FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

Accounting for Intergenerational Educational Immobility in Low- and Middle-Income Countries

Fabian Koenings | Jakob Schwab

JENA ECONOMICS RESEARCH PAPERS · # 2025 – 009

The JENA ECONOMICS RESEARCH PAPERS
is a publication of the Friedrich Schiller University Jena, Germany (www.jenecon.de).

ACCOUNTING FOR INTERGENERATIONAL EDUCATIONAL IMMOBILITY IN LOW- AND MIDDLE-INCOME COUNTRIES*†

Fabian Koenings‡ *Jakob Schwab*§

Accepted for publication at the Journal of Human Resources

Abstract

This study investigates intergenerational educational immobility and its transmission channels in four developing countries: Ethiopia, India, Peru, and Vietnam. From data elicited throughout children's childhood, we extract latent factors of children's attributes and their environments. We decompose educational immobility by analyzing the extent to which these factors mediate intergenerational persistence. The findings show that relevant channels in developed countries are also important in these developing countries. Additionally, developing-country specific factors, such as starting a family while underage and performing child labor, play a role. The factors' importance differs moderately between the countries. Other factors – most notably non-cognitive skills – play no role.

***JEL* Codes**: I24, J62, O15.
Keywords: Intergenerational social mobility, transmission channels, low- and middle-income countries, decomposition, mediation analysis, factor analysis

---

*The external data used in this article is and can be obtained from the Young Lives Project at the Oxford Department of International Development via the UK Data Service (https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000060). Additional replication materials are provided in the Online Appendix.

†There are no conflicts of interest and nothing to disclose by either author.

‡Fabian Koenings is an associate researcher at the Friedrich Schiller University Jena.

§Jakob Schwab is an associate researcher at the German Institute of Development and Sustainability (IDOS), Email: jkbschwab@googlemail.com (corresponding author). We are indebted to Silke Uebelmesser, Eva Berger, and Florian Hett, for their valuable suggestions and guidance. We thank the involved research assistants for their essential support. In addition, we are grateful for the very helpful comments of three anonymous referees and for discussions with the participants at the Paris School of Economics Applied Economics Lunch Seminar, the Development Economics Seminar in Göttingen, the Lüneburg Microeconomic Workshop 2018, the Nordic Conference on Development Economics Helsinki 2018, the Annual Conference of the DSA 2018 in Manchester, the Annual Meeting of the VfS 2018 in Freiburg, the INFER Annual Conference 2018 in Göttingen, and the PEGNet Conference 2018 in Cotonou, which greatly improved the paper.

1. Introduction

The study of intergenerational social mobility has recently gained increasing attention (Chetty and Hendren 2018a; Güell et al. 2018). It is well established that parental socioeconomic status substantially determines children's socioeconomic outcomes. The mechanisms that underlie these associations are also understood more and more.[1] However, there is only little evidence on the channels of immobility for developing countries. This is problematic: Just as countries with higher economic inequality show greater levels of social immobility (Corak 2013; Becker et al. 2018; Leone 2018), poorer countries tend to have higher levels of immobility (van der Weide et al. 2024), which implies higher inequality of opportunity. Because the degree of immobility in poorer countries is different from that in richer countries, the channels of immobility in developing countries are also likely to differ. Theoretically, this may be because at lower development levels, also parenting styles are likely to differ (Doepke and Zilibotti 2017). Moreover, the institutional setting can influence how the parental background presents itself in the ways children grow up and which characteristics are important for their later success.

In this study we explore the factors of children's development paths which are influenced by parental education levels and determine children's schooling outcomes. These *pathway factors* of intergenerational educational immobility are examined in four developing countries: Ethiopia, India, Peru, and Vietnam. We use longitudinal data from surveys of children and their families. From the rich information about the children's development and environment, we extract latent pathway factors from thematic groups of childhood characteristics by means of factor analysis. We then use mediation analysis to decompose the correlation between the children's and their parents' education levels into the absolute and relative contributions of the different pathway factors in each of the countries.

Of the four countries under study, our results indicate that educational immobility is largest in Vietnam, where one additional year of parental schooling increases children's schooling by 0.41 years, followed by Ethiopia (0.34) and Peru (0.3), while India displays the highest level of educational mobility (0.25). The findings on the channels of this immobility show that children's individual characteristics play an important role in the transmission of educational status, as has been shown for developed countries: cognitive skills are an important pathway factor in all four countries, accounting for between 11 % (Vietnam) and 29 % (India) of immobility. Childhood aspirations are a very important factor in all countries but Ethiopia. They are even the most important factor in Vietnam, where they account for 23 % of immobility. Uniformly, the non-cognitive skills surveyed do not play a role in any of the developing countries studied, which is different from findings in developed countries (Blanden et al. 2007).

[1] In developed countries, it has been shown that both the individual characteristics of children (Blanden et al. 2007) and their local environment (Chetty and Hendren 2018a,b; Chetty et al. 2016, 2014; Acciari et al. 2022) contribute to the fact that children of disadvantaged households fare worse in their adult lives than children from more privileged backgrounds.

Additionally, the findings reveal particular patterns for each of the four countries: In Ethiopia, differences in the local school infrastructure as well as poorer children's higher likelihood of starting a family while underage and having to perform child labor contribute to the transmission of the socioeconomic status across generations, each accounting for 6-7 % of educational immobility. In India, the channels analyzed can explain much of the immobility, where besides children's cognitive skills, aspirations, underage family engagement, and child labor, also the monetary spending of parents for their children's education manifests differences in educational levels between families. Peru is the only country in our study in which poorer children are not more likely to start a family underage, so that differences in cognitive ability, aspirations, and the likelihood for child labor constitute the relevant pathways. In Vietnam, child labor is only a minor pathway, but besides cognitive ability and aspirations, starting a family underage is a strong factor in children from low socioeconomic backgrounds to also achieve lower educational outcomes. In all countries where this is a relevant factor (Ethiopia, India, Vietnam), it is so only for girls but not for boys.

This study contributes to the literature in several ways. First and foremost, it is the first to identify the pathway factors of intergenerational educational immobility in developing countries. Several studies have analyzed which pathways are responsible for the transmission of socioeconomic status across generations in developed countries. Bowles and Gintis (2002) are among the first to attempt to empirically identify the underlying pathways of intergenerational immobility by means of mediation analysis. Using survey data from the United States (US), they find evidence for the importance of wealth, race, and schooling when investigating the intergenerational transmission of income and earnings. Adopting the same approach, Blanden et al. (2007) evaluate the role of cognitive and non-cognitive skills, education, and labor market experience in the transmission of the parents' socioeconomic status, measured by family income, to their sons' earnings in the United Kingdom (UK). The authors show that cognitive and non-cognitive skills are important factors, but much of their effect operates through education and labor market attachment. All of these factors together account for half of the association between parents' and sons' socioeconomic statuses. Similarly, Blanden et al. (2014), using data on fathers and sons from both the UK and the US, analyze the mediating role of early marriage, health status, labor market attachment, occupations, and education. The results confirm that a large part of the association is transmitted through education.[2] Considering the same age cohort, Schad (2015) replicates the analysis undertaken by Blanden et al. (2014) for Germany and obtains similar results. In our study, we are able to include a much broader set of potential pathway factors. Some of the characteristics are specific to the low- and middle-income country context. Our analysis concentrates on children's education as the relevant socioeconomic outcome variable and not as a pathway factor, because children's educational success is closely related to their later socioeconomic status and suffers less from life-cycle bias than other indicators for socioeconomic status.

[2]This is supported by Björklund et al. (2017), who find that the differences in the degree of immobility between the UK and the US can be explained by the fact that skills, particularly schooling performance, are more strongly linked to the children's socioeconomic outcome in the UK, not because the link of schooling performance to the parental income is more pronounced.

