile increases substantially in India, from 0.30 to 0.40 SD in the preferred specification in column (4).

Table 2. OLS Regressions, Four YL Countries, Majority Language Sample

PERU									
	PPVT age 5				**PPVT age 8**				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Expq5	0.63***	0.62***	0.60***	0.59***	0.54***	0.53***	0.51***	0.54***	0.50***
	(0.099)	(0.103)	(0.099)	(0.102)	(0.105)	(0.108)	(0.101)	(0.107)	(0.106)
Preschool		0.25***		0.22***		0.21**			0.17*
		(0.062)		(0.063)		(0.088)			(0.088)
HAZ age 5			0.12***	0.12***			0.14***		0.13***
			(0.030)	(0.031)			(0.029)		(0.030)
In school								0.57*	0.52*
								(0.281)	(0.288)
Constant	2.03**	2.11**	3.50***	3.50***	-2.06*	-1.84*	-0.51	-2.64**	-0.91
	(0.847)	(0.852)	(0.961)	(0.978)	(1.059)	(1.052)	(1.050)	(1.080)	(1.072)
Observations	1,562	1,562	1,562	1,562	1,562	1,562	1,562	1,562	1,562
R-squared	0.367	0.373	0.379	0.384	0.358	0.362	0.372	0.359	0.376
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ETHIOPIA									
	PPVT age 5				**PPVT age 8**				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Expq5	0.51***	0.47**	0.51***	0.47**	0.41***	0.37***	0.40***	0.41***	0.37***
	(0.168)	(0.164)	(0.169)	(0.165)	(0.092)	(0.087)	(0.093)	(0.086)	(0.080)
Preschool		0.34*		0.34*		0.34***			0.31**
		(0.166)		(0.168)		(0.101)			(0.106)
HAZ age 5			-0.00	-0.00			-0.00**		-0.00**
			(0.001)	(0.001)			(0.001)		(0.001)
In school								0.34***	0.31***
								(0.083)	(0.085)
Constant	0.15	0.27	0.59	0.73	-0.97	-0.78	0.47	-0.93	0.65
	(0.804)	(0.822)	(0.986)	(0.959)	(1.233)	(1.209)	(1.447)	(1.259)	(1.379)
Observations	623	623	623	623	623	623	623	623	623
R-squared	0.339	0.347	0.340	0.348	0.493	0.502	0.494	0.505	0.513
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
INDIA									
	PPVT age 5				**PPVT age 8**				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Expq5	0.34**	0.34**	0.30*	0.30*	0.45***	0.45***	0.41***	0.44***	0.41***
	(0.157)	(0.155)	(0.161)	(0.160)	(0.116)	(0.116)	(0.108)	(0.116)	(0.108)
Preschool		0.09		0.09		0.01			0.01
		(0.102)		(0.100)		(0.088)			(0.087)
HAZ age 5			0.12***	0.12***			0.12***		0.12***
			(0.033)	(0.033)			(0.023)		(0.023)
In school								0.26**	0.19
								(0.119)	(0.120)
Constant	0.61	0.56	1.48*	1.44*	-0.42	-0.43	0.38	-0.62	0.22
	(0.693)	(0.706)	(0.814)	(0.821)	(1.136)	(1.136)	(1.067)	(1.122)	(1.050)
Observations	1,191	1,191	1,191	1,191	1,191	1,191	1,191	1,191	1,191
R-squared	0.133	0.134	0.145	0.146	0.202	0.202	0.216	0.203	0.216
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 2. OLS Regressions, Four YL Countries, Majority Language Sample (continued)

VIETNAM									
	PPVT age 5				PPVT age 8				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Expq5	0.53***	0.52***	0.42***	0.41***	0.40***	0.39***	0.32***	0.37***	0.30***
	(0.107)	(0.106)	(0.117)	(0.117)	(0.112)	(0.111)	(0.103)	(0.106)	(0.099)
Preschool		0.00		0.00		0.00			0.00
		(0.003)		(0.003)		(0.003)			(0.003)
HAZ age 5			0.18***	0.18***			0.12***		0.12***
			(0.050)	(0.050)			(0.034)		(0.034)
In school								0.92***	0.89***
								(0.184)	(0.182)
Constant	-0.53	-0.53	1.16	1.15	0.22	0.23	1.37	-0.48	0.66
	(0.949)	(0.950)	(0.737)	(0.741)	(1.244)	(1.245)	(1.160)	(1.181)	(1.121)
Observations	1,149	1,149	1,149	1,149	1,149	1,149	1,149	1,149	1,149
R-squared	0.182	0.182	0.203	0.204	0.141	0.142	0.152	0.149	0.159
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: OLS regressions. Sex, age, birth-order, region, urban dummy, household size, expenditure quintiles, education of the mother and father, and mother's height are included as controls in all columns.
Robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; * significant at 1%

Table 3. OLS Regressions, Four YL Countries, Full Sample

	PPVT age 5				PPVT age 8				
PERU									
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Expq5	0.65***	0.63***	0.61***	0.59***	0.57***	0.55***	0.52***	0.56***	0.50***
	(0.081)	(0.085)	(0.081)	(0.084)	(0.094)	(0.096)	(0.089)	(0.097)	(0.092)
No majority language	-0.32**	-0.32**	-0.33**	-0.32**	-0.32***	-0.32***	-0.31***	-0.32***	-0.31***
	(0.135)	(0.113)	(0.137)	(0.116)	(0.097)	(0.109)	(0.097)	(0.099)	(0.108)
Preschool		0.26***		0.23***		0.21**			0.18*
		(0.058)		(0.059)		(0.092)			(0.094)
HAZ age 5			0.10***	0.10***			0.10***		0.10***
			(0.027)	(0.028)			(0.026)		(0.027)
In school								0.40*	0.34
								(0.212)	(0.229)
Constant	2.67***	2.71***	2.98***	2.99***	-1.70**	-1.51**	-1.43**	-2.08***	-1.61**
	(0.381)	(0.374)	(0.352)	(0.347)	(0.620)	(0.601)	(0.588)	(0.661)	(0.622)
Observations	1,745	1,745	1,745	1,745	1,745	1,745	1,745	1,745	1,745
R-squared	0.358	0.365	0.367	0.373	0.408	0.413	0.417	0.409	0.422
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ETHIOPIA									
	PPVT age 5				**PPVT age 8**				
Expq5	0.47***	0.42***	0.47***	0.42***	0.48***	0.42***	0.47***	0.42***	0.37***
	(0.108)	(0.098)	(0.108)	(0.098)	(0.126)	(0.115)	(0.125)	(0.113)	(0.098)
No majority language	-0.36***	-0.17*	-0.36***	-0.17*	-0.85***	-0.62***	-0.85***	-0.83***	-0.63***
	(0.089)	(0.094)	(0.089)	(0.095)	(0.235)	(0.193)	(0.235)	(0.231)	(0.191)
Preschool		0.41***		0.41***		0.49***			0.45***
		(0.115)		(0.115)		(0.105)			(0.113)
HAZ age 5			-0.00	-0.00			-0.00		-0.00
			(0.000)	(0.000)			(0.001)		(0.001)
In school								0.38***	0.36***
								(0.110)	(0.113)
Constant	0.33	0.13	0.40	0.18	1.74***	1.55***	2.49***	1.79***	2.43***
	(0.327)	(0.324)	(0.448)	(0.436)	(0.517)	(0.494)	(0.756)	(0.554)	(0.716)
Observations	1,710	1,710	1,710	1,710	1,710	1,710	1,710	1,710	1,710
R-squared	0.245	0.257	0.245	0.257	0.355	0.372	0.356	0.378	0.394
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
INDIA									
	PPVT age 5				**PPVT age 8**				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Expq5	0.40***	0.40***	0.35**	0.35**	0.49***	0.49***	0.45***	0.48***	0.45***
	(0.135)	(0.134)	(0.136)	(0.135)	(0.124)	(0.123)	(0.121)	(0.125)	(0.121)
No majority language	-0.35***	0.35***	-0.31***	0.31***	-0.06	-0.06	-0.10	-0.06	-0.10
	(0.085)	(0.083)	(0.076)	(0.074)	(0.116)	(0.114)	(0.119)	(0.117)	(0.117)
Preschool		0.03		0.02		0.01			0.01
		(0.101)		(0.102)		(0.077)			(0.076)
HAZ age 5			0.13***	0.13***			0.11***		0.11***
			(0.032)	(0.032)			(0.022)		(0.023)
In school								0.24*	0.17
								(0.123)	(0.122)
Constant	-0.14	-0.17	0.03	0.01	-0.01	-0.02	-0.00	-0.22	-0.16
	(0.470)	(0.415)	(0.502)	(0.453)	(0.702)	(0.718)	(0.671)	(0.695)	(0.672)
Observations	1,454	1,454	1,454	1,454	1,454	1,454	1,454	1,454	1,454
R-squared	0.181	0.181	0.195	0.195	0.189	0.189	0.201	0.189	0.201
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 3. OLS Regressions, Four YL Countries, Full Sample (continued)

