

**Ethnic Group Disparities in Academic Skill Development across
Four Low-Income Countries**

Jessica Heckert¹

Graduate Student, Dual-degree Program in Human Development and Family Studies and
Demography
Pennsylvania State University
E-mail: jmh635@psu.edu

Rukmalie Jayakody

Associate Director, Population Research Institute
Associate Professor, Department of Human Development and Family Studies
Pennsylvania State University
E-mail: jayakody@pop.psu.edu

¹ The author wishes to acknowledge support by the Eunice Kennedy Shriver National Institute of Child Health and Human Development, Family Demography Training Grant (No. T-32HD007514) to the Pennsylvania State University Population Research Institute.

Abstract

Many countries experienced dramatic economic growth in recent decades, but within country growth was unequally distributed, and disparities persist by race, ethnicity, and social class.

Although education should promote equity across social groups, this idealization falls short when educational resources are unequally distributed. Instead, educational disparities perpetuate within country inequality. We examine disparities in academic skill development by ethnicity in

Ethiopia and Vietnam, caste in India, and race in Peru using data from the Young Lives Study, a longitudinal study of childhood poverty. Three mechanisms may help explain this achievement gap. First, marginalized families have fewer economic and educational resources. Second, they are geographically concentrated in urban slums or inaccessible rural regions. Finally, they often prefer traditional languages not used in the educational system. We will decompose academic skill indicators to highlight the particular factors contributing to observed disparities.

Understanding the role of these factors will elucidate policy intervention points.

Introduction

Due to dramatic economic expansions in recent decades, many countries experienced poverty declines and improved living standards (Ranis, Stewart, Ramirez, 2000). However, within-country growth was unequally distributed, and disparities persist by race, ethnicity, social class, and location (Cornia, 2003). Accurately evaluating social improvements requires that researchers compare progress across social groups to determine if benefits of social development are equally distributed across social groups.

Academic skill acquisition, a primary developmental task of middle childhood, is a critical component of future productivity and life circumstances (Psacharopoulos & Patrinos, 2004). Although education should promote equity across social groups, this idealization falls short when resources promoting education are unequally distributed. Social group membership imposes hierarchies where particular groups are undervalued, while others maintain powerful influences, and current educational practices perpetuate these hierarchies, ensuring future inequality (Bechmann & Hannum, 2001). This study examines academic skill acquisition during middle childhood in four low-income countries: Ethiopia, Peru, Vietnam, and India's Andhra Pradesh state. Specifically, it examines the aspects of social group membership that perpetuate academic skill disparities.

The four countries examined in this study experienced unequal economic and social development, but the specific social stratification mechanisms differ by country. Ethiopia and Vietnam exemplify social group divisions by ethnicity, India by caste, and Peru by race. Across these four settings, children born into marginalized social groups encounter unequal opportunities and resources, particularly by education investments. Discriminatory practices in

schools and the labor market further reinforce their marginalization. These phenomena sustain poverty's persistence across generations.

Following the *Declaration of the Rights of the Child* in 1959, the world witnessed large improvements in childhood conditions. Child mortality rates declined (Ahmad, Lopez, & Inoue, 2000); universal primary education expanded (Burnett, 2008); and severe forms of child labor drew attention (Psarcharopoulos, 1997). There is a global understanding that national economic well-being necessitates childhood investments. The international community recognizes universal primary education access as a necessity, and many now realize that enrollment indicators alone are insufficient for understanding academic achievement disparities (Ben-Arieh, Kaufman, Andrews, Goerge, Lee, & Aber, 2001; Burnett 2008). Aggregate enrollment indicators mask between group disparities and cannot confirm that school based learning occurs. For example, we observe in this study's sample that over 95% of the children, even in the most marginalized groups, have enrolled in school at some point. However, maintained school enrollment and educational quality remain a concern.

Given differential investment in education along social group lines, all enrolled children fail to experience equitable educational gains. Implementing commitments to primary education requires ensuring that basic educational skills are acquired during primary school and that marginalized social groups share equitable gains. Therefore the current study aims to understand educational disparities and their modifiable components.

The following literature review describes educational disparities and the social factors that perpetuate them. It also depicts how social group divisions occur across the four national settings included in this study.

Marginalized Groups and Academic Achievement

Educational disparities by social group exist across diverse contexts. Directly relevant to the Peruvian setting, indigenous status across Latin America predicts late school entry and grade repetition (Patrinos & Psacharopoulos, 1996), and although differences in completed schooling years between indigenous and non-indigenous Peruvians shrank between 1994 and 2004, a 2.3 year disparity persists (Hall & Patrinos, 2005). Using other indicators, achievement test scores in both Bolivia and Chile demonstrate that indigenous children in the same grade fall behind their non-indigenous peers (McEwan, 2004).

Similarly, ethnic minority status in Vietnam reveals substantial educational disparities. While only 3 % of majority households contain no literate members, the percentage rises to 12% of ethnic minority households. Additionally, while majority households live an average 0.2 kilometers away from the closest lower secondary school and 6.1 kilometers from an upper secondary schools, minority households live 2.4 and 10.3 kilometers away (van de Walle & Gunewardena, 2001). Net primary school enrollment rates for ethnic minorities are a sixth below the national average and net lower secondary school rates are half the national average (Baulch, Chuyen, Haughton, & Haughton, 2002).

Caste differences from India reveal that those from lower castes report fewer completed years of education, even in Kerala considered one of India's more egalitarian states (Deshpande, 2000). National statistics reveal that among higher castes 58% were literate, but among lower castes and tribal groups, only 37% and 30% were literate (Dreze & Loh, 1995). Despite engaging multiple strategies, educational reform for lower castes were largely insufficient, and they remain at notably lower educational attainment rates across indicators (Chanana, 1993; Rao, Cheng, and Narain, 2003).

Turning to Ethiopia, where history is beset with ethnic inequality, previous research has not extensively examined educational inequality. However, evidence of ethnic disparities in health and economic indicators across Africa, (Brockerhoff & Hewitt, 2000; Charasse-Pouélé & Fournier, 2006) suggest that ethnic disparities in educational outcomes are likely.

McEwan (2004) suggests three primary domains explain educational achievement gaps. First, parents often have fewer economic and educational endowments. Second, geographic concentration limits resources and power. Third, traditional languages are undervalued and underused in current educational systems. Additionally, children in poorer families work more, which may reduce school investments (Hall & Patrinos, 2005) and experience poor cognitive development via early childhood nutritional deficiencies (Grantham-McGregor, Cheung, Cueto, Glewwe, Richter, Strupp, 2007). Furthermore there is concern that minority families expect fewer returns to educational investments or undervalue its usefulness (van de Walle & Gunewardena, 2001). The following section elaborates on the processes occurring within these domains and provides examples from different national settings.

Household Economic and Educational Endowments

One explanation for educational disparities is simply that marginalized social groups shoulder a disproportionate poverty burden, and children from poor families acquire less education because their families cannot afford the direct and indirect costs. Poverty rates are higher in Peruvian indigenous households, ethnic minority groups in Vietnam, lower castes in India, and among more marginalized ethnic groups in Ethiopia (Trivelli, 2005; Baulch, Chuyen, Haughton, & Haughton, 2002). Direct schooling costs, including books, materials, uniforms, transportation, and school meals strain resource limited families. There is a strong a positive

association between household income and child educational attainment, even in Vietnam, where the government has emphasized educational equality and progressive school fees, but fee reduction effects are limited due to the prohibitive cost of uniforms, books, school supplies, and transportation (Behrman & Knowles, 1999).

Deeply intertwined in household resources are the number of children and parental education. Higher fertility may mean that limited family resources are spread across more offspring. The fertility rate among ethnic minorities in Vietnam is 25% higher than the majority Kinh (Baulch, Chuyen, Haughton & Haughton, 2002), and high fertility persists among Latin American Amazonian groups, despite regional fertility decreases (McSweeney & Arps, 2005). Evidence finds that Vietnamese household with more children have lower child educational attainment, but effects are strong only for secondary and tertiary education and for relatively large families containing four or more children (Anh, Knodel, Lam, & Friedman, 1998).

