DISCUSSION PAPER SERIES

DP17344

The Rise and Fall of SES Gradients in Heights around the World

Elisabetta Aurino, Adriana Lleras-Muney, Alessandro Tarozzi and Brendan Tinoco

DEVELOPMENT ECONOMICS

CEPR

ISSN 0265-8003

The Rise and Fall of SES Gradients in Heights around the World

Elisabetta Aurino, Adriana Lleras-Muney, Alessandro Tarozzi and Brendan Tinoco

Discussion Paper DP17344
Published 31 May 2022
Submitted 30 May 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

The Rise and Fall of SES Gradients in Heights around the World

Abstract

We use data from a large sample of low and middle income countries to study the association (or "gradient") between child height and maternal education. While the strong positive association between child health and measures of parental socio-economic status (SES) is well established, we uncover novel results regarding the evolution of this gradient as children age. The association is small at birth, rises throughout childhood and declines in adolescence as girls and boys approach puberty. This pattern is consistent with a degree of catch up in height among children of low SES families, in partial contrast to the argument that height deficits cannot be overcome after the early years of life. This catch up is partly explained by the association between SES and the timing of puberty and therefore of the adolescent growth spurt. By contrast, we do not find evidence in support of the role of behavioral responses in driving the inverted U-shape of the gradient.

JEL Classification: I14, I15, O15

Keywords: height, socio-economic status, Maternal Schooling

Elisabetta Aurino - e.aurino@ub.edu
Universitat de Barcelona

Adriana Lleras-Muney - alleras@econ.ucla.edu
UCLA

Alessandro Tarozzi - alessandro.tarozzi@gmail.com
European University Institute, Universitat Pompeu Fabra and CEPR

Brendan Tinoco - btinoco@nd.edu
Wilson Sheehan Lab for Economic Opportunities, University of Notre Dame

The Rise and Fall of SES Gradients in Heights around the World

Elisabetta Aurino
Adriana Lleras-Muney
Alessandro Tarozzi
Brendan Tinoco*

May 2022

Abstract

We use data from a large sample of low and middle income countries to study the association (or "gradient") between child height and maternal education. While the strong positive association between child health and measures of parental socio-economic status (SES) is well established, we uncover novel results regarding the evolution of this gradient as children age. The association is small at birth, rises throughout childhood and declines in adolescence as girls and boys approach puberty. This pattern is consistent with a degree of catch up in height among children of low SES families, in partial contrast to the argument that height deficits cannot be overcome after the early years of life. This catch up is partly explained by the association between SES and the timing of puberty and therefore of the adolescent growth spurt. By contrast, we do not find evidence in support of the role of behavioral responses in driving the inverted U-shape of the gradient.

JEL: I14, I15, O15

Key words: Height, Socio-Economic Status, Maternal Schooling

*We are very grateful to Bo Honoré and to several participants of conferences and seminars for comments, to the research and survey teams who created the CLHNS, DHS, and YLS data sets for allowing free access to their outstanding data and to Sungwoo Cho for excellent research assistance. Please note that we do not hold the right to share the data, but upon request we can share all codes necessary to (a) re-create all the necessary datasets starting from the raw data that need to be obtained from the original data repositories and (b) to replicate all results, including figures and tables. All errors and omissions are our own.

Contacts: Aurino, University of Barcelona, E-mail e.aurino@ub.edu; Lleras-Muney, UCLA and NBER, allerasmuney@gmail.com; Tarozzi: European University Institute, Universitat Pompeu Fabra, Barcelona School of Economics and CEPR, alessandro.tarozzi@gmail.com; Tinoco: Wilson Sheehan Lab for Economic Opportunities, University of Notre Dame btinoco@nd.edu.

1 Introduction

A well-established literature documents the ubiquitous strong association (or "gradient") between different individual measures of health and socio-economic status (SES), both within and across countries (Strauss and Thomas 1998, 2008, Cutler et al. 2006). Richer and more educated individuals are healthier and live substantially longer lives, and children of high-SES parents enjoy better health and lower mortality rates in both rich and poor countries. Key questions in this literature are when these gradients emerge, how they evolve over the lifetime, and, most importantly, whether they are malleable—i.e. the extent to which children that are born and grow up in disadvantaged backgrounds can partially or fully catch up in terms of their health outcomes (Case et al. 2002, Martorell et al. 1994).

In this paper, we study the relationship between parental SES and children's heights and how it evolves from birth into young adulthood, using high-quality individual-level data from a large number of low and middle income countries (LMICs). We offer the first evidence of an inverted U-shaped age profile of the height-SES profile during childhood and adolescence. We find that SES-based differences in height are small at birth, but they become progressively larger during childhood. However, while remaining positive, the gradient *decreases* during the adolescent years, which suggests a degree of catch-up of low-SES children relative to high-SES ones. Using a novel empirical model of human growth from early childhood to early adulthood, we show that the inverted U-shape can be explained by variation across SES groups in the timing of the onset of adolescence—when the second growth spurt occur—as well as in when adult height is achieved.

We focus on height, instead of other commonly used measures of health used in the literature such as self-reported status, or presence of conditions. Aside from genetic factors, height is primarily determined by the availability and diversity of nutrients, and the prevalence of disease (Tanner 1989, Martorell and Habicht 1986, Steckel 1995). Indeed, economic historians have often used adult height as an indicator of economic or human development (Fogel 1994, Steckel 1995, 2009). As a health indicator, height has multiple advantages. First, it is relatively easy to measure objectively, and does not suffer from reporting biases. Second, height is a widely available health indicator for both children and adults in LMICs, and importantly it is easily comparable across all age groups. Third, height is a good measure of overall health that correlates with other objective measures of health, such as disease incidence and mortality (Fogel 1994, Steckel 1995, 2009, Perkins et al. 2016). Finally, height is an important predictor of economic outcomes in adulthood and across generations. On average, taller individuals have more human capital and earn higher wages, an association that is likely mediated by several determinants, including physical strength (Haddad and Bouis 1991, Strauss and Thomas 1998), social factors (Persico et al. 2004) and cognitive ability (Case and Paxson 2008). In addition, transmission of low height from parents to their children has been identified as one of the drivers of substantial persistence in SES inequalities in human capital across generations in both high- and low-income settings (Akbulut-Yuksel and Kugler 2016, Behrman et al. 2017).

We use maternal education as an indicator of parental SES. In LMICs parental education is both more readily available than income or consumption data, and sometimes more reliable as a measure of long-term resources than income measures. In economies with a large share of employment in the agricultural and informal sectors, measures of income are often not available and when they are, they can be measured with considerable error (Deaton and Grosh 2000). Maternal education is a well known correlate of child health (Caldwell 1986, Heath and Jayachandran 2017). Studies in LMICs find support for the notion that this association is often causal, see for example Grépin and Bharadwaj (2015) or Andriano and Monden (2019).

We start our investigation of the gradient by using data from the Demographic and Health Surveys (DHS) on about 1.6 million children under five years of age born in 1981-2018 in 73 LMICs. In these data, the cross sectional association between child height and maternal schooling is small and insignificant at birth but it increases steeply between birth and five years of age. Although DHS data do not include height for children older than five, most surveys also record height of women from the age of 15 onward and for adolescents who have not yet left their family of origin, so it is possible to link their height to their mother's education. In this (potentially selected) sample of adolescents, the association between height and maternal education, while still substantive and statistically significant, is much smaller than for children around five years of age. This suggests that the gradient increases monotonically until a certain age but then declines.

To better evaluate the age-profile of the height-SES gradient we use panel data from five LMICs where we can follow cohorts of individuals from birth until young adulthood. We use data from two cohorts in Ethiopia, India, Peru and Vietnam from the Young Lives study (YLS hereafter, Barnett et al. 2013), and from the Philippines' Cebu Longitudinal Health and Nutrition Survey (CLHNS, Adair et al. 2010). These data confirm the existence of a consistently positive relationship between height and maternal education. As in the DHS data, the strength of the association has an inverted U-shape, increasing first but then decreasing in adolescence, with the decline taking place earlier for girls and later for boys. Our findings are very similar if we use alternative measures of SES, or height z-scores.

Next, we investigate biological and behavioral explanations for these patterns. Regarding the potential role of biological factors, we hypothesize that, for both boys and girls, the inverted U-shape of the gradient can be explained by the link between SES and the onset and duration of the adolescent growth spurt (AGS). This hypothesis is based on two documented patterns. First, in many countries there has been a well-established secular decline in the age at menarche among girls linked to overall improvements in socio-economic conditions and health (Hauspie et al. 1996, de Muinck Keizer-Shrama and Mul 2001, de La Rochebrochard 2000). The same considerations suggest the existence of a cross-sectional negative association between age at menarche and SES in low-income settings, an association that indeed has been documented in the Philippines (Adair 2001) and is confirmed in our data. Low-SES children will thus reach the peak of their AGS when high-SES are already past theirs, allowing them a degree of catch-up. Second, it has been observed that low-SES children achieve on average their adult height at older ages (Steckel 1986, Bozzoli et al. 2009). Based on these insights, we propose and estimate a growth model that rationalizes differential SES profiles of human growth, and we show that the results strongly support this hypothesis.

We also explore the potential role of behavioral responses in driving the inverted U-shape of the gradient. We hypothesize that taller adolescents may start adult life earlier in ways that may be detrimental to further growth in stature. For instance, taller boys may start working at younger ages, and taller, sexually mature girls may marry and have children at a younger age. Indeed previous research documents that the age at menarche predicts marriage rates and education levels among girls (Field and Ambrus 2008, Khanna 2020). Such behavioral responses may impose a 'nutritional cost' that could be detrimental to physical growth, especially if adolescents are still far from having achieved their adult height, and could explain the inverted U-shape pattern if they matter differently for children of low vs. high-SES. However, when we use longitudinal data from YLS we find limited evidence in support of such mechanisms.

Our results add to different strands of literature. First, much research has studied the emergence and evolution of the SES health gradient. In a seminal paper, Case et al. (2002) documented that, in the United States, the correlation between indicators of general health status and a measure of long-term income originates in childhood and becomes progressively steeper into adulthood. Similar results have been found for Canada (Currie and Stabile 2003), Australia (Khanam et al. 2009), the Czech Republic (Borga et al. 2021), and in other US data sets (Murasko 2008, Fletcher and Wolfe 2014). However, using data from the UK Census, West (1997) shows that SES-gradients in various health measures flatten among children or sometimes reverse for adolescents, only to re-appear in adulthood (see also Ward and Viner 2018 and references therein). In Germany, Reinhold and Jürges (2012) find that the gradient decreases while children are ages 9-12. Several studies in LMICs do not find that the gradient increases with age though it is not clear why (e.g. see Cameron and Williams 2009 and Park 2010 for Indonesia, and Sepehri and Guliani 2015 for Vietnam). In addition, several studies report different patterns around adolescence when the gradient appears to flatten, at least in some contexts (Currie et al. 2007, Propper et al. 2007, West 1997, Reinhold and Jürges 2012). We contribute to this literature with data from many LMICs, documenting an inverted U-shape in the age profile of the SES gradient in height (rather than self reported measures), and exploring mechanisms.[1]

Second, we present a novel methodology to estimate the shape of the growth curve with longitudinal data where height is only measured at infrequent intervals, such as YLS or CLHNS. The model links (unobserved) growth velocity at high frequency to (observed) height measured at low frequency. The model approximates the typical pattern of growth velocity in humans, which is highly non-linear, see Tanner et al. (1966, Fig. 8). We approximate this pattern with a piece-wise continuous linear function, where the kinks coincide with key transitions in growth velocity (such as the beginning of the AGS, or its peak), and may depend on SES. We show that this model can be estimated using constrained ordinary least squares, with the location of the kinks determined by a simple algorithm in the spirit of Hansen (2017). Our method differs from alternative non-linear models that have been proposed in the literature, see Preece and Baines (1978), Sayers et al. (2013), Beath (2007) and Cole et al. (2010).

[1]The growing association between height and SES during early childhood is also broadly consistent with the findings in Aiyar and Cummins (2021), who find that the gradient between stature and GDP at birth starts very small but then increases with age among young children in pooled data from the DHS.

These approaches are best suited to model individual growth patterns with longitudinal data that include height measurements taken with high frequency, which are rare and expensive to collect. In addition, such models have been validated for the description of height growth velocity around the timing of puberty, while we are interested in the whole age profile of growth velocity, including the early years and the time when adult height is achieved.

Third, we contribute to a growing literature on the potential role of adolescence as a window of opportunity for catch-up in health outcomes. A large body of research has found that the timing of health investments and events during the life cycle matters, with circumstances in early childhood or *in utero* having long-run effects on health, human capital and incomes (Strauss and Thomas 2008, Almond et al. 2018). The potential for catch-up in linear growth retardation after the first 1,000 days is widely considered to be limited, although most studies do not follow children until adulthood (Martorell et al. 1994, Leroy et al. 2020). We show that height gaps between low- and high-SES children in our study populations narrow during and after adolescence. This is consistent with Martorell et al. (1994)'s view that the potential for catch-up growth increases with delayed maturation and a longer growth period. This is also consistent with a recent but growing literature from both economics and developmental science that documents the importance of adolescence for height and human capital more generally, see Akresh et al. (2012), van den Berg et al. (2014), Carneiro et al. (2019), and Andersen et al. (2021).

The rest of the paper is organized as follows. Section 2 describes the data; Section 3 describes the results; Section 4 explores mechanisms; and finally, Section 5 concludes.

2 Data and Measurement

We use data from a large number of surveys that broadly belong to three separate data collection initiatives: the DHS, YLS, and the CLHNS. In this section we provide some details on these data sources, and describe our main variables of interest.

2.1 Data

Demographic and Health Surveys (DHS). The primary purpose of these cross-sectional household surveys is to provide a detailed snapshot of each country surveyed, with a focus on demography, health, and fertility choices and preferences. Data are typically nationally representative and comparable across surveys. The primary respondents are women—in some cases only if ever married—'of fertility age', defined as 15-49. Detailed information is also available for their children under the age of five years, often including measurements on weight and height taken by trained enumerators.[2] Several of the more recent surveys also include detailed information on adult men.

We make use of all data available at the time of writing that contain information on child height. For children under five we drop less than 0.2% of observations for which height was $<$ 30 cms or

[2]In a small number of cases there is some variation in the target population. For instance, the 2004 Bangladesh DHS interviewed ever-married women 13-49, while in India only children below 4 were included in 1992-93 and only the last two births below three years of age were included in 1998-99. We ignore these differences.

> 1.4 m, that is, very likely measured with error. Table A.1 in the Appendix includes a complete list of all the surveys we use together with selected summary statistics on height. We restrict attention to children with non-missing anthropometric measures and maternal education. Overall, our data include height measurements for about 1.6 million children born in 1981-2018 from 245 surveys and 73 countries.

Young Lives (YLS). YLS is an international longitudinal study of childhood poverty conducted in four countries: Ethiopia, India (only in the state of Andhra Pradesh, part of which in 2014 was separated into a new state, Telangana), Peru and Vietnam. While the sample was not designed to be nationally representative (or, in the case of India, state representative), comparison of key child outcomes or socio-economic variables to those collected in nationally representative surveys show similar patterns and variations (Barnett et al. 2013).

The study follows two cohorts of children in each country since 2002, totalling roughly 12,000 children, over 15 years. Children in the younger cohort were first sampled in 2002 at ages 6-18 months and subsequently surveyed and measured in 2006, 2009, 2013 and 2016, at about 5, 8, 12 and 15 years of age, respectively. The older cohort was around 8 years of age in 2002, and then about 12, 15, 19 and 22 years old at the following survey rounds of in-person data collection. Attrition in this panel is low, around 10% over 15 years, with some variation across cohorts (younger cohort: 8%; older cohort: 16.5%) and countries (Ethiopia: 14%; India: 7%; Peru: 14%; Vietnam: 9%).[3] We limit our sample to individuals that were present at all rounds, but results are very similar when we consider the full cross-sectional sample in each round. We drop individuals with any missing data in any of the waves for heights (3% of the panel sample). The final analysis sample contains around 7,195 children for the Younger Cohort, and 2,991 children for the Older cohort. Panel A in Appendix Table A.2 shows summary statistics for these data.

Cebu Longitudinal Health and Nutrition Survey (CLHNS). The CLHNS is a panel data set of mothers and children from the Philippines' Metropolitan Cebu area originally designed to study how different infant feeding patterns in early life directly affect various health and socioeconomic outcomes in the lives of the mother, child, and household (Adair et al. 2010). The CLHNS surveyed—using a clustered design—a cohort of women sampled from both urban and rural communities (or *barangays*) who gave birth between May 1983 and April 1984. The baseline survey collected information about the mother's behaviors during pregnancy, demographics, socioeconomic status, as well as information on other household members. The initial sample included 3,080 non-twin live births. These children were measured at birth, then regularly at the end of every subsequent two-months period following their birth up until roughly 2 years of age. Over time, the study added more follow-up surveys and evolved into a longitudinal study of the long-term health outcomes of the children. The children were measured again in 1991, 1994, 1998, 2002 and 2005, when they were roughly 7, 10, 14, 18 and 21 years

[3] Socio-economic variables such as household wealth index, parental education, household size or child height-for-age z-scores at round 1 are not predictive of attrition, and the only variable that is significantly and negatively associated with the probability of being in the panel after 15 years is being urban in the first round of data collection.

of age respectively.[4] The rate of attrition was higher than in the YLS, at 33% from birth until 2005.[5] Again we limit our sample to children with non-missing maternal education and height measurements in all waves. Summary statistics are displayed in Panel B of Appendix Table A.2.

2.2 Height, SES and other variables of interest

Heights. The key dependent variable in all our regressions is height measured in centimeters. Much of the literature on child height uses 'z-scores', that is, measures of height standardized relative to a reference population. We prefer employing raw height in our estimates given that our focus is on the evolution of the gradient from childhood to early adulthood, and z-scores for height can only be constructed for individuals up to age 19 years.[6] Nevertheless, we check the robustness of our results by using z-scores for children. Table A.1 in the Appendix shows that a large fraction of children in the countries we study are shorter than children in the reference populations, leading to high prevalence of stunting, see also Ssentongo et al. (2021).

Education as a measure of SES. Maternal education, as reported by the mother herself in all surveys, is our main proxy of SES and long-term resources. While this is a coarse measure, it offers the advantage of being simple, fairly comparable across years and countries, and measured in all our data sources. Additionally, other measures such as income or consumption only capture resources and SES at a given point in time. Indeed, Case et al. (2002) use average income over a period of time as a proxy for 'permanent income'. More importantly, there are no consistent measures of income or consumption in our surveys. The DHS and YLS do include a wealth index, constructed using principal component techniques from data on asset ownership and availability of services such as electricity, improved toilets, and so on. This index could potentially capture permanent resources (Filmer and Pritchett 2001), but it is calculated separately in each survey and the list of assets is not identical across all surveys, so the resulting measures are not directly comparable between countries or, in the case of DHS, even within country over time. Moreover, maternal education is significantly correlated with other measures of resources or SES in surveys where different measures are available.[7]

[4]Two more surveys were conducted in 2007 and 2009, but children's heights were not measured, and so data from these rounds are not used in this paper.

[5]Similar to YLS, being in an urban community was significantly and negatively associated with the probability of being in the panel after 21 years. Unlike YLS, father's level of education is also associated with higher attrition, albeit with low predictive power.

[6]An additional issue that may potentially hamper the suitability of z-scores for our specific purposes relates to the change in growth references used to compute z-scores for children aged 0-5 and children aged 5-19 years when using WHO growth charts. For the former group, we use the 2006 WHO growth charts (World Health Organization 2008), which are based on a sample of children from different countries and ethnicity. The 2006 WHO growth charts were developed to replace the US National Center for Health Statistics (NCHS)/WHO standard, which had been recommended for international use since the late 1970s, when it became clear that this standard did not adequately represent the growth patterns of children from diverse ethnic backgrounds and cultural settings. By contrast, for children aged 5-19 years, we use charts from the WHO Reference 2007, which are based on the 1977 US NCHS/WHO reference. Although these are adapted to ensure a smooth transition around age 5, as described in de Onis et al. (2007), it is not clear how our results may be sensitive to a change in the reference standard from a multi-country to a US-based population.

[7]In the YLS data, the correlation of maternal education with total real *per capita* consumption expenditure is 0.2

We measure maternal education by constructing an indicator of whether the mother has completed at least secondary school. DHS measures both completed schooling and the number of years of schooling for each household member, so we define SES as a binary variable = 1 if the mother completed at least secondary schooling. In contrast, YLS only records the last grade completed, while CLHNS records the number of years completed in the most recent schooling level (i.e. three years of primary, four years of secondary, etc.). We use these variables to construct a SES indicator comparable to DHS, based on the number of years of schooling that each country requires for graduation from high school. In YLS, the binary variable for secondary education is thus set = 1 when the mother has completed a minimum of 10 years of schooling in Ethiopia, 12 in India, 11 in Peru, and 9 in Vietnam. In CLHNS, "secondary school" is = 1 if the mother has completed at least four years of secondary school at the time of the first survey wave. About 19, 18 and 23 percent of women have completed at least secondary education in the DHS, YLS and CLHNS respectively. Our results are robust to using years of education instead as well as other indicators of SES.