VIETNAM									
	PPVT age 5				PPVT age 8				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Expq5	0.53***	0.53***	0.39***	0.39***	0.41***	0.41***	0.31***	0.38***	0.29***
	(0.104)	(0.104)	(0.110)	(0.111)	(0.105)	(0.105)	(0.093)	(0.099)	(0.089)
No majority language	-0.69***	-0.68***	-0.61***	-0.60***	-1.05***	-1.04***	-0.99***	-0.69***	-0.64***
	(0.144)	(0.145)	(0.127)	(0.127)	(0.251)	(0.251)	(0.275)	(0.141)	(0.164)
Preschool		0.00		0.00		0.00			0.00
		(0.003)		(0.003)		(0.003)			(0.002)
HAZ age 5			0.18***	0.18***			0.12***		0.12***
			(0.052)	(0.052)			(0.034)		(0.034)
In school								0.97***	0.95***
								(0.165)	(0.157)
Constant	1.03**	1.04**	1.31***	1.32***	2.20***	2.22***	2.39***	1.32*	1.53**
	(0.392)	(0.395)	(0.367)	(0.371)	(0.660)	(0.668)	(0.659)	(0.684)	(0.685)
Observations	1,180	1,180	1,180	1,180	1,180	1,180	1,180	1,180	1,180
R-squared	0.199	0.200	0.223	0.224	0.165	0.165	0.177	0.177	0.189
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; * significant at 1%

3.2 The SES Gradient for Changes in PPVT Scores across Rounds

In this section, a *value-added* specification is estimated, taking advantage of the longitudinal data. Given my interest in the pattern of the PPVT over time and SES, the coefficient of interest is now that of the lagged PPVT, and again the fifth household expenditure quintile.

The first columns of each country in Table 4 show the same specification as in column (9) of Table 2. The second column of each country shows value added regression of the PPVT at age 8 on the PPVT at age 5 with child and household controls. Here the two coefficients of interest are positive, as expected. The coefficient on lagged PPVT is largest for Peru (0.42) and Vietnam (0.25). This means that an increase of the PPVT at age 5 of one SD will be associated with an increase in PPVT at age 8 of 0.42 SD in Peru and of 0.25 SD in Vietnam, 0.20 SD in India, and 0.13 SD in Ethiopia. These coefficients are interpreted in the education literature as "persistence" or depreciation rates of human capital. In this sense, Peru has the highest persistence, and policies aimed at increasing performance on vocabulary tests should be more effective (on average) in Peru than elsewhere.

The coefficient on the fifth quintile dummy is still significant, but decreases substantially, and is now largest in India and Ethiopia (0.35 and 0.31 SD, respectively), showing that in terms of *changes* in PPVT, the gradient is more pronounced in these two countries. Peru's coefficient (0.28 SD) is then followed by Vietnam (0.20 SD). This means that being richer increases the chances of moving up in the ranking of the distribution of the PPVT more in India and Ethiopia, although the coefficients are not that different from the Peruvian one. In the third columns for each country, I take out the expenditures quintile dummies and the lagged PPVT coefficient stay nearly the same in the four countries. The latter implies that lagged PPVT has indeed an independent and significant effect on PPVT at age 8, above and beyond the effect of SES.

Another interesting comparison is to interpret the lagged PPVT as a mediator and therefore compare the SES coefficient in columns (1) and (2). Lagged PPVT is indeed an important mediator, as the SES coefficient falls by 22, 6, 6, and 10 percentage points in Peru, Ethiopia, India, and Vietnam, respectively.

In the context of the value-added specifications, a Peruvian child who finds himself below the mean (z-score between -2 and 0) will improve less than children in the other countries, while a child above the mean will see his z-score decrease in relative terms but not as much as in the other three countries. In other words, convergence between groups is going to be (*ceteris paribus*) around twice as slow in Peru as in the other three countries.

However, the coefficient of lagged PPVT should be interpreted with caution and not in a way that reflects a causal relationship because, by construction, it is correlated with fixed unobserved child characteristics. This suggests that the estimated coefficient will conflate the effect of unobserved innate abilities of the child.

Last but not least, I reject the assumption of lagged PPVT being a sufficient statistic for all historical inputs, because if it were true, one would not expect that factors such as lagged HAZ would have a significant association with PPVT at age 8.[13] It seems that in both Peru and India, there are *additional* effects of nutritional status at age 5—a finding that deserves further investigation.

Table 4. Value-Added Regressions, Four YL Countries, Majority Language Sample

	PERU			ETHIOPIA		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
PPVT age 5 (z-score)		0.42***	0.43***		0.13***	0.16***
		(0.039)	(0.038)		(0.040)	(0.041)
Expq5	0.50***	0.28***		0.37***	0.31***	
	(0.106)	(0.088)		(0.080)	(0.084)	
Observations	1,562	1,562	1,562	623	623	623
R-squared	0.376	0.487	0.478	0.513	0.524	0.513
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes
Expenditure	Yes	Yes	No	Yes	Yes	No

	INDIA			VIETNAM		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
PPVT age 5 (z-score)		0.20***	0.20***		0.25***	0.25***
		(0.028)	(0.027)		(0.034)	(0.036)
Expq5	0.41***	0.35***		0.30***	0.20**	
	(0.108)	(0.104)		(0.099)	(0.083)	
Observations	1,191	1,191	1,191	1,149	1,149	1,149
R-squared	0.216	0.249	0.239	0.159	0.210	0.203
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes
Expenditure	Yes	Yes	No	Yes	Yes	No

Note: Robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; * significant at 1%

Table 5 replicates Table 4 for the full sample. The ranking of countries in term of the coefficients remain similar, although the coefficient for past PPVT decreases in Peru, but not in the other countries. The dummy that indicates whether the language used by child during administration was different from the majority language is negative and significant in all specifications, except the India one.

[13] Other determinants of HAZ and PPVT, such as parental investments at birth, age 5, and age 8, are presented in Appendix B as background.

Table 5. Value-Added Regressions, Four YL Countries, Full Sample

	PERU (1)	PERU (2)	PERU (3)	ETHIOPIA (4)	ETHIOPIA (5)	ETHIOPIA (6)
PPVT age 5 (z-score)		0.37***	0.39***		0.16***	0.18***
		(0.039)	(0.039)		(0.026)	(0.026)
Expq5	0.50***	0.31***		0.37***	0.30***	
	(0.092)	(0.076)		(0.098)	(0.094)	
No majority language	-0.31***	-0.45***	-0.43***	-0.63***	-0.60***	-0.59***
	(0.108)	(0.125)	(0.124)	(0.191)	(0.198)	(0.193)
Observations	1,745	1,745	1,745	1,710	1,710	1,710
R-squared	0.422	0.511	0.503	0.394	0.413	0.403
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes
Expenditure	Yes	Yes	No	Yes	Yes	No

	INDIA (1)	INDIA (2)	INDIA (3)	VIETNAM (4)	VIETNAM (5)	VIETNAM (6)
PPVT age 5 (z-score)		0.20***	0.21***		0.24***	0.25***
		(0.022)	(0.020)		(0.032)	(0.034)
Expq5	0.45***	0.38***		0.29***	0.20**	
	(0.121)	(0.118)		(0.089)	(0.075)	
No majority language	-0.10	-0.16	-0.12	-0.64***	-0.55***	-0.56***
	(0.117)	(0.116)	(0.121)	(0.164)	(0.188)	(0.185)
Observations	1,454	1,454	1,454	1,180	1,180	1,180
R-squared	0.201	0.233	0.222	0.189	0.237	0.230
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes
Expenditure	Yes	Yes	No	Yes	Yes	No

Note: Robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; * significant at 1%

4. Discussion of Results and Conclusions

This paper uses unique comparable data from the four YL countries to show that there are important differences in early language development between children in richer and poorer households that persist as children enter the early school years. It seems that differences in income levels and in other measures of well-being that are apparent in adulthood arise early in children's lives. Although gradients in cognitive development arise in all countries, they do arise more starkly in Peru, *vis à vis* the other three countries.

The main contribution of this study is that it is the first multicontinent comparison of SES gradients in cognitive development for young children in the developing world over critical periods of their lives, based on a common outcome measure for four countries. The second contribution is the use of the longitudinal structure of the data to analyze how deficits in receptive language ability observed at young ages evolve as children enter the early school years and how persistent these are.

Conditional analysis show that SES disparities seem to decline over time and that there is no mediator (either preschool, early nutrition, or schooling) that changes the SES cognition gradients—although the magnitude of the gradient decreases when all of them are included. Still, preschool seems an important mediator in Ethiopia, particularly at age 5, while early nutrition seems to be an important mediator in Vietnam.

Peru has both the largest gradient and the highest PPVT persistence in the value added specification. The latter finding is important from the point of view ofpolicy making, as these factors might be preventing a faster catch-up between disadvantaged and better-off children. There is no difference in this persistence between boys and girls, but there are differences between urban and rural households in Peru and Vietnam, and between the rainforest (*Selva*) region and other regions in Peru (see Appendix Tables A2 and A3).