In addition to limited economic resources, poor families often have limited parental education. Higher grade repetition occurs among children in families with lower parental education, and teachers may discriminate against students with uneducated parents (Patrinos & Psacharopoulos, 1996). Less education may mean parents experience difficulty negotiating the educational system or may find school does not meet intended family goals. For example, female education may have limited value if a daughter is groomed primarily for marriage (Chanana, 1993). Moreover, group membership is associated with lower educational returns and labor market discrimination, suggesting that parents may correctly perceive education as less valuable (Trivelli, 2005; Zoninsein, 2001; van de Walle & Gunewardena, 2001).

Beyond educational costs, families may also suffer opportunity costs. Families will no longer benefit from child labor contributions when school attendance replaces child production

in both labor market participation and family centered production such as farm work and household chores. Indigenous children throughout Latin America work at higher rates, and children engaged in economic activities acquire fewer academic skills for the amount of time spent in school (Hall & Patrinos, 2005; Psacharopoulos, 1997). However, evidence from Ethiopia argues that children are engaged in economic or home production activities regardless of school enrollment, suggesting these activities do not replace school directly (Rose & Al-Samarrai, 2001). Yet, household and labor responsibilities expend children's energy and compete with school tasks, and thus limit investment in classroom engagement and homework.

One way which limited parental resources leads to poor cognitive development is through undernourishment in early childhood. Extensive research on nutrition and cognitive development across multiple middle and low-income countries finds that early undernourishment, most prevalent among poor families, leads to inadequate brain development, poorer cognitive skills, fewer years of education, decreased knowledge accumulation per year of school, and lower earnings in adulthood (Grantham-McGregor et al., 2007). Furthermore physical stunting or exceptionally short stature for age also reflects early nutritional experiences and has a strong positive correlation with cognitive development (Grantham-McGregor et al., 2007)

Geography

Due to geopolitical pressures, marginalized groups are often concentrated in regions that are difficult to access, have limited production potential, and lack sufficient educational and health infrastructure. For example, Vietnamese ethnic minorities, indigenous Peruvians, and tribal groups in India are more heavily concentrated in difficult to access mountainous areas. Limited infrastructure also characterizes the homelands of Ethiopia's marginalized ethnic

groups. Geographic inequality is also prevalent within urban areas. Lower castes in urban India and rural to urban migrants, exemplified by indigenous Peruvians, are geographically concentrated in urban slums with limited infrastructure and resources.

Rural areas have traditionally lagged behind in primary education access, creating higher costs to families education and attracting less well trained teachers (Ilon & Moock, 1991; McEwan, 2004). Even amidst high primary education enrollment rates, rural schools receive fewer resources, compromising educational quality (Trivelli, 2005). Furthermore, national policy makers often overlook indigenous regions, and even as their economic resources increase, national resources are not directed to areas with high indigenous concentrations (Hall & Patrinos, 2005).

Language

Language creates additional barriers for marginalized social groups preferring minority languages. Spanish in Peru and Vietnamese in Vietnam dominate nationally and are not traditionally spoken by indigenous peoples. In Ethiopia, Amhara dominated nationally until the early 1990s. Currently, primary education occurs in regional languages, but educational resources in languages other than Amhara are limited (Tronvoll, 2000). In India's multilinguistic context, Telugu is the local majority language and non-Telugu speakers must negotiate a primarily Telugu environment, which translates to schools and the work place.

Children's preferred languages may have lower status and dominant languages may be the primary means of student-teacher communication. Schools historically discouraged, even punished, local language use. Though instructional language preference and bilingual education are increasing, they are still relatively rare and linguistic barriers still exists (McEwan, 2004).

Finding well trained teachers who speak local languages is challenging and supporting materials in local languages are underdeveloped (Aikman & Pridmore, 2001; Wagaw, 1999). Furthermore language preferences may limit parental navigation of the educational system.

Finally learning additional language increases schooling burdens. Ethiopian children in non-Amharic speaking regions receive language instruction in the regional language, later adding Amharic, the de facto national language, and then English, the secondary education instructional language (Wagaw, 1999).

Social Groups by Country

Describing social group divisions in the four target countries provides a basis for understanding how social group membership generates educational achievement disparities. The following section explains social groups divisions, describes the population distribution, introduces how empirical analyses will incorporate social group membership, and highlights events influencing social groups dynamics within each country. Table 1 displays the population distributions for each country and the sample we use.

Ethiopia

In the late 19th century the Amhara colonized their southern and western neighbors, creating current Ethiopia (Tronvoll, 2000). Ethiopia contains over 80 documented ethnic groups, many of which are quite small, but we limit this study to the Amhara and the Oromo. The Oromo comprise 32% of the population and are politically and socially marginalized. The next largest, the Amhara, 30% of the population, have politically and socially dominated Ethiopia.

Prior to 1974 Amhara was the official state culture and language, and other regional languages were forbidden in formal activities (Mengisteab, 2001). Acceptance of non-Amharic

ethno-linguistic groups improved in the subsequent revolutions, which first aimed to end class-based and later ethnic discrimination (Young, 1996). Under the new government, power has shifted to regional ethnic based authorities, and previously subordinate ethnic groups have reclaimed aspects of their identity that were suppressed under previous political system (Mengisteab, 2001). Furthermore, the educational curriculum and instructional language are determined by regional authorities, but linguistically appropriate teachers and resources are limited (Tronvoll, 2000).

India, Andhra Pradesh state

The Indian Constitution outlawed the caste system, but its vestiges still permeate Indian society. Birth determines caste membership, dictating employment and resource distribution, and the belief that lower class children do not deserve an education is deeply rooted in India's middle class, leading to discrimination in schools (Rao, et al., 2003). Scheduled Tribes (ST) and Scheduled Castes (SC) are the most marginalized groups, followed by the Other Backwards Classes (OBC). STs are geographically isolated and more often reside in rural areas, including in Andhra Pradesh. Traditionally they relied on natural resources, but the colonial period reduced their access and land rights. SCs include untouchables and are the lowest position of the Hindu hierarchy. They are more urbanized and hold more political power than STs but still hold a low social position (Chanana, 1993). According to the 2001 Census, 16.2, 8.2, and 52.0 percent of the population are members of SCs, STs, and OBCs respectively.

India's educational expansion in recent decades focused largely on higher education. University seats were reserved for SCs and STs, but they went unfilled when many failed to complete primary school (Chanana, 1993). The government opened new schools, but additional educational costs to families remained high, and education for lower classes did not improve;

more recently, the government increased the number of teachers from lower classes, aimed at improved retention (Rao, et al., 2003).

Peru

Peru's racial composition results from of 500 years of *mestizaje*, mixing of Spanish colonizers, Indigenous inhabitants, African slaves, other immigrants, and their descendents. Estimates suggest the population is approximately 45% Indigenous, 37% *mestizo*, 15% white, and 3% other (e.g., African, Chinese, Japanese), though race is not specifically reported in the national census. European descendents have typically

Identifying Indigenous households with survey data poses obstacles when Peruvians of vastly different origins self-identify as *mestizo* (Ñopo, Saavedra, & Torero, 2007; Trivelli, 2005). This allows the user to avoid extremes in a cultural context where race conversations are subdued and there is a preference for both lighter skin and European facial features (Drzewieniecki, 2004). Often language identifies Indigenous households, underestimating Indigenousness, particularly now that growing migration to urban centers has increased the tendency to self-identify as Spanish speaking (Trivelli, 2005). Using an innovative data collection strategy to address this issue, researchers used independent observers' phenotype ratings to conclude that labor market discrimination occurs on the basis of Indigenous physical features even for those who do not self-identify as Indigenous (Ñopo, Saavedra, & Torero, 2007).

Vietnam

Like Ethiopia, social group divisions in Vietnam developed along ethnic divisions. The Kinh comprise 86% of the population and inhabit the fertile low lands and river deltas with more access to resources and infrastructure. The 53 ethnic minority groups inhabit isolated rural areas

and speak local languages. They were largely excluded from recent economic growth, maintain higher fertility rates, and have poor access to health care services (van de Walle & Gunewardena, 2001). Ethnic minorities in Vietnam include those inhabiting the region for centuries, such as the Tay, and more recent geopolitical migrants. The Hao, or ethnic Chinese, are an exception; they largely assimilated with the urban Kinh and flourished economically (Baulch, et al., 2004).

Vietnam's Confucian values advocate education as an accepted form of upward social mobility (Rao, et al., 2003), but this idyllic view may not apply to ethnic minorities, who demonstrate lower returns to education (van de Walle & Gunewardena, 2001). The government has increased its interest in improving their living standards, but poverty alleviation strategies in the highland areas where ethnic minorities are concentrated have done little to improve ethnic minorities' living standards; instead, their rural Kinh neighbors benefitted most from the investments (van de Walle & Gunewardena, 2001).