Our paper also adds to the literature on the degree of educational immobility in developing countries. First, it provides estimates of educational mobility across a more recent period of time than any other study that we know of in the countries studied. Whereas the most recent data typically includes the generations born before 1985, the cohort in our design is born in 1994/1995. The study design furthermore allows the levels of relative immobility to be compared across developing regions using recent data from a single survey. van der Weide et al. (2024) in a new database (World Bank 2023) and Hertz et al. (2007) present estimates of relative intergenerational educational immobility across many more countries, but at the cost of using different sources of data and not including control variables on the individual level. The most recent estimates are those on the cohort born between 1980 and 1989 in World Bank (2023), and in Emran and Shilpi (2015) for India. Bossuroy and Cogneau (2013), Balán et al. (1973), Emran and Shilpi (2011), Haile (2018), Hnatkovska et al. (2013), Lambert et al. (2014), and Mohammed (2019) estimate levels of intergenerational social immobility for single developing countries / regions, using different measures of socioeconomic status. More recently, works by Alesina et al. (2021) and Munoz (2021) analyze upward mobility in a number of African and Latin American countries, respectively (and discuss the role of the local environment). Although the time trends identified constitute an interesting comparison to our findings, the measures are not directly comparable to the measurement of social mobility that we use. We present estimates of relative immobility because this allows to distinguish the persistence of inequality across generations from common factors like economic growth. Asher et al. (2024) reports differently measured estimates of relative immobility in India, also deriving the latest estimation from the 1980s birth cohort.

By using novel and comprehensive data, we are able to mitigate many problems often associated with estimating relative intergenerational mobility in developing countries: We use information on education levels – elicited directly from the parents and children –, which reduces concerns of life-cycle fluctuations of income levels and avoids issues with taking data in retrospective. Moreover, the sample is not plagued by co-residency bias which otherwise tends to produce a downward bias in the estimations of intergenerational mobility in developing countries (for an overview, see Emran et al. 2018, and Emran and Shilpi 2019). In the data we use, children are tracked also after they have moved out of the household.

Finally, we also advance the literature methodologically. We extend the mediation analysis approach by Bowles and Gintis (2002) and Blanden et al. (2014). Pathway factors are extracted from a battery of data on childhood characteristics by means of factor analysis. This allows us to use the abundant data without encountering issues of multicollinearity and measurement error, which are common concerns when analyzing survey data. We additionally correct for typical forms of measurement error in the measures of non-cognitive skills in developing countries in particular (Laajaj and Macours 2021). The mediation analysis then picks up on recent discussions and advancements in the literature on the methodology. Notably, we also allow for mediating relationships between the pathway factors themselves, following the methodology in Bolt et al. (2021).

The remainder of this paper is structured as follows: The next section presents some country background information, the dataset used and the methodology to extract latent factors of the

children's development from the survey data. In Section 3, we present the estimation strategy. Section 4 shows the results and discusses them, including some extensions and robustness checks. Section 5 concludes.

2. Data

For our analysis, we use external data retrieved from the Young Lives (YL) project, a longitudinal and multi-dimensional survey investigating causes and consequences of childhood poverty. The data comprises five rounds across four countries: Ethiopia, India (in the states of Andhra Pradesh and Telangana), Peru, and Vietnam.[3] These countries were selected for the survey because they represent the four major regions of the developing world, both low- and middle-income countries, and diverse socioeconomic and political systems (Young Lives 2017). Table 1 shows information on key economic variables and the youth in each of the four countries at the time of the surveys. Peru is (and has been) the richest of the four countries, and despite having the highest inequality levels in terms of the Gini coefficient, also displays the lowest poverty rate. Primary school enrollment in all countries except for Ethiopia is almost universal, but only Peru has over 80 % of the relevant age group enrolled in secondary school, again with Ethiopia falling short. Enrollment rates in tertiary education differ to some extent between the four countries (10.7 % to 32.8 %), but are much lower than in developed countries (US: 81 % in 2005). Returns to education have been found to be substantial in all four countries, with around 10 % increase in income with each year of schooling. However, all four countries have, or previously had, substantial numbers of children working in employed labor, although India has reduced this number substantially around the turn of the century. But these official numbers do not include children working in their homes or family businesses and, thus, neglect an important form of child labor.[4] Child marriage is not legal in any of the four countries studied. Nonetheless, like in many developing countries, it remains widespread, particularly for girls (Duflo 2012). Child marriage rates are particularly high in Ethiopia and India, but also substantial in Peru and Vietnam. Childhood stunting is also prevalent in all four countries but, again, especially in Ethiopia and India. Crime rates, here proxied by the number of intentional homicides because of data availability, are highest in Ethiopia and Peru. These numbers provide some first indication what may hold back children with poorer backgrounds from being successful in school in each of the four countries.

The YL survey has several advantages for the study of intergenerational mobility. First, it is the most comprehensive data collection available on children in developing countries and the

[3]The data is based on the following datasets for the different rounds: Jones and Huttly (2018), Boyden (2018a,b), Woldehanna et al. (2018), and Sanchez et al. (2018). In addition, we use the dataset from Boyden (2018c), which collects selected items across all five rounds.

[4]The UN Convention on the Rights of the Child includes all work that is "likely [...] to interfere with the child's education" (United Nations 1989, Article 32). Evidence for India shows that with better control and regulation of child labor, domestic work is becoming the more prevalent form (see UNICEF 2018a).

Table 1. *Country Information*

	Ethiopia	India	Peru	Vietnam
GDP per capita in 2005, in US$	162.4	714.9	2729.5	687.5
GDP per capita in 1970, in US$[(1)]	203.6	112.4	552.2	231.5
Gini Coefficient in 2004	29.8	34.4	49.9	36.8
Poverty rate in 2004 (at 1.90 US$ PPP per day)	38.5	39.9	13.6	25.8
Fertility Rate (Births per Woman) in 2005	5.8	3.0	2.7	1.9
Primary school enrolment rate in %, 2005 gross (net)	79.0 (60.4)	110.0 (91.1)	113.7 (96.6)	97.7 (90.7)
Secondary school enrolment rate in %, 2005 gross	25.1	54.0	82.2	58.3
Tertiary school enrolment rate in %, 2005 gross	28.1	10.7	32.8	16.2
Returns to education (in % per year of schooling)[(2)]	12.5	8.5	10.7	8.8
Child labor rate in 2015[(3)]	45	4	15	13
Child Marriage Rate in 2015[(4)]	59.0	47.0	17.3	10.4
Childhood Stunting in 2015[(5)]	50.0	47.8	29.2	33.2
Intentional Homicide, per 100,000 people, in 2005[(6)]	9.4	3.6	11.1	1.25

Notes: This table shows macroeconomic, social, and education-relevant variables for the four countries included in the study, for the times most relevant for the study, or the closest possible if not available in other years. The source is World Bank (2021), if not indicated otherwise: [(1)]Data for Vietnam is for South Vietnam. [(2)]Source: Psacharopoulos and Patrinos (2018). [(3)]Sources: UNICEF (2021b,a). [(4)]Sources: UNICEF (2015a,b), World Bank (2016), and UNICEF (2017). [(5)]Source: Nutrition Accountability Framework (2021). [(6)]Source: UNODC (2017).

hardships they may face. Second, survey questions are nearly universally harmonized across the four countries, which makes it easy to compare the countries and avoid biases arising from the way questions are framed. Third, attrition is extremely low due to immense efforts to monitor children throughout their lives (Aurino and Burchi 2017, p. 304).[5] Lastly, the data captures the diversity of children in the countries and allows to compare poor and better-off children. However, those included in the survey to this end were selected by applying a "'pro-poor' and multi-stage sampling procedure" (Aurino and Burchi 2017, p. 294) and thus selection was not fully random. First, 20 sites were non-randomly predetermined for each country with the purpose of including poorer areas in the survey. Children of a certain age were randomly selected within these pre-selected sites.

To assess the national representativeness of the YL sample used, we follow Singh (2020) and compare the means of selected key statistics for the sample and corresponding variables from the respective national-representative Demographic and Health Surveys (DHS) in Table A1 of Online Appendix A. Consistent with the results by Singh (2020), this comparison shows that the YL data for India, Peru, and Vietnam are nationally representative on most dimensions observed.[6] However, the Ethiopian YL sample is better-off than the national representative

[5]From the first to the fifth observation round there is an attrition of 186 observations in Ethiopia, 86 in India, 106 in Peru, and 90 in Vietnam. This corresponds to an attrition rate of 13 % across all countries over the 15 years of the study.

[6]A small exception is the parents in India who have fewer years of education than the nationally representative parents. However, this somewhat worse-off family background is not reflected by other socioeconomic status

households in terms of socioeconomic indicators, such as parental education and housing amenities. Singh (2020) shows that this is due to a more difficult surveying in rural areas in Ethiopia and a resulting oversampling of urban households in the YL data, such that the proportion of urban to rural households is three times larger than in the DHS data. Thus, in the present analysis with the YL data, the estimation results are nationally relatively representative for India, Peru, and Vietnam along the observed dimensions.[7] The results for Ethiopia might be more representative for better-off households in Ethiopia and, hence, for households from urban areas with potentially better access to and quality of education. Although the YL data has limits for making general statements about the four countries under study due to the sampling strategy chosen by the researchers, the comparison shows that the benefits of using this data for the analysis outweigh the disadvantages. These benefits include the amount of elicited characteristics of children and their environment, the follow-up over long time horizons even after they leave the households, and its currentness.