One possible venue that could be investigated (although I am not aware that the data exist) is how much of the higher persistence in Peru is explained by the correlation between SES of the family of a child and the quality of the school quality this child attend. A cross-country study (Heynemann and Loxley, 1983) found that this correlation was 0.25 for Peru and 0.06 for India (with no data for either Vietnam or Ethiopia).[14] If new data of this kind become available, this seems a promising venue to investigate the patterns found in this paper further.

Moreover, the analysis for Peru in relation to other countries (and continents) is particularly timely, as the Humala administration is making Early Childhood Development (ECD) a priority in its social policy agenda, and has already created a new program of integrated ECD services for poor households. The results in this paper would directly inform the debate in Peru and, more generally, in many countries in Latin America that are considering similar policy reforms. They reinforce the importance of programs directed at poor children in developing countries emphasized in a prominent recent survey (Engle et al., 2011), but with much more direct evidence. A number of interventions have been shown to affect the cognitive development of young children in Latin America and the Caribbean, including preschool in Argentina (Berlinski et al., 2009), parenting interventions in Jamaica (Walker et al., 2011), nutritional supplements in Guatemala (Pollitt et al., 1993), and cash transfers in Nicaragua (Macours et al., 2012). Extending the coverage of these and other programs is likely to be an important policy priority in Latin American countries.

[14] According to the YL data, 98 percent of Vietnamese children, 83 percent of Peruvian children, 73 percent of Indian children, and 68 percent of Ethiopian children attend public schools.

References

Alderman, H. 2000. "Anthropometrics." In: Margaret Grosh and Paul Glewe, editors. *Designing HH Survey Questionnaire for Developing Countries: Lessons from 10 Years of LSMS Experience,* 251–72, Oxford: Oxford University Press.

Behrman, J.R., M.C. Calderon, S. Preston, J. Hoddinott, R. Martorell, and A.D. Stein. 2009. "Nutritional Supplementation of Girls Influences the Growth of their Children: Prospective Study in Guatemala." *American Journal of Clinical Nutrition* 90: 1372–79.

Berlinski, S., S. Galiani, and P. Gertler. 2009. "The Effect of Pre-Primary Education on Primary School Performance." *Journal of Public Economics* 93(1–2): 219–34.

Black, S.E., and P.J. Devereux. 2010. *Recent Developments in Intergenerational Mobility*. Cambridge, Mass: National Bureau of Economic Research.

Blanden, J., P. Gregg, and L. Macmillan. 2007. "Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education." *The Economic Journal* 117 (519): C43–C60.

Case, A., and C. Paxson. 2008. "Stature and Status: Height, Ability, and Labor Market Outcomes." *Journal of Political Economy* 116(3):499–532.

Case, A., D. Lubotsky, and C. Paxson. 2002. "Economic Status and Health in Childhood: The Origins of the Gradient." *American Economic Review* 92(5): 1308–34.

Cueto, S., J. Leon, G. Guerrero, and I. Munoz. 2009. *Psychometric Characteristics of Cognitive Development and Achievement Instruments in Round 2 of Young Lives*. Technical Note 15. Oxford: Young Lives.

Currie, J. 2009. "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development." *Journal of Economic Literature,* American Economic Association 47(1): 87–122.

Currie, J., and E. Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence From College Openings." *The Quarterly Journal of Economics,* MIT Press 118(4): 1495–1532.

Currie, J., and D. Thomas. 2000. "School Quality and the Longer-Term Effects of Head Start." *Journal of Human Resources* 35(4): 755–74.

Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Johns Hopkins University Press for the World Bank.

Efron, B., and R.J. Tibshirani. 1993. *An Introduction to the Bootstrap*. New York: Chapman & Hall.

Engle, P.L., L.H. Fernald, H. Alderman, J., Behrman, C. O'Gara, A. Yousafzai, M.C. de Mello, M. Hidrobo, N. Ulkuer I. Ertem, and S. Iltus. 2011. "Strategies for Reducing Inequalities and Improving Developmental Outcomes for Young Children in Low- and Middle-income Countries." *Lancet* 378(9799): 1339–53.

Fan, J., and I. Gijbels. 1996. *Local Polynomial Modeling and its Applications.* London: Chapman and Hall.

Feinstein, L. 2003. "Inequality in the Early Cognitive Development of British Children in the 1970 Cohort." *Economica* 70 (277): 73–97.

Fernald, L.C.H., A. Weber, E. Galasso, and L. Ratsifandrihamanana. 2011. "Socioeconomic Gradients and Child Development in a Very Low-income Population: Evidence from Madagascar." *Developmental Science* 14(4): 832–47.

Fleming, S., M. Thompson, R. Stevens, C. Heneghan, A. Plüddemann, I. Maconochie, L. Tarassenko, and D. Mant. 2011. "Normal Ranges of Heart Rate and Respiratory Rate in Children from Birth to 18 Years of Age: A Systematic Review of Observational Studies." *Lancet* 377(9770): 1011–18.

Gasparini, L., and N. Lustig. 2011. "The Rise and Fall of Income Inequality in Latin America." Working Paper 0118. CEDLAS, Universidad Nacional de La Plata.

Harpham, T. 2002. "Measuring Child Poverty and Health: A New International Study." Young Lives Working Paper Series No. 2.

Heckman, J., S. Moon, R. Pinto, P.A. Sazvelyev, and A. Yavitz. . 2010. "The Rate of Return to the High Scope Perry Preschool Program." *Journal of Public Economics* 94(1–2): 114–28.

Hertz, T., T. Jayasundera, P. Piraino, S. Selcuk, N. Smith, and A. Verashchagina. 2008. "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends." *The B.E. Journal of Economic Analysis & Policy* 7.2.

Heynemann, S., and S. Loxley. 1983. "The Effect of Primary School Quality on Academic Achievement across Twenty-nine High- and Low-income Countries." *American Journal of Sociology* 88(6).

Hoddinott, J., J.A. Maluccio, J.R. Behrman, R. Flores, and R. Martorell. 2008. "Effect of a Nutrition Intervention during Early Childhood on Economic Productivity in Guatemalan Adults." *Lancet* 371(9610): 411–16.

IDB (Inter-American Development Bank). 2011. "Estrategia social para la equidad y productividad: América Latina y el Caribe." Documento de Estrategia, http://www.iadb.org/en/publications/publication-detail,7101.html?id=32906.

Macours, K., N. Schady, and R. Vakis. 2012. "Cash Transfers, Behavioral Changes, and the Cognitive Development of Young Children: Evidence from a Randomized Experiment." *American Economic Journal: Applied Economics 4(2):247-273.*

Maluccio, J.A., J. Hoddinott, J.R. Behrman, J.R., R. Martorelli, A.R. Quisumbing, and A.D. Stein. 2009. "The Impact of Experimental Nutritional Interventions on Education into Adulthood in Rural Guatemala." *Economic Journal* 119(537): 734–63.

Moscicki, A., S. Shiboski, N.K. Hills, K.J. Powell, N. Jay, E.N. Hanson, S. Miller, L. Canjura-Clayton, S. Farhat, J.M. Broering, T.M, Darragh. "Regression of Low-grade Squamous Intraepithelial Lesions in Young Women." *Lancet* 364(9446): 1678–83.

Naudeau, S., S. Martínez, P. Premand, and D. Filmer. 2011. "Cognitive Development among Young Children in Low-income countries." In: H. Alderman, editor. *No Small Matter: The Impact of Poverty, Shocks, and Human Capital Investments in Early Childhood Development*. Washington, DC: World Bank.

Park, Y., D.J. Hunter, D. Spiegelman, L. Bergkvist, F. Berrino, P.A. van den Brandi, J.E. Buring, G.A. Colditz, J.L. Freudenheim, C.S. Fuchs, E. Giovannucii, R.A. Goldbohm, S. Graham, L. Harnack, A.M. Hartman, D.R. Jacobs, Jr., I. Kato, V. Krogh, M.F. Leitzmann, M.L. McCullough, A.B. Miller, P. Pierinen, T.E. Rohan, A. Schatzkin, W.C. Willet, A. Wolk, A. Zeleniuch-Jacquotte, S.M. Zhang, and S.A. Smith-Warner. 2005. "Dietary Fiber Intake and Risk of Colorectal Cancer: A Pooled Analysis of Prospective Cohort Studies." *Journal of the American Medical Association* 294(22): 2849–857.

Paxson, C., and N. Schady. 2007. "Cognitive Development among Young Children in Ecuador: The Roles of Wealth, Health, and Parenting." *Journal of Human Resources* 42(1): 49–84.

Pollitt, E., K.S. Gorman, P. Engle, R. Martorell, and J. Rivera. 1993. "Early Supplementary Feeding and Cognition: Effects over Two Decades." *Monograms in Social Research in Child Development* 58(7): 1–99.

Schady, N. 2011. "Parental Education, Vocabulary, and Cognitive Development in Early Childhood: Longitudinal Evidence from Ecuador." *American Journal of Public Health* 101(12): 2299–307.

Schady, N., J. Behrman, M.C. Araujo, R. Azuero, R. Bernal, D. Bravo, F. Lopez-Boo, K. Macours, D. Marshall, C. Paxson, and R. Vakis. 2011. "Wealth Gradients in Early Childhood Cognitive Development in Five Latin American Countries." Unpublished.