Data

We employ data from the Young Lives Study, a unique longitudinal study of childhood poverty conducted in Ethiopia, India, Peru, and Vietnam (Huttly, Jones, & Boyden, 2009). Target children were selected and initially interviewed in 2002 at age 8 (between 7.5 and 8.5 years old) and again in 2006 at age 12. At both time points the child and his or her primary caregiver were interviewed regarding household resources, family member characteristics, attitudes and perceptions, and the child's schooling, health, and activities. Child participants also completed a battery of basic academic skills. We draw on responses from both rounds.

Respondents were selected using a sentinel site sampling approach. Within each country 20 sites were semi-purposefully selected to best represent each country's regions. Approximately

50 households within each site were randomly selected from among all households with a child between 7.5 and 8.5 years old. Post-stratification analyses comparing the Young Lives sample to large nationally representative surveys determined that although true random sampling was not employed that each country's sample adequately represents the national population (Escobal & Flores, 2008; Kumra, 2008; Outes-Leon & Sanchez, 2008). Attrition between rounds and incomplete cases were minimal, and we exclude incomplete cases from the analyses.

Country specific circumstances reduced the eligible sample. First, due to extensive ethnic diversity in Ethiopia we limit our sample to include only the Oromo and Amhara ethnic groups. The Amhara and Oromo reflect the two largest ethnic groups within the country and are both well represented in the data.

Second, beginning in the late 1990s, before the index children began school, the Andhra Pradesh state initiated a large scale, multi-systemic approach to improve human development indicators for Scheduled Castes (SC) and Scheduled Tribes (ST) (Government of Andhra Pradesh, 2003). Specifically, they targeted individual districts where these groups were concentrated, and that overlap with the sample (World Bank OPCS, 2005). As a result, children from the most vulnerable circumstances excelled on both health and educational indicators to the extent that they are out performing even the most privileged sectors. This is documented in both our sample and reports by the Andhra Pradesh State which documented students in targeted schools surpassing mean standardized exam scores (Government of Andhra Pradesh, 2003). Though these improvements are promising, it presents a data analysis problem. The current study aims to understand inequalities, and SC and ST children were not heavily targeted outside these areas. Therefore, we exclude from our analyses the four sampling clusters where over half the respondents are members of SC and ST and revisit this topic in the discussion.

Finally, the Vietnamese central government requires residency permits. As a result, ethnic minorities reside almost exclusively in rural mountain areas. To provide relevant conclusions, we analyze the rural sample in the Northern Uplands and Central Coastal regions, where both ethnic minorities and the Kinh majority are sampled.

Measures

Academic achievement

To assess literacy, interviewers asked children to read and write two separate sentences in the child's selected language. Interviewers scored responses according to whether the child read letters, words, complete sentence, or nothing and whether the child wrote easily, with difficulty, or not at all. The combined literacy score is scaled from 0 to 1. Literacy scores at age 8 demonstrate sufficient reliability for the subgroups within each country ($\alpha_{Et,Am} = 0.87$, $\alpha_{Et,Or} = 0.69$, $\alpha_{In,S} = 0.66$, $\alpha_{In,O} = 0.72$, $\alpha_{Pe,In} = 0.78$, $\alpha_{Pe,NI} = 0.65$, $\alpha_{Vi,EM} = 0.93$, $\alpha_{Vi,Ki} = 0.70$). The literacy inventory was repeated at age 12 and demonstrates poor reliability due to a ceiling effect, where nearly all respondents in the in the higher achieving groups received the highest score in every county but Ethiopia ($\alpha_{Et,Am} = 0.68$, $\alpha_{Et,Or} = 0.71$). Therefore the age 12 literacy scores are useful for demonstrating absolute differences between groups. However, they do not differentiate among top scorers, underestimate the achievement gap, and are not useful as dependent variables in a regression equation.

Children completed a ten-item arithmetic inventory at age 12. We excluded items with differential item functioning by gender or test administration language. This follows the analyses and advice of previous work (Cueto, Leon, Guerrero, & Muñoz, 2009). We then used Item Response Theory (IRT) techniques to estimate a standardized ability score for each respondent

using PARAM-3PL Version 0.89 (Rudner, 2007). An advantage of IRT over an additive scale is that it incorporates item difficulty, item differentiation, and guessing parameters into the estimate of math ability. It also includes estimates for respondents with incomplete response series without assuming that unanswered items are incorrect.

The total school grades is the current school grade for children who are currently enrolled in school and the last completed grade for children who are not currently enrolled in school.

Social group category

The adult respondents reported ethnicity, race, or caste information for the index children and their parents. We use responses from the second round of data collection, capitalizing on their improved specificity. Our selected method compares disparities between two discrete groups. Therefore we divide respondents in each country into two groups.

In the Ethiopian sample, we include only those who identify as Amhara or Oromo. In the Indian sample we divide the respondents into those who identify as SC or ST and all others. In the case of Vietnam, we aggregate all ethnic minorities groups and compare them to the Kinh, the majority ethnic group. Their limited sample size not allow for further disaggregation.

Classifying race in Peru is particularly problematic for reasons explained previously. Therefore, we create two groups, which we label ‘more-indigenous’ and ‘more-European.’ The more-indigenous group contains all children who first spoke an indigenous language, identify as indigenous, have at least one parent whose mother tongue is an indigenous language, or have at least one parent who identifies as indigenous. A small number of Afro-Peruvian descendents are in the sample, and we classify them with the more-indigenous group for analytic purposes due to their historically marginalized status. All other children are classified as more-European.

Predictor variables

The first set of predictor variables reflects family endowment characteristics. Adult respondents provided information on housing quality, utilities access, and consumer durables ownership. Using this information, separate wealth indices were constructed for each country at each time. The adult respondent also reported maternal education years during the second round, and we classified responses according to the educational systems of each country. We categorized adult literacy program participation as not completing primary education, because few could read, and it is not equivalent to completed primary education. An adult respondent also reported the total number of living children who had been born to the child's biological mother by the time the child was aged 8 years. Furthermore, due to the high proportion of Ethiopian children who had lost a parent, we include an indicator for at least one deceased parent.

The second set of predictor variables reflects geographical factors. Region and rural residence are classified using national census information. Children reported distance to school referencing either the current school or the most recent one attended if they were no longer enrolled.

To examine language for Ethiopian respondents we include an indicator variable for whether the child's mother tongue is the dominant language where they reside. For the Indian sample we include an indicator for whether the child's primary language is Telugu, the dominant language of Andhra Pradesh. In the case of Peru, we include an indicator variable for whether the child first spoke Spanish. Nearly all Vietnamese ethnic minority children first spoke their

minority language. Therefore we use an indicator for whether the child fluently speaks Vietnamese.

We also include a set of indicator variables describing attitudes towards school and education. The surveys were not consistent in content and meaning across country. Therefore we attempt to maintain consistency across the four settings but include slightly different variables when necessary. For the Ethiopian sample we include a measure of parental school efficacy where parents reported if they believe they can help their child do better in school. Nearly all parents reported they would like their children to complete high levels of education, but fewer actually expected that their children would reach this goal. We include a variable indicating whether parents believed the child would achieve this goal. Child respondents reported whether they had missed more than a week of school during the past month. Those who did are considered to have high school absences. For India and Peru we use the same three items used in the Ethiopian analyses. For Vietnam we include variables indicating that parents believe education is essential, whether they expect their children to receive a high level of education, and if children participated in extra classes during the previous six months.

Standardized height for age at age eight is used as a proxy for early childhood nutrition. Extremely short stature, even in adulthood, reflects early childhood undernourishment. Child height and exact age in months, were transformed into height for age z-scores based on updated standards from the World Health Organization's Multicentre Growth Reference Study (Borghini, de Onis, Garza, Van den Broeck, Frongillio, Grummer-Strawn, Pan, Molinari, Matorell, Onyango, & Martines 2006).

Child respondents reported the daily hours contributed to the family livelihood strategy, including working for money and non-paid contributions such as farming activities, household

chores, and caring for others. We combined these contributions into a single variable to reflect the experiences of boys and girls and because these activities equally detract from school and study time.

We also include child gender and the child's age in months centered at exactly age eight as control variables.