The first survey round was conducted in 2002, when the children were seven or eight years old. From then on, interviews were repeated with the same children (and their families) approximately every three years – in 2006, 2009, 2013, and 2016 – thus, the children were 21 or 22 years old when they were last interviewed. The survey asked a range of questions about the socioeconomic background of the families and their offspring's situation each time. Approximately 1,000 children were surveyed per country, with five data points – one in each survey round – for each child.[8] However, due to different response rates between the countries for some of the survey items that we use in our analysis, the number of observations slightly differs between the countries, with 921 observations from India, 906 from Vietnam, 752 from Ethiopia, and 582 from Peru.[9] The deleted observations do not differ significantly from the rest with regard to the observed characteristics, as can be seen in Table A2 in Online Appendix A. This table provides mean statistics for some key variables for all available observations and those included in the final sample for all countries taken together.[10]

---

variables such as the assets and amenities the households have. Furthermore, the YL data in Peru over-represents urban housing which is mirrored by the fact that more homes have drinking water that is piped into their dwelling. However, the socioeconomic status in terms of parental education is very much in line with the nationally representative DHS data for Peru.

[7]It has, however, to be noted that the Indian sample is limited to the states of Andhra Pradesh and Telangana, such that local peculiarities not related to the observable household dimensions may feed into the analysis.

[8]The original sample for Peru includes only 714 children because fewer children from provincial sites were recruited.

[9]We discuss the sample restrictions after introducing the relevant variables used in the analysis and their measurement.

[10]The educational achievements of the children that were not included in the sample are slightly lower than of those in the sample. This, however, only refers to a rather small number of children for whom information on the education level was available but who were dropped from the estimation because other information was

Education

We operationalize children's educational outcome by the years of schooling achieved by the child in the last observation round at age 21–22 years.[11] As the survey only provides information on children's first 12 years of schooling by years, we deduce any additional years of schooling from information about the highest educational degree obtained. We use the concordance by van der Weide et al. (2024, p. 5) to transpose levels used by the International Standard Classification of Education (ISCED, United Nations 2016) into the corresponding years of schooling.[12] The resulting number of years of schooling – and, hence, the main dependent variable – ranges from zero to 18. The average years of children's schooling in the sample are 9.61 (Ethiopia), 11.92 (India), 12.97 (Peru), and 11.63 (Vietnam).

To assess a measure of parental education, we follow the same procedure as for the children's education. We use the maximum number of either the father's or mother's years of schooling (or the household head's years of schooling if this is not the mother or father) from the first observation period as the main explanatory variable. The resulting average parental years of schooling are 3.34 (Ethiopia), 3.83 (India), 9.1 (Peru), and 7.0 (Vietnam). These numbers are considerably lower than those of the children, which implies substantial absolute upward mobility in terms of education across all four countries, but particularly in Ethiopia and India.

---

missing.

[11] The dataset does not allow to use other common socioeconomic status indicators, such as children's wages, because the surveyed children were on average only 22 years old in the last observation period and in many cases had not yet begun formal employment. Additionally, literature has shown that wages and earnings in young adulthood are often not representative of later socioeconomic status (see, for example, Chetty et al. 2014). We instead employ information about the children's educational achievements because it is the best observable measure of their socioeconomic status. Earlier studies on the intergenerational transmission of income show that educational outcome is a good predictor of later economic success (see Lambert et al. 2014 and Blanden et al. 2014), and the close association between education and lifetime earnings is well documented (starting with Mincer 1974). Wantchekon et al. (2015) demonstrate the social and economic benefits of education in a developing country setting.

[12] Some children were still enrolled in the education system in the latest survey round at this still young age, mostly in tertiary education. In order to still exploit the information, we assign these children the completion of this education level as their educational degree. Ignoring this variation and using only yet completed schooling as outcome variable would have small effects on the estimation of the degree of immobility (lowering its estimates by construction), but not on that of the corresponding pathways. Even using the information on current enrollment, however, it cannot be ruled out that children who left the education system will return to education at later points in their life. Thus, issues with life-cycle fluctuations may even be present with education as the relevant outcome.

Pathways

As for the information on the circumstances of the children's youth, many survey items aim to elicit similar information. For example, the children's nutritional and/or health status is captured by measuring their weight, height, and malnutrition. If all measures were included in the analysis at the same time, it would raise issues of multicollinearity and lead to underestimating their joint effect. Furthermore, survey data is prone to measurement error. To address this issue, we make use of the richness of the information in the data by extracting latent factors from grouped survey measures by means of factor analysis. This reduces the dimensionality of the data and at the same time addresses measurement error, which we will explain below in this section. The analyzed factors are chosen based on their potential as transmission pathways, as discussed in the literature. They can be specifically relevant for developing countries or shown to be influential in general. In the YL survey, we identify measures relating to ten such underlying factors. These factors need to be potentially correlated with parental background and, at the same time, also likely to affect their offspring's educational perspectives. These ten latent factors are[13]:

(i) *Child labor*: It is likely that poorer children in developing countries have to work, either inside or outside the home. It is also likely that this will negatively impact their chances of succeeding in school (Keane et al. 2022; Woldehanna and Gebremedhin 2015; Putnick and Bornstein 2015; Emerson and Souza 2003). Measures for this factor are comprised of hours of paid and unpaid work and whether injuries occurred at work.

(ii) *Infrastructure*: The public (or private) infrastructure where a child lives may critically determine their likelihood of attending and being successful in school (Vuri 2008; Kazeem et al. 2010). Poorer parents are more likely to live in rural areas or where it is more difficult to get to school, which are also the measures for this factor.

(iii) *Education spending*: Poorer parents have fewer financial resources to invest in their offspring's education (Kornrich and Furstenberg 2013), which affects educational outcomes (Singh 2015). Related measures are the monetary expenditures for school fees and learning materials as well as whether children attend a private school.

(iv) *Underage family engagement*: Children from poorer backgrounds are more inclined to start a family (either forced or unforced) while still at school age (Wodon et al. 2017), which in turn decreases their chances of finishing school when they otherwise would have (Nguyen and Wodon 2014).[14] The factor is measured by whether children actually get married or become parents while still a minor. Also, the previous expectations of the parents for their child to marry or become a parent when underage are considered.

(v) *Parental attentiveness*: The time and energy that parents devote to fostering and educating

[13]The factors can be broadly distinguished as "opportunity" pathways (i-iii), "social" pathways (iv-vi), and "individual" pathways (vii-x), although this classification is not important for the further analysis.

[14]Whereas cross-country differences in tendencies to start a family while underage may be driven by cross-cultural differences, within countries, poverty is a central determinant.

each of their children is an important determinant of their outcomes and is linked to the parents' socioeconomic status and decisions regarding the family structure (Gould et al. 2020; Chetty et al. 2014; Darroch and Singh 2013; Goodman et al. 2012; Black et al. 2005). Measures of this factor include the number of children living in the household and self-assessments of the children about their personal exchange with their parents.

(vi) *Social environment*: The area's social structure and the children's peer groups can determine how well they fare in school. They are influenced by parents' potentially resource- or status-related location decisions (Sacerdote 2014, 2011). The factor is measured by assessments of both the parents and the children about the children's social environment and the area they live in.

(vii) *Health*: Children's health status is a major focus of the YL study. Woodhead et al. (2014, p.13), using data from the younger cohort of the dataset, find that "in Peru over 50 % of the children from households in the poorest quintile were stunted in 2006, compared to just under 10 % in the wealthiest quintile." Childhood health is also found to affect children's educational attainment in developed countries (Case et al. 2005). Standard measures of weight, height, and undernourishment at young age elicit this factor.

(viii) *Aspirations*: What children want to achieve in life (or at school) is an important determinant of actual outcomes (Figlio et al. 2019; Genicot and Ray 2017). The reason poorer children's aspirations are lower than those of richer children may be due to differences in preferences (that is, mostly, imitation: see also Agupusi 2019), missing information (Hoxby and Avery 2013; Hoxby and Turner 2015; Jensen 2010), or awareness of actual group constraints (Dalton et al. 2016).[15] The factor of aspirations is measured by the expressed professional goals of the children and the level of schooling they wish to complete as well as the level their parents wish for them to complete, all elicited at age 11 or 12.