Stein, A.D., M. Wang A. DiGirolamo, R. Grajeda, U. Ramakrishnan, M. Ramirez-Zea, K. Yount, and R. Martorell. 2008. "Nutritional Supplementation in Early Childhood, Schooling, and Intellectual Functioning in Adulthood. A Prospective Study in Guatemala." *Archives of Pediatrics and Adolescent Medicine*, 162(7): 612–18.

Todd, P.E, and K.I. Wolpin. 2007. "The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps." *Journal of Human Capital* 1(1): 91–136.

Victora, C.G., A. Wagstaff, J.A. Schellenberg, D. Swatkin, M. Claeson, and J-P Habicht. 2003. "Applying an Equity Lens to Child Health Mortality: More of the Same Is Not Enough." *Lancet* 362: 233–41.

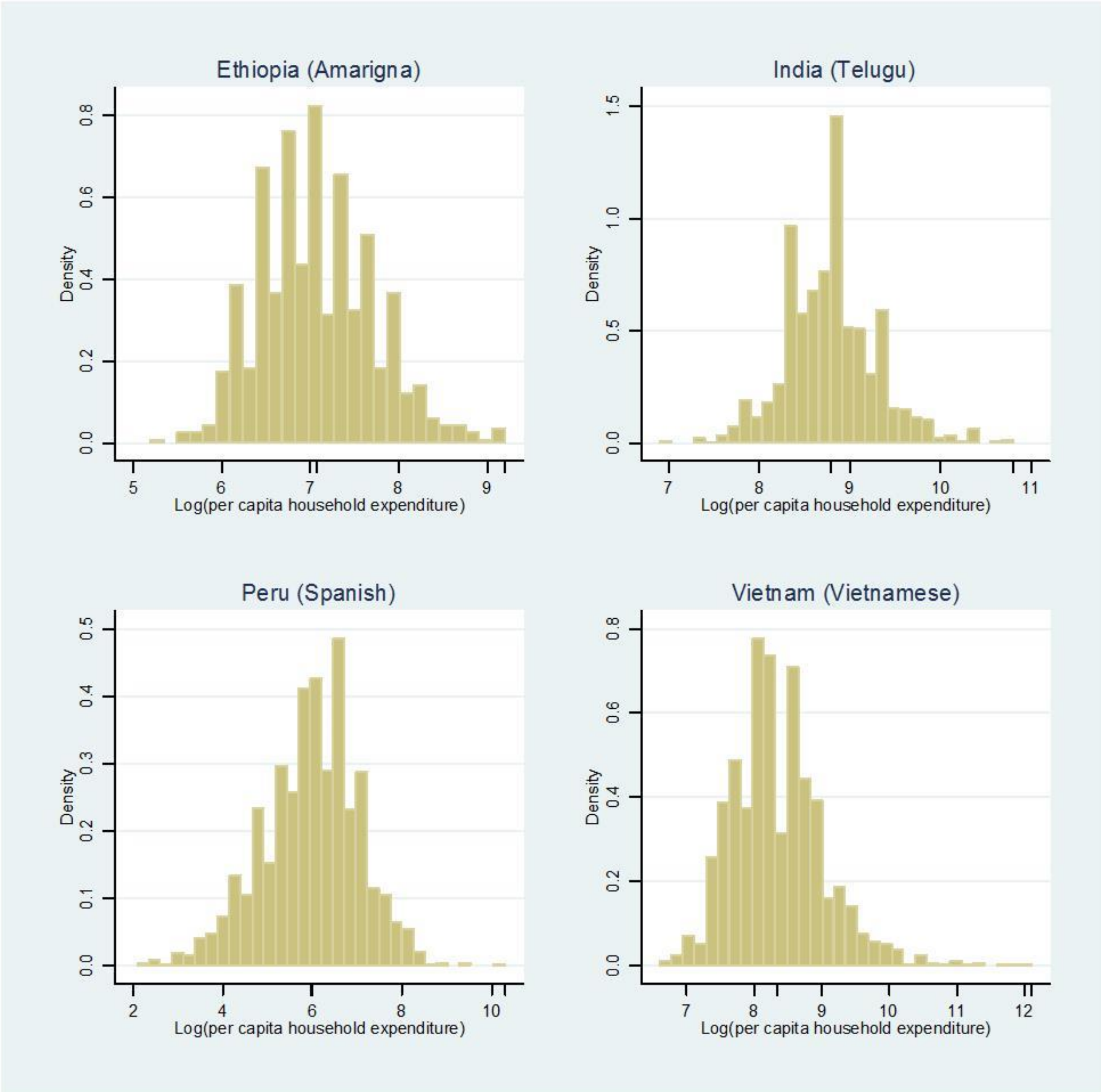
Walker, S.P., S.M. Chang, M. Vera-Hernández, and S. Grantham-McGregor. 2011. "Early Childhood Stimulation with Stunted Children Benefits Adult Competence and Reduces Violent Behavior." *Pediatrics* 127(5): 849–57.

WHO (World Health Organization). 2000. *WHO/NCHS Reference Standards: WHO Working Group: Use and Interpretation of Anthropometric Indicators on Nutritional Status*. Geneva: WHO Press.

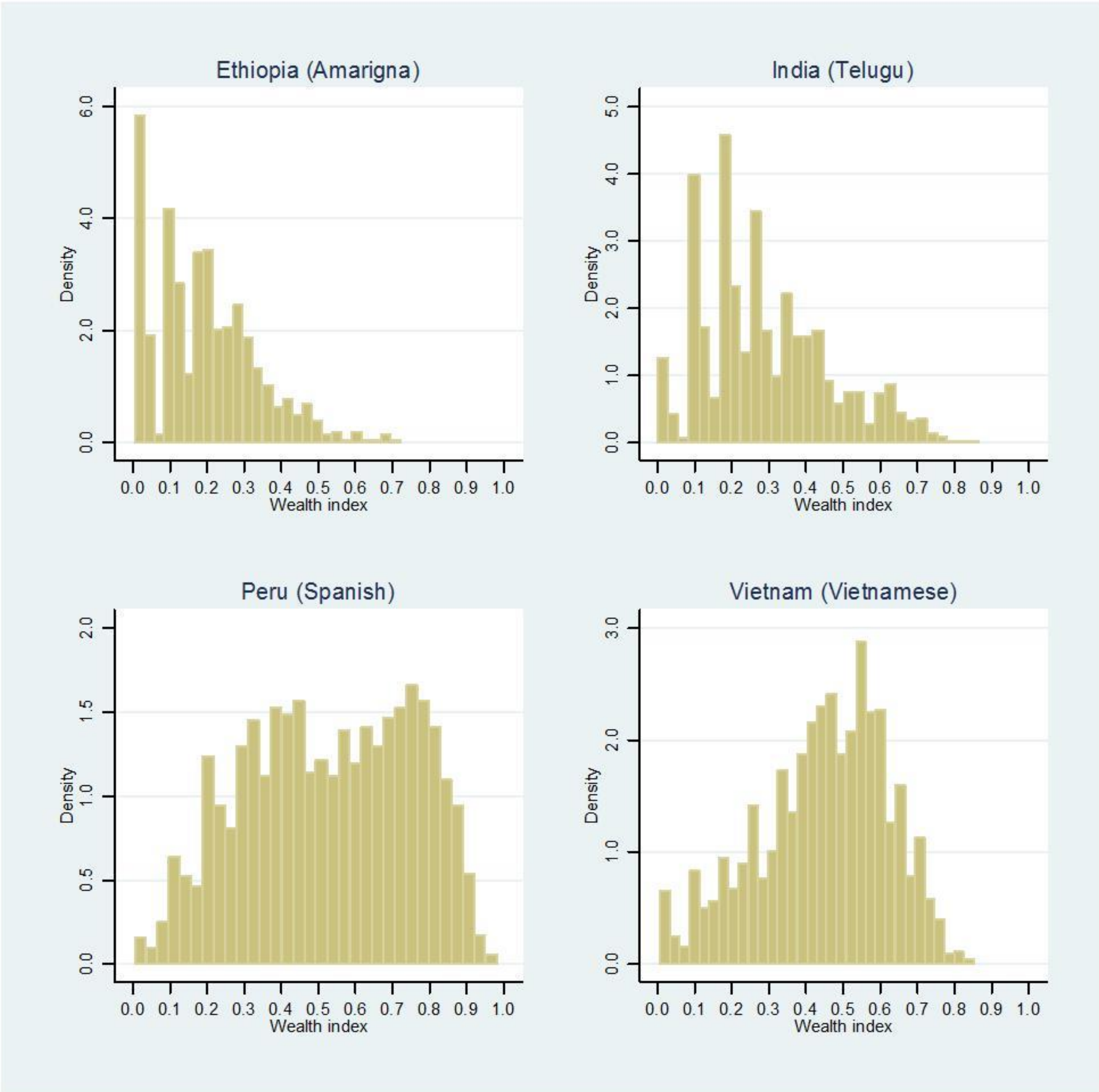
World Bank. 2005. World Development Indicators. World Bank, Washington, DC.