Other data We use information on age at menarche (the first occurrence of menstruation) from the panel data and from four countries in the DHS that report it: Gabon (2000), Ghana (1998), India (2015-16), and Turkey (2013). Appendix A.1 has more details on why data limitations in the DHS only allow us to focus on these countries, and on how we construct the samples for the analysis of age at menarche. Lastly, we use information in the YLS on behaviors during adolescence. Specifically we look at whether adolescents marry or have children, whether they sleep enough, work a lot, have a diverse diet or undertake risky behaviors (drinking and smoking). Appendix A.2 has more details on how we construct these variables.

3 Results

We start by documenting the key empirical pattern motivating our analysis: the steep rise of SES gradients during childhood and their subsequent decline around puberty in low- and middle-income countries. We show first the results using cross-sectional data from DHS, before moving to longitudinal data from YLS and CLHNS.

3.1 Empirical Strategy

For children of a given age, we estimate the SES gradient by estimating the following equation:

$$height_{iac} = \alpha_a + \beta_a \times MomEd_{iac} + \gamma_c + e_{iac} \tag{1}$$

where $height_{iac}$ is the height of child i at age a in country c, and $MomEd_a$ is an indicator equal to one if the mother has completed at least secondary education. We estimate this equation separately

($p < 0.001$). The correlation is also strong (0.47, $p < 0.001$) with a wealth index constructed as a composite indicator of asset ownership, access to services and housing quality. In the DHS surveys where a wealth index (variable `v191`) is reported, the correlation between the index and maternal schooling ranges between 0.22 and 0.30.

for each age a. We include dummy variables for each country (γ_c) but no other controls. The standard errors are clustered at each survey primary stage unit.

The coefficient of interest is β_a. It captures the SES gradient at a given age a, estimated as the difference in height between children whose mothers have at least secondary education and those whose mothers do not. We can estimate this gradient by pooling all countries, and also separately for each country.

3.2 Cross-sectional Results for Children under 5 from the DHS

Before turning to the regression results, we show how the non-parametric relationship between years of schooling and height changes with age, measured in years. Figure 1 presents age-specific associations between average height of boys and girls and maternal schooling. The categorical variable for maternal schooling distinguishes between no education, incomplete primary, complete primary, incomplete secondary, complete secondary, or higher. The figure shows two salient patterns. First, for both genders there is a clear positive association between average height and maternal schooling. Second, this pattern is much stronger at age 5 than at age 2, and it is barely visible at age 1. The lines rotate counterclockwise (become steeper), indicating that the association becomes stronger as children grow older.

We confirm these patterns by estimating our main model (eq. 1) for every age in months, and using a dummy for completed secondary schooling as the measure of education. Figure 2 plots the point estimates of the gradient together with 95% confidence intervals. The results are very similar between genders, with the gradient increasing almost monotonically with age. At birth the association between maternal education and height is small (less than 1cm) and not, or barely statistically significant. But one-year old children of mothers with secondary education are already about 1.5 cms taller than those born of mother with less schooling (95% C.I. [1.16, 1.98] for boys and [1.11,1.49] for girls). The gap increases to more than 2 cms at age 2 (95% C.I. [1.9,2.55] for boys and [2,3.15] for girls), and to almost 3 cms at age 3 (95% C.I. [2.45,3.53] for boys and [2.47,3.33] for girls). The gradient flattens out thereafter, especially for girls, though the data is noisier. The results are very similar if we use maternal years of education (Appendix Figure A.1) or alternative indicators of SES, such as paternal education (Appendix Figure A.2) or maternal height (Appendix Figure A.3), although both these alternative indicators reduce the number of observations with non-missing values.

The pattern of gradients increasing with age is also observed *within* countries. In Figure 3, we show box-plots of age and gender-specific coefficients estimated separately for each country. These graphs do not show confidence intervals as in Figure 2 but rather describe the distribution of the 73 country-level coefficients we estimate for each age and gender. The diamonds show the median coefficients while the darker central sections of the vertical lines show the inter-quartile ranges. The broader thinner lines show the whole variation excluding outliers, which are shown separately. The pattern of these box plots is similar to that for the estimated OLS slopes, and it also shows that the variation in coefficients increases with age. The median gradients start close to zero but then steadily increase until they reach about 4 cms by age 5.

The increase in the gradient with age is not a mechanical product of the increased scale of the dependent variable (height) when age increases.[8] In fact, the patterns remain almost identical if we use the logarithm of height as dependent variable, in which case the slope can be interpreted as the predicted proportional change in height associated with having a mother with at least secondary school (see Figure A.4 in the appendix).

The patterns described above remain similar, although somewhat flatter after age two, if we use 'height-for-age' z-scores instead of raw height as the dependent variable (see Appendix Figure A.5). Z-scores are a standardized measure that use, as a reference group, populations of children that received optimal nutrition and health inputs during pregnancy and in the first five years of life (World Health Organization 2006). The fact that our patterns remain robust to the use of z-scores is consistent with the well-known and typical age profile in low-SES populations of child height-for-age z-scores, which decline with age until about two years of age, and somewhat stabilize after that (Shrimpton et al. 2001): sub-optimal growth conditions generate a growth gap relative to the reference population that accumulates over time, especially during the first two years of life. A similar age profile has also been shown for the association between height z-scores and GDP at birth, see Aiyar and Cummins (2021).

Overall, the increase in the gradient with age for children under age five is very robust to how we define SES or the dependent variable. These results are broadly consistent with the findings in the seminal Case et al. (2002) study for the US, despite some important differences in how health and SES are measured in our study, and the very different economic context of the countries we investigate. Another key difference relative to Case et al. (2002) is that so far we have looked a narrow age range that only includes early childhood.

3.3 Results for Adolescents in the DHS

We now investigate if the gradients continue to increase after age 5. Ideally we would have height measured for all children and adults in the surveys. However, the DHS only measure heights for children under 5 and for women (in most surveys) and men (in some surveys) between 15 and 49 years old. In principle, this allows the analysis of the age profile of the gradient at age 15 or higher. In practice, this is only possible for very young individuals, because parental education is only recorded if the individual still co-resides with the parents. In addition, several DHS do not include identifiers for parents, and those that do almost exclusively do it for boys and girls younger than 18. This generates an obvious selection problem. Selection, however, should not be too severe among the youngest individuals, the large majority of which are still co-resident.[9]

[8] In a simple univariate OLS regression, if the scale of the dependent variable increases the slope will increase even if the correlation between the dependent variable and the regressor stays the same, as long as the standard deviation of the regressor does not change.

[9] In DHS surveys where young women and men can be linked to their mother (which is only possible in case of co-residency), maternal education is missing for 26-36% of observations. Among older individuals, maternal education is available for less than 10% of observations.

With these caveats in mind, in Table 1 we show the coefficients for maternal education for adolescents 15, 16 and 17 years old, separately by gender. For perspective, we also report estimates for children under five. The latter figures are estimated using the same sample used in Figure 2, but measuring age in years rather than months. When we look at teenagers, all but one of the estimated gradients are large and very precisely estimated, with magnitudes above 2 cm among both boys and girls (and standard errors around 0.1). The only exception is the coefficient for 15-year old boys, where the slope is 0.7 and not significant at standard levels. This result is apparently driven by the very low prevalence of high-SES mothers in this sub-sample (only 37 of 9,940), which generates very noisy estimates. With this exception, the age profile is fairly flat among both boys and girls.

Most interestingly, the estimated slopes are *smaller* (again, with the exception of 15-year old boys) than the corresponding coefficients for children age 4. This suggests a decline in the gradient in adolescence, but these comparisons are also complicated by the fact that not all DHS have data on adult heights, so comparisons between age groups may, in fact, be driven by differences in the countries or cohorts represented in each survey. However, the age profiles for children under 5 remain very similar if we only include observations from DHS where height was recorded for children as well as adults of both genders (results not shown). Perhaps more importantly, comparisons in the gradients between children 0-5 and adolescents are complicated by the cross sectional nature of these estimates. This implies that composition effects could in principle explain the differences in the findings.

3.4 Evidence from Panel Data

To better investigate whether the decline in the gradients during adolescence is real we now turn to longitudinal data. While such data allow us to follow the same children over time, they also have the drawback of forcing us to focus on a limited number of LMICs for which such data are available, and on a limited number of birth cohorts.[10] Nevertheless, panel data allow us to rule out composition effects and to investigate other aspects of the evolution of the gradient by age.

Tables 2 and 3 report estimates of the gradient by age using data from YLS (Ethiopia, India, Peru and Vietnam) and CLHNS (the Philippines), for girls and boys, respectively. Panels A.1 and A.2 show estimates for the younger and older cohorts of YLS, respectively at around ages 1, 5, 8, 12 and 15 years for the younger cohort, and 8, 12, 15, 19 and 22 years for the older cohort. The regression includes country dummies (as in model 1). In this specification we also control for age-in-months dummies to account for the fact that children were interviewed at slightly different ages in each wave.

Consistent with the results using DHS data, the patterns in the YLS show that the gradient has an inverted U-shape with age. In the younger cohort the gradient increases from 1.6 cm (about 2% of the average height) to 3.6 cm (about 3.4% of average height) between age one and five for children of the younger cohort. The gradient then continues to increase until 12 years of age reaching around 5cm for both boys and girls, something that we could not observe in the DHS due to the lack of height measurements in this age range. All slopes are estimated precisely, with standard errors in the

[10] We did not use other existing longitudinal data sets either because of small sample size or because the data are not made publicly available.

0.2-0.6 range, and all are statistically significant at the 1 percent level. Most interestingly, we now also observe that there is a sudden and substantial drop from 4.7 cm at 12 years to 2.3 cm at 15 years for girls, while the coefficient remains relatively stable for younger cohort boys. A similar patterns is also apparent in the older cohort, where the slope of the gradient increases monotonically between 8 and 12 years for both genders, but then declines from 3.4 cm to 2 cm for girls between 12 and 15 years. By contrast, it remains fairly stable among boys. After 15 years, the gradient keeps decreasing for girls, reaching 1.4 cm at age 22 years. For boys in the older cohort, the decline occurs after 15 years, with the coefficient moving from 5.1 cm to 3.1 cm between 15 and 19 years, and declines further to 2.7 cm at age 22. Estimates are similar but less precise if we estimate the regressions separately by country.[11,12]

The same pattern of inverted-U shapes is also evident in the CLHNS data from the Philippines, as shown in Panel B of Tables 2 and 3. In this sample, there is a monotone increase in the SES-gradient up to age 11 for both boys and girls, followed by a decline for girls from 3.8 cm at 11 years to 2.1 cm at 15 years and from 3.9 cm at 15 years of age to 2.7 cm at 18 years for boys. By age 21, when the large majority of individuals have reached their final adult height, the gradient is about 100 percent larger for boys as compared to girls but still significant for both.[13] For both genders, the gradient at age 21 is substantially smaller than the gradient at the onset of adolescence, when it reaches its peak. Given that girls, on average, reach sexual maturity earlier than boys – in LMICs, pubertal development occurs on average at age 13.5-15.5 among girls and about 2 years later among boys (Thomas et al. 2001) – these results suggest that the timing of the inversion of the age profile of the gradient takes place around puberty.[14]

4 Why Does the Gradient Decline in Adolescence?

We now discuss two hypotheses to explain why the gradients fall in adolescence. The first relates to the physiology of human growth: if high-SES children have an earlier adolescent growth spurt and stop growing earlier, then low-SES children may catch up. The second is that the onset of adolescence may lead to behavioral changes that affect later growth, and may do so differentially by SES. We now discuss these in more detail and provide some evidence for each. We find strong evidence in favor of the former hypothesis but not of the latter.

[11] Given that maternal years of education is time invariant, the slope in these regressions should not change once the child has achieved adult height, as long as height is measured consistently and the sample itself does not change due to attrition. However, in LMICs adult height is often achieved after age 20.

[12] Note that there is no reason to expect the gradients in Panels A and B to be identical conditional on age, given that the same age is reached in different years for the two cohorts. For instance, children in the young cohort were about 8 in 2009, while those in the older cohort were this age at the time of their first measurement, in year 2000.

[13] This larger gradient for boys is consistent with evidence suggesting that mortality among males is higher than for females during crises or conditions of extreme hardship, underlying a potential higher sensitivity of males— especially infant boys—to environmental inputs, see e.g. Drevenstedt et al. (2008) and Zarulli et al. (2018).

[14] The inverted U-shape pattern is also apparent if we use height-for-age z-scores instead of raw height, see Appendix Table A.3.

4.1 Pubertal maturation, SES and the Age Profile of the Gradient

The first hypothesis is that the increasing and then decreasing association between height and maternal education may be explained by the physiology of human growth, and SES-based variation in the timing and duration of such growth. Among girls, it is well known that the adolescent growth spurt precedes menarche—the onset of menstruation—by about one year, and that growth stops within the following year or two (Gluckman et al. 2016). It has also been observed that as economic conditions improve and nutritional intakes and dietary diversity increase, the onset of menarche occurs earlier (de Muinck Keizer-Shrama and Mul 2001, Lam 2021). Consistent with this, Thomas et al. (2001), summarizing results from 67 countries, find a strong negative association between average age at menarche and different measures of development, including female life expectancy and literacy rates. Simondon et al. (1998) use longitudinal data from 1,650 children in Senegal and show that girls who were stunted before schooling age had menarche later than non-stunted girls but their height grew faster—leading to some catch-up—in late adolescence.

Compared with girls, there is less evidence available for boys on the relationship between pubertal maturation and SES in LMICs. This is partly due to the greater challenges in measuring pubertal timing for boys in the absence of a clearly defined marker of pubertal maturation such as menarche.[15]

Given these insights, if within LMICs there is a negative association between age at pubertal maturation and measures of material well-being, [16] high-SES children will grow—on average—faster than their low-SES cohort peers both before and during the adolescent growth spurt, which they will reach, on average, sooner. At this point the gap between the high and low-SES children may reach a maximum. However, once low-SES children reach the adolescent growth spurt, a degree of catch up may take place, especially if physical growth continues well after adolescence or if pubertal maturation occurs very late. Indeed it has also been shown that poor or poorly fed populations grow more slowly and reach their final height at later ages (Steckel 1986). This "catch-up" mechanism may thus lead to a reduction in the height-SES gradient after puberty.

To investigate the association between SES and age at pubertal maturation, we start by examining data on age at menarche in DHS data using the four countries (Gabon, Ghana, India, and Turkey) where the data allow it.[17] For each of these countries, we estimate models such as eq. (1) but where the binary dependent variable is equal to one if the girl had menarche before age 13. Although these

[15]This evidence gap is equally marked for high-income settings. The only paper we are aware of is Sun et al. (2017), which documents an inverse relationship between socioeconomic disadvantage and pubertal maturation among boys in an Australian cohort. This is consistent with a wide body of evidence showing that in high-income settings *lower* SES predicts earlier maturation, the opposite of what we find in LMICs.

[16]From a biological perspective, the relationship between SES and pubertal onset and tempo may be mediated by recently-uncovered mutations in brain receptors that are activated by caloric deprivations in childhood (Lam 2021). In turn, these mutations are associated with delayed pubertal onset and reduced linear growth rate throughout childhood and adolescence, which are then partially offset by a longer period of limb growth due to a later pubertal onset, allowing for an extended period of growth.

[17]Age at menarche is available for several other DHS countries, but they cannot be linked to maternal SES due to the data structure. See Appendix A.1 for details.

countries differ considerably in their level of development, there is a *positive* association between early menarche and maternal education in all of them, although the coefficient is only statistically significant in India, and its magnitude is small for Turkey (Table 4, Panel A). In India, high maternal education increases the predicted probability of early menarche by 3 percentage points (95% CI [0.016,0.044]), relative to the mean (20%). In Gabon and Ghana, both very poor countries where fewer girls have already reached menarche at 12, the association is even stronger, although very imprecisely estimated and thus not significant at standard levels: in Gabon high maternal education predicts a 100% increase in the probability (from 20 to 39%, 95% CI of the change [-0.061,0.445]), while in Ghana is predicts a 228% increase (from 6.4 to 21%, 95% CI of the change [-0.04,0.333]). In Turkey, a wealthier context compared to the others in this sample, where average female education is also higher, the association is still positive but much weaker and not significant at standard levels.

The negative association between SES and age at menarche is confirmed when we use the longitudinal data from both cohorts from YLS and CHLNS, as reported in Table 4, Panel B. With the exception of Ethiopia, where the association is weak and not significant at standard levels, early menarche is substantively more likely among daughters of high-SES mothers, with point estimates ranging from 0.11 in Peru to 0.21 in Vietnam (95% C.I. [-0.03, 0.41] for India, [0.01,0.2] for Peru, [0.12, 0.31] for Vietnam, and [0.11, 0.25] for the Philippines). These coefficients are large relative to the respective means. These simple associations are of course not necessarily causal, but they are consistent with the hypothesis that high-SES girls grow faster and stop growing sooner. Ethiopia may be an exception due to the very low prevalence of early menarche among sample girls, at less than 4%.

The associations between SES and early menarche we observe in both DHS and panel data may be driven at least in part by a higher prevalence of overweight among high-SES girls, as excess adiposity in childhood is an important factor associated with earlier pubertal onset (Marcovecchio and Chiarelli 2013). Overweight girls in the pre-pubertal phase tend to grow faster than leaner peers, but this advantage in growth tends to decline during puberty, when overweight girls display a reduced growth spurt. This, again, could lead to a degree of catch-up in height among poorer girls, who are less likely to be overweight. We check whether taking into account overweight and obesity changes the point estimates for maternal education in both the DHS and YLS panel data for girls, but we do not find evidence that this is the case.

This evidence presented so far is, however, incomplete because we cannot link the onset of adolescence directly to the SES gradients in heights at various ages. To do this we now estimate a model of growth separately for boys and girls by SES.

4.1.1 A Model of the Age Profile of Growth Velocity and SES

In this sub-section we describe and estimate a simple model where both the timing and speed of height growth is a function of SES. We model the growth rate of heights assuming that it follows the well-known patterns described in the literature (e.g. see Tanner et al. 1966, Fig. 8, or Gluckman et al. 2016, Fig. 5.8). We estimate a model where the parameters are the growth rates in different developmental periods and the age at which each period starts. In this model there are four periods: early childhood (before age 2-3), childhood (roughly ages 3-10), the adolescent growth spurt (sometime after age 10), and adulthood (once growth is completed). The exact duration of each period varies across time and place, and may depend on SES.

Growth is typically highest at birth, and falls rapidly during early childhood. During childhood velocity declines slowly until the adolescent growth spurt (AGS). At this point growth velocity increases, reaches a peak and then declines at a steady rate until adult height is achieved. The shape of the velocity curves is thus well approximated by a piece-wise continuous linear function, with three slope changes: a first change at the end of the fastest growth period in early childhood, a second at the beginning of the AGS, and a third at its peak. In Figure 4 we illustrate the typical velocity curves for boys and girls, as illustrated for instance in Tanner et al. (1966, Fig. 8). We superimpose on the figure an illustration of the model we estimate, with labels corresponding to the corresponding parameters, described below.

Formally, let t_1, t_2, and t_3 denote the timing of the kinks in the piece-wise linear velocity curve, and let t_4 be the time when adult height is achieved. Let also h_t denote height of an individual at age t (measured in months). For an individual who has not yet achieved adult height (that, is for $t < t_4$), growth between $t-1$ and t can be written as

$$\begin{aligned} h_t - h_{t-1} &= \alpha + \beta_1 \left(\min\{t, t_1\} - 1\right) + 1\left(t > t_1\right) \beta_2 \left(\min\{t, t_2\} - t_1\right) \\ &\quad + 1\left(t > t_2\right) \beta_3 \left(\min\{t, t_3\} - t_2\right) + 1\left(t > t_3\right) \beta_4 \left(\min\{t, t_4\} - t_3\right), \end{aligned} \tag{2}$$

where the coefficients β_1, β_2, β_3, and β_4 are thus the slopes of the four linear intervals. Because adult height is achieved at $t = 4$, growth must be equal to zero at this time, so that the following constraint must hold:

$$\alpha + \beta_1 \left(t_1 - 1\right) + \beta_2 \left(t_2 - t_1\right) + \beta_3 \left(t_3 - t_2\right) + \beta_4 \left(t_4 - t_3\right) = 0, \tag{3}$$

This model cannot be estimated directly in our data, given that for the same child we never observe height measured in two consecutive months. However, in Appendix A.3 we show that equation (2) can be used in an iterative fashion to write down height at age t as:

$$h_t = h_0 + \alpha 1(t \leq t_4)t + \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3 + \beta_4 v_4 + \delta 1(t > t_4), \tag{4}$$

where the v functions are deterministic and known functions of age and/or the location of the kinks such that

$$\begin{aligned} v_1 &= 1(t \leq t_4)\frac{\min(t,t_1)(\min(t,t_1)-1)}{2} + 1(t_1 < t \leq t_4)(t-t_1)(t_1-1) \\ v_2 &= 1(t_1 < t \leq t_4)\frac{(\min(t,t_2)-t_1)(\min(t,t_2)-t_1+1)}{2} + 1(t_2 < t \leq t_4)(t-t_2)(t_2-t_1) \\ v_3 &= 1(t_2 < t \leq t_4)\frac{(\min(t,t_3)-t_2)(\min(t,t_3)-t_2+1)}{2} + 1(t_3 < t \leq t_4)(t-t_3)(t_3-t_2) \\ v_4 &= 1(t_3 < t \leq t_4)\frac{(t-t_3)(t-t_3+1)}{2} \end{aligned}$$

and where in addition to constraint (3) the following should also hold

$$\delta = t_4\alpha + \left[\frac{t_1(t_1-1)}{2} + (t_4-t_1)(t_1-1)\right]\beta_1 + \left[\frac{(t_2-t_1)(t_2-t_1+1)}{2} + (t_4-t_2)(t_2-t_1)\right]\beta_2 + \left[\frac{(t_3-t_2)(t_3-t_2+1)}{2} + (t_4-t_3)(t_3-t_2)\right]\beta_3 + \frac{(t_4-t_3)(t_4-t_3+1)}{2}\beta_4. \quad (5)$$

This second constraint imposes that height be constant once adult height is achieved (at time $t = t_4$). Both these constraints are linear in parameters, and given that in our data we observe both height and age for each child, the coefficients in (4) can be estimated in a straightforward way using constrained OLS, *once the location of the kinks is known.* Given that such location is actually unobserved, we use an approach analogous to that developed in Hansen (2017) for the estimation of regression kink models with an unknown threshold. First we set the positions of the kinks t_1, t_2, t_3, and t_4. Then we estimate (4) using constrained OLS, and we calculate and store the corresponding sum of squared residuals (SSR). Finally, we choose the estimates that minimize the SSR over the whole grid. Because the kinks are naturally ordered, we always impose $t_1 < t_2 < t_3 < t_4$, but we also impose a minimum of twelve months between t_2 and t_3, that is, between the beginning and the peak of the AGS. This is because, due to the timing of the height measurement, the number of children measured around this period is sometimes small, and this leads to estimates of the duration of the AGS that are unreasonably short when compared to what suggested by the literature on human growth.[18]

4.1.2 Model estimation results

In order to increase precision, we pool together data from each of the four YLS countries and cohorts. YLS data include measurements of the same individuals at different ages. The frequency of measurements is too sparse to allow estimating individual growth velocity at frequent intervals, but there is sufficient variation in the exact age at measurement around the mean age that we can use the model described above to estimate the age profile of growth velocity around ages 1, 5, 8, 12, 15, 19 and 22. We do not include data from CLHNS because the timing of the measurements only partly overlaps with YLS, and it was undesirable to have different sets of countries driving results over different ages ranges.