(ix) *Cognitive ability*: In developed countries, cognitive ability, whether transmitted by nature or nurture[16], has been found to be one of the central pathway factors of the intergenerational transmission of socioeconomic status (see, for example, Blanden et al. 2007). Cognitive ability is measured by math and literacy tests, which were conducted together with the YL survey, as well as a Peabody Picture Vocabulary Test (PPVT), which aims at measuring crystallized intelligence. Although these measures are commonly used in the literature as measures of cognitive ability[17], there is a risk that they capture schooling success up to that age and, thus, part of the outcome variable, that is, the years of schooling completed as a young adult.

[15]It is also argued that too high aspirations can dampen the efforts of children coming from low socioeconomic backgrounds, because the "aspiration gap" may be too large (Ray 2016). However, since our measures of aspirations rather target the lower end of the ambition distribution, we do not explicitly account for this possibility.

[16]See Black et al. (2020) for a recent discussion of the distinction between the roles of nature and nurture in the process of intergenerational transmission of the socioeconomic status in general and of human capital in particular.

[17]For example, Blanden et al. (2007) use math and literacy tests at age 10 as well as a predecessor of the PPVT

Therefore, as alternative factor variables for cognitive ability in two robustness tests, we use only the PPVT and only the Raven's Progressive Matrices (RPM) test. The latter aims at eliciting fluid intelligence.

(x) *Non-cognitive ability*: Along with cognitive skills, non-cognitive skills have been increasingly spotlighted as determinants of lifetime (and educational) outcomes in developed countries (see Kautz et al. 2014 for an overview). As with cognitive ability, children's non-cognitive skills are strongly associated with the socioeconomic status of their parents (Kosse et al. 2020).[18] In the YL data, there are several groupings of measures that aim to identify different non-cognitive skills. The literature is clear, however, that there is no common latent factor for non-cognitive skills. Each skill that falls under this definition stands on its own and is not assumed to be strongly correlated with the others (Borghans et al. 2008). In order to not include too many non-cognitive factors in the presentation, we report only the estimations including grit as a factor of non-cognitive skills in the main text. Grit is defined as perseverance and the passion for long-term goals and has been shown to be key to educational and lifetime outcomes in developed countries (Duckworth et al. 2007).[19] However, including other factors of non-cognitive ability in our analysis, such as agency, pride, trust, and inclusion, either serially or all at the same time, yields the same results as grit. Thus, the results for grit are representative of all factors of non-cognitive ability elicited in the surveys.

Factor Analysis

The relationship between the measures and the factors is laid out in the following. We assume that each group of measures is associated with only one factor. The set of measures for each factor i is denoted by $\mathcal{M}^i$ and the set of factors by $\mathcal{F}$. In order to more easily interpret the relationship between measures and the respective factors, all measures have been standardized to have a mean of zero and a standard deviation of one. The relationship between the measures $m_j^{i,h}$, $j \in \mathcal{M}^i$, associated with factor $i \in \mathcal{F}$, is given by:

$$m_j^{i,h} = \psi_j^i PW_{i,h} + \eta_j^{i,h} \tag{1}$$

where $PW_{i,h}$ is the factor score of factor i for child h. $\eta_j^{i,h}$ is an error term with a mean of zero. Since the measures are standardized and the factors are scale free, the measurement system does not include a non-zero intercept. ψ_j^i is the factor loading for measure j of factor

as measures of cognitive ability. Bolt et al. (2021) use math and reading tests as well as teacher-assessed math and language ability at age 16.

[18] Black et al. (2017) analyze the roles of nature versus nurture for non-cognitive skills, and find a genetic component of the transmission of non-cognitive ability.

[19] The measures of grit are the only measures of a pathway factor that we use from the fifth survey round, because they were only elicited in the last round. We must therefore assume at least some stability of this factor over time. All other non-cognitive factor measures were obtained in earlier survey rounds.

i. Estimating the measurement system yields both the factor loadings ψ_j^i, $j \in \mathcal{M}^i$, $i \in \mathcal{F}$, and the factor scores for the latent factors $PW_{i,h}$, $i \in \mathcal{F}$. Intuitively, the estimation of the measurement system with only one factor per set of measures extracts as the factor scores PW_i of the principal factor i the common variation of its measures m_j^i, $j \in \mathcal{M}^i$ which can explain most of these measures' common variation. For detailed discussions of this approach, see Heckman et al. (2013), Gorsuch (1983), and Kim and Mueller (1978). We restrict the included measures for all factors to those that are generally available for almost all children in the sample in all four countries. In our resulting sample, the Kaiser-Meyer-Olkin (KMO) statistic (Kaiser 1974) is well above 0.7 and the Bartlett-test of sphericity (Bartlett 1951) is significant at the 1 %-level, which confirms the adequacy of the sample for a factor analysis. Table 2 lists all included measures for the factors, their respective factor loadings, the children's age at which they were elicited, and the mean values of each measure for each country. Tables A3 and A4 in Online Appendix A detail the respective survey items for each measure.

Table 2. *Factor Measures*

Factor	Measure	Data Source	Observ. Round (Child's Age)	Loading	Mean Ethiopia	India	Peru	Vietnam	All
Child Labor	Working Hours	Child quest.	round 2 (11/12)	.431	4.53	1.93	4.71	2.23	3.15
	Activity for Money	Child quest.	round 2 (11/12)	.469	0.07	0.21	0.27	0.06	0.14
	Work Injury	Child quest.	round 2 (11/12)	.263	0.07	0.06	0.08	0.05	0.06
Infrastructure	Access to Education	Household quest.	round 2 (11/12)	.089	0.99	0.99	1.00	0.97	0.99
	Type of Living Site (Urban)	Fieldworker	round 2 (11/12)	.329	0.34	0.24	0.77	0.19	0.35
	Time to School *	Child quest.	round 2 (11/12)	.313	25.26	17.15	13.97	16.16	18.16
Education Spending	Spent on Schooling Fees	Household quest.	round 2 (11/12)	.712	2943.77	21029.71	9507.22	6791.46	10508.30
	Spent on School Books	Household quest.	round 2 (11/12)	.558	3239.45	7763.12	12707.58	6855.20	7405.30
	Private Schooling	Child quest.	round 2 (11/12)	.533	0.04	0.24	0.09	0.00	0.09
Underage Family	Child Marriage	Child quest.	round 4 (18/19)	.568	0.06	0.17	0.01	0.08	0.09
	Child Parent	Child quest.	round 4 (18/19)	.582	0.05	0.10	0.13	0.04	0.08
	Expected Child Marriage	Household quest.	round 2 (11/12)	.438	0.01	0.02	0.01	0.00	0.01
	Expected Child Parent	Household quest.	round 2 (11/12)	.361	0.00	0.01	0.00	0.00	0.00
Parental Attentiveness	Without Father *	Household quest.	round 1 (7/8)	.105	0.27	0.03	0.20	0.04	0.12
	Number of Children *	Child quest.	round 2 (11/12)	.223	2.68	1.5	1.90	1.35	1.83
	Attention received	Child quest.	round 2 (11/12)	.451	3.25	3.59	3.57	3.43	3.46
	Love received	Child quest.	round 2 (11/12)	.477	3.75	3.86	3.85	3.79	3.81
	Conversation	Child quest.	round 2 (11/12)	.387	2.26	2.90	2.37	2.66	2.58
	Supported by Parents	Child quest.	round 2 (11/12)	.430	3.07	3.39	3.14	3.25	3.23
	Free Speech	Child quest.	round 2 (11/12)	.407	3.33	3.68	3.41	3.35	3.45
Social Environment	Safe Area	Child quest.	round 1 (7/8)	.177	0.77	0.91	0.69	0.88	0.82
	Friends with Alcohol *	Child quest.	round 3 (14/15)	.499	1.59	1.19	1.86	1.57	1.51
	Friends with Trouble *	Child quest.	round 3 (14/15)	.285	1.32	1.42	1.79	1.41	1.48
	Friends who Smoke *	Child quest.	round 3 (14/15)	.496	1.02	1.17	1.64	1.44	1.30
	Shock-theft *	Household quest.	round 2 (11/12)	.133	0.01	0.02	0.06	0.03	0.03
	Trust in Neighborhood	Household quest.	round 2 (11/12)	.220	3.37	3.57	3.06	3.55	3.42
	Safe to go Outside	Household quest.	round 2 (11/12)	.357	3.15	3.73	1.93	2.65	2.95
Health	Weight-for-age	Fieldworker	round 1 (7/8)	.820	-2.05	-1.95	-0.50	-1.73	-1.64
	Height-for-age	Fieldworker	round 1 (7/8)	.645	-1.55	-1.56	-1.40	-1.47	-1.50
	Thinness Indicator *	Fieldworker	round 1 (7/8)	.375	1.29	1.32	1.01	1.21	1.22
Aspirations	Job Aspirations Child	Child quest.	round 2 (11/12)	.369	2.73	2.56	2.49	2.57	2.59
	School Aspirations Child	Child quest.	round 2 (11/12)	.578	4.98	4.88	5.73	5.21	5.16
	School Aspirations Parents	Household quest.	round 2 (11/12)	.571	5.14	4.26	5.81	5.16	5.02
Cognitive Ability	PPVT Test Score	Test	round 2 (11/12)	.440	-0.05	0.02	0.02	-0.01	-0.01
	Math Score	Test	round 2 (11/12)	.577	-0.10	0.05	0.08	0.02	0.01
	Reading Level	Test	round 2 (11/12)	.692	3.23	3.67	3.93	3.95	3.69
	Writing Level	Test	round 2 (11/12)	.876	2.43	2.65	2.85	2.93	2.72
	Literacy	Test	round 2 (11/12)	.840	0.44	0.66	0.85	0.93	0.72
Non-Cognitive Ability: Grit	GRIT 1 *	Child quest.	round 5 (21/22)	.335	2.62	2.94	3,03	2.69	2.81
	GRIT 2	Child quest.	round 5 (21/22)	.193	3.35	3.28	3.23	3.38	3.31
	GRIT 3 *	Child quest.	round 5 (21/22)	.493	2.56	2.75	2.68	2.76	2.70
	GRIT 4	Child quest.	round 5 (21/22)	.598	4.10	4.20	3.86	3.83	4.01
	GRIT 5 *	Child quest.	round 5 (21/22)	.502	2.76	2.96	2.96	2.59	3.81
	GRIT 6 *	Child quest.	round 5 (21/22)	.495	2.70	2.94	2.67	2.62	2.74
	GRIT 7	Child quest.	round 5 (21/22)	.652	4.02	4.26	3.86	3.88	4.02
	GRIT 8	Child quest.	round 5 (21/22)	.664	4.03	4.32	3.84	3.77	4.00