Appendix A. Additional Figures and Tables

Appendix Figure A.1. Distribution of Expenditure in Four YL Countries



Appendix Figure A.2 Distribution of Wealth in Four YL Countries



Appendix Table A1. OLS Regressions, Four YL Countries, Majority Language Sample

	PERU								
	PPVT age 5				PPVT age 8				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female	-0.05	-0.05	-0.04	-0.04	-0.07	-0.07	-0.06	-0.06	-0.05
	(0.043)	(0.042)	(0.041)	(0.040)	(0.044)	(0.044)	(0.042)	(0.045)	(0.043)
Age (in months)	-0.06***	-0.06***	-0.06***	-0.06***	0.01	0.01	0.01	0.01	0.01
	(0.006)	(0.006)	(0.005)	(0.005)	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)
R1firstborn	0.01	0.06	0.04	0.08	0.58***	0.62***	0.61***	0.58***	0.64***
	(0.197)	(0.172)	(0.198)	(0.176)	(0.165)	(0.171)	(0.190)	(0.167)	(0.197)
R1lastborn	0.07*	0.08**	0.07*	0.08**	-0.02	-0.01	-0.01	-0.02	-0.00
	(0.036)	(0.035)	(0.037)	(0.036)	(0.041)	(0.040)	(0.044)	(0.040)	(0.041)
Region 1	0.10	0.05	0.08	0.04	-0.12	-0.17	-0.14	-0.13	-0.19*
	(0.123)	(0.120)	(0.120)	(0.117)	(0.095)	(0.099)	(0.094)	(0.095)	(0.095)
Region 2	0.05	0.00	0.07	0.02	-0.06	-0.10	-0.04	-0.06	-0.08
	(0.088)	(0.089)	(0.087)	(0.088)	(0.077)	(0.085)	(0.077)	(0.077)	(0.084)
Urban	0.44***	0.44***	0.40***	0.41***	0.37***	0.37***	0.33***	0.37***	0.33***
	(0.077)	(0.071)	(0.075)	(0.071)	(0.093)	(0.090)	(0.088)	(0.092)	(0.085)
HHSIZE	-0.04***	-0.04***	-0.04***	-0.04***	-0.03**	-0.03**	-0.02*	-0.03**	-0.02*
	(0.010)	(0.011)	(0.010)	(0.010)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Expq2	0.12	0.12	0.11	0.11	0.34***	0.34***	0.33***	0.34***	0.33***
	(0.086)	(0.087)	(0.089)	(0.090)	(0.077)	(0.077)	(0.078)	(0.078)	(0.078)
Expq3	0.27***	0.26***	0.25***	0.24***	0.33***	0.32***	0.31***	0.33***	0.30***
	(0.081)	(0.084)	(0.080)	(0.083)	(0.088)	(0.088)	(0.086)	(0.090)	(0.088)
Expq4	0.47***	0.45***	0.44***	0.43***	0.52***	0.51***	0.50***	0.52***	0.48***
	(0.080)	(0.085)	(0.082)	(0.086)	(0.088)	(0.091)	(0.084)	(0.090)	(0.088)
Expq5	0.63***	0.62***	0.60***	0.59***	0.54***	0.53***	0.51***	0.54***	0.50***
	(0.099)	(0.103)	(0.099)	(0.102)	(0.105)	(0.108)	(0.101)	(0.107)	(0.106)
Caregiver's edu	0.05***	0.05***	0.05***	0.04***	0.05***	0.05***	0.05***	0.05***	0.05***
	(0.011)	(0.011)	(0.010)	(0.010)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)
Father's edu	0.05***	0.04***	0.04***	0.04***	0.05***	0.05***	0.05***	0.05***	0.05***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)	(0.008)	(0.009)	(0.009)	(0.008)
Mother's height	0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.01*	0.00	-0.01*
	(0.005)	(0.005)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Preschool		0.25***		0.22***		0.21**			0.17*
		(0.062)		(0.063)		(0.088)			(0.088)
HAZ age 5			0.12***	0.12***			0.14***		0.13***
			(0.030)	(0.031)			(0.029)		(0.030)
In school								0.57*	0.52*
								(0.281)	(0.288)
Constant	2.03**	2.11**	3.50***	3.50***	-2.06*	-1.84*	-0.51	-2.64**	-0.91
	(0.847)	(0.852)	(0.961)	(0.978)	(1.059)	(1.052)	(1.050)	(1.080)	(1.072)
Observations	1,562	1,562	1,562	1,562	1,562	1,562	1,562	1,562	1,562
R-squared	0.367	0.373	0.379	0.384	0.358	0.362	0.372	0.359	0.376
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Appendix Table A1. OLS Regressions, Four YL Countries, Majority Language Sample (continued)

ETHIOPIA									
	PPVT age 5				PPVT age 8				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female	-0.06	-0.05	-0.06	-0.05	0.04	0.05	0.04	0.02	0.04
	(0.084)	(0.086)	(0.084)	(0.086)	(0.078)	(0.082)	(0.078)	(0.076)	(0.081)
Age (in months)	0.00	0.00	-0.00	-0.01	-0.00	-0.00	-0.02	-0.00	-0.02
	(0.009)	(0.009)	(0.013)	(0.013)	(0.008)	(0.008)	(0.012)	(0.008)	(0.012)
R1firstborn	1.11***	1.14***	1.13***	1.15***	1.40***	1.43***	1.43***	1.38***	1.43***
	(0.317)	(0.320)	(0.325)	(0.328)	(0.245)	(0.240)	(0.257)	(0.227)	(0.233)
R1lastborn	0.05	0.08	0.05	0.08	0.03	0.06	0.03	0.03	0.06
	(0.065)	(0.069)	(0.066)	(0.070)	(0.082)	(0.085)	(0.084)	(0.082)	(0.088)
Region 1	0.26*	0.20	0.27*	0.21	-0.04	-0.11	-0.04	-0.08	-0.14
	(0.147)	(0.146)	(0.152)	(0.152)	(0.176)	(0.164)	(0.180)	(0.174)	(0.169)
Region 2	-0.51***	-0.43***	-0.51***	-0.43***	-0.34*	-0.26	-0.35*	-0.38*	-0.32*
	(0.063)	(0.070)	(0.063)	(0.071)	(0.188)	(0.156)	(0.188)	(0.187)	(0.155)
Urban	0.26***	0.10	0.25***	0.10	0.95***	0.79***	0.94***	0.89***	0.75***
	(0.065)	(0.121)	(0.068)	(0.125)	(0.149)	(0.145)	(0.146)	(0.148)	(0.148)
HHSIZE	-0.01	-0.01	-0.01	-0.01	-0.04**	-0.03**	-0.04**	-0.03**	-0.03**
	(0.017)	(0.016)	(0.018)	(0.018)	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)
Expq2	-0.02	-0.02	-0.02	-0.02	-0.06	-0.06	-0.05	-0.04	-0.04
	(0.071)	(0.068)	(0.072)	(0.070)	(0.085)	(0.086)	(0.085)	(0.090)	(0.091)
Expq3	-0.11	-0.10	-0.11	-0.10	-0.04	-0.03	-0.04	-0.03	-0.02
	(0.095)	(0.087)	(0.096)	(0.088)	(0.090)	(0.092)	(0.090)	(0.089)	(0.091)
Expq4	0.11	0.09	0.11	0.09	0.10	0.08	0.10	0.12	0.10
	(0.107)	(0.100)	(0.109)	(0.101)	(0.077)	(0.078)	(0.076)	(0.081)	(0.080)
Expq5	0.51***	0.47**	0.51***	0.47**	0.41***	0.37***	0.40***	0.41***	0.37***
	(0.168)	(0.164)	(0.169)	(0.165)	(0.092)	(0.087)	(0.093)	(0.086)	(0.080)
Caregiver's edu	0.01**	0.01**	0.01**	0.01**	0.01*	0.01	0.01*	0.01*	0.01*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Father's edu	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Mother's height	-0.00	-0.00	-0.00	-0.00	0.00	0.00	0.01	0.00	0.00
	(0.003)	(0.003)	(0.003)	(0.003)	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)
Preschool		0.34*		0.34*		0.34***			0.31**
		(0.166)		(0.168)		(0.101)			(0.106)
HAZ age 5			-0.00	-0.00			-0.00**		-0.00**
			(0.001)	(0.001)			(0.001)		(0.001)
In school								0.34***	0.31***
								(0.083)	(0.085)
Constant	0.15	0.27	0.59	0.73	-0.97	-0.78	0.47	-0.93	0.65
	(0.804)	(0.822)	(0.986)	(0.959)	(1.233)	(1.209)	(1.447)	(1.259)	(1.379)
Observations	623	623	623	623	623	623	623	623	623
R-squared	0.339	0.347	0.340	0.348	0.493	0.502	0.494	0.505	0.513
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Appendix Table A1. OLS Regressions, Four YL Countries, Majority Language Sample (continued)