We show graphically the results of the estimation in Figure 5, while the details of the estimations are in Table 5.[19] As expected, the AGS takes place significantly sooner among girls relative to boys, and girls achieve their final height earlier than boys. And, perhaps unsurprisingly given our earlier results, there are visible differences in growth velocity by maternal education.

Three patterns are apparent. First, growth velocity is faster among high-SES children until a few months after the AGS peak: among boys the gap is small but persistent until the start of the AGS

[18]The small number of observations at these two kinks means that the SSR obtained with or without imposing such minimum duration are very close, and so the choice between constrained and unconstrained estimates lead to very similar values of the objective function (the SSR) but to quite different estimates.

[19]In Appendix Figure A.6 we also report the country-specific patterns. The country-specific point estimates and standard errors of the slopes are available upon request from the authors.

(t_2), while among girls it is large especially between 1 and 3 years of age and after t_2. Second, the AGS starts sooner among high-SES children, especially among girls where it takes place about one year sooner. Third, growth continues for a longer period among low-SES children, especially among boys.

The model-implied SES gradient, shown in Figure 6, rises until adolescence and then falls. The average height gap between high and low-SES increases gradually with age, opens up further when high-SES have their AGS, but then low-SES catch up both because their AGS peak occurs when growth is already slowing down for high-SES children and because they achieve their adult height at an older age. This indicates a degree of catch up, although this is only partial. Indeed the parameter estimates in Table 5 show that average adult height ($h_6 + \delta$) is 167 cms among low-SES boys and 168 cms among high-SES boys, while among girls the two estimates are 154.9 cm and 155.9 cm, respectively.[20]

4.2 Is there a greater cost of 'early adulthood' for high-SES children?

A complementary hypothesis that could explain the fall of the gradient during adolescence is that in this period children may start to engage in behaviors that could hamper their growth. Adolescence is a period of great biological, economic and social changes, as children transition into adulthood. Thus, it is plausible that higher-SES children—which are more likely to reach pubertal development before lower-SES peers—may also start earlier to engage in behaviors that may harm their growth. In turn, this would reduce their height advantage over children from more economically-disadvantaged backgrounds.

We base this hypothesis on the observation that adolescents whose physical maturity is apparent are more likely to engage in behaviours that may potentially harm their growth, as documented by previous literature. For instance, girls who have an earlier menarche are more likely to drop out of school, marry early and have children early. Field and Ambrus (2008) show that marriage rates in Bangladesh after age 13 were strongly and positively correlated with the onset of puberty among girls. Khanna (2020) finds that, in India, girls who reach menarche before twelve (controlling for several indicators of SES) have 13% lower school enrolment. By the same token, children that undergo their pubertal growth spurt earlier, may be more likely to be engaged in physically demanding labor as compared to peers that have a delayed pubertal growth spurt. Both early childbearing and increased work may impose a 'height cost' by increasing a child's nutritional expenditures and slowing down growth (Johnson and Moore 2016). Decreased nutritional investments for children that appear taller than their peers could be another potential behavioral explanation. Evidence from Guatemala shows that parents may invest less (more) in their children's health and nutrition if they perceive them to be tall (small) by local standards (Wang et al. 2020). Earlier age of puberty may also disrupt sleep patterns and lead to lower sleeping hours. As the growth hormones are produced during sleep, this

[20] These figure suggest a height gap between high vs. low-SES adults that is smaller than the ones of 1.4-2.7 cms documented in Tables 2 and 3. This is likely due to the approximation induced by the piece-wise continuous shape of the growth velocity curve that we impose for the estimation.

can lead to lower height. Finally, earlier pubertal timing has been shown to predict higher sensation seeking and engagement in risky behaviors (Steinberg et al. 2008), which in turn may decelerate adolescents' subsequent growth.

However, to help explaining the drop in the gradient, these behaviors need to be more frequent or more costly among high-SES children, and especially so around the timing of pubertal development for boys and girls. We test this hypothesis by using the rich data available in the YLS which collected information on these adolescent behaviors.

We proceed in two steps. First, we investigate if these behaviors—marriage, childbearing, low sleep, high work, poor quality diets, and engagement in health risk behaviors—are negatively associated to height at 22 conditional on height at age 8, and whether they do so differentially by SES. This would be consistent with these behaviors being in fact detrimental to growth during the adolescent years. Then, we investigate if these behaviors are more or less prevalent among high SES groups. Consistent with our previous findings, we show results separately by girls and boys, also because engagement in some of these risky behaviors are highly gendered in LMICs.

The results regarding the predictive role of behaviors on growth are shown in Table 6. For each gender we estimate two regressions: one where we predict adult height using behaviors for both high and low-SES pooled, and a second one where the behaviors are interacted with maternal education to assess whether the associations between these behaviors and adult height vary across SES groups.

The results show that boys and girls that are taller at age 8 end up as taller adults, confirming that much of the variation in adult height is explained by growth in early childhood. Second, most of the behaviors we observe in adolescence do indeed predict lower adult height: all the coefficients are negative as hypothesized, although all but one are insignificant at standard levels. The exception is adolescent marriage and childbearing, which is a significant predictor of lower final height for boys. However, and most crucially to explain the inverted U-shape, the deleterious associations between early marriage and fertility and adulthood height do not appear to vary by SES: most of the interactions are not significant at standard level. Moreover, the only two significant interactions (marriage/childbearing and risky behavior for boys) are positive rather than negative, and even larger than the main effects. That is, among high-SES individuals these behaviors actually *increase* predicted adult height.

We then turn to examining whether the incidence of these behaviors is different among high and low-SES adolescents. Even if the effect of the behavior is the same, the reversal of the gradient could occur if the behavior is more frequent among high SES children. We focus on marriage, the only observable behavior that predicts lower growth during adolescence in our sample. Table 7 shows that, as expected, girls with early menarche are more likely to marry early (before age 17), and that daughters of secondary-educated women are less likely to marry early (column 1). However, column 2 shows that the interaction between maternal education and early menarche is negative: in other words, early menarche *increases* the gap between SES groups, rather than decreasing it. For boys, the interaction term is small and statistically insignificant. This results are similar if we use height at age 8 as a measure for the onset of adolescence. In sum, while we confirm that there are behaviors in adolescence that are negatively associated with growth during this period (in particular early marriage), we find no evidence that these behavioral differences can account for the decline in the SES gradients in adolescence. In fact, if anything, we find the opposite. However, these results have a silver lining in that they suggest that catch up could be larger among low-SES children if marriage and childbearing during adolescence could be avoided.

5 Discussions and Conclusions

Using a large number of LMICs countries and cohorts we have shown that the association between height (a measure of long-term health) and maternal education (a proxy for SES, and a well-known determinant of child health) follows an age profile with an inverted U-shape. We show that such profile is likely mediated by the physiology of human growth, and the role that SES plays as a determinant of the timing and duration of puberty. In LMICs populations, children from high-SES families start their adolescent growth spurt earlier, on average, than children from low-SES families. This, together with the fact that low-SES children achieve their adult height at older ages, allows such children to partly compensate the height disadvantage they have accumulated during children. In contrast, we show that behavioral responses are unlikely to explain the observed patterns in the data.

Thus, our results are consistent with the hypothesis that adolescence may provide a window of opportunity for catch up for otherwise deprived children. This is consistent with a recent but growing literature that finds that shocks and investments during adolescence may be important for height and human capital more generally. Note also that the decline in the gradient occurs in a 'business-as-usual' scenario, in the absence of nutritional or other interventions that may further decrease height inequalities that have emerged earlier in childhood. On the one hand, this evidence opens up opportunities for future research examining the effectiveness of interventions that support the resource environments of low-SES children during the pubertal growth period in reducing inequalities in health. On the other hand, our analysis is descriptive. Providing an overview of which specific interventions to support catch-up growth (aside, perhaps, from preventing early marriages) is an important area for future research.

Our results also suggest that height, often used as an indicator of long-term health, is not an equally good indicator at different ages, and is a particularly poor indicator of SES around birth. The decline in the gradient at older ages could also help explaining the weak association documented by Deaton (2007) in DHS data between adult height of women and GDP at birth. Deaton (2007) and Bozzoli et al. (2009) argue that another key contributing factors may be mortality selection. That is, in poor countries where infant mortality is high, increases in GDP at birth predict not only improvements in SES, but also a decrease in mortality. However, the latter decrease likely lead to the survival of individuals of poor health and likely smaller height, who would have died under less favorable conditions. Such decline in 'harvesting' will then weaken the cross-sectional association between GDP at birth and the average height of the surviving adults.[21] In this paper, we provided

[21] Although this is beyond the scope of this paper, we find that, in DHS data, child height is very weakly associated with GDP at birth at age 0, but the correlation increases substantively with age. These results are available upon request.

another explanation for why gradients among adults in developing countries are smaller than among children: there is some amount of catch up during adolescence. While the catch up is not complete, it is possible that a better understanding of the factors that increase catch up can both help explaining the 'Deaton puzzle' and provide avenues for interventions that would lower SES gradients in height. Future research in this area should further investigate these.

References

Adair, L. S. (2001). Size at birth predicts age at menarche. *Pediatrics 107*(4), 1–7.

Adair, L. S., B. M. Popkin, J. S. Akin, D. K. Guilkey, S. Gultiano, J. Borja, L. Perez, C. W. Kuzawa, T. McDade, and M. J. Hindin (2010). Cohort Profile: The Cebu Longitudinal Health and Nutrition Survey. *International Journal of Epidemiology 40*(3), 619–625.

Aiyar, A. and J. R. Cummins (2021). An age profile perspective on two puzzles in global child health: The Indian enigma and economic growth. *Journal of Development Economics 148*, 102569.

Akbulut-Yuksel, M. and A. Kugler (2016). Intergenerational persistence of health: Do immigrants get healthier as they remain in the u.s. for more generations. *Economics and Human Biology 23*, 136–48.

Akresh, R., S. Bhalotra, M. Leone, and U. O. Osili (2012). War and stature: Growing up during the Nigerian civil war. *American Economic Review Papers and Proceedings 102*(3), 273–277.

Almond, D., J. Currie, and V. Duque (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature 56*(4), 1360–1446.

Andersen, S. H., L. Steinberg, and J. Belsky (2021). Beyond early years versus adolescence: The interactive effect of adversity in both periods on life-course development. *Developmental psychology 57*(11), 1958.

Andriano, L. and C. W. Monden (2019). The causal effect of maternal education on child mortality: Evidence from a quasi-experiment in Malawi and Uganda. *Demography 56*(5), 1765–1790.

Barnett, I., P. Ariana, S. Petrou, M. E. Penny, L. T. Duc, S. Galab, T. Woldehanna, J. A. Escobal, E. Plugge, and J. Boyden (2013). Cohort Profile: The Young Lives Study. *International Journal of Epidemiology 42*(3), 701–708.

Beath, K. J. (2007). Infant growth modelling using a shape invariant model with random effects. *Statistics in Medicine 26*(12), 2547–2564.

Behrman, J. R., W. Schott, S. Mani, B. T. Crookston, K. Dearden, L. T. Duc, L. C. Fernald, and A. D. Stein (2017). Intergenerational transmission of poverty and inequality: parental resources and schooling attainment and children's human capital in Ethiopia, India, Peru, and Vietnam. *Economic development and cultural change 65*(4), 657–697.

Borga, L. G., D. Münich, and L. Kukla (2021). The socioeconomic gradient in child health and noncognitive skills: Evidence from the Czech Republic. *Economics & Human Biology 43*, 101075.

Bozzoli, C., A. Deaton, and C. Quintana-Domeque (2009). Adult height and childhood disease. *Demography 46*(4), 647–669.

Caldwell, J. C. (1986). Routes to low mortality in poor countries. *Population and development review 12*(2), 171–220.

Cameron, L. and J. Williams (2009). Is the relationship between socioeconomic status and health stronger for older children in developing countries? *Demography 46*(2), 303–324.

Carneiro, P., I. L. Garcia, K. G. Salvanes, and E. Tominey (2019). Intergenerational mobility and the timing of parental income. Working Paper.

Case, A., D. Lubotsky, and C. Paxson (2002). Economic status and health in childhood: the origins of the gradient. *American Economic Review 92*(5), 1308–1334.

Case, A. and C. Paxson (2008). Stature and status: height, ability, and labor market outcomes. *Journal of Political Economy 116*(3), 499–532.

Cole, T., M. Donaldson, and Y. Ben-Shlomo (2010). SITAR–a useful instrument for growth curve analysis. *International Journal of Epidemiology 39*(6), 1558–66.

Currie, A., M. A. Shields, and S. W. Price (2007). The child health/family income gradient: Evidence from England. *Journal of health economics 26*(2), 213–232.

Currie, J. and M. Stabile (2003). Socioeconomic status and child health: why is the relationship stronger for older children? *American Economic Review 93*(5), 1813–1823.

Cutler, D., A. Deaton, and A. Lleras-Muney (2006). The determinants of mortality. *Journal of Economic Perspectives 20*(3), 97–120.

de La Rochebrochard, E. (2000). Age at puberty of girls and boys in France: Measurements from a survey on adolescent sexuality. *Population: An English Selection 12*, 51–79.

de Muinck Keizer-Shrama, S. M. and D. Mul (2001). Trends in pubertal development in Europe. *Human Reproduction Update 7*(3), 287–291.

de Onis, M., A. W. Onyango, E. Borghi, A. Siyam, C. Nishida, and J. Siekmann (2007). Development of a WHO growth reference for school-aged children and adolescents. *Bulletin of the World health Organization 85*(9), 660–667.

Deaton, A. (2007). Height, health, and development. *Proceedings of the National Academy of Sciences 104*(33), 13232–13237.

Deaton, A. and M. Grosh (2000). Consumption. In M. Grosh and P. Glewwe (Eds.), *Designing household survey questionnaires for developing countries: lessons from 15 years of the Living Standards Measurement Study*, Volume 1, Chapter 5, pp. 91–133. Oxford University Press for the World Bank.

Drevenstedt, G. L., E. M. Crimmins, S. Vasunilashorn, and C. E. Finch (2008). The rise and fall of excess male infant mortality. *Proceedings of the National Academy of Sciences 105*(13), 5016–5021.

Field, E. and A. Ambrus (2008). Early marriage, age of menarche, and female schooling attainment in Bangladesh. *Journal of Political Economy 116*(5), 881–930.

Filmer, D. and L. Pritchett (2001). Estimating Wealth Effects Without Expenditure Data - or Tears: an Application to Educational Enrollments in States of India. *Demography 38*(1), 115–132.

Fletcher, J. and B. Wolfe (2014). Increasing our understanding of the health-income gradient in children. *Health economics 23*(4), 473–486.

Fogel, R. W. (1994). Economic growth, population theory, and physiology: The bearing of long-term processes on the making of economic policy. *American Economic Review 84*(3), 369–395.

Gluckman, P., A. Beedle, T. Buklijas, F. Low, and M. Hanson (2016). *Principles of Evolutionary Medicine*. Oxford University Press.

Grépin, K. A. and P. Bharadwaj (2015). Maternal education and child mortality in zimbabwe. *Journal of health economics 44*, 97–117.

Haddad, L. J. and H. E. Bouis (1991). The impact of nutritional status on agricultural productivity: wage evidence from the Philippines. *Oxford Bulletin of Economics and Statistics 53*(1), 45–68.

Hansen, B. (2017). Regression kink with an unknown threshold. *Journal of Business and Economic Statistics 35*(2), 228–240.

Hauspie, R. C., M. Vercauteren, and C. Susanne (1996). Secular changes in growth and maturation: An update. *Hormones 45*(Suppl. 2), 8–17.

Heath, R. and S. Jayachandran (2017). The causes and consequences of increased female education and labor force participation in developing countries. In *The Oxford Handbook of Women and the Economy*, pp. 345–367. Oxford University Press.

Hirshkowitz, M., K. Whiton, S. M. Albert, C. Alessi, O. Bruni, L. DonCarlos, N. Hazen, J. Herman, E. S. Katz, L. Kheirandish-Gozal, et al. (2015). National sleep foundation's sleep time duration recommendations: methodology and results summary. *Sleep health 1*(1), 40–43.

Johnson, W. and S. Moore (2016). Adolescent pregnancy, nutrition, and health outcomes in low-and middle-income countries: what we know and what we don't know. *BJOG 123*(10), 1589–1592.

Khanam, R., H. S. Nghiem, and L. B. Connelly (2009). Child health and the income gradient: evidence from Australia. *Journal of Health Economics 28*(4), 805–817.

Khanna, M. (2020). The precocious period: Impact of early menarche on schooling in India. Working Paper.

Lam, B. Y. H. e. a. (2021). MC3R links nutritional state to childhood growth and the timing of puberty. *Nature 599*, 436–41.

Leroy, J. L., E. A. Frongillo, P. Dewan, M. M. Black, and R. A. Waterland (2020). Can children catch up from the consequences of undernourishment? Evidence from child linear growth, developmental epigenetics, and brain and neurocognitive development. *Advances in Nutrition 11*(4), 1032–1041.

Ma, C., P. Bovet, L. Yang, M. Zhao, Y. Liang, and B. Xi (2018). Alcohol use among young adolescents in low-income and middle-income countries: a population-based study. *Lancet Child Adolescent Health 2*, 415–429.

Marcovecchio, M. L. and F. Chiarelli (2013). Obesity and growth during childhood and puberty. *World Review of Nutrition and Dietetics 106*, 135–141.

Martorell, R. and J.-P. Habicht (1986). Growth in early childhood in developing countries. In F. Falkner and J. M. Tanner (Eds.), *Human growth: a comprehensive treatise*, Volume 3. New York: Plenum Press.

Martorell, R., L. Kettel Khan, and D. Schroeder (1994). Reversibility of stunting: epidemiological findings in children from developing countries. *European journal of clinical nutrition 48*(1), 45–57.

Murasko, J. E. (2008). An evaluation of the age-profile in the relationship between household income and the health of children in the United States. *Journal of Health Economics 27*(6), 1489–1502.

Park, C. (2010). Children's health gradient in developing countries: evidence from Indonesia. *Journal of Economic Development 35*(4), 25.

Perkins, J. M., S. Subramanian, G. Davey Smith, and E. Özaltin (2016). Adult height, nutrition, and population health. *Nutrition Reviews 74*(3), 149–165.

Persico, N., A. Postlewaite, and D. Silverman (2004). The effect of adolescent experience on labor market outcomes: the case of height. *Journal of Political Economy 112*, 1019–1053.

Preece, M. and M. Baines (1978). A new family of mathematical models describing the human growth curve. *Annals of Human Biology 5*(1), 1–24.

Propper, C., J. Rigg, and S. Burgess (2007). Child health: evidence on the roles of family income and maternal mental health from a UK birth cohort. *Health economics 16*(11), 1245–1269.

Reinhold, S. and H. Jürges (2012). Parental income and child health in Germany. *Health economics 21*(5), 562–579.

Sayers, A., M. Baines, and K. Tilling (2013). A new family of mathematical models describing the human growth curve—Erratum: Direct calculation of peak height velocity, age at take-off and associated quantities. *Annals of Human Biology 40*(3), 298–299.