Notes: This table shows the measures (column 2) from the survey data by the factors (column 1) that they are associated to as presented in Section 2. For a more detailed description of the measures and the associated scales see the Tables A3 and A4 in Online Appendix A. Column (3) displays the data source and column (4) lists the observation round and respective age of the children when the information was elicited. The respective factor loadings of each measure are depicted in Column (5). The last five columns show the mean values for each measure by country and for all countries combined. * Measure enters factorization negatively. The PPVT test scores are de-meaned by the language and the math scores by the country.

In order to more straightforwardly interpret some of the estimation results, the factor variables themselves are also standardized after extracting them from the measures.[20] The ten factor scores $PW_{i,h}$ obtained for each child are considered to be potential pathway variables in the decomposition of the observed social immobility.[21]

Besides reducing dimensionality, employing factor analysis also deals with measurement error in the individual survey measures by extracting only the common variation from the measures (see Wansbeek and Meijer 2003). An alternative to factor analysis would be creating simple averages of the grouped measures. Besides ignoring the covariability between the measures and using arbitrary weighting, thereby giving weight to measures that are not correlated with other measures of an underlying factor of interest, this would correct for measurement error by simply averaging out error terms (Heckman et al. 2013). Factor analysis, in contrast, excludes the (uncorrelated) error terms when estimating the individual factor scores. Nonetheless, the quality of the factor scores depends on the accurateness of the underlying measures. The quality of the data is particularly questionable for survey measures of non-cognitive skills in developing countries, as Laajaj and Macours (2021) point out. As far as possible when using pre-existing data, we follow the recommendations of Laajaj and Macours (2021) in order to address potential shortcomings of the non-cognitive-skills data: First, in order to derive the factor scores, we create our own scale from the data by conducting the factor analysis described above instead of relying on existing scales obtained in different contexts. Second, in advance of the factor analysis, we correct the measures of non-cognitive skills for the interviewers' influence (by controlling for interviewer fixed effects)[22] and for acquiescence bias (by subtracting the acquiescence score). These procedures ensure that answering patterns that are generated solely by the survey design do not play a role for the measurement system. Online Appendix G elaborates on how we deal with non-cognitive-skills data.

We intuitively associate measures with underlying factors, in line with the structure of the surveys. This gives the researcher some degrees of freedom. A more purely data driven approach would be to conduct an exploratory factor analysis (EFA), which groups measures according to their actual relationship, identifying underlying factors by the data structure. The downside of such an approach is that it gives the same consideration to all measures without considering their conceptual importance, and hence tends to emphasize factors measured by many items. To affirm the groupings from intuitive approach, we also conduct an EFA,

[20] If a child or parent did not answer one of the 46 survey items used as measures, we assigned it the mean of the respective answer in order to use the information provided through one of the other measures of the respective factor. If information on all (three to eight) measures from one respective factor was missing, we dropped the observation from the sample.

[21] Summary Statistics for all variables in each country and all countries combined are provided in Tables A5 and A6 in Online Appendix A. Table A7 shows the number of observations for each variable and, accordingly, how each sample restriction drives the final sample size.

[22] The interviewer identification is not published with the YL data, but was made available for the purpose of this study.

which supports the general associations of measures (and the interpretation of factors as detailed above) presented in Table 2. The only difference is that selection by EFA (which also leaves many degrees of freedom to the researcher through the choice of factor retention criteria) tends to drop concepts from the analysis which have fewer measures. We report the results of the intuitive approach in the main text because it uses all the information given, with those measures with lower factor loadings contributing less but still informatively to the identification of the latent factors. Also, the intuitive grouping of measures allows to disentangle the role of related but independent concepts. This gives a more nuanced overview of the contributions of different factors, particularly those with fewer measures. The main results of the analysis remain the same, irrespective of the method to group the measures of factors. The results from employing an EFA are briefly discussed in Section 4 and presented in detail in Online Appendix I.

3. Empirical Strategy

We conduct the estimation of intergenerational immobility and its pathways in the cross-section, but we exploit the panel structure of the data as we regress later outcomes in life (last round) on initial starting positions (first round). Children's characteristics observed between those two points in time (first to fifth round) act as potential pathway characteristics. The decomposition is conducted through an estimation of how parental background determines the potential pathway factors and how these in turn correlate with the children's outcomes. The mediation analysis comprises the following four steps. (i) We estimate the extent to which the children's educational levels can be associated with their parents' education, that is, the degree of relative educational immobility, which is then decomposed. (ii) For the decomposition, we first estimate each pathway factor's correlation with the parents' education. (iii) Then the influence of all pathway factors on the children's educational outcomes is estimated by one joint regression. (iv) Finally, we combine steps (ii) and (iii) to elicit the overall contribution of each pathway factor to the immobility estimated in step (i). Figure B1 in Online Appendix B graphically illustrates the approach, which is laid out in detail in the following. We conduct all estimations separately for each of the four countries, and once for all countries combined.

The relationship between the parental background and children's outcomes, controlling for other characteristics, gives us the degree of immobility. The baseline regression to elicit this degree reads:

$$Ed_h^C = \alpha_1 + \beta Ed_h^P + X_h{}'\zeta_1 + \epsilon_{h,1} \tag{2}$$

where Ed_h^C is the years of schooling of child h in round five and Ed_h^P is child h's parents' years of schooling. Our coefficient of interest, β, is the measure of the degree of educational immobility, representing the relation between one additional year of parental schooling and the child's probability of obtaining one additional year of schooling. The vector X_h contains child h-specific control variables that are not expected to be correlated with parental background but could be correlated with their educational success. $\epsilon_{h,1}$ is the error term, with $\mathbb{E}[\epsilon_{h,1}|Ed_h^P, X_h] = 0$.

As control variables X_h, we incorporate a set of variables typically used in other analyses of intergenerational immobility (see Solon 1999). These include the child's gender, birth rank, the household head's age and a squared term of that. All the control variables are observed during the first round of the survey. The child order, that is, each child's birth rank, is a categorical variable valued from 1 if child h is the first born to 5 for the fifth born. The variable takes a value of 6 if the child is the sixth or later born child. The birth order is considered because it has been theoretically and empirically shown that the level of children's education differs in intra-household comparison between the siblings born earlier and those born later (see, for instance, Ejrnæs and Pörtner 2004, and Black et al. 2005 for developed countries).[23] Including the age of the household head and its quadratic term controls for accumulation over age, which is particularly relevant for robustness tests using parental wealth as a proxy for the parents' socioeconomic status. Controlling for the age of the children is not necessary since the children in the sample have almost the same age due to the sampling design.