Method

We begin by using regression analysis with a single variable for group membership to examine the achievement gaps within each country. We then proceed with bivariate analyses exploring how the two groups within each country differ on the indicators we expect to be associated with academic achievement.

To decompose observed academic achievement disparities within each country, we use a version of the Oaxaca-Blinder decomposition (Blinder, 1978; Oaxaca, 1978). This method was initially developed to understand wage differentials by gender and has been applied to other disparities, such as test score gaps between indigenous and non-indigenous students in Bolivia and Chile (McEwan, 2004), Maori and European descended students in New Zealand (Lock & Gibson, 2008), and Black and White students in the US (Myers, Kim, Mandala, 2004).

A drawback of using linear regression with a variable indicating group membership is that it constrains the remaining coefficients to operate equally in both groups. The Oaxaca-Blinder decomposition permits varying effects between subgroups, which corresponds with previous findings suggesting that ethnic minority groups often respond differentially to poverty alleviation strategies and may experience differential returns to educational investments. Furthermore this technique allows us to identify the portion of the gap attributable to the

different variables or groups of variables. Therefore we group predictor variables into the following categories: family endowment, geographic factors, language, family education attitudes, hours spent working, standardized height for age, and child specific control variables.

Regression decomposition uses least squares regression estimates for each subgroup to break down the gap into an explained portion, attributable to differences in characteristics, and an unexplained portion, attributable to differing effects of the predictor variables. Specifically, we specify a pooled model, an extension of the Oaxaca-Blinder decomposition proposed by Neumark (1988), which pools regression estimates over the two models as a source of comparison. The pooled regression decomposition model is expressed as follows:

$$Y^1 - Y^2 = \Delta X B^{\text{pooled}} + [X^1(B^1 - B^{\text{pooled}}) + X^2(B^{\text{pooled}} - B^2)]$$

We fit all models using the `oaxaca` command (Jann, 2008; 2010) and adjust the standard errors to account for clustering in STATA 11.0 (StataCorp, 2009). Also, because the decomposition results may be affected by the choice of omitted base category, we include the base categories and estimate categorical variables as deviation from the grand mean (Jann, 2008). We then calculate the portion of the achievement gap explained by group differences and their differing effects for each group of predictor variables.

Results

Unadjusted achievement gap

Table 1 describes the observed academic achievement gaps and reveals the extent of disparities within each country. In all cases results are in the anticipated direction, and the more privileged groups within each country out perform the marginalized groups. The largest gaps are

present in Vietnam. The smallest gaps are observed in Ethiopia, where the age 8 literacy gap is not significant.

Overall, the age 12 literacy gaps are relatively small. This is due at least partially to the ceiling effect of the inventory used at this age, which underestimates ability for the higher performers. However, the results still note a significant margin in every case except India. We do not conduct further analyses with the age 12 literacy gaps where reliability is poor.

Descriptive results

We observe that families from marginalized groups in India, Peru, and Vietnam have considerably fewer material resources. Ethiopia is an exception, where we observe only trivial wealth differences between the Amhara and Ormo. In all countries mothers from the marginalized groups were more likely to have never attended school or not complete primary school, while their counterparts in non-marginalized groups were more likely to complete primary and secondary school. Children from the marginalized groups were also more likely to grow up in families with larger numbers of children, likely stretching parental resources. In the Ethiopian sample, more Amhara children had lost a parent.

In Ethiopia, India, and Peru, children from marginalized groups more heavily concentrated in rural areas, while the entire Vietnamese sample resides in rural areas. On average marginalized children reported traveling farther to arrive at school. The proportion of respondents from each group matches closely with the population distributions described in national data. An exception is the coastal region of Andhra Pradesh, where we excluded a portion of the sample due to external factors influencing our primary outcomes.

No majority group children speak minority languages in either Peru or Vietnam. A higher proportion of sample children from SC and ST speak Telugu, the locally dominant language, as their mother tongue, and a similar proportion of Amhara and Oromo children speak the regionally dominant language.

Differences in families' educational attitudes between Amhara and Oromo respondents were minimal. Slightly more Amhara respondents reported that they could act to improve their children's educational outcomes and assert that their children would achieve a high level of education. Indian children who were not members of SC and ST were more likely to attend private schools and their families more frequently reported that they could act to improve their children's educational outcomes and assert that their children would achieve a high level of education. More-indigenous Peruvian families were less likely to believe they could act to improve their children's educational outcomes and were less likely to assert that their children would achieve a high level of education. Differences in the number of children who reported missing more than a week of school during the past month were trivial in Ethiopia, India, and Peru. Fewer Vietnamese ethnic minority families believe school is essential and fewer children participate in extra classes. However, parents are more likely to believe their children will achieve high levels of education.

Children from marginalized groups also dedicate more time to family livelihood strategies, such as working for pay and household tasks. In Peru and Vietnam these differences are quite large; whereas in Ethiopia and India, the difference is considerably smaller.

Children from marginalized groups in India, Peru, and Vietnam are smaller than their non-marginalized counterparts, reflecting poorer early childhood nutrition. However, in Ethiopia, the reverse is true, and Oromo children are taller than Amhara children.

Multivariate Results

We continue with OLS regression estimates for each subgroup. It is worth noting when and how regression coefficients for the same predictor variable differ between the two subgroups in magnitude or direction. This is a justification for regression decomposition methods.

Examining these results across Ethiopia, (see Table 3) we find that the association with maternal education is relatively strong across these models. Regional residence indicators have strong predictive power, though these variables operate differently between the Amhara and Oromo. The amount of time spent working and standardized height are also significant predictors in these models.

As we move to the results from India, we find that household wealth is significantly associated with achievement for the ‘other’ group. In contrast wealth does not significantly predict achievement for SC/ST, the poorer of the two groups. Maternal secondary education is a stronger achievement predictor for SC/ST. Parental attitudes towards school are also significantly associated with achievement in several models.

Considering the results from Peru, both higher wealth and maternal education are associated with higher achievement. Though regional residence is not significant, it appears to be operating in different directions where non-coastal residence is associated with better outcomes for less-indigenous children and worse outcomes for more-indigenous children

Finally, in the case of Vietnam, it is first interesting to note that these models have substantially more predictive power, as indicated by the R-squared values, for ethnic minorities than for the Kinh. This suggests that overall Kinh children have similar achievement outcomes regardless of the factors we include in the model and that opportunities for children within the

Kinh ethnic group are more equitably distributed. Both wealth and maternal education are associated with achievement, but these relationships vary across the models. Regional residence, as in both Peru and Ethiopia, again appears to be associated with achievement in different directions for the two ethnic groups. The ability to speak Vietnamese is positively associated with achievement for ethnic minorities across these models. Standardized height for age is also positively associated with achievement, though more so for ethnic minorities.

Regression Decomposition

Regression decomposition identifies the variables or variable groups with the largest contributions to the observed achievement gaps. The preliminary decomposition results are reported in Table 7, where we report their relative importance compared to the other components of the model as minimal, moderate, or large.

Across all four settings, family endowments, specifically wealth, family size, and maternal education are associated with the observed academic achievement gaps between marginalized children and their non-marginalized counterparts. The contribution of family endowment is weakest in Ethiopia, where measured wealth differences between the Amhara and Oromo are not significant.

Instead we find in Ethiopia that geographical factors are most strongly associated with academic achievement disparities. This is not surprising given the country's history of allocating resources according to the homelands of particular ethnic groups. Geographic variables are also moderately associated with achievement gap in Vietnam, while the ability to speak Vietnamese also makes a notable contribution. However, we find the contribution of language is minimal in the other three countries.

Family educational attitudes make have only a trivial association with the achievement gap in every case except India, where contributions are still only moderate.. The amount of time children spend working is differentially associated with outcomes and settings. For example the strongest associations are total grades in India and math achievement scores in Peru and Vietnam.

We find standardized height, appears to play a larger role in the total number of grades achieved than in indicators of academic skill. This contradicts previous findings which attribute large deficits in cognitive ability to factors associated with early childhood nutrition. It instead suggests that children may be deemed ready for school and socialized based on their size and that nutritional contributions to cognitive ability have been overestimated.

Discussion

At the turn of the century the United Nations (UN) established the Millennium Development Goals (MDGs), specific objectives for improving social and economic conditions in the poorest countries by 2015 (UN, 2000). While the MDGs target gender equity in primary education, comparable goals advancing ethnic minorities and other marginalized groups are absent (Kabeer, 2006).