Sepehri, A. and H. Guliani (2015). Socioeconomic status and children's health: Evidence from a low-income country. *Social Science & Medicine 130*, 23–31.

Shrimpton, R., C. G. Victora, M. de Onis, R. Costa Lima, M. Blössner, and G. Clugston (2001). Worldwide timing of growth faltering: implications for nutritional interventions. *Pediatrics 107*(5:E75), 1–7.

Simondon, K., F. Simondon, I. Simon, A. Diallo, E. Bénéfice, P. Traissac, and B. Maire (1998). Preschool stunting, age at menarche and adolescent height: a longitudinal study in rural Senegal. *European journal of clinical nutrition 52*(6), 412–418.

Ssentongo, P., A. E. Ssentongo, D. M. Ba, J. E. Ericson, M. Na, X. Gao, C. Fronterre, V. M. Chinchilli, and S. J. Schiff (2021). Global, regional and national epidemiology and prevalence of child stunting, wasting and underweight in low-and middle-income countries, 2006–2018. *Scientific reports 11*(1), 1–12.

Steckel, R. H. (1986). A peculiar population: The nutrition, health, and mortality of American slaves from childhood to maturity. *The Journal of Economic History 46*(3), 721–741.

Steckel, R. H. (1995). Stature and the standard of living. *Journal of Economic Literature 33*(4), 1903–1940.

Steckel, R. H. (2009). Heights and human welfare: Recent developments and new directions. *Explorations in Economic History 46*(1), 1–23.

Steinberg, L., D. Albert, E. Cauffman, M. Banich, S. Graham, and J. Woolard (2008). Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: evidence for a dual systems model. *Developmental psychology 44*(6), 1764.

Strauss, J. and D. Thomas (1998). Health, nutrition, and economic development. *Journal of Economic Literature 36*(2), 766–817.

Strauss, J. and D. Thomas (2008). Health over the life course. In T. P. Schultz and J. Strauss (Eds.), *Handbook of Development Economics*, Volume IV, Chapter 54, pp. 3375–3474. Amsterdam: Elsevier Science.

Sun, Y., F. K. Mensah, P. Azzopardi, G. C. Patton, and M. Wake (2017). Childhood social disadvantage and pubertal timing: a national birth cohort from Australia. *Pediatrics 139*(6), e201640.

Tanner, J., R. Whitehouse, and M. Takaishi (1966). Standards from birth to maturity for height, weight, height velocity, and weight velocity: British children, 1965. I. *Archives of Disease in Childhood 41*(219), 454–471.

Tanner, J. M. (1989). *Fetus into man. Physical Growth from Conception to Maturity* (Second ed.). Cambridge, Massachussetts: Harvard University Press.

Thomas, F., F. Renaud, E. Benefice, T. De Meeüs, and J.-F. Guegan (2001). International variability of ages at menarche and menopause: patterns and main determinants. *Human Biology 73*(2), 271–90.

van den Berg, G., P. Lundborg, P. Nystedt, and D.-O. Rooth (2014). Critical periods during childhood and adolescence. *Journal of the European Economic Association 12*(6), 1521–1557.

Wang, F., E. Puentes, J. Behrman, and F. Cunha (2020). You are what your parents think: Height and local reference points. Penn Institute of Economic Research Working Paper 18-007.

Ward, J. L. and R. M. Viner (2018). Investigating equalisation of health inequalities during adolescence in four low-income and middle-income countries: an analysis of the Young Lives cohort study. *BMJ open 8*(9), e022114.

West, P. (1997). Health inequalities in the early years: is there equalisation in youth? *Social science & medicine 44*(6), 833–858.

World Health Organization (2006). WHO child growth standards length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. Technical report, World Health Organization, Geneva.

World Health Organization (2008). Global database on child growth and malnutrition. Internet Database. Available at http://www.who.int/nutgrowthdb/database/en/.

World Health Organization (2017). Global nutrition monitoring framework: operational guidance for tracking progress in meeting targets for 2025. Technical report, World Health Organization, Global Malaria Programme, Geneva.

Zarulli, V., J. A. B. Jones, A. Oksuzyan, R. Lindahl-Jacobsen, K. Christensen, and J. W. Vaupel (2018). Women live longer than men even during severe famines and epidemics. *Proceedings of the National Academy of Sciences 115*(4), E832–E840.

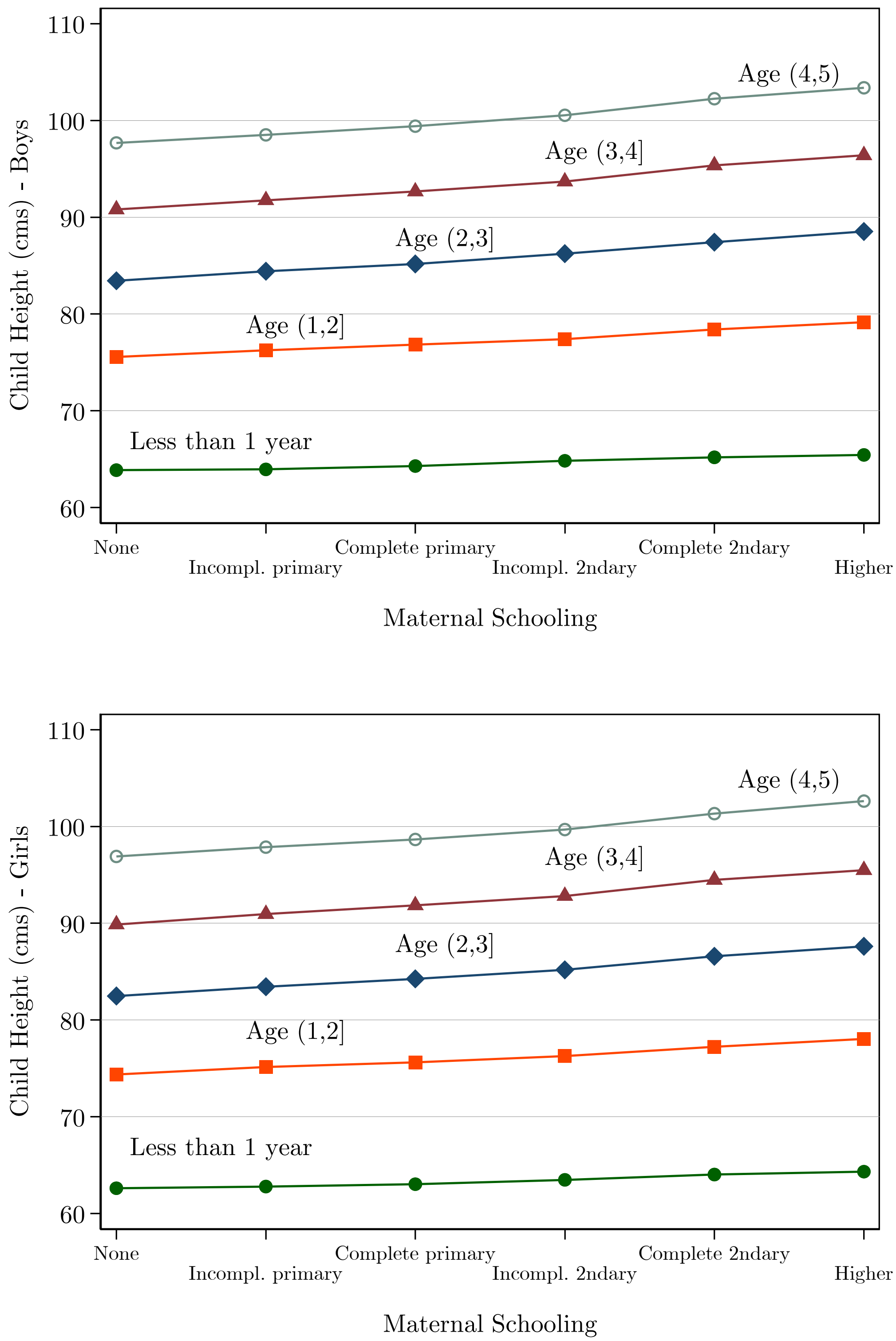

Figure 1: DHS: SES Gradient by Child Age

Source: Authors' calculations from DHS data. For each age interval, each line shows the relationship between average height and maternal schooling.

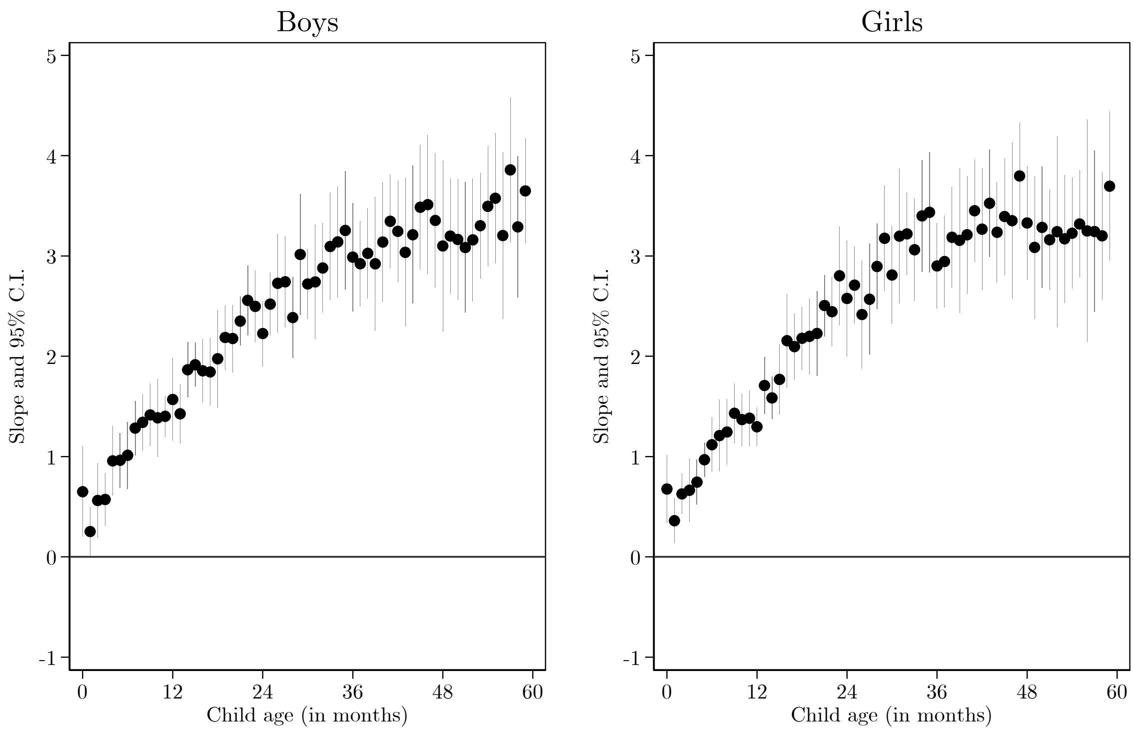

Figure 2: DHS: Child Height vs. Maternal Education

Source: Authors' calculations from DHS data. For each age (in months) the figure shows the point estimate and a 95% confidence interval of the slope of a regression, estimated with OLS, of child height (in cms) on a dummy variable equal to one if the mother has completed at least secondary education. All estimates do not use sampling weights, and the confidence intervals are calculated allowing for correlation of residuals within each survey primary stage unit. Total sample size is $n = 1,556,034$.

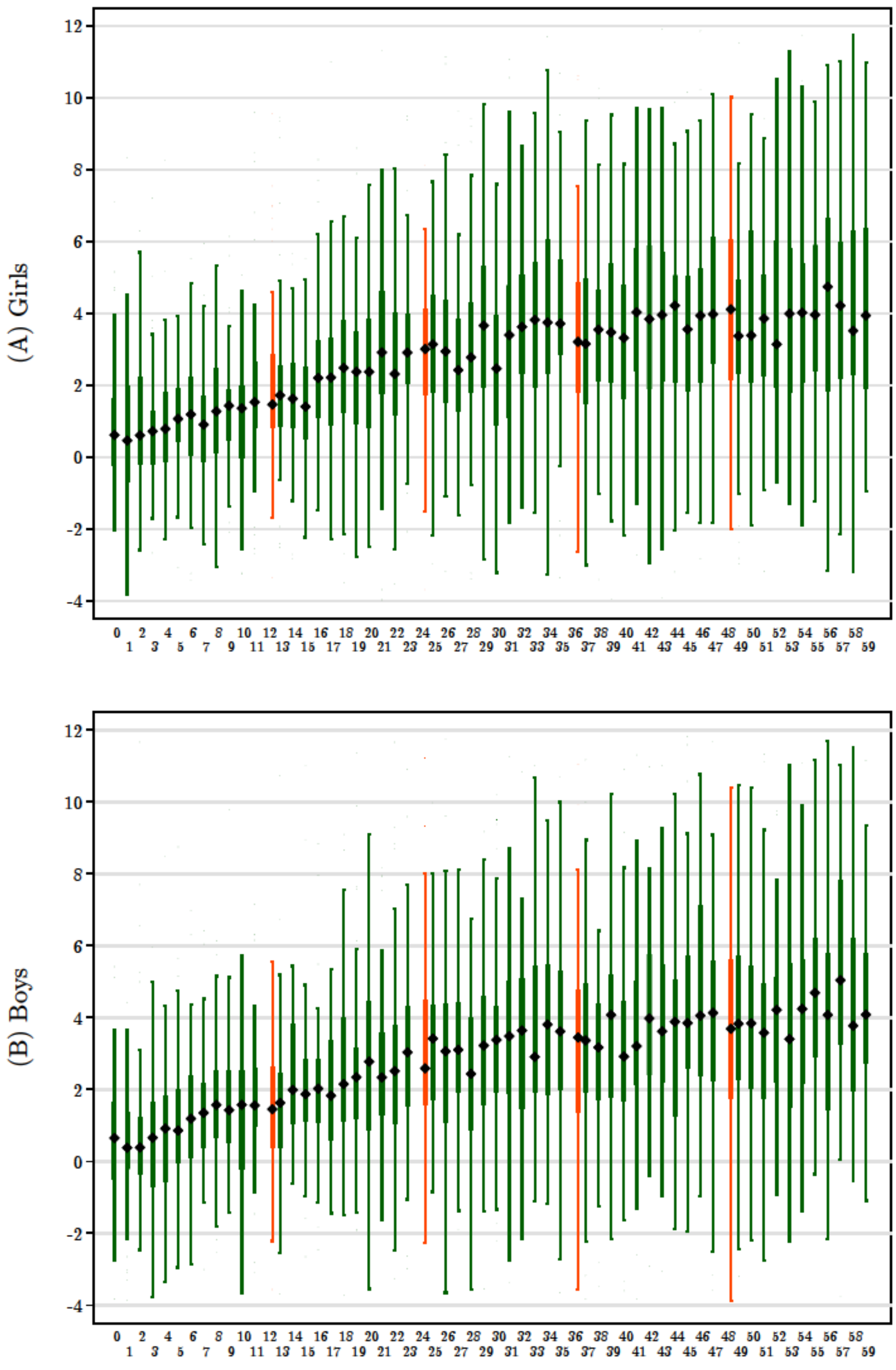

Figure 3: DHS: Child Height vs. Maternal Education

Source: Authors' calculations from DHS data. For each age (in months) the figure shows a box plot of the estimated country-specific OLS slopes of regressions, estimated with OLS, of child height (in cm) on a dummy variable equal to one if the mother has completed at least secondary education. All estimates do not use sampling weights. If more than one DHS was completed for a given country all observations were pooled together.

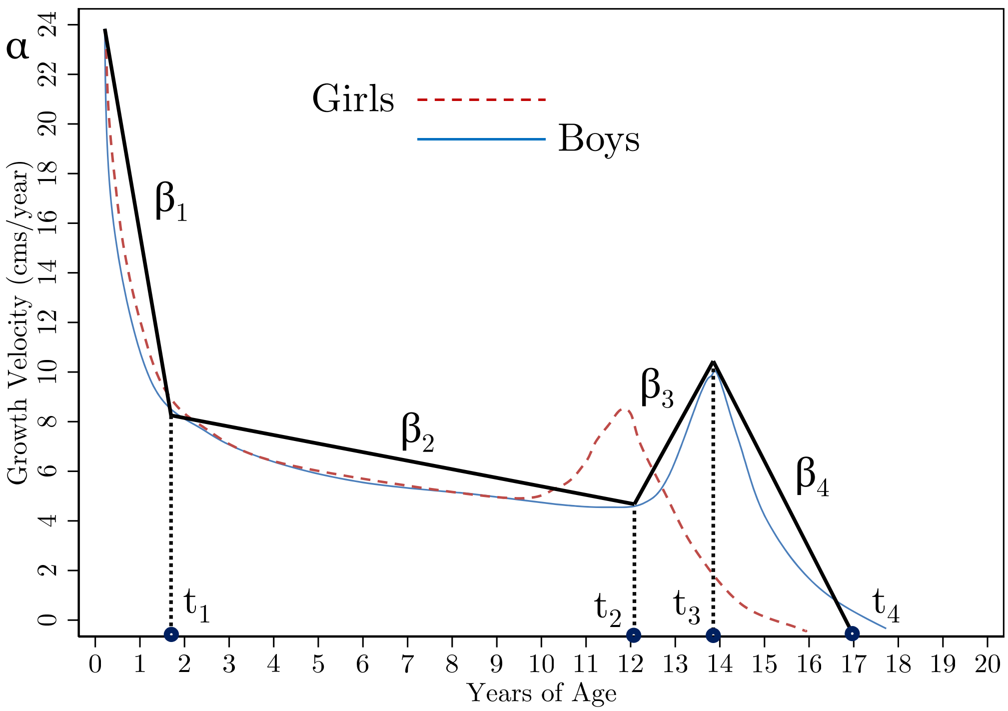

Figure 4: Growth Velocity

Source: Authors' elaboration from Tanner et al. (1966, Fig. 8). The labels indicate the parameters estimated for boys using the procedure described in Section 4.1.1: α is growth velocity at birth; t_1, t_2, t_3, and t_4 show the age of the most salient changes in growth velocity, while β_1, β_2, β_3, and β_4 are the slopes of the piecewise linear curve in each interval.

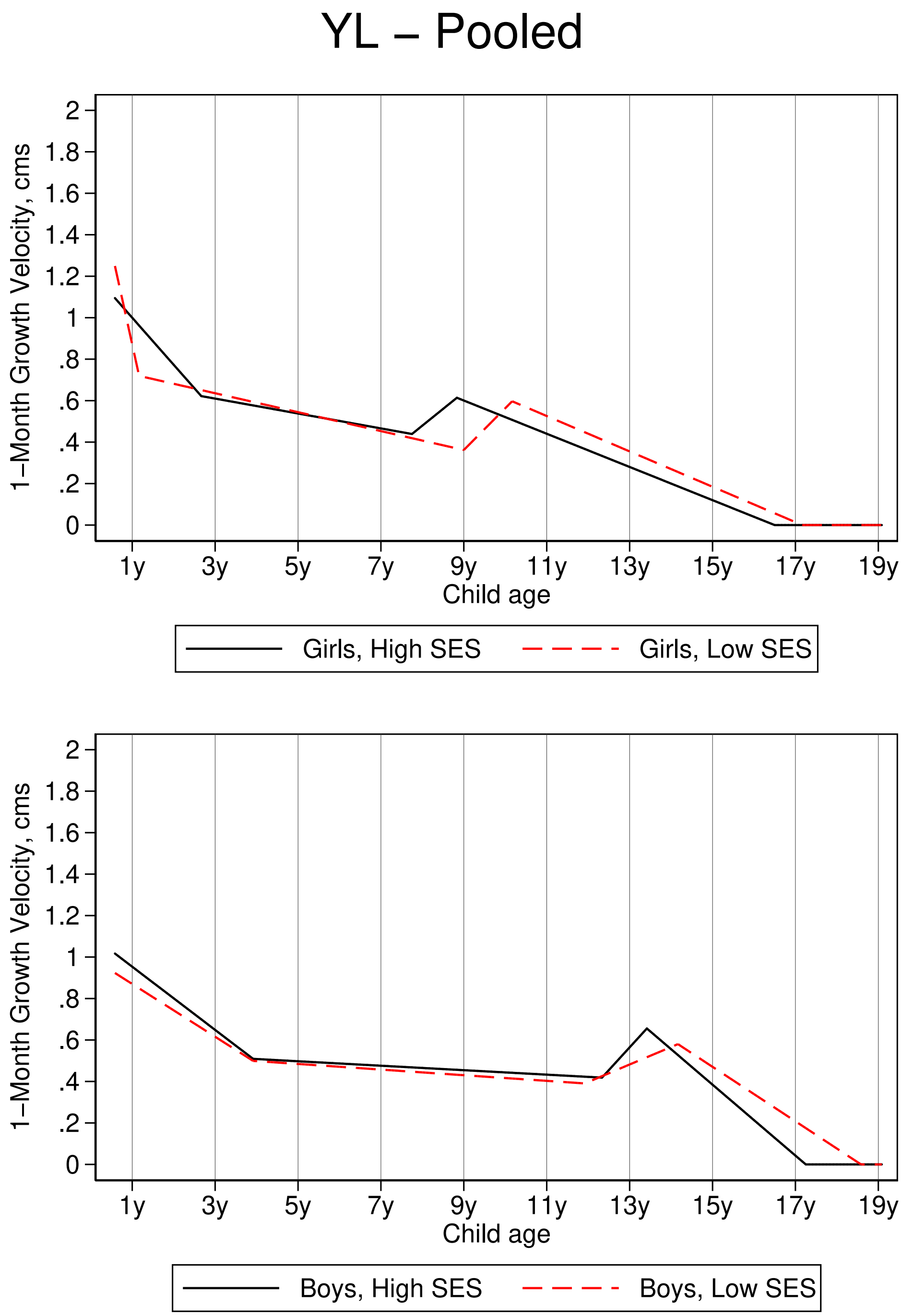

Figure 5: YLS: Growth Velocity and SES

Source: Authors' calculations from YLS data. The lines show height growth velocity predicted by the piece-wise continuous regression model described in Section 4.1.1, estimated separately for boys and girls, and by SES. High-SES is binary and equal to one when the mother has completed at least secondary schooling.