	INDIA								
	PPVT age 5				PPVT age 8				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female	-0.05	-0.05	-0.06	-0.06	-0.24***	-0.24***	-0.25***	-0.23***	-0.25***
	(0.047)	(0.047)	(0.047)	(0.047)	(0.063)	(0.063)	(0.063)	(0.063)	(0.063)
Age (in months)	-0.00	-0.00	-0.00	-0.00	0.00	0.00	0.00	0.00	0.00
	(0.006)	(0.006)	(0.007)	(0.007)	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)
R1firstborn	0.22	0.21	0.23	0.22	-0.23	-0.23	-0.22	-0.23	-0.22
	(0.421)	(0.416)	(0.434)	(0.429)	(0.975)	(0.977)	(0.880)	(0.975)	(0.883)
R1lastborn	-0.05	-0.05	-0.04	-0.04	-0.18***	-0.18***	-0.17***	-0.18***	-0.17***
	(0.067)	(0.066)	(0.066)	(0.066)	(0.044)	(0.044)	(0.042)	(0.044)	(0.042)
Region 1	0.49*	0.49*	0.46*	0.46*	0.21**	0.21**	0.18**	0.21**	0.18**
	(0.266)	(0.266)	(0.263)	(0.263)	(0.084)	(0.083)	(0.077)	(0.083)	(0.077)
Region 2	0.06	0.06	0.03	0.03	0.58***	0.58***	0.55***	0.58***	0.55***
	(0.071)	(0.073)	(0.080)	(0.081)	(0.125)	(0.125)	(0.115)	(0.125)	(0.116)
Urban	0.27	0.27	0.25	0.25	0.42***	0.42***	0.40***	0.42***	0.40***
	(0.224)	(0.223)	(0.228)	(0.227)	(0.092)	(0.091)	(0.084)	(0.092)	(0.083)
HHSIZE	-0.04*	-0.03*	-0.03	-0.03	-0.04***	-0.04***	-0.04**	-0.04***	-0.03**
	(0.019)	(0.018)	(0.019)	(0.019)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Expq2	0.09	0.10	0.08	0.08	0.17	0.17	0.16	0.17	0.16
	(0.106)	(0.107)	(0.105)	(0.106)	(0.118)	(0.118)	(0.118)	(0.118)	(0.118)
Expq3	0.15	0.14	0.14	0.13	0.16*	0.16*	0.15*	0.16*	0.15*
	(0.142)	(0.143)	(0.144)	(0.144)	(0.085)	(0.085)	(0.082)	(0.085)	(0.082)
Expq4	0.21	0.20	0.19	0.18	0.31***	0.31***	0.28**	0.30***	0.28**
	(0.121)	(0.123)	(0.122)	(0.123)	(0.106)	(0.106)	(0.104)	(0.107)	(0.103)
Expq5	0.34**	0.34**	0.30*	0.30*	0.45***	0.45***	0.41***	0.44***	0.41***
	(0.157)	(0.155)	(0.161)	(0.160)	(0.116)	(0.116)	(0.108)	(0.116)	(0.108)
Caregiver's edu	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***	0.01***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Father's edu	0.02**	0.02**	0.02**	0.02**	0.01**	0.01**	0.01**	0.01**	0.01**
	(0.008)	(0.008)	(0.008)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Mother's height	-0.01	-0.01	-0.01**	-0.01**	0.00	0.00	-0.00	0.00	-0.00
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)
Preschool		0.09		0.09		0.01			0.01
		(0.102)		(0.100)		(0.088)			(0.087)
HAZ age 5			0.12***	0.12***			0.12***		0.12***
			(0.033)	(0.033)			(0.023)		(0.023)
In school								0.26**	0.19
								(0.119)	(0.120)
Constant	0.61	0.56	1.48*	1.44*	-0.42	-0.43	0.38	-0.62	0.22
	(0.693)	(0.706)	(0.814)	(0.821)	(1.136)	(1.136)	(1.067)	(1.122)	(1.050)
Observations	1,191	1,191	1,191	1,191	1,191	1,191	1,191	1,191	1,191
R-squared	0.133	0.134	0.145	0.146	0.202	0.202	0.216	0.203	0.216
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Appendix Table A1. OLS Regressions, Four YL Countries, Majority Language Sample (continued)

VIETNAM									
	PPVT age 5				PPVT age 8				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.04	-0.04
	(0.070)	(0.070)	(0.068)	(0.067)	(0.057)	(0.057)	(0.057)	(0.055)	(0.055)
Age (in months)	-0.03***	-0.03***	-0.02***	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***
	(0.007)	(0.007)	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
R1firstborn	0.20	0.20	0.10	0.10	0.13	0.13	0.06	0.11	0.05
	(0.366)	(0.366)	(0.325)	(0.326)	(0.157)	(0.156)	(0.133)	(0.153)	(0.129)
R1lastborn	-0.10	-0.10	-0.07	-0.07	-0.14*	-0.14*	-0.12*	-0.14*	-0.12*
	(0.059)	(0.059)	(0.057)	(0.057)	(0.066)	(0.066)	(0.064)	(0.069)	(0.067)
Region 1	0.11	0.11	0.21	0.21	-0.01	-0.01	0.06	0.01	0.07
	(0.165)	(0.163)	(0.155)	(0.153)	(0.154)	(0.154)	(0.153)	(0.149)	(0.149)
Region 2	0.39	0.39	0.44	0.44	0.52***	0.52***	0.56***	0.53***	0.56***
	(0.282)	(0.282)	(0.282)	(0.282)	(0.140)	(0.140)	(0.146)	(0.139)	(0.145)
Urban	0.74***	0.74***	0.70***	0.70***	0.36*	0.36*	0.34	0.37*	0.34*
	(0.167)	(0.168)	(0.152)	(0.153)	(0.209)	(0.208)	(0.196)	(0.209)	(0.196)
HHSIZE	-0.03	-0.03	-0.02	-0.02	-0.02	-0.02	-0.01	-0.01	-0.01
	(0.028)	(0.028)	(0.026)	(0.026)	(0.021)	(0.021)	(0.019)	(0.020)	(0.018)
Expq2	0.30**	0.30**	0.26**	0.26**	0.11	0.11	0.08	0.09	0.07
	(0.119)	(0.119)	(0.124)	(0.125)	(0.124)	(0.124)	(0.124)	(0.115)	(0.116)
Expq3	0.25*	0.24*	0.20	0.20	0.08	0.08	0.05	0.06	0.03
	(0.122)	(0.122)	(0.125)	(0.125)	(0.130)	(0.130)	(0.125)	(0.127)	(0.121)
Expq4	0.30**	0.30**	0.21	0.22	0.32**	0.32**	0.26**	0.29**	0.24*
	(0.135)	(0.136)	(0.135)	(0.136)	(0.129)	(0.130)	(0.122)	(0.122)	(0.116)
Expq5	0.53***	0.52***	0.42***	0.41***	0.40***	0.39***	0.32***	0.37***	0.30***
	(0.107)	(0.106)	(0.117)	(0.117)	(0.112)	(0.111)	(0.103)	(0.106)	(0.099)
Caregiver's edu	0.02*	0.02*	0.02*	0.02*	0.02	0.02	0.02	0.02	0.02
	(0.012)	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)
Father's edu	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
Mother's height	0.01*	0.01*	0.00	0.00	0.01**	0.01**	0.01	0.01**	0.01
	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.006)	(0.005)
Preschool		0.00		0.00		0.00			0.00
		(0.003)		(0.003)		(0.003)			(0.003)
HAZ age 5			0.18***	0.18***			0.12***		0.12***
			(0.050)	(0.050)			(0.034)		(0.034)
In school								0.92***	0.89***
								(0.184)	(0.182)
Constant	-0.53	-0.53	1.16	1.15	0.22	0.23	1.37	-0.48	0.66
	(0.949)	(0.950)	(0.737)	(0.741)	(1.244)	(1.245)	(1.160)	(1.181)	(1.121)
Observations	1,149	1,149	1,149	1,149	1,149	1,149	1,149	1,149	1,149
R-squared	0.182	0.182	0.203	0.204	0.141	0.142	0.152	0.149	0.159
Child and HH controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: OLS regressions. Sex, age, birth-order, region, urban dummy, household size, expenditure quintiles, education of caregiver, and father and mother's height are included as controls in all columns.
Robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; * significant at 1%