In line with most of the literature on educational immobility we conduct all estimations by employing an ordinary least squares (OLS) model.[24] Because we run separate estimations for each country, we use heteroscedasticity robust standard errors without further clustering them in our estimations. In the cross-section, the only possible additional clustering of standard errors would be on the very fine-grained level of survey sites, which leads to few observations per cluster because sites comprise as little as 19 observations. This can equally bias the estimations of the standard errors as having too few clusters (for a discussion, see Cameron and Miller 2015). We show below that clustering at the site level does not affect any of our results.

When we use the raw number of the parents' and children's years of schooling, the estimation yields what is typically referred to as the intergenerational regression coefficient (IGRC). If we instead standardize the parents' and children's years of schooling to have a mean of zero and a standard deviation of one, respectively, we obtain the intergenerational correlation (IGC; for a discussion of the differences between the IGRC and the IGC, see Emran and Shilpi 2019). Estimating the IGC mitigates concerns of co-residency bias (Emran et al. 2018) and avoids capturing patterns of general upward mobility across the population when comparing estimations for different countries or regions. The IGC estimates are larger than the IGRC estimates if the variance of parental education is larger than the variance in children's education (which is the case in our data) and vice versa. For the analysis of the pathways of the measured immobility, using the IGRC or the IGC is isomorphic. Thus, the choice of the relevant measure of immobility does not influence our central results regarding the pathway factors.

[23] Birth rank may be correlated with parental socioeconomic background because of decisions regarding the family structure. However, family structure and the attention that parents can devote to each child will be included as a measure of a pathway variable later, so that, conditional on these, the rank itself can be treated as exogenous.

[24] The decomposition of the different pathways further below also requires linear estimation.

For the analysis of the pathway factors, we are interested in (a) the potential influence of parental background on these variables and (b) their effect on the probability of being successful in school. The set of pathway factors obtained from the factor analysis is given by $\mathcal{F}$. We test relationship a) and investigate whether the pathways $PW_{i,h}$, $i \in \mathcal{F}$, are correlated with the parents' education. The respective regression equation for each pathway $i \in \mathcal{F}$ reads:

$$PW_{i,h} = \alpha_{i,2} + \lambda_i Ed_h^P + X_h{}'\zeta_{i,2} + \epsilon_{i,h,2} \tag{3}$$

We estimate separate regressions of Equation (3) for each pathway factor. The same set of variables X_h from Equation (2) remains to be controlled for, and $\mathbb{E}[\epsilon_{i,h,2}|Ed_h^P, X_h] = 0, \forall i \in \mathcal{F}$.

After investigating the effect of parental background on the potential pathway factors, we test relationship b), analyzing whether each of the pathway factors is also correlated with the children's educational outcomes. To this end, we estimate the following regression:

$$Ed_h^C = \alpha_3 + \sum_i \rho_i PW_{i,h} + \gamma_{Ed^P} Ed_h^P + X_h{}'\zeta_3 + \epsilon_{h,3} \tag{4}$$

where the different ρ_i's indicate the relationship of each pathway i with child h's educational outcome, and $\mathbb{E}[\epsilon_{h,3}|\{PW_{i,h} : i \in \mathcal{F}\}, Ed_h^P, X_h] = 0$.[25] In this step, all mediating factors are included in the regression. This eliminates their common variation from the estimation, which may raise concerns of multicollinearity between the factors. We return to the issue of potential multicollinearity and the relationships between the pathway factors further below after discussing the main results in Section 4. Because of the potential correlation between the pathway factor variables, it is important to include them jointly in one regression. Otherwise, the estimations of the ρ_i's would suffer from omitted variable bias. We will also return to this issue in Section 4.

In addition to the pathway factors, we include parental education as an explanatory variable in the estimation. γ_{Ed^P} represents the *direct relation* between parental education and child h's educational outcome, which cannot be explained by means of the included pathway factors (see also Conti et al. 2016).[26]

In the last step, we apply the decomposition approach by Bowles and Gintis (2002) and extend it in the spirit of Blanden et al. (2014) to analyze the significance of each pathway factor in the transmission of socioeconomic status between generations. The procedure can be understood through the following relationships. Inserting Equations (3) into Equation (4)

[25] In the estimations on the sample of all countries combined, we furthermore include country fixed effects to capture differences in the level of education between the four countries. The corresponding estimation equations are laid out in Online Appendix C.

[26] Including the parental background as an explanatory variable also further reduces the risk of omitted variable bias in the estimation of the relationship between the pathway factors and educational outcomes. It can capture, for example, the effects of networks and other children's characteristics that cannot be measured.

yields

$$\begin{aligned} Ed_h^C =&(\gamma_{Ed^P} + \sum_i \rho_i\lambda_i)Ed_h^P + {X_h}'(\zeta_3 + \sum_i \rho_i\zeta_{i,2}) \\ &+ \alpha_3 + \sum_i \rho_i\alpha_{i,2} + \sum_i \rho_i\epsilon_{i,h,2} + \epsilon_{h,3} \end{aligned} \tag{5}$$

Comparing this with Equation (2) reveals that $\zeta_3 + \sum_i \rho_i\zeta_{i,2} = \zeta_{i,1}$, and $\alpha_3 + \sum_i \rho_i\alpha_{i,2} = \alpha_1$. Most importantly, $\gamma_{Ed^P} + \sum_i \rho_i\lambda_i = \beta$, where γ_{Ed^P} is the direct relation between parental education and children's education, and $\sum_i \rho_i\lambda_i$ is the indirect relationship mediated by the pathway factors.[27] In order to obtain the mediated relationship between parental socioeconomic status and the educational outcome by each pathway variable, the results from the estimations of the respective ρ_i and λ_i by Equations (3) and (4) are multiplied. Hence, the mediation analysis combines the estimation results of the two previous steps into one result. The estimated share that each pathway i contributes to the immobility is given by $\frac{\hat{\rho}_i\hat{\lambda}_i}{\hat{\beta}}$. Whether the findings from the first of these steps, estimating the relationship between parental background and the pathway factors, express a causal relationship is a matter of perspective. In the succeeding step, it is plausible that for some of the pathways the observed correlation captures a reverse causal effect from the expected children's schooling outcome on the pathways. We therefore abstract from any causal interpretations throughout the paper.[28]

[27] See Gelbach (2016) for a discussion of the connection between the direct and indirect relations. The method we used to elicit the contribution of the pathway variables is also suggested there.

[28] It is central to the literature on intergenerational mobility that the influence of the background on children is multidimensional and that it is the parents' characteristics, not only their education or income, that affect children's outcomes. Thus, our observable measure of socioeconomic status is a proxy for the socioeconomic status in general. From this perspective, reverse causality can be excluded for the "effects" of the background on both, children's outcomes and on their and their youth's characteristics, as the time structure rules out that children's characteristics and outcomes influence their former background. This perspective underlines the ties between background in an abstract sense on the one hand and the offspring's and their youth's characteristics on the other hand as inherent tendencies. The relationship between pathway factor variables and children's outcomes, estimated by Equation (4), can even from this perspective be bidirectional if the knowledge of lower outcomes incentivizes children or their families to adjust their decisions and attitudes accordingly. This is also common to other studies that analyze the mediating factors of the persistence of socioeconomic status, such as Blanden et al. (2014, 2007), or those that analyze the mediating factors of randomized interventions, such as Heckman et al. (2013). Our study design alleviates this concern by using data on children's characteristics elicited early in the children's lives. Nonetheless, in order to be consistent in terminology, and in line with the majority of literature on intergenerational mobility, we in all steps restrict the interpretation of relationships to one of correlation.

In order to test for the statistical significance of the combinations of estimated parameters (particularly that of $\frac{\hat{\rho}_i\hat{\lambda}_i}{\hat{\beta}}$), we bootstrap standard errors with 1,000 replications. Bootstrapping is preferred over computing standard errors by the delta method, because it allows any assumptions regarding the underlying distribution to be dropped (Preacher and Hayes 2008).