Achieving equity for marginalized social groups has typically lagged behind gender equity in UN discourse. For example, the UN General Assembly adopted the *Declaration of the Elimination of Discrimination against Women* in 1967, whereas the *Declaration on the Rights of Persons Belonging to National or Ethnic, Religious and Linguistic Minorities* and the *Declaration on the Rights of Indigenous Peoples* were not adopted until 1992 and 2007 respectively.

As initially presented, the main purpose of this paper is to examine within country disparities by social group. Our findings reveal a clear trend that educational disparities persist by social groups across the four settings examined in this study. These findings provide strong evidence that there is a need to develop large global efforts that target educational equity by social group.

Additionally, our findings also reveal information about the processes contributing to the observed disparities. Overall, differences in family endowments of wealth and education are most strongly associated with academic achievement disparities. Furthermore in certain circumstances, family endowments are operating through early childhood nutritional deficiencies and the amount of time children dedicate to family livelihood strategies.

Geographic residence is the largest contributor in Ethiopia, where the Oromo homelands experiences less investment in schools and infrastructure. The ability to speak the predominant Vietnamese language is also strongly associated with academic achievement disparities within Vietnam. There is very little evidence that family attitudes towards education contribute to educational disparities.

Implications

The current findings suggest the need for immediate action aimed at reducing educational disparities between social groups. Education's importance for predicting lifetime productivity will likely only increase in the coming decades as technological inputs increase and the global economy becomes more knowledge based. If educational disparities are further neglected, economic and social disparities will likely persist.

Previously, Knodel and Jones (1996) demonstrated that educational disparities by relative income status within countries exceeded the educational disparities found between boys and girls. They argued that universal primary education goals needed to shift their emphasis away from gender and target children by poverty status. However, gender equity remained the principal focus of primary education goals. An opposition was that targeting poverty status is a difficult task, especially in countries where accurate income documentation is rare.

An obvious solution to the disparities that persist within countries is to target policies and services to marginalized groups. This strategy seems like the most effective means for maximizing the effects of limited resources. A number of strategies, such as conditional cash transfers or improved schooling environments have demonstrated their effectiveness in other situations and could easily be applied to programs that target specific groups.

However, there are drawbacks to social group targeting. Namely, there is concern that this method would serve to increase jealousy and conflict within already strained group dynamics. Knowing that one's equally poor neighbor from a different ethnic group was receiving a subsidy would likely create animosity and conflict before it could alleviate disparities. Community members in post-ethnic conflict of Nepal felt that targeted cash transfers by ethnic group were not favorable and would only increase community divisiveness (Köhler, Cali, & Stirbu, 2009).

Geographically targeted interventions also have potential. Marginalized groups are often geographically concentrated. Targeting regions with a high proportion of minority group members has the potential to benefit marginalized groups and avoid the divisiveness that would likely occur when targeting individuals. Regional targeting appears to have been highly effective in Andhra Pradesh, India. Previously, we discussed the need to exclude several clusters from our

analyses due to the high possibility that they were targeted by multisystemic poverty alleviation strategies. These efforts emphasized improved economic opportunities, access to health facilities, and a specific emphasis on improving educational quality (Government of Andhra Pradesh, 2003; World Bank OPCS, 2005). Consequently, children from Scheduled Tribes in these areas were exceeding state-wide performance averages (Government of Andhra Pradesh, 2003).

However, a risk of geographically targeted efforts is that marginalized groups may not respond to efforts that do not meet their specific needs. There are two perspectives for addressing ethnic inequalities. One assumes that the majority model will improve the outcomes of the minority group if it is applied to the minority group, and the other deems it necessary to work within a framework relevant to the disadvantaged group (van de Walle & Gunewardena, 2001). Therefore, efforts should consider the specific needs and behaviors of the targeted group.

Furthermore, when assuming the majority model, residents from the majority ethnic group may be better able to utilize of the increased access to resources. This has been the case in rural Vietnam. A number of government policies have emphasized poverty alleviation in the rural mountain regions that are home to ethnic minority groups. However, the Kinh in these areas benefitted more than the ethnic minorities, because the poverty alleviation model assumed that ethnic minorities would respond similarly to the Kinh (van de Walle & Gunewardena, 2001).

Contributions and Limitations

This paper makes several unique contributions. We address within country educational achievement gaps from an internationally comparative perspective, across four distinctly different settings. This design grants the opportunity to draw broader conclusions about the nature of achievement disparities.

Additionally we analyze dependent variables, math ability and literacy, that measure academic skill acquisition. We use these in addition to the child's grade in school. Given that the schools where children from marginalized social groups attend often receive fewer investments than the schools where other children within the same country attend school, it is important to use indicators that evaluate whether learning actually occurs. Using these skill indicators allows for a more nuanced understanding of academic achievement.

However, the paper also has several limitations. The greatest concern is testing bias. Despite extensive efforts to reduce testing bias in both the initial instrument design and the subsequent analyses, we cannot ignore the possibility of testing biases. Any testing biases would overestimate the achievement gap and cause parameter estimation errors.

A further weakness is that we are not able to measure change in the same indicators over time. Examining growth in particular academic domains would provide a clearer understanding of how academic skill disparities develop throughout childhood.

- Ahmad, Omar, Alan Lopez, and Mie Inoue. 2000. "The decline in child mortality: a reappraisal." *Bulletin of the World Health Organization* 78:1175-1191.
- Aikman, Sheila, and Pat Pridmore. 2001. "Multigrade schooling in remote areas of Vietnam." *International Journal of Educational Development* 21:521-536.
- Anh, Truong Si, John Knodel, David Lam, and Jed Friedman. 1998. "Family Size and Children's Education in Vietnam." *Demography* 35:57-70.
- Baulch, B, K Chuyen, D Haughton, and J Haughton. 2002. "Economic growth, poverty, and household welfare in Vietnam.".
- Baulch, B, T. T. K Chuyen, D Haughton, and J Haughton. 2004. "Ethnic minority development in Vietnam: A socioeconomic perspective." in P. Glewwe, N. Agrawal, & D. Dollar (Eds.), *Economic growth, poverty, and household welfare in Vietnam*. The International Bank for Reconstruction and Development, The World Bank, Washington, DC.
- Behrman, Jere R., and James C. Knowles. 1999. "Household Income and Child Schooling in Vietnam." *World Bank Econ Rev* 13:211-256.
- Ben-Arieh, Asher et al. 2001. *Measuring and Monitoring Children's Well-Being*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Borghi, E. et al. 2006. "Construction of the World Health Organization child growth standards: selection of methods for attained growth curves." *Statistics in Medicine* 25:247-265.
- Brockerhoff, M, and P Hewitt. 2000. "Inequality of child mortality among ethnic groups in Sub-Saharan Africa." *Bulletin of the World Health Organization* 70:30-41.
- Buchmann, Claudia, and Emily Hannum. 2001. "Education and Stratification in Developing Countries: A Review of Theories and Research." *Annual Review of Sociology* 27:77-102.

- Burnett, N. 2008. "Education for All: An Imperative for Reducing Poverty." New York Academy Sciences Annals 1136:269-275.
- Chanana, Karuna. 1993. "Accessing Higher Education: The Dilemma of Schooling Women, Minorities, Scheduled Castes and Scheduled Tribes in Contemporary India." Higher Education 26:69-92.
- Charasse-Pouélé, Cécile, and Martin Fournier. 2006. "Health disparities between racial groups in South Africa: a decomposition analysis." Social Science & Medicine 62:2897-2914.
- Cornia, Giovanni Andrea. 2003. "The Impact of Liberalisation and Globalisation on Within-country Income Inequality." CESifo Economic Studies 49:581-616.
- Cueto, Santiago, Juan Leon, Gabriela Guerrero, and Muñoz Ismael. 2009. "Psychometric characteristics of cognitive development and achievement instruments in Round 2 of Young Lives." <http://www.younglives.org.uk/pdf/publication-section-pdfs/technical-notes-pdfs/YL-TN15-Cueto.pdf> (Accessed March 2, 2010).
- Deshpande, Ashwini. 2000. "Does Caste Still Define Disparity? A Look at Inequality in Kerala, India." The American Economic Review 90:322-325.
- Dreze, Jean, and Jackie Loh. 1995. "Literacy in India and China." Economic and Political Weekly 30:2868-2878.
- Drzewieniecki, Joanna. 2004. "Peruvian youth and racism: The category of 'race' remains strong." Las Vegas, Nevada http://dlc.dlib.indiana.edu/dlc/bitstream/handle/10535/1886/Peruvian_Youth_and_Racism.pdf?sequence=1 (Accessed September 21, 2009).
- Government of Andhra Pradesh. 2003. "Strategy Paper on Welfare of Scheduled Caste and Scheduled Tribe Population." <http://unpan1.un.org/intradoc/groups/public/documents/APCITY/UNPAN015043.pdf> (Accessed March 2, 2010).