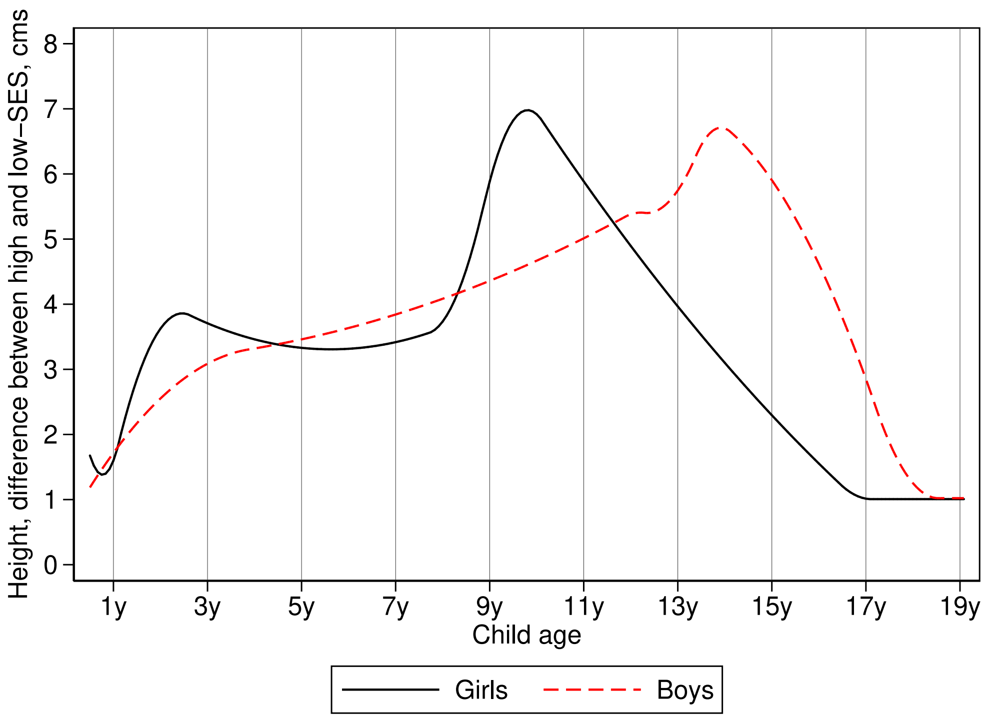

Figure 6: YLS: Height high vs. low-SES gap

Source: Authors' calculations from YLS data. The lines show differences in height growth velocity between high-SES and low-SES children, as predicted by the piece-wise continuous regression model described in Section 4.1.1, estimated separately for boys and girls. High-SES is binary and equal to one when the mother has completed at least secondary schooling.

Table 1: Height vs. maternal schooling, DHS, Girls and Boys 0-4 and 15-17

	Age (years)							
	0	1	2	3	4	15	16	17
	Girls							
Mother at least secondary	1.22	2.11	2.96	3.29	3.25	2.48	2.3	2.16
(s.e.)	(0.164)	(0.139)	(0.223)	(0.242)	(0.313)	(0.106)	(0.106)	(0.11)
R^2	0.014	0.048	0.069	0.082	0.087	0.157	0.177	0.192
Obs.	167,615	162,950	155,190	144,159	133,560	54,531	51,199	43,246
Mean dependent variable	63.1	75.5	84	91.6	98.5	152.7	153.6	154
% maternal education missing	0.04	0.05	0.05	0.02	0.02	0.26	0.30	0.37
	Boys							
Mother at least secondary	1.12	2.06	2.8	3.21	3.32	0.68	2.11	2.58
(s.e.)	(0.192)	(0.137)	(0.19)	(0.269)	(0.287)	(1.333)	(0.589)	(0.335)
R^2	0.0149	0.0494	0.07	0.0832	0.0894	0.111	0.077	0.105
Obs.	173,344	169,923	160,510	149,340	139,443	9,940	9,780	8,124
Mean dependent variable	64.4	76.7	85	92.4	99.3	159.9	162.4	164.5
% maternal education missing	0.04	0.04	0.05	0.02	0.02	0.25	0.29	0.36

Source: Authors' calculations from DHS data.
Notes: For each age (in years) the table reports estimates and standard errors of the slope of a regression, estimated with OLS, of height (in cms) on a dummy variable equal to one if the mother has completed at least secondary education. Regressions for children under five include all children of a given age (in years) born of women of fertility age in the sample. Regressions for 15 to 17-year old boys and girls only include individuals who are still co-residing with their mother, and for whom maternal schooling can be identified through unique individual identifiers in the data, see text for additional details. All regressions include country FE and do not use sampling weights. Standard errors are calculated allowing for correlation of residuals within each survey primary stage unit.

Table 2: Girl height vs. maternal schooling, YLS and CLHNS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A.1	Young Lives: Younger Cohort						
	Age 1y	Age 5y	Age 8y	Age 12y	Age 15y		
Mother at least secondary	1.701***	3.554***	4.195***	4.727***	2.333***		
	[0.2266]	[0.4137]	[0.4429]	[0.6214]	[0.3135]		
Observations	3,433	3,433	3,433	3,433	3,433		
R-squared	0.4483	0.2260	0.1399	0.1289	0.0941		
Mean height	70.84	103.8	120	143	153.6		
Panel A.2			Young Lives: Older Cohort				
			Age 8y	Age 12y	Age 15y	Age 19y	Age 22y
Mother at least secondary			2.480***	3.356***	2.035***	1.454***	1.374***
			[0.4869]	[0.6130]	[0.4016]	[0.4123]	[0.3769]
Observations			1,494	1,494	1,494	1,494	1,494
R-squared			0.0758	0.0687	0.0872	0.1276	0.1971
Mean height			117.9	142.1	151.7	154.6	155.3
Panel B	CLHNS						
	Age 1y		Age 8y	Age 11y	Age 15y	Age 18y	Age 21y
Mother at least secondary	1.083***		3.282***	3.844***	2.095***	1.500***	1.655***
	(0.208)		(0.437)	(0.637)	(0.369)	(0.346)	(0.321)
Observations	677		677	677	677	677	677
R-squared	0.033		0.058	0.119	0.057	0.037	0.030
Mean height	69.99		117.6	135.2	149.1	151	151.3

Source: Authors' calculations from Young Lives and Cebu Longitudinal Health and Nutrition Survey.
Notes: This table presents OLS regression estimates of girl height (in cms) on a dummy variable equal to one if the mother has completed at least secondary education. All estimates do not use sampling weights and include dummies for country and age in months. Standard errors are clustered at the level of primary stage unit of residence in the first wave ('sentinel site' in YLS and *barangay*—district or village—in CLHNS). In YLS, secondary education is set = 1 when the mother has completed a number of years of schooling corresponding to the country-specific typical requirement, that is, 10 in Ethiopia, 12 in India, 11 in Peru and 9 in Vietnam. In CLHNS it is = 1 if the mother has completed at least 4 years of secondary school at the time of the first survey wave. Results in Panel A are estimated pooling all observations for girls from Ethiopia, India, Peru and Vietnam, for the Younger Cohort (born 2001/02, Panel A.1) and the Older Cohort (born 1994/95, Panel A.2).

Table 3: Boy height vs. maternal schooling, YLS and CLHNS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A.1	Young Lives: Younger Cohort						
	Age 1y	Age 5y	Age 8y	Age 12y	Age 15y		
Mother at least secondary	1.558***	3.614***	4.002***	5.281***	4.744***		
	[0.2664]	[0.4195]	[0.5316]	[0.6808]	[0.5107]		
Observations	3,762	3,762	3,762	3,762	3,762		
R-squared	0.4247	0.2235	0.1427	0.1582	0.1522		
Mean height	72.27	104.6	120.3	140.9	159		
Panel A.2			Young Lives: Older Cohort				
			Age 8y	Age 12y	Age 15y	Age 19y	Age 22y
Mother at least secondary			3.638***	4.778***	5.111***	3.182***	2.771***
			[0.4526]	[0.5961]	[0.5342]	[0.3900]	[0.3530]
Observations			1,497	1,497	1,497	1,497	1,497
R-squared			0.0950	0.1132	0.0992	0.1137	0.1323
Mean height			118.5	140	156.4	166.5	167.6
Panel B	CLHNS						
	Age 1y		Age 8y	Age 11y	Age 15y	Age 18y	Age 21y
Mother at least secondary	1.703***		3.443***	4.110***	3.899***	2.658***	2.579***
	(0.179)		(0.559)	(0.698)	(0.603)	(0.536)	(0.565)
Observations	748		748	748	748	748	748
R-squared	0.069		0.073	0.175	0.125	0.067	0.062
Mean height	71.46		117.6	132.1	158.1	162.3	162.8

Source: Authors' calculations from Young Lives and Cebu Longitudinal Health and Nutrition Survey.
Notes: This table presents OLS regression estimates of boy height (in cms) on a dummy variable equal to one if the mother has completed at least secondary education. All estimates do not use sampling weights and include dummies for country and age in months. Standard errors are clustered at the level of primary stage unit of residence in the first wave ('sentinel site' in YLS and *barangay*—district or village—in CLHNS). In YLS, secondary education is set = 1 when the mother has completed a number of years of schooling corresponding to the country-specific typical requirement, that is, 10 in Ethiopia, 12 in India, 11 in Peru and 9 in Vietnam. In CLHNS it is = 1 if the mother has completed at least 4 years of secondary school at the time of the first survey wave. Results in Panel A are estimated pooling all observations for boys from Ethiopia, India, Peru and Vietnam, for the Younger Cohort (born 2001/02, Panel A.1) and the Older Cohort (born 1994/95, Panel A.2).

Table 4: Association between early menarche and maternal schooling

	(1)	(2)	(3)	(4)	(5)
Panel A: DHS	India	Turkey	Gabon	Ghana	
Mother at least secondary	0.030***	0.017	0.192	0.146	
	(0.007)	(0.052)	(0.128)	(0.095)	
Observations	63,989	862	541	409	
R-squared	0.000	0.000	0.007	0.014	
Mean of dependent variable	0.200	0.306	0.197	0.064	
Age range	15-17	15-17	15-19	15-19	
Panel B: YLS and CLHNS	Ethiopia	India	Peru	Vietnam	Philippines
Mother at least secondary	0.018	0.184*	0.109**	0.205***	0.182***
	(0.027)	(0.105)	(0.046)	(0.043)	(0.034)
Observations	1,213	1,343	1,167	1,359	787
R-squared	0.018	0.056	0.091	0.125	0.154
Mean of dependent variable	0.02	0.15	0.27	0.17	0.41

Source: Authors' calculations from DHS, YLS (both cohorts), and CLHNS data.
Notes: The dependent variable is a dummy = 1 if the individual had menarche before 13 years of age. See Appendix A.1 for additional details on data construction for DHS. All estimates do not use sampling weights and include dummies for country and age in months. Standard errors are clustered at the level of primary stage unit (PSU) of residence (in DHS), or the PSU in the first wave ('sentinel site' in YLS and *barangay*—district or village—in CLHNS). In YLS, secondary education is set = 1 when the mother has completed a number of years of schooling corresponding to the country-specific typical requirement, that is, 10 in Ethiopia, 12 in India, 11 in Peru and 9 in Vietnam. In CLHNS it is = 1 if the mother has completed at least 4 years of secondary school at the time of the first survey wave.

Table 5: YLS: A Model of Growth Velocity and maternal schooling

	(1) Boys		(2) Girls	
	Low schooling	Sec. schooling	Low schooling	Sec. schooling
Intecept (Height at 6 months, h_6)	61.0718	61.6517	55.6757	59.4736
	(0.15201)	(0.26354)	(0.69938)	(0.47361)
Total growth up to adult height (δ)	105.8816	106.3238	99.2212	96.4303
	(0.18033)	(0.32247)	(0.70180)	(0.49974)
Initial growth velocity (α)	0.9869	1.0933	1.7036	1.2080
	(0.01432)	(0.02532)	(0.10908)	(0.05174)
Slope of velocity curve:				
- $t \leq t_1$: Early childhood (β_1)	-0.0106	-0.0127	-0.0757	-0.0189
	(0.00040)	(0.00071)	(0.00856)	(0.00202)
- $t_1 < t \leq t_2$: Before AGS (β_2)	-0.0011	-0.0009	-0.0038	-0.0030
	(0.00007)	(0.00013)	(0.00009)	(0.00031)
- $t_2 < t \leq t_3$: AGS (β_3)	0.0070	0.0182	0.0168	0.0135
	(0.00027)	(0.00120)	(0.00060)	(0.00083)
- $t_3 < t \leq t_4$: End of growth (β_4)	-0.0109	-0.0143	-0.0071	-0.0067
	(0.00009)	(0.00022)	(0.00005)	(0.00004)
Kinks (months)				
- t_1: End of early childhood	47	47	14	32
- t_2: Start of AGS	143	148	108	93
- t_3: AGS Peak	170	161	122	106
- t_4: Adult height	223	207	206	198
Root MSE	6.8750	6.1726	6.6904	5.5289
Observations	23,776	5,037	22,112	4,928
No. children	5,010	1,057	4,673	1,033

Source: Authors' calculations from pooled YLS data.
Notes: The table shows the estimates of the model described in Section 4.1.1, and illustrated graphically in Figure 5. AGS indicates the adolescent growth spurt.

Table 6: YLS: Behavioral determinants of growth during adolescence

	(1)	(2)	(3)	(4)
	Girls		Boys	
Height at age 8	0.420***	0.421***	0.491***	0.489***
	[0.0347]	[0.0346]	[0.0262]	[0.0261]
Mother at least secondary	0.251	-0.827	0.886**	-0.542
	[0.4064]	[0.9704]	[0.3713]	[0.9726]
Married, cohabitating or child before age 17	-1.048	-1.058	-2.390**	-2.790***
	[0.6131]	[0.6205]	[0.8657]	[0.7664]
Low sleep at age 12 and/or 15	-0.213	-0.260	-0.126	-0.170
	[0.3122]	[0.3677]	[0.4378]	[0.4425]
High work at age 12 and/or 15	-0.413	-0.588	-0.392	-0.534
	[0.4522]	[0.5420]	[0.4472]	[0.5289]
Low dietary diversity at age 12 and/or 15	-0.033	-0.185	-0.349	-0.336
	[0.3290]	[0.3171]	[0.3461]	[0.3566]
Risky behaviors at age 15	-0.132	-0.142	-0.135	-0.479
	[0.4216]	[0.4711]	[0.2904]	[0.3472]
Married or child before 17×Mother at least secondary		0.435		3.451***
		[1.7476]		[1.1873]
Low sleep×Mother at least secondary		0.319		0.389
		[0.6013]		[0.7517]
High Work×Mother at least secondary		0.883		0.986
		[0.8675]		[1.0457]
Low dietary diversity×Mother at least secondary		0.640		-0.080
		[0.6453]		[0.5773]
Risky behaviors×Mother at least secondary		0.073		1.796**
		[0.6997]		[0.8324]
Observations	1,494	1,494	1,497	1,497
R-squared	0.3540	0.3549	0.3561	0.3589
Mean dependent variable (height at age 22)	155.3	155.3	167.6	167.6

Source: Authors' calculations from the Young Lives older cohort. The dependent variable is height (in cms.) at 22y. All regressions also include fixed effects for child age (in months) and country. Standard errors clustered at community in the first round. 'Married or child before 17y' is a binary variable = 1 if the child was married or cohabiting, or a had a child before 17y. 'Low sleep' is a binary variable = 1 if the child sleeps on a typical weekday in the previous week less than the age-specific minimum recommended by the National Sleep Foundation society for recommended sleep time duration at different ages (Hirshkowitz et al. 2015). Such recommendations are 9-11 hours for school-age children 6-13y, and 8-10 hours for teenagers 14-17y. 'High work' is a binary variable = 1 if daily hours worked on a typical weekday in the previous week at least equal to the median of each round and cohort (that is, 2.25 hours for 12y, and 3hrs for 15y). This includes any type of work (self-employment, wage employment, housework). 'Low dietary diversity' is binary and = 1 if the child has not consumed in the previous day more than four food groups (excluding fats). The variable is defined based on WHO/UNICEF guidelines on minimum dietary diversity (World Health Organization 2017). 'Risky behaviours' is binary and = 1 if child engages at least once a month in drinking or smoking at 15 years. Asterisks denote statistical significance, with *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: YLS: Are high-SES children more likely to marry early if they reach adulthood young?

	(1)	(2)	(3)	(4)
	Girls		Boys	
Panel A: early menarche				
Early menarche/puberty	0.063**	0.073***	-0.012**	-0.012**
	[0.0227]	[0.0253]	[0.0048]	[0.0048]
Mother at least secondary	-0.071***	-0.057***	-0.001	-0.001
	[0.0114]	[0.0112]	[0.0053]	[0.0072]
Menarche/puberty×Mother at least secondary		-0.048		0.002
		[0.0314]		[0.0081]
Observations	1,494	1,494	1,497	1,497
R-squared	0.0650	0.0655	0.1335	0.1336
Early marriage (mean)	0.116	0.116	0.00601	0.00601
Early puberty (mean)	0.202	0.202	0.202	0.202
Panel B: height at age 8				
Prepubertal height (8 years)	0.004**	0.005***	0.000	0.000
	[0.0015]	[0.0016]	[0.0002]	[0.0002]
Mother at least secondary	-0.075***	0.759***	-0.002	-0.061
	[0.0111]	[0.2188]	[0.0050]	[0.0953]
Prepubertal height×Mother at least secondary		-0.007***		0.000
		[0.0018]		[0.0008]
Observations	1,494	1,494	1,497	1,497
R-squared	0.0660	0.0678	0.1305	0.1307
Early marriage (mean)	0.116	0.116	0.00601	0.00601
Height at 8 years (mean)	117.9	117.9	118.5	118.5

Source: Authors' calculations from Young Lives older cohort. The dependent variable is binary = 1 if the child was married or cohabiting, or a had a child before 17y. Early menarche is binary and = 1 if the girl had menarche before 13y. Early puberty is binary and = 1 if the boy had hair on chin or a voice break before 14y. All regressions also include fixed effects for child age in months and country. Standard errors clustered at community in the first round. Asterisks denote statistical significance, with *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A Appendix

A.1 Construction of DHS Data on Age at Menarche

In this section we describe the construction of the data used to produce the results in panel A of Table 4. Most of the DHS listed in Table A.1 do not include data on age at menarche. Among those that do, the question is usually only available for very young women. In addition, our preferred proxy of SES, maternal education, is only available if the woman still co-resides with her mother. To identify which DHS have data on age at menarche we used a list made available in 2018 in the DHS Program User Forum, see `https://userforum.dhsprogram.com/index.php?t=msg&th=5716`, (accessed June 6, 2020). We exclude surveys that did not measure women's height, given that the corresponding data were not used to produce the results we describe in the paper. In the end, we only use data from Gabon (2000), Ghana (1998), India (2015-16), and Turkey (2013).[22]

A.1.1 Gabon (2000)

Age at menarche was recorded for all women 15-49, but parental education is as usual only available for women cohabiting with their mothers. Data about age at menarche are included in variable `s252` in the 'individual recode' file (gair41dt.zip). In the household roster parents are identified only for girls below 15. In order to impute maternal schooling to young girls we thus use only information from unmarried daughters of the household head, dropping women 20 or older, or those from polygynous households (for whom this matching scheme cannot be used). Information on maternal schooling is then derived from the 'person recode' file (gapr41dt.zip).

A.1.2 Ghana (1998)

The data structure is similar to that of Gabon (2000). Hence, the regression is run for young girls < 20 years of age, unmarried and still cohabiting with their mother, and who are daughters of the head of a non-polygynous household. Data about age at menarche are included in variable `s520` in the 'individual recode' file (ghir41dt.zip).

A.1.3 India (2015-16)

Age at menarche was recorded for women 15-24 or younger, but parental education is only available for girls 15-17, and only if they were still cohabiting with their parents. This latter condition held for $> 90\%$ of them. Age at menarche is recorded in variable `s256` in the 'individual recode' file (iair74dt.zip), while maternal schooling is derived from the 'person recode' file (iapr74dt.zip), using the identifiers linking each household member to her/his parents. The identifiers for the father and mother are only present for women below the age of 18 who are still co-residing with them.

A.1.4 Turkey (2013)

Age at menarche was recorded for all women 15-49, but parental education is as usual only available for women cohabiting with their mothers. In the regressions we use only data from girls 15-17 for comparability and to reduce recall error and missing data on maternal schooling. Data about age at menarche are included in variable `s235` in the 'individual recode' file (trir62dt.zip), while information on maternal schooling is derived from the 'person recode' file (trpr62dt.zip), using the identifiers linking each household member to her/his parents (when cohabiting).