For the indirect relations – constituted by the products of the coefficients $\hat{\rho}_i\hat{\lambda}_i$ – to be estimated without bias (and consequently, also the direct relation, since $\hat{\beta} - \sum_i \hat{\rho}_i\hat{\lambda}_i = \hat{\gamma}_{Ed^P}$), the error terms from estimating Equations (3) and (4), $\epsilon_{i,h,2}$ and $\epsilon_{h,3}$, must be uncorrelated, that is, $\epsilon_{i,2} \perp \epsilon_3, \forall i \in \mathcal{F}$ (Blanden et al. 2014; Imai et al. 2010b). This cannot be tested (Imai et al. 2010a), but a number of sensitivity analyses can be conducted. We present these after discussing the results in the following section.

Furthermore, in an extension to the baseline model, we allow for interdependencies between factors. In the unstructured single level mediation analysis, all factors are treated on an equal footing and only their effects that are independent of each other are considered to estimate their share in contributing to educational immobility. To allow for interdependencies, we apply a structured multilevel mediation analysis, as in Bolt et al. (2021). Here, we explicitly consider that some factors will influence other factors, and that their common effect should be attributed to the one that influences the other. The results of the multilevel mediation analysis are presented after discussing the findings from the baseline estimation. The multilevel methodology is outlined in detail in Online Appendix J.

4. Results

The results of the estimations are shown in Table 3, both for each country separately and for the sample of all countries combined. The table includes the results from estimating Equations (2) and (5). It shows the degree of immobility, measured by the IGRC ($\hat{\beta}$) or the IGC, and the shares that each pathway contributes to this immobility, given by $\frac{\hat{\rho}_i\hat{\lambda}_i}{\hat{\beta}}$.[29] The explained component of $\hat{\beta}$ displays the extent to which all the pathways account for the observed degree of social immobility, that is $\frac{\sum_i \hat{\rho}_i\hat{\lambda}_i}{\hat{\beta}}$. The unexplained component of $\hat{\beta}$ is the residual part of the persistence of socioeconomic status, which cannot be accounted for by the pathway mechanisms included. This is the direct relation $\hat{\gamma}_{Ed^P}$ from Equation (4) as a share of immobility, $\frac{\hat{\gamma}_{Ed^P}}{\hat{\beta}}$ or $1 - \frac{\sum_i \hat{\rho}_i\hat{\lambda}_i}{\hat{\beta}}$. Bootstrapped standard errors are shown in parentheses.

In order for each share $\frac{\hat{\rho}_i\hat{\lambda}_i}{\hat{\beta}}$, reported in Table 3, to be significant and, thus, pathway factor i to play a significant role in the overall transmission process, both relationships – that is, a) between parental education and a pathway factor (step (ii)) and b) between a pathway factor and the children's educational outcome (step (iii)) – all other things equal, need to be economically and statistically significant. For those that play no role, one of the two

[29] These percentage values are identical when the IGC is decomposed, which is why we do not report these estimations separately.

Table 3. *Results*

	Ethiopia	India	Peru	Vietnam	All
IGRC ($\hat{\beta}$)	0.339***	0.250***	0.296***	0.405***	0.318***
	(0.033)	(0.025)	(0.037)	(0.027)	(0.015)
IGC	0.404***	0.297***	0.352***	0.482***	0.378***
	(0.040)	(0.030)	(0.044)	(0.032)	(0.018)
Pathways in percent of $\hat{\beta}$					
Child Labor	0.060***	0.157***	0.061**	0.021**	0.067***
	(0.021)	(0.034)	(0.030)	(0.010)	(0.011)
Infrastructure	0.069**	0.026	0.003	0.012	0.032***
	(0.034)	(0.026)	(0.030)	(0.013)	(0.012)
Education Spending	0.013	0.124***	0.020	0.041***	0.027**
	(0.023)	(0.040)	(0.020)	(0.014)	(0.012)
Child Engagement	0.065***	0.102***	0.038	0.091***	0.084***
	(0.020)	(0.021)	(0.024)	(0.018)	(0.010)
Parent Attentiveness	0.012	0.035**	0.036	0.012	0.026***
	(0.010)	(0.015)	(0.028)	(0.010)	(0.006)
Social Environment	−0.003	0.003	−0.008	0.001	−0.001
	(0.006)	(0.007)	(0.017)	(0.003)	(0.002)
Health	0.027*	−0.004	0.013	0.013	0.013*
	(0.015)	(0.016)	(0.019)	(0.011)	(0.007)
Aspirations	0.047*	0.259***	0.089***	0.234***	0.183***
	(0.026)	(0.041)	(0.032)	(0.032)	(0.018)
Cognitive Ability	0.188***	0.289***	0.153***	0.114***	0.200***
	(0.041)	(0.047)	(0.051)	(0.024)	(0.018)
Non-Cognitive: Grit	−0.001	−0.004	0.014	0.001	0.005*
	(0.006)	(0.011)	(0.017)	(0.006)	(0.003)
Explained component of $\hat{\beta}$	0.477***	0.987***	0.419***	0.539***	0.636***
	(0.067)	(0.102)	(0.080)	(0.049)	(0.036)
Unexplained component of $\hat{\beta}$	0.523***	0.013	0.581***	0.461***	0.364***
	(0.067)	(0.102)	(0.080)	(0.049)	(0.036)
Observations	752	921	582	906	3161

Notes: This table shows the results from a decomposition approach, as presented in Equation (5). Thereby, the respective coefficients from estimating Equations (3) and (4), as given in Tables D1-D3 and E1, are multiplied to elicit the mediating effect of the pathway variables in the transmission of the overall effect of parental education on the children's educational outcome ($\hat{\beta}$). The decomposition approach is conducted separately for each country individually and all countries combined. The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. The pathway factors mainly represent characteristics of the children between the ages 11 and 18. For detailed information on the pathway factors, see Table 2. The values provided for each pathway presented in the table give the relative share of this pathway factor for $\hat{\beta}$. Bootstrapped standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

relationships is not present. The results of the intermediate steps of estimating Equations (3) and (4) are shown in reduced form in Table 4.[30] The first column for each country (and for the full sample) shows the $\hat{\lambda}_i$'s from estimating Equation (3) for each $i \in \mathcal{F}$. The second column shows the $\hat{\rho}_i$'s from estimating Equation (4). All factor variables are standardized. Hence, the coefficients can be compared in size between factors, but they cannot be intuitively interpreted as they express the relationship between one standard deviation of the factor variables and one year of the parents' or children's schooling, respectively.

Common Results Across All Four Countries

The aggregated results depicted in Table 3 show that some observations are common to all four countries studied. First, all four countries exhibit a considerable amount of educational immobility, measured either by the IGRC (between 0.25 and 0.41) or the IGC (between 0.3 and 0.48). However, the levels of immobility are smaller than in most previous studies for all four countries. They are still at the upper bound of most estimations for developed countries, but low for developing countries. One of the most mobile countries in the world, Denmark, has an educational immobility (IGRC) of 0.2 and even the US, which is considered to be rather immobile, exhibits an IGRC of around 0.33 for children born in the 1980s (World Bank 2023). Our measures and methodology to elicit the immobility are similar to comparable studies, but the sampling period used here is more recent by at least a decade.[31]

Second, differences in cognitive ability and aspirations between children from high- and low-educated parents account for the largest part of the immobility in all four countries (with the exception of aspirations in Ethiopia). The result of the major role of cognitive skills is in line with findings for developed countries. They also hold in our sample when different measures of cognitive ability are used, the results being illustrated in Online Appendix F. This finding shows that also in developing countries, individual characteristics of the children seem to be more important pathway factors than factors relating to their opportunities or their environment. However, another type of individual characteristics, namely non-cognitive abilities, play no role in either of the four developing countries studied.

[30] Full expositions of both steps are shown in Appendices D and E, including full regression tables in Tables D1, D2, and E1.

[31] It has to be noted, however, that the samples are not fully representative, particularly in Ethiopia, and regionally also not for India.