- Grantham-McGregor, Sally et al. 2007. "Developmental potential in the first 5 years for children in developing countries." *The Lancet* 369:60-70.
- Hall, Gillette, and H. A Patrinos. 2005. "Indigenous peoples, poverty, and human development in Latin America: 1994-2004." http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2005/07/20/000112742_20050720171830/Rendered/PDF/330200ENGLISH01Poverty1Exec1Summ1en.pdf (Accessed September 21, 2009).
- Huttly, S., Jones, N. & Boyden, J., 2009. "Young Lives: an International Study of Childhood Poverty: Round 1, 2002 & Round 2, 2006 [computer file]." Colchester, Essex: UK Data Archive [distributor], August 2009. SN: 5307.
- Ilon, Lynn, and Peter Moock. 2001. "School attributes, household characteristics, and demand for schooling: A case study of rural Peru." *International Review of Education* 37:429-451.
- Jann, Ben. 2008. "The Blinder-Oaxaca decomposition for linear regression models." *The Stata Journal* 8:453-479.
- Knodel, John, and Gavin W. Jones. 1996. "Post-Cairo Population Policy: Does Promoting Girls' Schooling Miss the Mark?." *Population and Development Review* 22:683-702.
- Lock, Kelly, and John Gibson. 2008. "Explaining Maori Under-achievement in standardized reading tests: the role of social and individual characteristics." *Kotuitui: NZ Journal of Social Sciences Online* 3:1-13.
- McEwan, Patrick J. 2004. "The Indigenous Test Score Gap in Bolivia and Chile." *Economic Development and Cultural Change* 53:157-190.
- McSweeney, Kendra, and Shahna Arps. 2005. "A "Demographic Turnaround": The Rapid

- Growth of Indigenous Populations in Lowland Latin America.” *Latin American Research Review* 40:3-29.
- Mengisteab, Kidane. 2001. “Ethiopia's Ethnic-Based Federalism: 10 Years after.” *African Issues* 29:20-25.
- Neumark, David. 1988. “Employers' Discriminatory Behavior and the Estimation of Wage Discrimination.” *The Journal of Human Resources* 23:279-295.
- Ñopo, Hugo, Jaime Saavedra, and Máximo Torero. 2007. “Ethnicity and Earnings in a Mixed-Race Labor Market.” *Economic Development and Cultural Change* 55:709-734.
- Patrinos, Harry Anthony, and George Psacharopoulos. 1996. “Socioeconomic and ethnic determinants of age-grade distortion in Bolivian and Guatemalan primary schools.” *International Journal of Educational Development* 16:3-14.
- Psacharopoulos, George. 1997. “Child labor versus educational attainment: Some evidence from Latin America.” *Journal of Population Economics* 10:377-386.
- Psacharopoulos, George, and Harry Anthony Patrinos. 2004. “Returns to Investment in Education: A Further Update.” *Education Economics* 12:111-134.
- Ranis, Gustav, Frances Stewart, and Alejandro Ramirez. 2000. “Economic Growth and Human Development.” *World Development* 28:197-219.
- Rao, Nirmala, Kai-Ming Cheng, and Kirti Narain. 2003. “Primary Schooling in China and India: Understanding How Socio-Contextual Factors Moderate the Role of the State.” *International Review of Education* 49:153-176.
- Rose, Pauline, and Samer Al-Samarrai. 2001. “Household Constraints on Schooling by Gender: Empirical Evidence from Ethiopia.” *Comparative Education Review* 45:36-63.
- Rudner, Lawrence. (2007). PARAM-3PL Calibration Software for the 3 Parameter Logistic IRT

- Model (freeware). <http://edres.org/irt/param>.
- Trivelli, Carolina. 2005. “*Una mirada cuantitativa a la situación de pobreza de los hogares indígenas en el Perú.*”
- Tronvoll, Kjetil. 2000. “Ethiopia: A new start?.” <http://www.addisvoice.com/resources/Ethiopia-new-start.pdf> (Accessed September 21, 2009).
- Wagaw, Teshome G. 1999. “Conflict of Ethnic Identity and the Language of Education Policy in Contemporary Ethiopia.” *Northeast African Studies* 6:75-88.
- van de Walle, Dominique, and Dileni Gunewardena. 2001. “Sources of ethnic inequality in Viet Nam.” *Journal of Development Economics* 65:177-207.
- World Bank, Operations Policy and Country Services (OPCS). 2005. “Status of Projects in Execution - FY05 SOPE, Region: South Asia, Country: India.” http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2009/09/15/000333037_20090915024840/Rendered/PDF/498220AR0India0050Box341974B01PUBLIC1.pdf (Accessed March 2, 2010).
- Young, John. 1996. “Ethnicity and Power in Ethiopia.” *Review of African Political Economy* 23:531-542.
- Zoninsein, Jonas. 2001. “The economic case for combating racial and ethnic exclusion in Latin American and Caribbean countries.” <http://www5.iadb.org/sds/doc/soc-ZoninsteinJonasi.pdf> (Accessed September 21, 2009).

Table 1. Unadjusted Achievement Gaps

	Ethiopia		India		Peru		Vietnam	
	Oromo	Amhara	SC/ST	Other	More-Indig.	Less-Indig.	Ethnic Min.	Kinh
Literacy age 8 (0 to 1)	0.31	0.35	-	0.57	0.70	0.71	0.85	0.91
Math ability age 12 N~(0,1)	-0.49	-0.15	*	0.20	0.66	0.06	0.65	1.08
Literacy age 12 (0 to 1)	0.71	0.82	***	0.83	0.87	†	0.96	0.98
Total grades completed age 12	4.1	4.6	***	6.2	6.6	**	6.0	6.6

*** p ≤ 0.001, ** p ≤ 0.01, * p ≤ 0.05, † p ≤ 0.10

Table 2. Mean Values by Within Country Group

	Ethiopia		India		Peru		Vietnam	
	Amhara	Oromo	Other	SC/ST	Less-indig.	More-indig.	Kinh	Ethnic min.
Family endowments								
Wealth index at age 8	0.17	0.17	0.38	0.28	0.56	0.45	0.44	0.21
Wealth index at age 12	0.16	0.18	0.37	0.26	0.60	0.46	0.46	0.26
Mom never attended school	0.43	0.35	--	--	--	--	--	--
Mom did not complete primary school	0.40	0.52	0.63	0.89	0.16	0.55	0.21	0.83
Mom completed primary school	0.17	0.13	0.24	0.08	0.40	0.28	0.47	0.13
Mom completed lower secondary school	--	--	0.13	0.04	0.44	0.17	0.32	0.05
Children in family = 1 or 2	0.25	0.13	0.45	0.35	0.44	0.29	0.45	0.18
Children in family = 3	0.18	0.18	0.36	0.35	0.25	0.21	0.28	0.32
Children in family = 4+	0.57	0.69	0.19	0.30	0.31	0.50	0.27	0.50
Parental death before age 8	0.15	0.10	--	--	--	--	--	--
Parental death before age 12	0.25	0.17	--	--	--	--	--	--
Attends a private school	--	--	0.31	0.14	--	--	--	--
Geography								
Ethiopia: Addis Abba	0.18	0.17	--	--	--	--	--	--
Amhara	0.68	0.00	--	--	--	--	--	--
Oromia	0.09	0.79	--	--	--	--	--	--
SNNP	0.05	0.04	--	--	--	--	--	--
India AP: Coastal region	--	--	0.33	0.20	--	--	--	--
Rayalaseema	--	--	0.31	0.34	--	--	--	--
Telangana	--	--	0.35	0.46	--	--	--	--
Peru: Coastal region	--	--	--	--	0.52	0.23	--	--
Mountain region	--	--	--	--	0.31	0.64	--	--
Jungle region	--	--	--	--	0.17	0.13	--	--
Vietnam: Central Coast	--	--	--	--	--	--	0.64	0.17
Northern Highlands	--	--	--	--	--	--	0.36	0.83
Rural area	0.90	0.89	0.68	0.80	0.17	0.38	1.00	1.00
More than 15 minutes trip to school	--	--	0.29	0.36	0.24	0.30	0.40	0.61
Language: Speaks national language or dor	0.90	0.89	0.83	0.91	1.00	0.71	--	--
Family attitudes towards school								
Parental school efficacy	0.55	0.51	0.77	0.71	0.73	0.69	--	--
Believes child will achieve high ed	0.96	0.92	0.86	0.82	0.97	0.92	0.72	0.78
Child had high school absences	0.13	0.12	0.17	0.16	0.09	0.06	--	--
Believes school is essential	--	--	--	--	--	--	0.94	0.74
Participates in extra classes	--	--	--	--	--	--	0.39	0.12
Hours spent working	4.18	4.69	1.85	2.23	4.11	5.52	2.66	3.57
Standardized Height for Age	-1.53	-1.21	-1.50	-1.70	-1.19	-1.61	-1.57	-2.26
Child specific characteristics								
Boy	0.49	0.55	0.51	0.50	0.53	0.54	0.51	0.49
Age in months centered	-1.35	-1.08	-0.30	-0.04	-0.90	-0.79	-0.41	-0.22