[22] Age at menarche is also recorded in Kyrgyz Republic (1998), Morocco (2003-04), and Yemen (2013). However, in the Kyrgyz Republic, it was recorded only for women 15 or above, while the mother's identifier was recorded only for girls below 15. In Morocco and Yemen, age at menarche was recorded only for ever married women. Given that married women can be linked to maternal education only if they are still cohabiting—and very few are—we do not use data from these surveys.

A.2 Construction of variables related to behavioral mechanisms

This section describes the construction of the data used to estimate the results in Tables 6 and 7. The data are pooled across countries data and only include the older cohort of Young Lives. 'Married or child before 17 years' is a binary variable = 1 if the the child was married or cohabiting, or a had a child before 17 years. 'Low sleep' is a binary variable = 1 if the child at 12 years and/or 15 years slept on a typical weekday in the previous week less than the age-specific minimum recommended by the *National Sleep Foundation society for recommended sleep time duration at different ages* (Hirshkowitz et al. 2015). Such recommendations are 9-11 hours for school-age children between 6-13 years, and 8-10 hours for adolescents aged 14-17 years. 'High work' is a binary variable = 1 if daily hours worked on a typical weekday in the previous week at 12 years and/or 15 years are at least equal to the median of each round and cohort (that is, 2.25 hours for 12-year-olds, and 3hrs for 15-year-olds). Child work includes any type of work, including self-employment in the family farm or business, wage employment, and housework and care activities. 'Low dietary diversity' is binary and = 1 if the child at 12 and/or 15 years has not consumed in the previous day more than four food groups (excluding fats) out of seven food groups. The variable is defined based on WHO/UNICEF guidelines on minimum dietary diversity for children (World Health Organization 2017). 'Risky behaviors' is binary and = 1 if child engaged at least once in a month in either drinking or smoking at 15 years. Data on these indicators were collected through a self-administered questionnaire to avoid under-reporting and increase confidentiality. The risky behavior variable is constructed from information on whether the adolescent drinks every day, at least once a week, or at least once a month, or smokes every day, every week, or sometimes. The cutoff of engaging in these behaviors at least once a month (as opposed to hardly ever and never for smoking, and on special occasions, hardly ever, and never for alcohol consumption) is based on the cutoff used by the WHO in its Global Youth Tobacco Surveys `https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/global-youth-tobacco-survey`, and relevant literature on alcohol consumption among adolescents in LMICs (Ma et al. 2018).

A.3 Derivation of the Model for Height

Let $t_0 = 0$ (that is, the beginning of the first period is at birth, or zero months), and let h_0 denote length at birth. Using equation (2), height at age t, $t \leq t_1$ can thus be written as

$$\begin{aligned} h_1 &= h_0 + \alpha + \beta_1 (\min\{1, t_1\} - 1) = h_0 + \alpha \\ h_2 &= h_1 + \alpha + \beta_1 (\min\{2, t_1\} - 1) = h_0 + \alpha + \alpha + \beta_1 = h_0 + 2\alpha + \beta_1 \\ h_3 &= h_2 + \alpha + \beta_1 (\min\{3, t_1\} - 1) = h_0 + 2\alpha + \beta_1 + \alpha + 2\beta_1 = h_0 + 3\alpha + (1+2)\beta_1 \\ &\quad \dots \qquad (6) \\ h_t &= h_0 + t\alpha + \beta_1 \sum_{s=1}^{t-1} s = h_0 + t\alpha + \frac{t(t-1)}{2}\beta_1, \quad t \leq t_1, \qquad (7) \end{aligned}$$

where the last step follows from the property that the sum of the first m integers can be written as $m(m-1)/2$. Height measured at the time of the end of the fast growth period in early childhood can thus be written as

$$h_{t_1} = h_0 + t_1\alpha + \frac{t_1(t_1 - 1)}{2}\beta_1. \qquad (8)$$

Next, using this result together with equation (2), we can write down height in the first period of the interval between t_1 and t_2 as

$$\begin{aligned} h_{t_1+1} &= h_{t_1} + \alpha + \beta_1 (\min\{t_1+1, t_1\} - 1) + \beta_2 (\min\{t_1+1, t_2\} - t_1) \\ &= h_0 + t_1\alpha + \frac{t_1 (t_1-1)}{2}\beta_1 + \alpha + \beta_1 (t_1 - 1) + \beta_2 (t_1 + 1 - t_1) \\ &= h_0 + (t_1+1)\alpha + \left[\frac{t_1 (t_1-1)}{2} + (t_1 - 1)\right]\beta_1 + \beta_2, \end{aligned}$$

while at time $t_1 + 2$:

$$\begin{aligned} h_{t_1+2} &= h_{t_1+1} + \alpha + \beta_1 (\min\{t_1+2, t_1\} - 1) + \beta_2 (\min\{t_1+2, t_2\} - t_1) \\ &= h_0 + (t_1+1)\alpha + \left[\frac{t_1 (t_1-1)}{2} + (t_1 - 1)\right]\beta_1 + \beta_2 + \alpha + \beta_1 (t_1-1) + 2\beta_2 \\ &= h_0 + (t_1+2)\alpha + \left[\frac{t_1 (t_1-1)}{2} + \underbrace{2}_{=t-t_1}(t_1 - 1)\right]\beta_1 + (1+2)\beta_2 \\ &= h_0 + \left(\underbrace{t_1+2}_{=t}\right)\alpha + \left[\frac{t_1 (t_1-1)}{2} + (t-t_1)(t_1 - 1)\right]\beta_1 + (1+2)\beta_2. \end{aligned}$$

Iterating further it is straightforward (if tedious) to see that for $t_1 < t \leq t_2$

$$h_t = h_0 + t\alpha + \left[\frac{t_1 (t_1-1)}{2} + (t-t_1)(t_1 - 1)\right]\beta_1 + \frac{(t-t_1)(t-t_1+1)}{2}\beta_2 \tag{9}$$

and in particular

$$h_{t_2} = h_0 + t_2\alpha + \left[\frac{t_1 (t_1-1)}{2} + (t_2-t_1)(t_1 - 1)\right]\beta_1 + \frac{(t_2-t_1)(t_2-t_1+1)}{2}\beta_2. \tag{10}$$

From equations (2) and (10) we can now see that in the first month of the third interval we have

$$\begin{aligned} h_{t_2+1} &= h_{t_2} + \alpha + \beta_1 (\min\{t_2+1, t_1\} - 1) + \beta_2 (\min\{t_2+1, t_2\} - t_1) + \beta_3 (\min\{t_2+1, t_3\} - t_2) \\ &= h_0 + t_2\alpha + \left[\frac{t_1 (t_1-1)}{2} + (t_2-t_1)(t_1 - 1)\right]\beta_1 + \frac{(t_2-t_1)(t_2-t_1+1)}{2}\beta_2 \\ &\quad +\alpha + \beta_1 (t_1-1) + \beta_2 (t_2 - t_1) + \beta_3 (t - t_2) \\ &= h_0 + (t_2+1)\alpha + \left[\frac{t_1 (t_1-1)}{2} + (t_2+1-t_1)(t_1 - 1)\right]\beta_1 \\ &\quad + \left[\frac{(t_2-t_1)(t_2-t_1+1)}{2} + (t_2 - t_1)\right]\beta_2 + \beta_3 \\ &= h_0 + t\alpha + \left[\frac{t_1 (t_1-1)}{2} + (t-t_1)(t_1 - 1)\right]\beta_1 \\ &\quad + \left[\frac{(t_2-t_1)(t_2-t_1+1)}{2} + (t_2 - t_1)\right]\beta_2 + \beta_3 \end{aligned}$$

while in the next period $t = t_2 + 2$

$$
\begin{aligned}
h_{t_2+2} &= h_{t_2+1} + \alpha + \beta_1(\min\{t_2+2, t_1\} - 1) + \beta_2\left(\min\{t_2+2, t_2\} - t_1\right) + \beta_3\left(\min\{t_2+2, t_3\} - t_2\right) \\
&= h_0 + (t_2+1)\,\alpha + \left[\frac{t_1\,(t_1-1)}{2} + (t_2+1-t_1)\,(t_1-1)\right]\beta_1 \\
&\quad + \left[\frac{(t_2-t_1)\,(t_2-t_1+1)}{2} + (t_2-t_1)\right]\beta_2 + \beta_3 + \alpha + \beta_1\,(t_1-1) + \beta_2\,(t_2-t_1) + \beta_3\,(t_2+2-t_2) \\
&= h_0 + t\alpha + \left[\frac{t_1\,(t_1-1)}{2} + (t-t_1)\,(t_1-1)\right]\beta_1 \\
&\quad + \left[\frac{(t_2-t_1)\,(t_2-t_1+1)}{2} + 2\,(t_2-t_1)\right]\beta_2 + (1+2)\,\beta_3.
\end{aligned}
$$

Continuing the iteration, height at age t, with $t_2 < t \leq t_3$ can be written as

$$
\begin{aligned}
h_t &= h_0 + t\alpha + \left[\frac{t_1\,(t_1-1)}{2} + (t-t_1)\,(t_1-1)\right]\beta_1 + \left[\frac{(t_2-t_1)\,(t_2-t_1+1)}{2} + (t-t_2)\,(t_2-t_1)\right]\beta_2 \\
&\quad + \frac{(t-t_2)\,(t-t_2+1)}{2}\beta_3 \qquad (11)
\end{aligned}
$$

and in the last month of the third period (that is, at the peak of the AGS) we have

$$
\begin{aligned}
h_{t_3} &= h_0 + t_3\alpha + \left[\frac{t_1\,(t_1-1)}{2} + (t_3-t_1)\,(t_1-1)\right]\beta_1 + \left[\frac{(t_2-t_1)\,(t_2-t_1+1)}{2} + (t_3-t_2)\,(t_2-t_1)\right]\beta_2 \\
&\quad + \frac{(t_3-t_2)\,(t_3-t_2+1)}{2}\beta_3. \qquad (12)
\end{aligned}
$$

Using a similar procedure, we can see that during the last interval, for $t_3 < t \leq t_4$, we have

$$
\begin{aligned}
h_t &= h_0 + t\alpha + \left[\frac{t_1\,(t_1-1)}{2} + (t-t_1)\,(t_1-1)\right]\beta_1 + \left[\frac{(t_2-t_1)\,(t_2-t_1+1)}{2} + (t-t_2)\,(t_2-t_1)\right]\beta_2 \\
&\quad + \left[\frac{(t_3-t_2)\,(t_3-t_2+1)}{2} + (t-t_3)\,(t_3-t_2)\right]\beta_3 + \frac{(t-t_3)\,(t-t_3+1)}{2}\beta_4, \qquad (13)
\end{aligned}
$$

so that at t_4, when adult height is achieved we have

$$
\begin{aligned}
h_{t_4} &= h_0 + t_4\alpha + \left[\frac{t_1\,(t_1-1)}{2} + (t_4-t_1)\,(t_1-1)\right]\beta_1 \\
&\quad + \left[\frac{(t_2-t_1)\,(t_2-t_1+1)}{2} + (t_4-t_2)\,(t_2-t_1)\right]\beta_2 \\
&\quad + \left[\frac{(t_3-t_2)\,(t_3-t_2+1)}{2} + (t_4-t_3)\,(t_3-t_2)\right]\beta_3 + \frac{(t_4-t_3)\,(t_4-t_3+1)}{2}\beta_4.
\end{aligned}
$$

This also implies that for individuals who have already achieved adult height we have

$$
h_{t_4} = h_0 + \delta
$$

where

$$
\begin{aligned}
\delta &= t_4\alpha + \left[\frac{t_1\,(t_1-1)}{2} + (t_4-t_1)\,(t_1-1)\right]\beta_1 + \left[\frac{(t_2-t_1)\,(t_2-t_1+1)}{2} + (t_4-t_2)\,(t_2-t_1)\right]\beta_2 \\
&\quad + \left[\frac{(t_3-t_2)\,(t_3-t_2+1)}{2} + (t_4-t_3)\,(t_3-t_2)\right]\beta_3 + \frac{(t_4-t_3)\,(t_4-t_3+1)}{2}\beta_4.
\end{aligned}
$$

From a comparisions of equations (7), (9), (11), and (13), it follows that height at any age can be written as

$$h_t = h_0 + \alpha 1(t \leq t_4)t + \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3 + \beta_4 v_4 + \delta 1(t > t_4),$$

where the v functions are deterministic functions of age and/or the location of the kinks:

$$\begin{aligned}
v_1 &= 1(t \leq t_4)\frac{\min(t,t_1)(\min(t,t_1)-1)}{2} + 1(t_1 < t \leq t_4)(t-t_1)(t_1-1) \\
v_2 &= 1(t_1 < t \leq t_4)\frac{(\min(t,t_2)-t_1)(\min(t,t_2)-t_1+1)}{2} + 1(t_2 < t \leq t_4)(t-t_2)(t_2-t_1) \\
v_3 &= 1(t_2 < t \leq t_4)\frac{(\min(t,t_3)-t_2)(\min(t,t_3)-t_2+1)}{2} + 1(t_3 < t \leq t_4)(t-t_3)(t_3-t_2) \\
v_4 &= 1(t_3 < t \leq t_4)\frac{(t-t_3)(t-t_3+1)}{2}
\end{aligned}$$

and where the two following constraints must hold:

$$\begin{aligned}
& \alpha + \beta_1(t_1-1) + \beta_2(t_2-t_1) + \beta_3(t_3-t_2) + \beta_4(t_4-t_3) = 0, \\
\delta &= t_4\alpha + \left[\frac{t_1(t_1-1)}{2} + (t_4-t_1)(t_1-1)\right]\beta_1 + \left[\frac{(t_2-t_1)(t_2-t_1+1)}{2} + (t_4-t_2)(t_2-t_1)\right]\beta_2 \\
& + \left[\frac{(t_3-t_2)(t_3-t_2+1)}{2} + (t_4-t_3)(t_3-t_2)\right]\beta_3 + \frac{(t_4-t_3)(t_4-t_3+1)}{2}\beta_4.
\end{aligned}$$

The first constraint imposes that growth must be equal to zero when adult height is reached at time $t = t_4$, while the second imposes that height is constant (and equal to adult height) for any age larger than t_4.

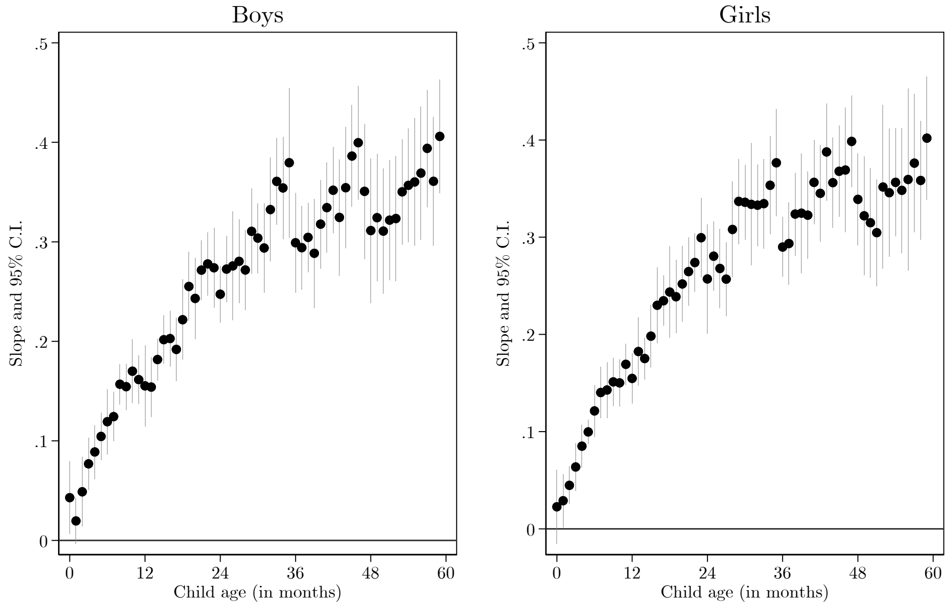

Figure A.1: DHS Child height vs maternal years of education

Source: Authors' calculations from DHS data. For each age (in months) the figure shows the point estimate and a 95% confidence interval of the slope of a regression, estimated with OLS, of child height (in cms) on the number of years of schooling of the mother. All estimates do not use sampling weights, and the confidence intervals are calculated allowing for correlation of residuals within each survey primary stage unit. Total sample size is $n = 1,598,229$.

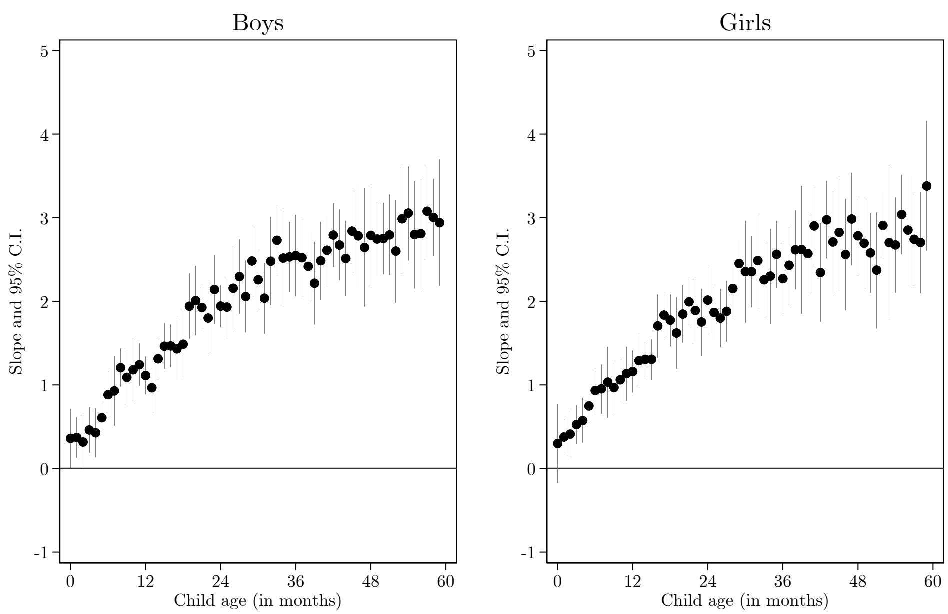

Figure A.2: DHS: Child height vs. paternal education

Source: Authors' calculations from DHS data. For each age (in months) the figure shows the point estimate and a 95% confidence interval of the slope of a regression, estimated with OLS, of child height (in cms) on a dummy variable equal to one if the father has at least a secondary education. All estimates do not use sampling weights, and the confidence intervals are calculated allowing for correlation of residuals within each survey primary stage unit. Total sample size is $n = 1,301,601$.

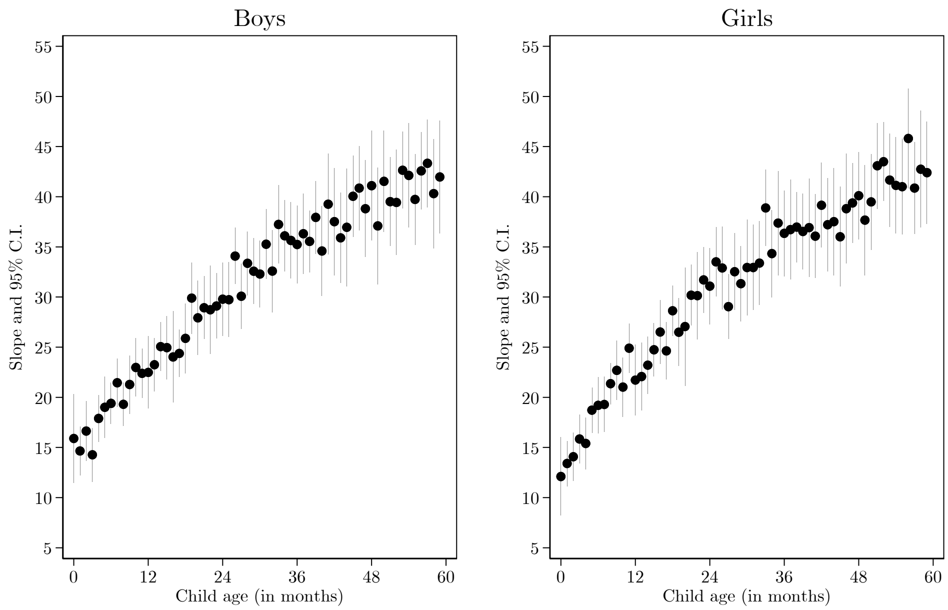

Figure A.3: DHS: Child height vs. log-maternal height

Source: Authors' calculations from DHS data. For each age (in months) the figure shows the point estimate and a 95% confidence interval of the slope of a regression, estimated with OLS, of child height (in cms) on the logarithm of maternal height (in cms.) as well as country fixed effects. All estimates do not use sampling weights, and the confidence intervals are calculated allowing for correlation of residuals within each survey primary stage unit. Total sample size is $n = 1,598,799$.