Appendix Table A2. Value-added Regressions: Urban vs. Rural Households

	PERU		ETHIOPIA		INDIA		VIETNAM	
	Urban	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
PPVT age 5 (z-score)	0.49***	0.33***	0.14**	0.13**	0.20**	0.20***	0.13*	0.27***
	(0.043)	(0.048)	(0.046)	(0.045)	(0.082)	(0.027)	(0.058)	(0.039)
Female	0.01	-0.11	0.11	-0.07	-0.48***	-0.17**	-0.01	-0.03
	(0.050)	(0.084)	(0.100)	(0.070)	(0.111)	(0.063)	(0.069)	(0.066)
Age (in months)	0.04***	0.04***	-0.01	-0.02*	-0.06**	0.01	-0.05**	-0.02**
	(0.007)	(0.008)	(0.021)	(0.011)	(0.021)	(0.009)	(0.020)	(0.007)
R1firstborn	0.48*	0.92***	0.77***	1.78***		-0.22	-0.42	0.06
	(0.272)	(0.200)	(0.161)	(0.175)		(0.901)	(0.284)	(0.096)
R1lastborn	-0.06	-0.01	-0.00	0.14	-0.07	-0.16***	-0.25***	-0.06
	(0.049)	(0.085)	(0.123)	(0.097)	(0.129)	(0.044)	(0.078)	(0.071)
Region 1	0.11**	-0.04	-0.17		-0.21	0.13**	0.50	0.01
	(0.045)	(0.135)	(0.163)		(0.202)	(0.062)	(0.370)	(0.146)
Region 2	0.23**	-0.17***	-0.46**	0.04	0.53**	0.59***	0.37	0.42***
	(0.082)	(0.038)	(0.171)	(0.150)	(0.239)	(0.131)	(0.389)	(0.114)
HHSIZE	0.01	-0.03	-0.02*	-0.05*	-0.06	-0.02*	0.04*	-0.02
	(0.009)	(0.019)	(0.012)	(0.022)	(0.056)	(0.012)	(0.019)	(0.019)
Expq2	0.11	0.30***	-0.33*	0.12	0.28	0.14	0.26	0.01
	(0.122)	(0.058)	(0.152)	(0.081)	(0.321)	(0.119)	(0.251)	(0.101)
Expq3	0.04	0.17	-0.16	0.09	0.19	0.12	0.25	-0.02
	(0.101)	(0.125)	(0.136)	(0.086)	(0.241)	(0.100)	(0.158)	(0.106)
Expq4	0.15	0.43***	0.01	0.13	0.12	0.26**	0.35**	0.20*
	(0.124)	(0.124)	(0.056)	(0.175)	(0.250)	(0.097)	(0.121)	(0.095)
Expq5	0.09	0.33	0.24*	0.15	0.20	0.38***	0.33	0.21**
	(0.132)	(0.241)	(0.100)	(0.205)	(0.218)	(0.122)	(0.250)	(0.090)
Caregiver's edu	0.01	0.04***	0.01	0.00	0.05**	0.01	0.04	0.01
	(0.011)	(0.008)	(0.004)	(0.004)	(0.017)	(0.004)	(0.029)	(0.008)
Father's edu	0.01	0.05***	0.00	0.00	0.04**	0.01	-0.00	0.00
	(0.010)	(0.010)	(0.003)	(0.003)	(0.015)	(0.004)	(0.004)	(0.004)
HAZ age 5	0.06*	0.12**	-0.00	-0.00**	0.05	0.11***	0.12	0.06*
	(0.031)	(0.042)	(0.001)	(0.001)	(0.074)	(0.028)	(0.074)	(0.032)
Mother's height	-0.00	-0.01	0.01	-0.01	0.02	-0.01	-0.01	0.01
	(0.004)	(0.007)	(0.009)	(0.005)	(0.015)	(0.004)	(0.004)	(0.005)
In school	1.06***	-0.01	0.32	0.36***	0.52**	0.06		0.69***
	(0.191)	(0.165)	(0.210)	(0.074)	(0.195)	(0.074)		(0.162)
Constant	-4.26***	-2.48	-0.35	2.25	2.07	-0.32	4.99**	-0.47
	(1.039)	(1.501)	(1.839)	(1.734)	(2.760)	(1.117)	(1.890)	(1.195)
Observations	957	605	364	259	203	988	253	896
R-squared	0.441	0.309	0.175	0.194	0.279	0.221	0.167	0.204

Note: Robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; * significant at 1%

Appendix Table A3. Value-added Regressions: Peru, by Region

	Coast	Mountain	Rainforest
PPVT age 5 (z-score)	0.47***	0.43***	0.27***
	(0.040)	(0.051)	(0.043)
Female	0.07*	-0.16**	-0.01
	(0.037)	(0.064)	(0.114)
Age (in months)	0.03***	0.05***	0.03
	(0.006)	(0.007)	(0.016)
R1firstborn	0.05	0.87***	0.82***
	(0.104)	(0.140)	(0.170)
R1lastborn	0.05	-0.12*	-0.08
	(0.045)	(0.062)	(0.091)
Urban	0.10	0.24**	0.14*
	(0.102)	(0.106)	(0.072)
HHSIZE	-0.00	-0.01	-0.03
	(0.009)	(0.019)	(0.025)
Expq2	0.19**	0.36***	0.04
	(0.084)	(0.086)	(0.099)
Expq3	0.16**	0.24*	-0.03
	(0.066)	(0.134)	(0.147)
Expq4	0.30***	0.34**	0.19***
	(0.078)	(0.149)	(0.052)
Expq5	0.23**	0.30	-0.03
	(0.096)	(0.176)	(0.078)
Caregiver's edu	0.01	0.04***	0.03
	(0.013)	(0.010)	(0.015)
Father's edu	0.02	0.04***	0.03
	(0.009)	(0.012)	(0.020)
HAZ age 5	0.08**	0.08**	0.14**
	(0.039)	(0.032)	(0.044)
Mother's height	-0.01	-0.00	-0.02
	(0.005)	(0.007)	(0.012)
In school	1.06***	0.70*	0.15
	(0.111)	(0.347)	(0.210)
Constant	-3.17***	-5.93***	0.02
	(1.030)	(1.278)	(1.349)
Observations	632	687	243
R-squared	0.447	0.508	0.375

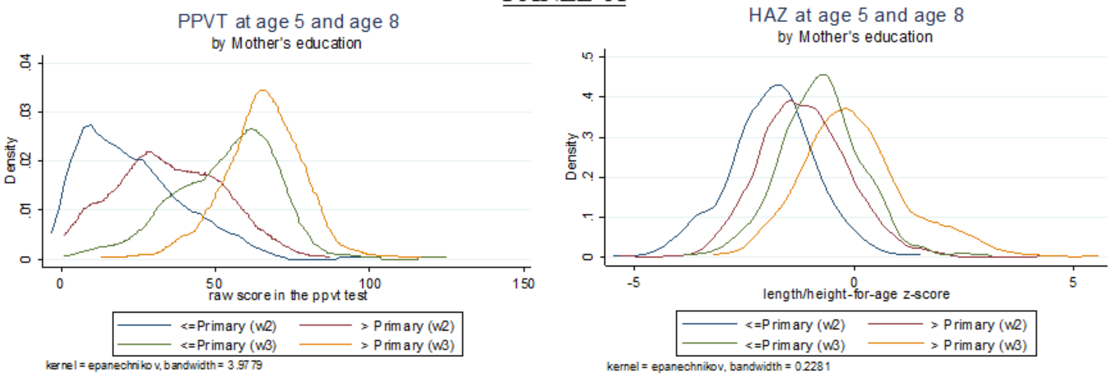
Note: Robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; * significant at 1%

Appendix B. A Further Look at Peru's Inequalities

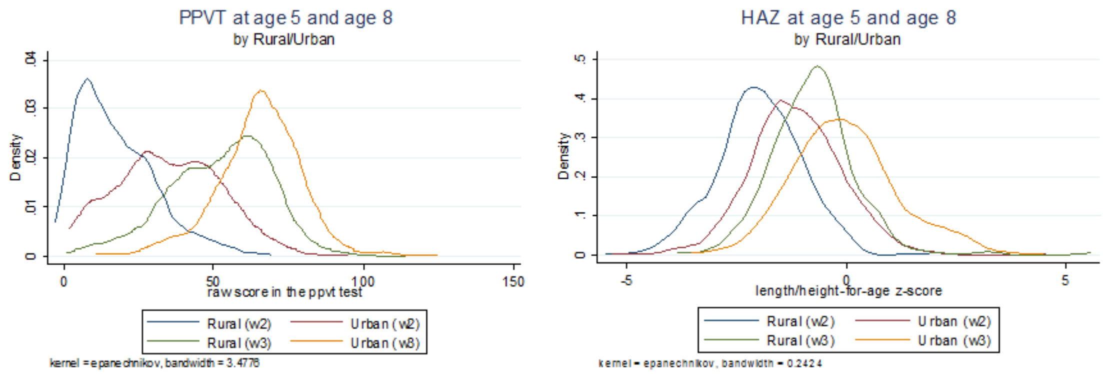
Having noticed that PPVT inequalities are larger in Peru, in the figures that follow, I show distributions for children at both age 5 and 8 evaluated at different proxies for household's SES, to see if the previous result still hold. These proxies are maternal education and rural versus urban status (Appendix Figure B1). I also look at these distributions by preschool enrollment status and birth weight (Appendix Figure B2), and by shocks *in utero* (Appendix Figure B2), which have been shown to influence our variable of interest, the PPVT.[15]

Appendix Figure B1. PPVT and HAZ at Age 5 and 8, by Mother's Education and Rural-Urban Area