Table 3. Ethiopia: Regression Models Predicting Academic Achievement by Group

	Literacy Age 8			Literacy Age 12			Math Ability Age 12			Total Grades Age 12						
	Amhara β	Oromo β	Rob SE	Amhara β	Oromo β	Rob SE	Amhara β	Oromo β	Rob SE	Amhara β	Oromo β	Rob SE				
Family endowments																
Wealth index	0.05	0.32	0.40	0.23	0.19 ^d	0.11	0.53	0.31	3.19 ^a	0.67	2.79	1.68	1.10	0.65	-0.05	0.93
Mom attended some school	0.09 ^b	0.03	0.05 ^d	0.02	0.03	0.03	0.10 ^d	0.05	0.32 ^c	0.15	0.58 ^a	0.10	0.11	0.23	0.45 ^b	0.12
Mom completed primary school	0.15 ^c	0.08	0.29 ^b	0.08	-0.07	0.06	0.09	0.07	0.49 ^d	0.25	0.89 ^c	0.33	0.14	0.32	1.22 ^d	0.69
Children in family = 3	0.10 ^d	0.05	-0.10	0.08	0.07 ^c	0.03	0.04	0.03	0.21	0.20	-0.57	0.31	0.36	0.27	0.07	0.41
Children in family = 4+	0.10	0.06	-0.03	0.06	-0.01	0.03	0.01	0.02	0.34 ^d	0.18	-0.19	0.26	0.00	0.27	0.05	0.40
Parental death	0.02	0.03	0.03	0.06	-0.08	0.05	-0.05	0.07	-0.12	0.18	0.07	0.19	-0.14	0.19	-0.15	0.17
Geography (Addis Abba is ref)																
Amhara	-0.11 ^d	0.06	-0.31 ^b	0.08	0.05	0.03	-0.03	0.08	-0.35 ^c	0.14	-2.75 ^b	0.62	0.88 ^a	0.17	-0.14	0.45
Oromia	-0.12	0.15	-0.20 ^b	0.06	0.00	0.19	-0.06 ^d	0.03	-0.41	0.56	-1.32 ^a	0.25	1.04	0.65	-0.03	0.26
SNNP	-0.05	0.06	-0.25 ^b	0.06	-0.03	0.04	-0.07	0.05	-0.32	0.22	-0.74 ^b	0.23	-0.12	0.30	-1.17 ^a	0.24
Rural area	-0.27 ^a	0.06	-0.04	0.05	-0.10 ^b	0.04	0.01	0.03	-0.14	0.10	0.06	0.30	-0.94 ^c	0.41	-1.32 ^b	0.32
Greater than 15 min to school	0.00	0.04	0.02	0.03	0.03	0.03	0.02	0.06	0.29	0.27	0.18 ^b	0.05	-0.34 ^b	0.09	0.06	0.15
Language																
Speaks dominant local language	0.12	0.14	0.03	0.05	0.19	0.19	-0.04	0.06	0.69	0.52	0.14	0.41	0.80	0.64	0.19	0.11
Family attitudes towards school																
Parental school efficacy	0.03	0.03	-0.01	0.04	0.10 ^a	0.02	0.03	0.04	0.06	0.15	0.16	0.26	0.23	0.18	0.19	0.20
Believes child will achieve high ed	0.02	0.08	0.00	0.07	0.12	0.18	0.11	0.08	-0.63 ^d	0.31	0.36	0.26	0.14	0.24	0.13	0.31
Child had high school absences	-0.01	0.04	-0.01	0.05	0.07	0.05	0.05	0.07	-0.15	0.29	0.04	0.25	-0.41	0.30	0.03	0.39
Hours spent working	-0.03 ^c	0.01	-0.02 ^d	0.01	-0.02 ^b	0.01	-0.01	0.01	-0.03	0.06	-0.05	0.09	-0.15 ^c	0.06	-0.21 ^a	0.05
Standardized Height for Age	0.06 ^a	0.01	0.04 ^d	0.02	0.03 ^b	0.01	0.03	0.03	0.17 ^b	0.05	0.10	0.12	0.50 ^a	0.06	0.33 ^b	0.08
Child specific characteristics																
Boy	-0.05	0.04	0.06	0.04	0.03	0.03	-0.01	0.06	0.34 ^d	0.17	0.12	0.22	-0.22	0.15	-0.12	0.27
Age in months centered	0.00	0.01	0.01	0.01	0.01	0.00	-0.01 ^c	0.00	0.03	0.02	-0.02	0.03	0.07	0.03	-0.03	0.01
Constant	0.57	0.17	0.47	0.21	0.60	0.29	0.60	0.22	-0.63	0.65	-0.54	0.83	4.96	0.89	5.48	0.48
N	250	195	195	254	193	193	193	251	251	195	195	258	199	199	199	199
R-squared	0.45	0.42	0.42	0.25	0.16	0.16	0.16	0.30	0.30	0.29	0.29	0.35	0.35	0.35	0.48	0.48

a p ≤ 0.001, b p ≤ 0.01, c p ≤ 0.05, d p ≤ 0.10

Table 4. India, Andhra Pradesh State: Regression Models Predicting Academic Achievement by Group

	Literacy Age 8			Math Ability Age 12			Total Grades Completed Age 12				
	Other	SC/ST	Other	Other	SC/ST	Other	Other	SC/ST	Other		
	β	Rob SE	β	β	Rob SE	β	Rob SE	β	Rob SE		
Family endowments											
Wealth index	0.31 ^b	0.10	-0.15	0.22	1.62 ^c	0.67	0.72	1.48	0.24	0.76	1.00
Mom completed primary school	0.06	0.04	-0.07	0.08	0.23	0.25	0.58	0.41	0.09	-0.37	0.46
Mom completed secondary school	0.08	0.05	0.29 ^b	0.10	0.43 ^d	0.22	2.46 ^b	0.68	0.10	1.18	0.78
Children in family = 3	-0.03	0.03	0.08	0.05	-0.14	0.10	0.46 ^c	0.21	0.08	0.02	0.23
Children in family = 4+	-0.06	0.04	-0.06	0.08	-0.27 ^c	0.13	-0.31	0.38	0.12	0.21	0.26
Attended private school	0.05	0.04	0.10	0.09	0.12	0.25	0.20	0.36	0.13	-1.01 ^c	0.43
Geography (Coastal region is reference)											
Rayalaseema	-0.04	0.05	-0.03	0.09	0.03	0.29	0.04	0.39	0.23	0.00	0.28
Telangana	0.02	0.05	-0.04	0.07	-0.04	0.35	0.12	0.40	0.20	-0.82 ^c	0.33
Rural area	0.00	0.05	-0.10	0.14	-0.15	0.46	0.69	0.50	0.16	0.04	0.32
More than 15 minute trip to school	0.02	0.03	0.01	0.06	0.13	0.14	0.26	0.24	0.09	0.15	0.26
Language											
Telugu	0.05	0.03	0.11	0.17	0.36 ^d	0.21	-0.47 ^d	0.24	0.15	0.56	0.63
Family attitudes towards school											
Parental school efficacy	0.01	0.03	0.04	0.07	0.79 ^a	0.16	1.12 ^b	0.35	0.12	0.07	0.28
Believes child will achieve high education	0.03	0.05	0.15 ^b	0.05	0.51 ^c	0.21	0.06	0.40	0.22	0.71 ^c	0.29
Child had high school absences	-0.02	0.03	-0.06	0.05	0.12	0.11	-0.39	0.31	0.09	0.15	0.21
Hours spent working	-0.02 ^b	0.01	-0.01	0.01	-0.07 ^d	0.04	-0.02	0.07	0.03	-0.24 ^b	0.08
Standardized Height for Age	0.02	0.01	0.06	0.04	0.01	0.08	0.08	0.15	0.03	0.09	0.10
Child specific characteristics											
Boy	0.04	0.03	-0.06	0.05	0.27 ^d	0.15	0.46	0.28	0.08	-0.05	0.15
Age in months centered	0.01 ^c	0.00	0.00	0.01	0.02	0.02	0.00	0.04	0.01	0.03	0.03
Constant	0.54	0.12	0.58	0.26	-1.33	0.74	-1.40	0.95	0.38	5.93	0.80
N	595	153	599	148	620	158	0.43				
R-squared	0.24	0.19	0.24	0.27	0.55	0.43					