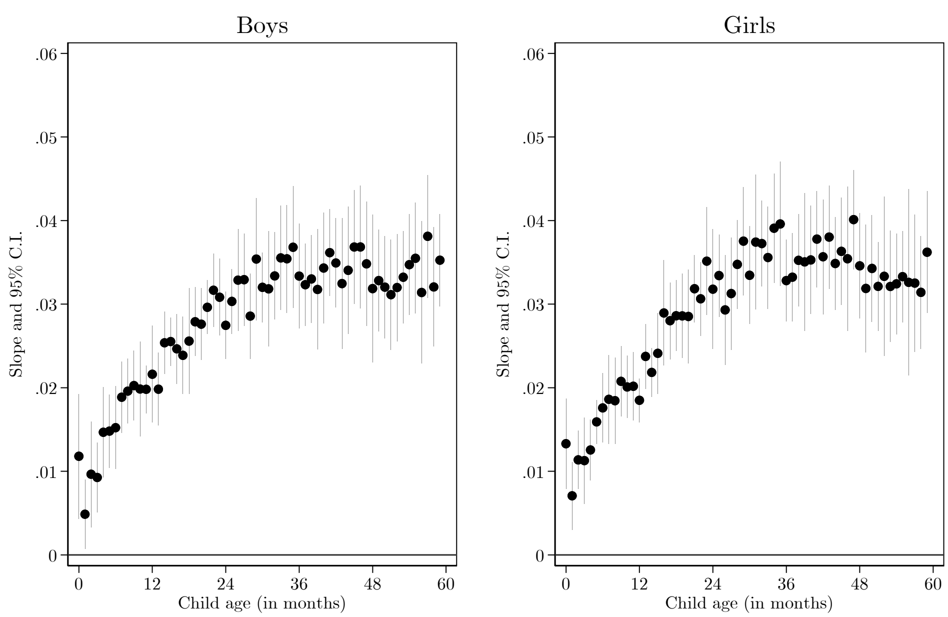

Figure A.4: DHS: Child (log) height vs. maternal education

Source: Authors' calculations from DHS data. For each age (in months) the figure shows the point estimate and a 95% confidence interval of the slope of a regression, estimated with OLS, of the logarithm of child height (in cms) on a dummy variable equal to one if the mother has completed at least secondary education. All estimates do not use sampling weights, and the confidence intervals are calculated allowing for correlation of residuals within each survey primary stage unit. Total sample size is $n = 1,409,607$.

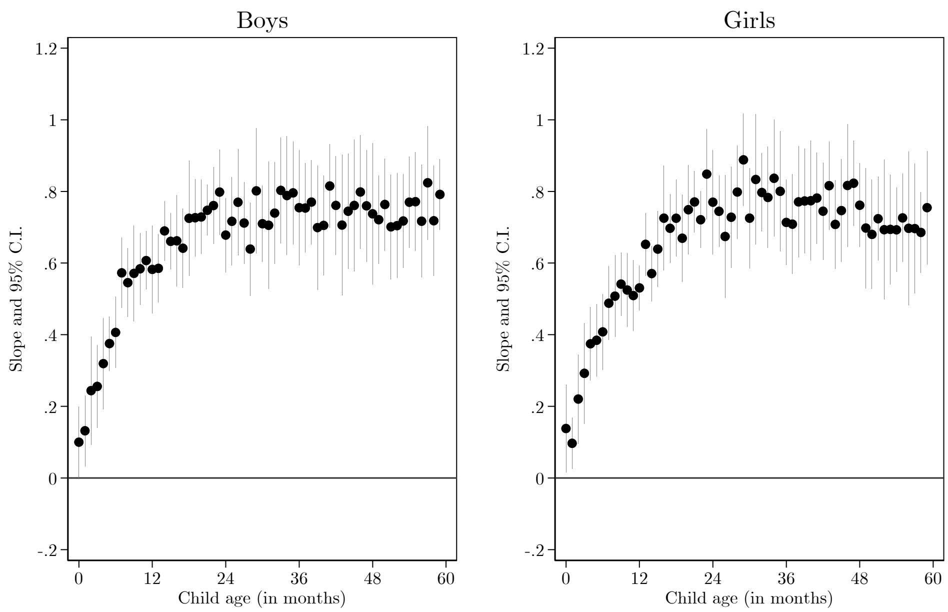

Figure A.5: DHS: Child HAZ vs. maternal education

Source: Authors' calculations from DHS data. For each age (in months) the figure shows the point estimate and a 95% confidence interval of the slope of a regression, estimated with OLS, of child height-for-age z-scores (HAZ) on a dummy variable equal to one if the mother has completed at least secondary education. HAZ are stored in variable hw5 in DHS data. All estimates do not use sampling weights, and the confidence intervals are calculated allowing for correlation of residuals within each survey primary stage unit. Total sample size is $n = 1,481,455$.

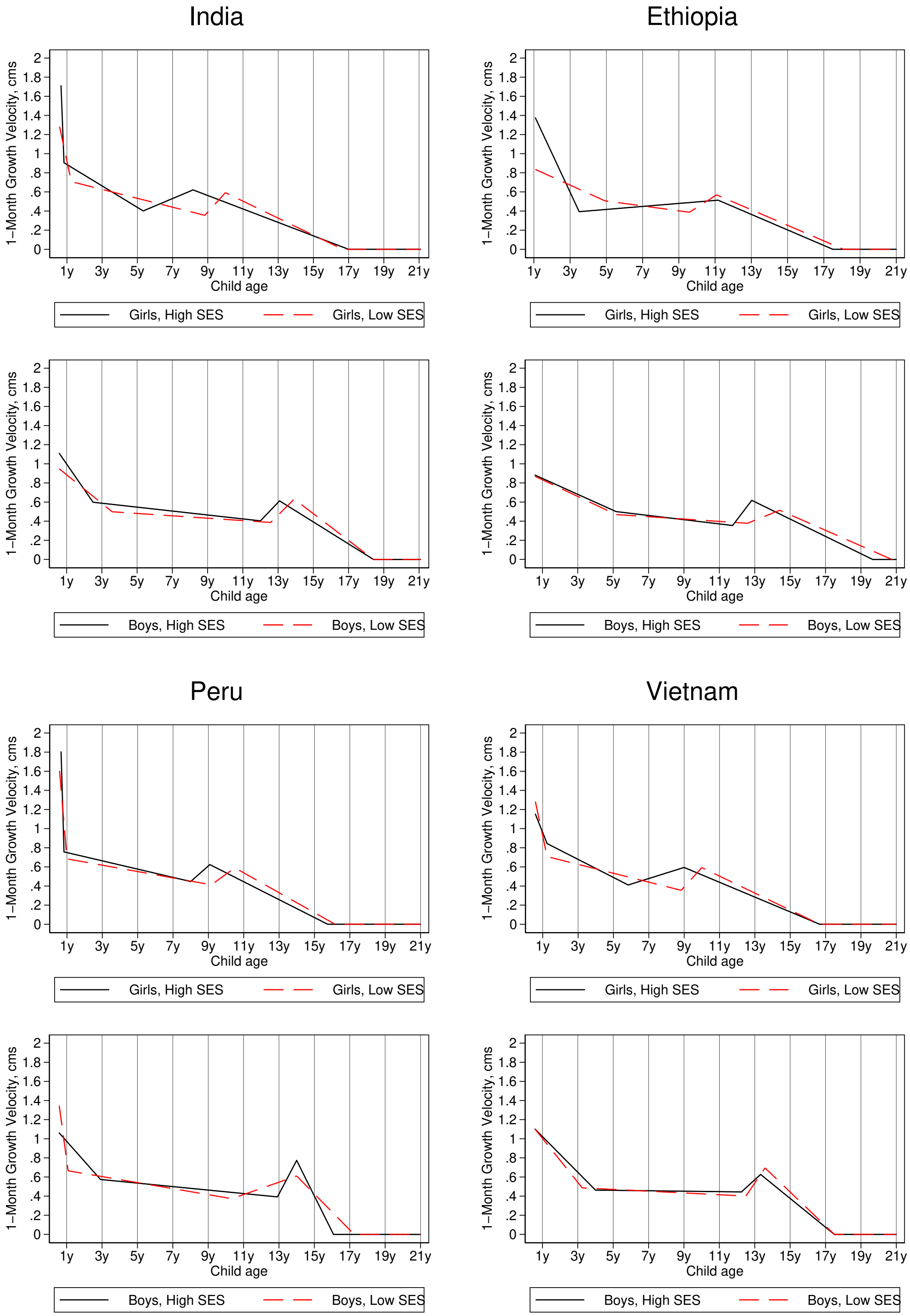

Figure A.6: YLS: Growth velocity curves by country, gender and maternal education

Source: Authors' calculations from YLS data. The lines show differences in height growth velocity between high-SES and low-SES children, as predicted by the piece-wise continuous regression model described in Section 4.1.1, estimated separately for boys and girls and for each YLS country. High-SES is binary and equal to one when the mother has completed at least secondary schooling.

Table A.1: **List of Demographic and Health Surveys Used in Analysis**

Country	(1) Survey years	(2) File	(3) Age	(4) Obs.	(5) Height	(6) s.d.	(7) HAZ	(8) Stun -ted	(9) WHO HAZ ref.	(10) Mother ≥ sec.	(11) Mother Years sch.	(12) Sam -ple
Albania	2008-2009	al50	0-59	1520	87.5	16.6	-.55	.22	New	.26	9.6	
Albania	2017-2018	al71	0-59	2510	87.6	14.5	-.39	.13	New	.22	11.9	
Angola	2015-2016	ao71	0-59	6556	81.7	14.3	-1.53	.37	New	.04	3.9	
Armenia	2000	am42	0-59	1539	86.5	14	-.66	.13		.87	11.3	
Armenia	2005	am54	0-59	1300	86.2	16	-.42	.11		.48	9.2	
Armenia	2010	am61	0-59	1406	83.9	15	-.78	.21	New	.87	11.9	
Armenia	2015-2016	am72	0-59	1603	88	15.7	-.22	.11	New	.88	11.9	
Azerbaijan	2006	az52	0-59	2089	82.8	14.7	-1.1	.26	New	.21	10.6	
Bangladesh	1996-1997	bd3a	0-59	5066	79.4	13.5	-2.14	.55		.02	2.4	EM
Bangladesh	1999-2000	bd41	0-59	5526	80.2	13.9	-1.81	.44		.04	3.2	EM
Bangladesh	2004	bd4j	0-59	6048	81.1	13.2	-1.75	.42		.03	3.8	EM
Bangladesh	2007	bd51	0-59	5397	82.1	13.2	-1.72	.42	New	.07	4.9	EM
Bangladesh	2011	bd61	0-59	7865	83	13.9	-1.67	.41	New	.06	5.5	EM
Bangladesh	2014	bd72	0-59	7134	82.9	13.5	-1.54	.37	New	.07	6.1	EM
Benin	1996	bj31	0-35	2652	73.4	10.3	-1.08	.26		0	.8	
Benin	2001	bj41	0-59	4518	81.4	13.5	-1.29	.31		0	1.3	
Benin	2006	bj51	0-59	13429	80.1	14.3	-1.7	.43	New	0	1.3	
Benin	2011-2012	bj61	0-59	11372	81.1	14.8	-1.57	.44	New	.01	1.6	
Benin	2017-2018	bj71	0-59	12089	82.3	14.1	-1.43	.32	New	0	2.2	
Bolivia	1989	bo01	3-36	2682	76	9	-1.46	.35		.32	5.1	
Bolivia	1993-1994	bo31	0-35	3015	74.7	10.2	-1.15	.27		.09	5.6	
Bolivia	1998	bo3b	0-59	6374	83.1	14.1	-1.27	.3		.1	5.7	
Bolivia	2003-2004	bo41	0-59	9333	83.9	13.7	-1.24	.27		.14	6.5	
Bolivia	2008	bo51	0-59	7817	83.7	13.8	-1.21	.26	New	.17	7.5	
Brazil	1986	br01	0-59	1180	83.5	13.8	-1.31	.29		.11	3.4	
Brazil	1996	br31	0-59	4179	85.6	15.1	-.49	.12		.13	5.5	
Burkina Faso	1992-1993	bf21	0-59	4576	81.5	13.6	-1.2	.3		0	1	
Burkina Faso	1998-1999	bf31	0-59	4763	80.6	13.5	-1.42	.36		.01	.7	
Burkina Faso	2003	bf43	0-59	8789	80.9	13.5	-1.5	.38		0	.8	
Burkina Faso	2010	bf62	0-59	6723	82.4	13.8	-1.38	.34	New	0	1	
Burundi	1987	bu01	3-36	1936	75.1	9	-1.8	.46		.06	1.3	
Burundi	2010-2011	bu61	0-59	3494	80.2	13	-2.11	.55	New	0	2.9	
Burundi	2016-2017	bu71	0-59	6062	80.8	13.1	-2.13	.55	New	.01	3.2	
Cambodia	2000	kh42	0-59	3772	81.2	13.8	-1.8	.46		.01	2.7	
Cambodia	2005-2006	kh51	0-59	3679	81.5	13	-1.85	.45	New	.01	3.1	
Cambodia	2010-2011	kh61	0-59	3806	82.4	13	-1.65	.4	New	.03	4.3	
Cambodia	2014	kh73	0-59	4427	82.5	13.7	-1.4	.33	New	.04	5.3	
Cameroon	1991	cm21	0-59	2688	81.8	13.9	-1.08	.23		.01	4.3	
Cameroon	1998	cm31	0-35	1871	73.7	10.4	-1.1	.28		.01	4.9	
Cameroon	2004	cm44	0-59	3329	81.5	14.3	-1.26	.31		.01	4.8	
Cameroon	2011	cm61	0-59	5184	82	13.8	-1.24	.32	New	.01	5.2	
Central Af Rep	1994-1995	cf31	0-35	2433	73.1	10.1	-1.39	.34		0	2.1	
Chad	1996-1997	td31	0-59	5852	80.4	13.8	-1.47	.38		0	1	
Chad	2004	td41	0-59	4650	81.3	14	-1.47	.39		.01	1.3	
Chad	2014-2015	td71	0-59	10422	82.2	13.8	-1.61	.43	New	.01	1.5	
Colombia	1986	co01	3-36	1332	77.8	9	-1.24	.26		.35	4.9	
Colombia	1995	co31	0-59	4561	84.5	13.8	-.88	.15		.15	6.4	
Colombia	2000	co41	0-59	4226	84.2	14.3	-.85	.14		.21	7	
Colombia	2004-2005	co53	0-59	12480	84.8	14.3	-.74	.12		.25	7.6	
Colombia	2009-2010	co61	0-59	16041	85.5	14.3	-.87	.15	New	.29	8.1	
Comoros	1996	km31	0-35	999	72.8	10.4	-1.39	.34		.01	2.3	
Comoros	2012	km61	0-59	2700	83.4	15.6	-1.06	.28	New	.04	4.4	

(Continued)

Country	(1) Survey years	(2) File	(3) Age	(4) Obs.	(5) Height	(6) s.d.	(7) HAZ	(8) Stun -ted	(9) WHO HAZ ref.	(10) Mother ≥ sec.	(11) Mother Years sch.	(12) Sam -ple
Congo, DR	2007	cd50	0-59	3647	80.5	14.1	-1.62	.45	New	.05	4.6	
Congo, DR	2013-2014	cd61	0-59	8391	81	13.5	-1.66	.44	New	.06	4.9	
Congo, Rep.	2005	cg51	0-59	4058	82.6	14.7	-1.05	.29	New	.02	6.8	
Congo, Rep.	2011-2012	cg60	0-59	4531	82.6	14.1	-1.13	.27	New	.02	6.1	
Cote d'Ivoire	1994	ci35	0-35	3507	74.1	10.6	-1.09	.24		0	1.8	
Cote d'Ivoire	1998-1999	ci3a	0-59	1589	81.8	13.8	-1.04	.23		.01	2.4	
Cote d'Ivoire	2011-2012	ci62	0-59	3297	81.9	13.9	-1.25	.3	New	.01	1.8	
Dominican Rep.	1986	dr01	6-36	1976	79.4	8.4	-1.01	.23		.25	5.3	
Dominican Rep.	1991	dr21	0-59	3276	83.4	14.6	-.97	.2		.12	6.8	
Dominican Rep.	1996	dr31	0-59	3811	85.3	14.2	-.66	.13		.12	6.7	
Dominican Rep.	2002	dr4a	0-59	9444	86	14.8	-.41	.09		.12	7.4	
Dominican Rep.	2007	dr5a	0-59	800	85.4	14.3	-.84	.17	New	.16	5.1	
Dominican Rep.	2007	dr52	0-59	9486	86.7	15	-.52	.11	New	.05	8.2	
Dominican Rep.	2013	dr6a	0-59	789	85.5	15.6	-.53	.1	New	.14	7.1	
Dominican Rep.	2013	dr61	0-59	3236	87	15.3	-.34	.08	New	.26	9.8	
Egypt	1988-1989	eg01	3-36	2080	76.3	9.4	-1.34	.3		.2	4.1	EM
Egypt	1992-1993	eg21	0-59	7713	84.8	14.7	-.97	.24		.17	4.2	EM
Egypt	1995-1996	eg33	0-59	10847	83.2	14.7	-1.23	.3		.2	4.6	EM
Egypt	2000	eg42	0-59	10719	83.7	15.1	-.85	.18		.29	5.9	EM
Egypt	2003	eg4a	0-59	6247	84.2	14.9	-.93	.16		.32	6.2	EM
Egypt	2005	eg51	0-59	13169	84.7	15.5	-.81	.21		.38	6.7	EM
Egypt	2008	eg5a	0-59	10454	83.2	15.3	-1	.29	New	.43	7.5	EM
Egypt	2014	eg61	0-59	15179	86.4	16.2	-.47	.2	New	.52	9	EM
Eswatini	2006-2007	sz51	0-59	2104	82.5	14.2	-1.18	.27	New	.12	7.6	
Ethiopia	2000	et41	0-59	9060	80.6	13	-1.9	.48		.02	1.1	
Ethiopia	2005	et51	0-59	4186	81.5	14.1	-1.65	.43		.01	1.2	
Ethiopia	2010-2011	et61	0-59	9879	82.1	13.5	-1.61	.42	New	.01	1.5	
Ethiopia	2016	et71	0-59	9061	82.8	13.7	-1.36	.36	New	.01	2.3	
Gabon	2000-2001	ga41	0-59	3569	81.9	14.5	-1.05	.24		.01	5.8	
Gabon	2012	ga60	0-59	3483	82.8	14.7	-.98	.23	New	.01	6.4	
Gambia, The	2013	gm60	0-59	3362	82.7	16.1	-1.08	.26	New	.04	2.8	
Ghana	1988	gh01	3-36	1989	76.2	8.4	-1.32	.3		.05	4.6	
Ghana	1993-1994	gh31	0-35	1964	74	10.1	-1.13	.26		.01	4.7	
Ghana	1998-1999	gh41	0-59	2838	82.4	13.3	-1.19	.28		.02	4.1	
Ghana	2003	gh4b	0-59	3200	82.2	13.2	-1.31	.31		.03	4	
Ghana	2008	gh5a	0-59	2521	83.3	14	-1.08	.28	New	.05	4.6	
Ghana	2014	gh72	0-59	2739	83.6	13.9	-.98	.19	New	.06	5	
Guatemala	1987	gu01	3-36	2251	73.1	8.7	-2.27	.58		.08	2.2	
Guatemala	1995	gu34	0-59	8792	79	13.1	-2.16	.56		.02	2.3	
Guatemala	1998-1999	gu41	0-59	4024	80.2	13.2	-2.06	.52		.03	2.6	
Guatemala	2014-2015	gu71	0-59	11787	81.8	13.7	-1.88	.46	New	.08	5.1	
Guinea	1999	gn41	0-59	4622	81.7	14.4	-1	.26		0	1	
Guinea	2005	gn52	0-59	2753	81.1	14.1	-1.3	.35		0	.7	
Guinea	2012	gn62	0-59	3216	82.9	14.1	-1.1	.31	New	.01	1.4	
Guinea	2018	gn71	0-59	3542	83.2	14.9	-1.13	.31	New	.02	1.8	
Guyana	2009	gy5i	0-59	1703	84.7	15	-1	.22	New	.24	8.2	
Haiti	1994-1995	ht31	0-59	2882	81.7	13.8	-1.35	.32		0	2.3	
Haiti	2000	ht42	0-59	5627	83.2	14	-1.06	.23		.01	2.6	
Haiti	2005-2006	ht52	0-59	2596	82.4	14.4	-1.12	.24		.01	3.6	
Haiti	2012	ht61	0-59	4042	83	14.2	-1.01	.22	New	.01	4.6	
Haiti	2016-2017	ht71	0-59	5648	84.4	14.8	-.94	.22	New	.02	5.5	
Honduras	2005-2006	hn52	0-59	9333	84.4	12.2	-1.51	.35	New	.06	4.8	
Honduras	2011-2012	hn62	0-59	10014	83.1	13.9	-1.22	.25	New	.09	6.1	

(Continued)