PANEL A



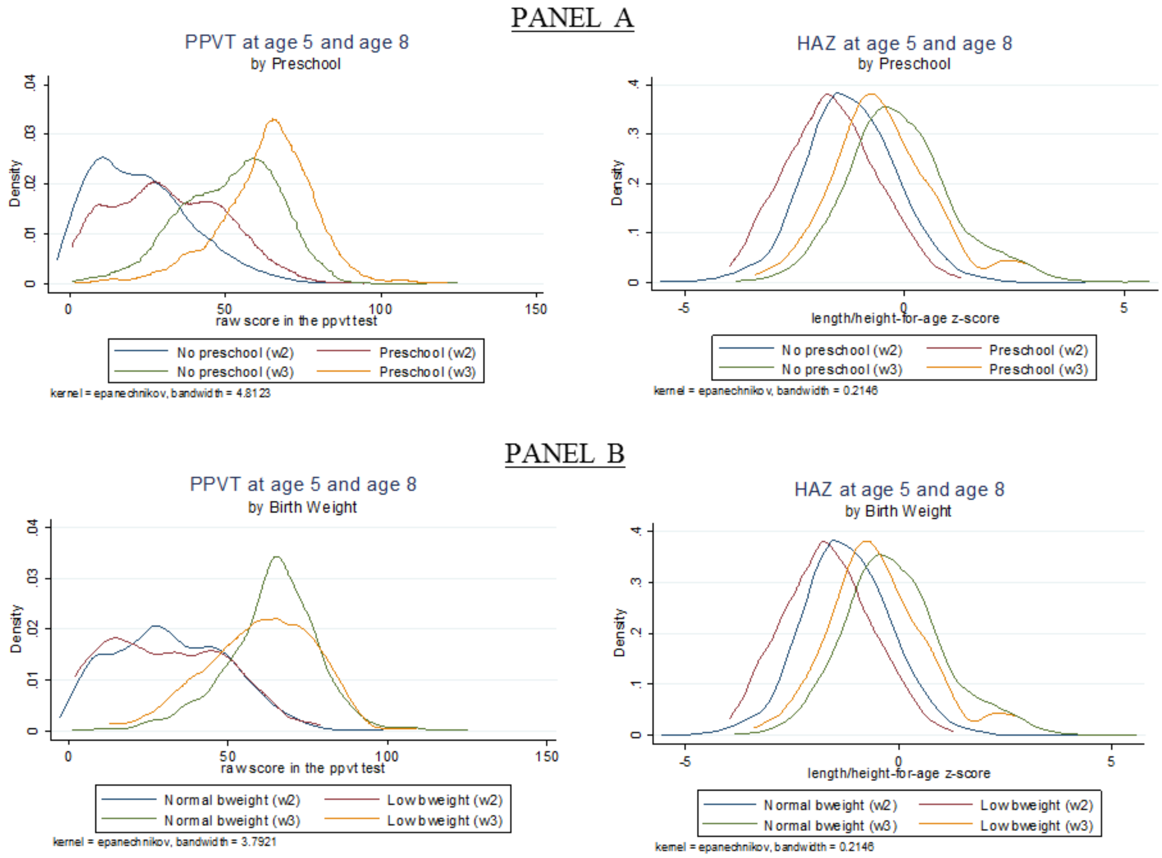
PANEL B



.

[15] Some argue that finding that birth weight matters for later outcomes could simply reflect the fact that it is correlated with some other component of family background that is the true causal factor in determining later outcomes. For example, "bad" parents may have children with lower birth weight, and this may be the factor that actually influences the future status of children with poor health in their early lives. Similarly, children with low birth weight may be more likely to attend lower quality schools or live in poorer neighborhoods. If I do not take in account of the potential biases due to selection on unobservables, the OLS estimates for either the PPVT score or HAZ would be biased and the direction of the potential bias is not clear.

Appendix Figure B2. PPVT and HAZ at Age 5 and 8, by Preschool Enrollment and Birth Weight



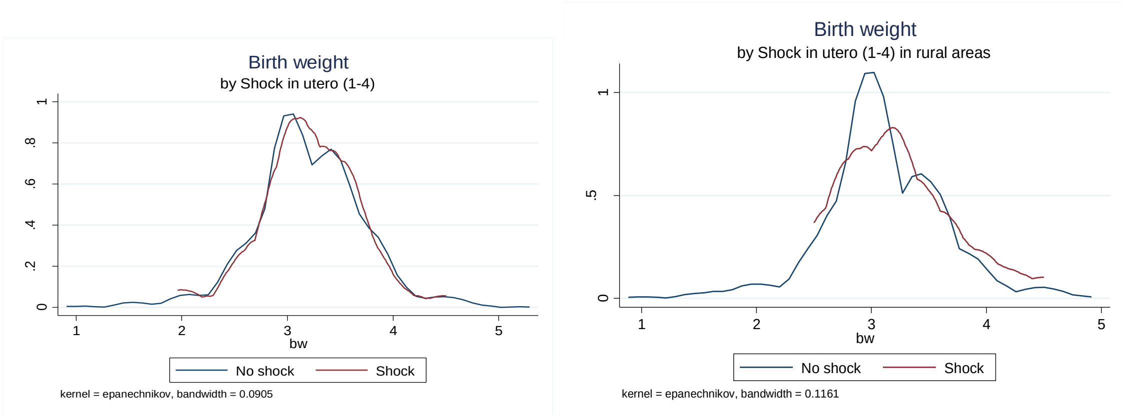
In Appendix Figure B1, I compare children whose mothers have no education or only primary education and those with higher education (Panel A) and children living in rural or urban area (Panel B). There are clear socioeconomic gradients in cognitive skills and nutritional status, using both maternal education and urban/rural as proxies of SES.

In Appendix Figure B2, I look at children who have been enrolled in preschool programs since age 3 and those who are not (Panel A). Finally, I compare children with poor initial health endowment measured by low birth weight with children with a birth weight higher than 2.500 grams (Panel B). Preschool seems positively associated with PPVT score at both ages, while low-birth-weight children have lower HAZ at both age 5 and 8 and slightly lower PPVT scores.

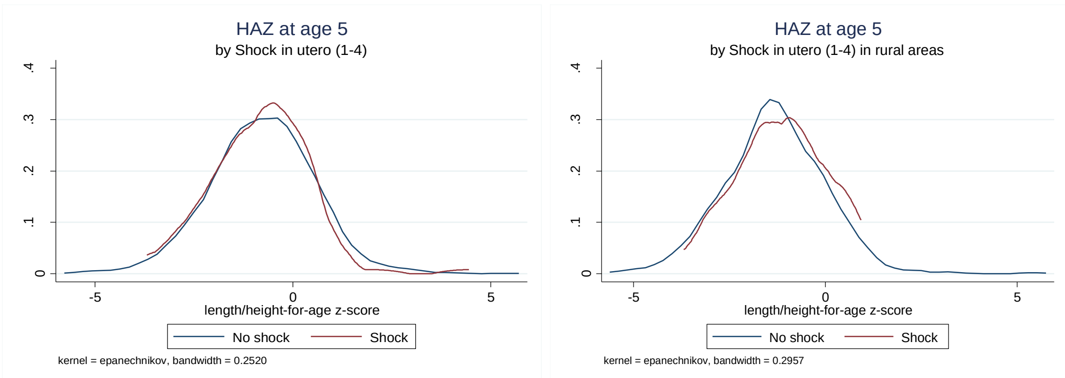
The definition of shock *in utero* is such as to identify those children who suffer an exogenous shock while in the womb. These are children whose mothers suffered a natural disaster, a decrease or change in food availability, or suffered death of livestock or crop failure.

Appendix Figure B3. Shock in Utero and Birth Weight and HAZ at Age 5 and 8

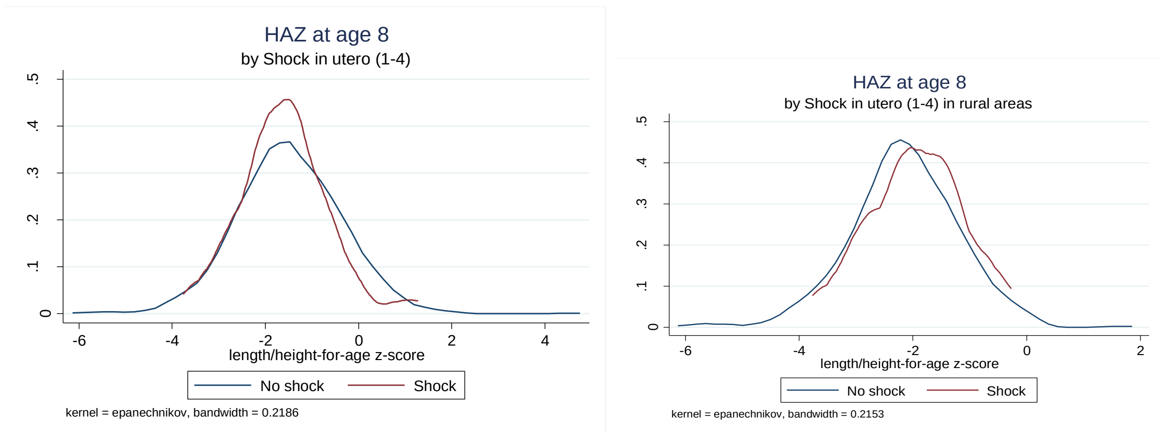
a. Birth weight



b. On HAZ at age 5



c. On HAZ at age 8



Appendix Figure B3 shows that these relationships are not significant. A negative shock *in utero* seems negatively correlated to birth weight and HAZ at age 5 only for those children living in rural/poorer areas. Negative shocks *in utero* completely disappear and do not seem to affect AZ at age 8, especially in rural areas.