^a p ≤ 0.001, ^b p ≤ 0.01, ^c p ≤ 0.05, ^d p ≤ 0.10

Table 5. Peru: Regression Models Predicting Academic Achievement by Group

	Literacy Age 8			Math Ability Age 12			Total Grades Completed Age 12					
	Less Indigenous More Indigenous	β	Rob SE	Less Indigenous More Indigenous	β	Rob SE	Less Indigenous More Indigenous	β	Rob SE			
Family endowments												
Wealth index	0.11	0.09	0.23 ^b	0.07	0.94 ^c	0.35	0.87	0.62	1.44 ^a	0.35	1.48 ^a	0.36
Mom completed primary school	0.01	0.04	0.12 ^c	0.06	0.08	0.19	-0.08	0.21	0.21	0.16	0.00	0.15
Mom completed secondary school	0.05	0.04	0.17 ^a	0.05	0.27	0.22	0.31 ^d	0.16	0.34 ^c	0.15	-0.06	0.18
Children in family = 3	-0.02	0.03	-0.08	0.06	-0.35 ^d	0.19	-0.43 ^c	0.18	-0.09	0.09	-0.39 ^c	0.18
Children in family = 4+	-0.08 ^b	0.03	-0.04	0.05	-0.31	0.19	-0.48 ^b	0.13	-0.25 ^c	0.12	-0.47 ^b	0.16
Geography (Coastal region is reference)												
Mountain region	0.03	0.04	-0.08	0.05	0.20	0.20	-0.14	0.21	0.06	0.13	0.10	0.14
Jungle region	0.07	0.05	0.00	0.07	0.28	0.17	-0.42	0.25	0.09	0.23	0.10	0.16
Rural area	-0.07	0.05	-0.01	0.10	-0.47 ^c	0.20	-0.20	0.32	0.07	0.26	-0.06	0.27
More than 15 minute trip to school	-0.02	0.03	-0.03	0.05	-0.15	0.17	0.03	0.17	-0.19	0.15	-0.08	0.15
Language												
Spanish was child's first language	---	---	-0.01	0.09	---	---	0.23	0.22	---	---	0.12	0.18
Family attitudes towards school												
Parental school efficacy	0.01	0.04	0.03	0.04	0.44 ^c	0.18	0.13	0.10	0.15	0.12	-0.03	0.13
Believes child will achieve high education	0.06	0.07	0.02	0.09	0.45	0.27	0.45 ^d	0.24	-0.02	0.26	0.49 ^d	0.28
Child had high school absences	-0.06	0.07	-0.02	0.09	-0.17	0.19	-0.16	0.30	-0.37	0.25	-0.39	0.30
Hours spent working	0.00	0.01	0.00	0.01	-0.06 ^d	0.04	-0.07 ^c	0.03	-0.02	0.03	-0.04	0.03
Standardized Height for Age	0.03 ^d	0.02	0.01	0.03	-0.04	0.07	0.08	0.64	0.17 ^b	0.05	0.03	0.06
Child specific characteristics												
Boy	-0.03	0.03	-0.02	0.03	-0.12 ^d	0.07	0.23	0.15	-0.23 ^c	0.10	-0.14	0.14
Age in months centered	0.01 ^b	0.00	0.01	0.01	-0.01	0.01	0.02	0.20	0.09 ^a	0.01	0.07 ^a	0.01
Constant	0.79	0.11	0.66	0.16	-0.36	0.45	-0.09	0.58	5.47	0.23	5.25	0.49
N	377		256		388		271		393		276	
R-squared	0.17		0.17		0.15		0.25		0.35		0.30	

^a p ≤ 0.001, ^b p ≤ 0.01, ^c p ≤ 0.05, ^d p ≤ 0.10

Table 6. Vietnam: Regression Models Predicting Academic Achievement by Group

	Literacy Age 8				Math Ability Age 12				Total Grades Completed Age 12			
	Kinh		Ethnic Minority		Kinh		Ethnic Minority		Kinh		Ethnic Minority	
	β	Rob SE	β	Rob SE	β	Rob SE	β	Rob SE	β	Rob SE	β	Rob SE
Family endowments												
Wealth index	0.15 ^b	0.04	0.84 ^c	0.30	-0.40	0.66	-2.05	1.42	0.81 ^c	0.32	1.40	0.87
Mom completed primary school	-0.04	0.02	0.27	0.14	0.30 ^d	0.31	0.72	0.40	-0.13	0.10	0.30	0.25
Mom completed lower secondary school	0.02	0.03	0.29 ^c	0.09	0.57 ^c	0.27	0.67	1.15	-0.07	0.06	0.52	0.29
Children in family = 3	0.01	0.02	-0.15	0.07	-0.47	0.18	0.11	0.23	0.21 ^c	0.07	0.19	0.23
Children in family = 4+	-0.05	0.05	-0.27	0.14	-0.24	0.29	-0.65	0.35	-0.06	0.09	-0.44	0.26
Geography (Central Coast is reference)												
Northern Highlands	0.04 ^b	0.01	-0.01	0.05	-0.30 ^d	0.16	0.75	0.42	0.18 ^c	0.07	0.82 ^c	0.19
More than 15 minutes to get to school	0.02	0.02	0.07	0.05	-0.26	0.19	0.61 ^c	0.20	0.01	0.09	0.46 ^c	0.11
Language												
Child speaks Vietnamese	---	---	0.18 ^b	0.06	---	---	1.2 ^d	0.51	---	---	0.57 ^c	0.13
Family Attitudes About School												
Believes school is essential	0.03	0.08	0.12	0.06	0.06	0.42	0.30	0.19	0.08	0.11	0.17	0.23
Believes child will achieve high education	0.03	0.03	-0.01	0.04	0.59 ^c	0.20	0.40	0.20	0.22 ^d	0.10	0.64	0.40
Participates in extra classes	0.01	0.02	0.07	0.05	-0.40	0.24	0.08	0.30	0.17 ^d	0.08	-0.15	0.20
Hours spent working	-0.02 ^c	0.01	-0.01	0.01	-0.14 ^c	0.06	-0.12	0.09	0.00	0.03	-0.03	0.10
Standardized Height for Age	0.04 ^d	0.02	0.04 ^d	0.01	0.12	0.13	0.29 ^b	0.05	0.08 ^c	0.03	0.39 ^d	0.17
Child specific characteristics												
Boy	-0.03	0.02	-0.04	0.06	-0.08	0.11	-0.19	0.28	0.05	0.07	0.53 ^c	0.18
Age in months centered	0.00	0.00	0.01	0.02	0.01	0.02	0.07	0.06	0.07 ^a	0.01	0.10 ^c	0.02
Constant	0.91	0.09	0.41	0.17	1.61	0.47	-0.66	0.77	6.07	0.33	4.41	0.36
N	267		107		263		103		264		110	
R-squared	0.16		0.53		0.11		0.41		0.27		0.41	

^a $p \leq 0.001$, ^b $p \leq 0.01$, ^c $p \leq 0.05$, ^d $p \leq 0.10$