Country	(1) Survey years	(2) File	(3) Age	(4) Obs.	(5) Height	(6) s.d.	(7) HAZ	(8) Stun -ted	(9) WHO HAZ ref.	(10) Mother	(11)	(12) Sam -ple
										≥ sec.	Years sch.	
India	1992-1993	ia23	0-47	29025	76.1	12.1	-1.9	.48		0	3.3	EM
India	1998-2000	ia42	0-35	27201	72.2	10.5	-1.73	.43		.08	4	EM
India	2005-2006	ia52	0-59	43582	82.4	13.8	-1.71	.44	New	.06	5.1	
India	2015-2016	ia74	0-59	236923	83.4	13.9	-1.48	.38	New	.09	6.2	
Jordan	1990	jo21	0-59	6887	84.5	14.2	-.8	.17		.14	7.1	EM
Jordan	1997	jo31	0-59	5675	86.3	14	-.56	.09		.28	9	EM
Jordan	2002	jo42	0-59	4936	85.6	14.5	-.57	.1		.83	10	EM
Jordan	2007	jo51	0-59	4764	85.7	14.9	-.52	.15	New	.2	10.7	EM
Jordan	2009	jo61	0-59	4429	85.9	14.3	-.52	.1	New	.18	11.2	EM
Jordan	2012	jo6c	0-59	6354	87.4	14.3	-.47	.09	New	.18	11.3	EM
Kazakhstan	1995	kk31	0-35	748	76.5	10.4	-.66	.17		.8	10.8	
Kazakhstan	1999	kk42	0-59	580	86.1	13.8	-.64	.11		.85	10.9	
Kenya	1993	ke31	0-59	5084	82.4	13.5	-1.39	.32		.02	5.7	
Kenya	1998	ke3a	0-35	3109	74.5	11.2	-1.22	.31		.14	6.8	
Kenya	2003	ke42	0-59	4873	82.3	14.2	-1.19	.29		.1	6.4	
Kenya	2008-2009	ke52	0-59	5333	83	14.5	-1.34	.34	New	.1	6.4	
Kenya	2014	ke72	0-59	18941	83.9	14	-1.17	.27	New	.11	6.5	
Kyrgyz Rep.	1997	ky31	0-35	991	75.1	10.5	-1.1	.25		.87	10.8	
Kyrgyz Rep.	2012	ky61	0-59	4068	83.9	13.7	-.82	.18	New	.82	12.2	
Lesotho	2004-2005	ls41	0-59	1463	79.3	14.4	-1.6	.38		.05	6.8	
Lesotho	2009-2010	ls61	0-59	1675	81.2	13.7	-1.54	.39	New	.05	7.2	
Lesotho	2014	ls71	0-59	1349	80.7	14.1	-1.49	.35	New	.09	7.7	
Liberia	2006-2007	lb51	0-59	4561	81.6	13.6	-1.49	.38	New	.03	2.7	
Liberia	2013	lb6a	0-59	3261	82.2	13.6	-1.27	.31	New	.02	2.7	
Madagascar	1992	md21	0-59	4230	78.4	13.1	-2.07	.53		.01	3.8	
Madagascar	1997	md31	0-35	3098	71.5	9.8	-1.83	.46		.01	3.4	
Madagascar	2003-2004	md42	0-59	4738	80.7	14.1	-1.71	.44		.05	4.7	
Madagascar	2008-2009	md51	0-59	5521	81.6	14.4	-1.77	.48	New	.01	3.1	
Malawi	1992	mw21	0-59	3361	78.7	13.4	-1.85	.46		.02	3.2	
Malawi	2000	mw41	0-59	9728	79	13.3	-1.81	.47		.03	3.7	
Malawi	2004-2005	mw4e	0-59	8674	79.4	13.9	-1.83	.48		.03	4.1	
Malawi	2010	mw61	0-59	4831	81.7	13.7	-1.78	.46	New	.04	4.8	
Malawi	2015-2016	mw7a	0-59	5247	83.1	13.5	-1.5	.35	New	.06	5.6	
Maldives	2009	mv52	0-59	2450	84.2	15.1	-.93	.19	New	.01	7.1	EM
Maldives	2016-2017	mv71	0-59	2417	86.7	15	-.85	.15	New	.03	9.7	
Mali	1987	ml01	3-36	1559	75.8	9.5	-1.05	.23		.01	1.2	
Mali	1995-1996	ml31	0-35	5001	72.9	10.2	-1.22	.3		0	.8	
Mali	2001	ml41	0-59	10006	80.2	14.3	-1.48	.38		0	.8	
Mali	2006	ml53	0-59	11638	81.5	14.3	-1.43	.38	New	0	.9	
Mali	2012-2013	ml6a	0-59	4591	84	13.6	-1.43	.38	New	.01	1.2	
Mali	2018	ml7h	0-59	8711	83.7	14.3	-1.09	.27	New	0	1.9	
Moldova	2005	mb53	0-59	1379	86.9	14.9	-.12	.08		.22	11.2	
Morocco	1987	ma01	0-59	5494	82.4	13.8	-1.25	.29		.06	.9	EM
Morocco	1992	ma21	0-59	4651	83.8	13.9	-1.12	.24		.01	1.2	
Morocco	2003-2004	ma43	0-59	5677	85.2	14.8	-.7	.19		.01	2.4	
Mozambique	1997	mz31	0-35	3583	71.9	10.4	-1.46	.36		0	2.3	
Mozambique	2003-2004	mz41	0-59	8286	80.1	13.9	-1.66	.39		0	2.3	
Mozambique	2011	mz62	0-59	9716	81.2	14.3	-1.58	.39	New	.02	3.4	
Myanmar	2015-2016	mm71	0-59	4231	83.4	13.1	-1.36	.31	New	.03	5.2	
Namibia	1992	nm21	0-59	2766	79.7	14.6	-1.27	.29		.04	5.2	
Namibia	2000	nm41	0-59	3035	81.6	14.6	-.97	.22		.09	6.5	
Namibia	2006-2007	nm51	0-59	3840	81.1	14.5	-1.24	.29	New	.1	7.1	
Namibia	2013	nm61	0-59	1879	82.1	14.5	-1.03	.23	New	.14	8	

(Continued)

Country	(1) Survey years	(2) File	(3) Age	(4) Obs.	(5) Height	(6) s.d.	(7) HAZ	(8) Stun -ted	(9) WHO HAZ ref.	(10) Mother ≥ sec.	(11) Mother Years sch.	(12) Sam -ple
Nepal	1996	np31	0-35	3812	72	9.5	-1.96	.49		.02	1.1	EM
Nepal	2001	np41	0-59	6252	80.6	12.7	-2.01	.5		.03	1.4	EM
Nepal	2006	np51	0-59	5283	81.8	12.6	-1.96	.5	New	.04	2.4	
Nepal	2011	np60	0-59	2359	82.5	12.9	-1.71	.42	New	.09	3.6	
Nepal	2016-2017	np7h	0-59	2379	83.1	13.2	-1.54	.36	New	.09	5.1	
Nicaragua	1997-1998	nc31	0-59	7164	83.4	14	-1.23	.27		.08	4.5	
Nicaragua	2001	nc41	0-59	6096	83.9	13.8	-1.01	.22		.08	4.7	
Niger	1992	ni21	0-59	4888	79.7	13.9	-1.46	.35		0	.9	
Niger	1998	ni31	0-35	4035	71.8	9.8	-1.61	.4		0	.9	
Niger	2006	ni51	0-59	3867	79.6	13.4	-1.79	.45		0	1	
Niger	2012	ni61	0-59	5146	81.2	14.4	-1.67	.42	New	0	1	
Nigeria	1990	ng21	0-59	6151	79.5	13.7	-1.62	.4		.07	3	
Nigeria	2003	ng4b	0-59	4786	80.2	14.6	-1.47	.38		.1	4.3	
Nigeria	2008	ng53	0-59	23034	80.3	15.9	-1.55	.42	New	.13	4.4	
Nigeria	2013	ng6a	0-59	26797	82.1	14.7	-1.34	.36	New	.17	5	
Nigeria	2018	ng7a	0-59	11474	82.5	14.1	-1.5	.36	New	.25	6.2	
Pakistan	1990-1991	pk21	0-59	4669	78.7	13.3	-2.07	.51		.02	1.9	EM
Pakistan	2012-2013	pk61	0-59	3626	80.6	15.2	-1.78	.45	New	.11	3.8	EM
Pakistan	2017-2018	pk71	0-59	4226	82.5	13.9	-1.55	.38	New	.12	4.4	EM
Paraguay	1990	py21	0-59	3682	84.3	14.9	-.81	.15		.07	5.5	
Peru	1991-1992	pe21	0-59	7870	82.5	13.5	-1.42	.33		.19	6.2	
Peru	1996	pe31	0-59	15258	83.3	13.8	-1.29	.3		.17	5.9	
Peru	2000	pe41	0-59	11794	83.7	13.6	-1.3	.29		.18	7.1	
Peru	2003-2008	pe51	0-59	10493	83.9	13.7	-1.36	.3	New	.24	8.1	
Peru	2009	pe5i	0-59	9406	84.1	14	-1.3	.27	New	.27	8.2	
Peru	2010	pe61	0-59	8804	84.3	13.9	-1.28	.26	New	.28	8.4	
Peru	2011	pe6a	0-59	8754	84.5	13.9	-1.22	.23	New	.28	8.4	
Peru	2012	pe6i	0-59	9228	84.8	14	-1.15	.2	New	.3	8.7	
Rwanda	1992	rw21	0-59	4414	80	12.9	-1.88	.47		.01	2.9	
Rwanda	2000	rw41	0-59	6378	80.8	13.8	-1.57	.41		.03	4	
Rwanda	2005	rw53	0-59	3781	79.8	13.5	-1.73	.44		.02	3.9	
Rwanda	2010-2011	rw61	0-59	4121	82.6	13	-1.74	.44	New	.02	3.9	
Rwanda	2014-2015	rw70	0-59	3601	81.9	13.6	-1.54	.37	New	.04	4.6	
S Tome and Pr.	2008-2009	st50	0-59	1705	83.2	15.6	-1.07	.28	New	0	4.6	
Senegal	1986	sn01	6-36	640	77.4	7.7	-1.17	.23		.05	1.2	
Senegal	1992-1993	sn21	0-59	4663	81.9	14.1	-1.1	.25		.01	1.1	
Senegal	2005	sn4h	0-59	2936	83	14.3	-.79	.17		.01	1.6	
Senegal	2010-2011	sn61	0-59	3927	82.8	14.5	-1.21	.29	New	0	1.5	
Senegal	2012-2013	sn6d	0-59	6067	83.7	14.5	-.97	.2	New	0	1.7	
Senegal	2014	sn70	0-59	6110	84.1	14.3	-1.05	.21	New	0	1.7	
Senegal	2015	sn7h	0-59	6235	83.6	14.4	-1.1	.22	New	0	1.8	
Senegal	2016	sn7i	0-59	6062	84.2	14.6	-1	.19	New	0	1.9	
Senegal	2017	sn7z	0-59	10831	84.6	14.4	-.97	.19	New	0	2.2	
Sierra Leone	2008	sl51	0-59	2282	81.3	14.7	-1.22	.34	New	.01	1.6	
Sierra Leone	2013	sl61	0-59	4696	82.4	15.9	-1.35	.38	New	.02	2.2	
South Africa	2016	za71	0-59	1126	84.6	15	-1.15	.26	New	.27	10.1	
Sri Lanka	1987	lk01	3-36	2024	76.8	8.7	-1.41	.3		.59	6.4	EM
Tajikistan	2012	tj61	0-59	4767	83.6	13.9	-1.06	.25	New	.46	9.8	
Tajikistan	2017	tj71	0-59	5913	85.5	13.4	-.82	.18	New	.44	10.3	
Tanzania	1991-1992	tz21	0-59	6745	79.4	13.7	-1.78	.44		0	3.9	
Tanzania	1996	tz3a	0-59	5575	79.5	13.6	-1.77	.44		0	4.6	
Tanzania	1999	tz41	0-59	2584	80.1	13.6	-1.68	.39		0	5	
Tanzania	2004-2005	tz4i	0-59	7302	80.6	13.5	-1.56	.36		0	4.9	

(Continued)

Country	(1) Survey years	(2) File	(3) Age	(4) Obs.	(5) Height	(6) s.d.	(7) HAZ	(8) Stun -ted	(9) WHO HAZ ref.	(10) Mother $\geq$ sec.	(11) Mother Years sch.	(12) Sam -ple
Tanzania	2009-2010	tz63	0-59	6953	81.3	13.9	-1.62	.4	New	.01	5	
Tanzania	2015-2016	tz7b	0-59	9049	82	13.8	-1.43	.34	New	.09	5.6	
Thailand	1987	th01	3-36	1862	77.8	9.1	-1.06	.19		.2	5.8	EM
Timor-Leste	2009-2010	tl61	0-59	8172	81.5	13.3	-2.08	.57	New	.14	5.7	
Timor-Leste	2016	tl71	0-59	6198	83.2	14.3	-1.53	.46	New	.27	7.2	
Togo	1988	tg01	0-36	1713	74.2	10.5	-1.29	.29		.09	1.8	
Togo	1998	tg31	0-35	3770	74	10.4	-1.02	.23		0	1.6	
Togo	2013-2014	tg61	0-59	3230	83.2	13.3	-1.26	.28	New	.01	3.1	
Trinidad Tob.	1987	tt01	3-36	847	80.8	10.1	-.25	.05		.5	7.7	
Tunisia	1988	tn01	3-36	2060	78.1	9.6	-.81	.18		.09	2.6	EM
Turkey	1993	tr31	0-59	3187	84.1	14	-.84	.19		.07	4.3	EM
Turkey	1998	tr41	0-59	2844	84.3	14.2	-.74	.17		.09	4.8	
Turkey	2003-2004	tr4a	0-59	4074	86.1	14.2	-.57	.14		.1	4.7	EM
Turkey	2013	tr62	0-59	2823	87.3	14.3	-.41	.1	New	.14	6.3	
Uganda	1988-1989	ug01	0-59	3737	79.4	13.8	-1.76	.43		.1	3.4	
Uganda	1995	ug33	0-47	4743	76.7	12	-1.47	.35		.06	4	
Uganda	2000-2001	ug41	0-59	5268	80.6	13.4	-1.58	.38		.01	4.2	
Uganda	2006	ug52	0-59	2420	81.6	13.7	-1.53	.38	New	0	4.1	
Uganda	2011	ug60	0-59	2107	82.1	13.9	-1.38	.32	New	.01	5.2	
Uganda	2016	ug7b	0-59	4455	83.4	14.1	-1.21	.28	New	.01	5.7	
Uzbekistan	1996	uz31	0-35	1092	75.4	11.6	-1.1	.31		.89	10.6	
Yemen, Rep.	1991-1992	ye21	0-59	2959	80	13.7	-1.5	.38		.02	1.3	EM
Yemen, Rep.	2013	ye61	0-59	14285	80.7	13.4	-1.83	.46	New			
Zambia	1992	zm21	0-59	5083	79.2	13.6	-1.7	.41		.02	5	
Zambia	1996-1997	zm31	0-59	5678	79.6	13.3	-1.77	.44		0	5.2	
Zambia	2001-2002	zm42	0-59	5643	79.5	13.4	-1.9	.47		.02	5	
Zambia	2007	zm51	0-59	5385	81.1	13.9	-1.65	.44	New	.03	5.5	
Zambia	2013-2014	zm61	0-59	11787	82.7	14	-1.57	.4	New	.06	6	
Zimbabwe	1988-1989	zw01	3-59	2477	84.4	13.1	-1.4	.29		.18	5	
Zimbabwe	1994	zw31	0-35	2143	74.4	10.6	-1.02	.22		0	6.4	
Zimbabwe	1999	zw42	0-59	2841	83	14.7	-1.07	.26		.21	7.3	
Zimbabwe	2005-2006	zw52	0-59	4206	82	14.7	-1.37	.34	New	0	7.6	
Zimbabwe	2010-2011	zw62	0-59	4409	80.5	14.1	-1.34	.31	New	.01	8.7	
Zimbabwe	2015	zw72	0-59	5001	83.6	14.5	-1.19	.26	New	.01	9.2	

Source: Authors' calculations from all DHS surveys available at the time of writing that contain measurements of child height.
Notes: each row shows summary statistics for children in a given country/survey. 'Survey years' (col. 1) refer to the years when the field work of the survey was completed. Col. 2 identifies the version of the 'child recode' used in the analysis (DHS data are sometimes updated, so updated versions may become available in the future). For instance, 'al50' indicates that data on child height from Albania in 2008-09 were extracted from file alkr50dt.zip. All files have been downloaded, after obtaining permission, from the DHS web site. Col. 3 shows the range of child ages (in months) whose height was measured, while Col. 4 shows the number of non-missing heights. Means and standard deviations of height (in centimeters) are shows in columns 5 and 6, respectively. Col. 7 shows the average height-for-age z-scores ('HAZ'), while Col. 8 shows the prevalence of stunting, that is, the proportion of children with HAZ< -2. Column 9 shows whether HAZ was calculated using the new WHO growth charts (World Health Organization 2008). Typically, older surveys only include HAZ calculated using older reference charts. Columns 10 and 11 show the fraction of children whose mother completed at least secondary schooling, and their average number of years of schooling. The majority of surveys targeted all women 'of fertility age' as the primary respondent, but Col. 11 indicates whether only ever-married women ('EM') were surveyed. All means and standard deviations are calculated without using sampling weights.

Table A.2: YLS and CLHNS Summary Statistics

	Mother at least secon -dary school	Complete obs. in panel	Age	Mean	s.d.	Mother at least secon -dary school	Complete obs. in panel	Age	Mean	s.d.
Panel A - YLS										
	Younger Cohort					Older Cohort				
Ethiopia	0.073	1,741	1	71.0	5.4	0.056	744	8	117.6	7.5
			5	103.8	5.4			12	140.2	9.0
			8	120.8	6.9			15	154.4	9.9
			12	140.8	7.2			19	164.5	8.9
			15	155.8	7.8			22	165.6	8.5
India	0.044	1,852	1	71.7	5.1	0.029	893	8	118.0	6.3
			5	104.0	5.0			12	140.9	11.3
			8	118.7	6.4			15	152.9	8.5
			12	140.0	7.9			19	158.7	12.3
			15	154.7	8.1			22	159.8	9.6
Peru	0.346	1,759	1	71.4	4.7	0.303	544	8	118.9	5.9
			5	104.2	6.4			12	141.7	8.9
			8	120.1	6.0			15	154.5	7.6
			12	142.6	8.0			19	158.9	8.2
			15	156.7	7.6			22	159.4	8.3
Vietnam	0.290	1,843	1	72.2	4.313	0.283	810	8	118.5	5.685
			5	104.8	5.215			12	141.6	7.942
			8	121.1	6.248			15	154.9	7.284
			12	144.1	8.317			19	160.1	7.707
			15	158.4	7.882			22	160.9	7.693
Panel B - CLHNS										
Philippines (Cebu)	0.231	1,686	1	70.8	2.9					
			8	117.7	5.5					
			11	133.6	7.4					
			15	154.0	7.8					
			18	157.1	8.1					
			21	157.5	8.2					

Source: Authors' calculations from YLS and CLHNS.
Notes: Child age is approximate given that, within each survey round, there is variation in the date of birth and in the date when the interview and the measurement took place.

Table A.3: Height-for-age vs. maternal schooling, YLS and CLHNS

	Girls						**Boys**					
Panel A. Younger Cohort YLS												
	Age 1	Age 5	Age 8	Age 12	Age 15		Age 1	Age 5	Age 8	Age 12	Age 15	
Mother at least secondary	0.640***	0.724***	0.729***	0.701***	0.341***		0.667***	0.760***	0.708***	0.742***	0.597***	
	[0.0928]	[0.0831]	[0.0769]	[0.0899]	[0.0456]		[0.1133]	[0.0889]	[0.0941]	[0.0957]	[0.0624]	
Observations	3,416	3,430	3,427	3,430	3,430		3,720	3,757	3,759	3,757	3,759	
Mean HAZ	-1.172	-1.474	-1.177	-1.253	-1.167		-1.414	-1.525	-1.260	-1.243	-1.301	
Panel B. Older Cohort YLS												
			Age 8	Age 12	Age 15	Age 19			Age 8	Age 12	Age 15	Age 19
Mother at least secondary			0.448***	0.460***	0.306***	0.202***			0.663***	0.684***	0.592***	0.434***
			[0.0870]	[0.0725]	[0.0590]	[0.0592]			[0.0771]	[0.0864]	[0.0621]	[0.0530]
Observations			1,485	1,483	1,486	1,477			1,490	1,490	1,493	1,488
Mean HAZ			-1.667	-1.372	-1.531	-1.284			-1.595	-1.426	-1.573	-1.402
Panel C. CLHNS												
	Age 1		Age 8	Age 11	Age 15	Age 18	Age 1		Age 8	Age 11	Age 15	Age 18
Mother at least secondary	0.434***		0.534***	0.549***	0.291***	0.268***	0.680***		0.586***	0.598***	0.484***	0.363***
	(0.0721)		(0.0679)	(0.0834)	(0.0456)	(0.0495)	(0.0764)		(0.0932)	(0.0954)	(0.0764)	(0.0742)
Observations	675		677	675	677	591	745		746	746	748	670
Mean HAZ	-1.429		-1.962	-1.900	-1.816	-1.856	-1.660		-2.079	-1.998	-1.897	-1.918

Source: Authors' calculations from YLS and CLHNS.

Notes: Height-for-age z-scores are calculated using WHO-recommended references. That is, we use 2006 WHO growth charts for children up to age 5 (World Health Organization 2008), while for older children (5-19) we use charts from the 1977 US National Center for Health Statistics adapted to ensure smooth transition around age 5, as described in de Onis et al. (2007). In these results we not include measurements taken at age 20 or above given that references are only available up to age 19.