Table 4. *Correlations between Factor Variables and Parental and Children's Education*

	Ethiopia		India		Peru		Vietnam		All	
	Ed^P	Ed^C	Ed^P	Ed^C	Ed^P	Ed^C	Ed^P	Ed^C	Ed^P	Ed^C
Child Labor	−0.038***	−0.538***	−0.059***	−0.666***	−0.062***	−0.293**	−0.021***	−0.403***	−0.045***	−0.468***
	(0.007)	(0.170)	(0.007)	(0.108)	(0.011)	(0.124)	(0.006)	(0.152)	(0.004)	(0.067)
Infrastructure	0.091***	0.260**	0.062***	0.105	0.052***	0.016	0.040***	0.117	0.063***	0.164***
	(0.009)	(0.122)	(0.007)	(0.100)	(0.008)	(0.168)	(0.007)	(0.122)	(0.004)	(0.062)
Education Spending	0.029***	0.153	0.136***	0.229***	0.055***	0.110	0.016***	1.024***	0.070***	0.120**
	(0.006)	(0.247)	(0.013)	(0.072)	(0.012)	(0.096)	(0.002)	(0.300)	(0.006)	(0.056)
Child Engagement	−0.025***	−0.877***	−0.051***	−0.496***	−0.011	−0.991***	−0.043***	−0.863***	−0.037***	−0.725***
	(0.006)	(0.206)	(0.008)	(0.082)	(0.007)	(0.265)	(0.008)	(0.111)	(0.004)	(0.071)
Parent Attentiveness	0.014	0.305***	0.026***	0.338**	0.067***	0.158	0.036***	0.133	0.033***	0.249***
	(0.010)	(0.112)	(0.006)	(0.116)	(0.012)	(0.113)	(0.008)	(0.110)	(0.004)	(0.057)
Social Environment	−0.004	0.282**	0.012**	0.057	−0.034***	0.074	0.002	0.139	−0.002	0.155**
	(0.007)	(0.137)	(0.005)	(0.136)	(0.011)	(0.131)	(0.008)	(0.100)	(0.004)	(0.062)
Health	0.037***	0.243**	0.034***	−0.028	0.034***	0.115	0.038***	0.136	0.037***	0.116**
	(0.010)	(0.112)	(0.007)	(0.105)	(0.008)	(0.164)	(0.008)	(0.109)	(0.004)	(0.059)
Aspirations	0.059***	0.272**	0.075***	0.862***	0.025***	1.057***	0.093***	1.018***	0.069***	0.843***
	(0.007)	(0.137)	(0.007)	(0.097)	(0.007)	(0.239)	(0.009)	(0.119)	(0.004)	(0.065)
Cognitive Ability	0.046***	1.380***	0.065***	1.106***	0.046***	0.974***	0.047***	0.982***	0.052***	1.211***
	(0.011)	(0.108)	(0.006)	(0.106)	(0.009)	(0.258)	(0.007)	(0.167)	(0.004)	(0.067)
Non-Cognitive: Grit	−0.002	0.215*	0.026***	−0.038	0.011	0.390***	0.021**	0.023	0.016***	0.107**
	(0.009)	(0.112)	(0.007)	(0.098)	(0.011)	(0.122)	(0.009)	(0.099)	(0.004)	(0.054)

Notes: This table shows the results for each step of the decomposition. The first column for each country (and the full sample) shows the relationship between parental education and each pathway factor as given by Equations (3). The full regression results are given in Tables D1-D3. The second column for each country (and the full sample) shows the relationship between the pathway factors and children's education, from estimating Equation (4). The full regression results are given in Table E1. Ed^P is the years of schooling completed by the parents. Ed^C is the years of schooling completed by the child at age 21–22. The pathway factors mainly represent characteristics of the children between the ages 11 and 18, and are standardized, such that the coefficients display the increase of one standard deviation of each pathway factor that is associated with one year more of Ed^P (first column for each country) and Ed^C (second column for each country). For detailed information on the pathway factors, see Table 2. Robust standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

An inspection of the disaggregated results of steps (ii) and (iii) in Table 4 shows that whereas non-cognitive skills are significantly related to parental education in India and Vietnam, they do not affect educational outcomes of the children in these countries. In Ethiopia and Peru, it is the other way around. This striking finding stands in contrast to studies of high-income countries but is in line with the results obtained by Nordman et al. (2015), who find that non-cognitive skills cannot explain wage differences in Bangladesh. We can only speculate about the underlying reason for this difference. One reason could be that non-cognitive skills are more important in determining differences in outcomes at a higher level of schooling, but tertiary education plays a lesser role in the developing countries studied. Including other non-cognitive skills in the estimation confirms that the available measures of non-cognitive skills, namely factors of agency (or locus of control), pride (or self-esteem), and inclusion play no role in the transmission of educational outcomes in these settings. We elaborate on these results in Online Appendix G.

However, the results also display notable differences between the four countries. In the following, we discuss the findings for each country individually.

Ethiopia

Ethiopia is the poorest of the four countries studied, and one of the poorest countries in the world. It is also the only country in this study in which real per capita income has even decreased since the 1970s (see Table 1). It has a weak educational system with comparably low completion rates of primary and secondary schooling. For the study of intergenerational educational mobility, it has been an impediment for the analysis of the interdependence, that the generation of the parents universally had almost no formal education. For example, the study by Hertz et al. (2007) found an IGRC of 0.75 (one of the highest in the world), but an IGC of only 0.1 (one of the lowest) for Ethiopia, where the difference can only be explained by anomalies that arise from missing variation on the side of parental education. This is different with our more recent and comprehensive, although not fully representative data, where the average parental education stands at a mean of 3.34 (which is the lowest in the sample, but not much different from India), with a standard deviation of 4.04 (which is similar to all three other countries, see tables A5 and A6). The resulting estimate of educational immobility is not the largest in the sample and not particularly high compared to other developing countries, including those in this study, with an IGRC of 0.34 (IGC of 0.4). This is in line with findings from Haile (2018), who find significant educational mobility in Ethiopia using other measures than the ones used here, revealing educational mobility being greater than mobility of earnings and occupational mobility. Our estimates of immobility are lower by 0.1 (IGC) to 0.3 (IGRC) compared to those of World Bank (2023) for the cohort of children born in the 1980s.

The channels responsible for the immobility can shed some light on the underlying structure of the identified immobility. Ethiopia is the only country in the sample in which the school infrastructure is relevant for immobility (with 6.9 % of the estimated immobility), implying that poorer children have less access to schools. To address this issue in Ethiopia, the Ministry

of Education has implemented an Alternative Basic Education Programme (UNICEF 2018b) in the early 2000s to improve access to primary schools for children from all backgrounds. Parental spending on education on the other hand plays no role, as the Ethiopian government abolished school fees in 1996. Given the still existing high levels of child labor and child marriage in the country, it is not surprising that these factors also contribute to the estimated Ethiopian immobility (with 6 % and 6.5 %, respectively). Ethiopia is furthermore the only country in the sample in which children's aspirations play only a relatively small role. Table 4 shows that aspirations predict actual schooling outcomes of the children in Ethiopia much less than in the other three countries. This suggests that in an environment in which economic activity is restricted, the plans of children (and their parents) have smaller chances of being realized and, thus, also play a smaller role. In contrast, Ethiopia with its high level of childhood stunting is the only country in which children's early life health status is a significant factor for the intergenerational persistence of educational success, although it is also small at 2.7 %, once all other factors are controlled for. We will return to the role of health for other pathway factors further below, when we explicitly model interactions between the pathways.

India

The estimated immobility in India is remarkably low with 0.25 (IGRC) and 0.3 (IGC). This is substantially lower than the estimates by Azam and Bhatt (2015) (applying the methodology of Hertz et al. 2007 for India), or those by World Bank (2023), or Emran and Shilpi (2015), who estimated either IGRC or IGC at between 0.5 and 0.6. It is, however in line with estimates by Jalan and Murgai (2008) and Maitra and Sharma (2009), who each estimate an IGRC for the father's education between 0.2 and 0.3. All these estimations are based on samples of children born before 1990 (and, in parts, back to the generations born in the 1950s). However, India underwent a radical economic liberalization process in the early 1990s. This was accompanied by an increase in inequality (World Bank 2011), but this does not necessarily imply an increase in immobility alike, and our results suggest that it, indeed, may not have.[32]

Despite the high levels of childhood stunting, children's health status does not play a role for intergenerational immobility in India. This may be due to the fact that in 2001, India established a program to provide free mid-day meals in public schools. Child labor and child engagement are relevant, explaining 15.7 % and 10.2 % of immobility, respectively. India's high child marriage rates make it unsurprising that this is a very relevant channel. And whereas the official numbers for child labor are low in India, it introduced a law prohibiting employment under the age of 14 only in 2016 and ratified the ILO Minimum Age Convention (ILO Convention 138) only in 2017. This is much later than most other countries, including all the ones studied here. It remains to be seen whether this will have an effect on the role of child labor and, with it, intergenerational mobility in India in general; particularly, given the noted prevalence of household work done by children. Furthermore, India has a

[32]Kundu and Sen (2023) also show that educational immobility between fathers and sons has decreased in India, looking a cohorts of sons born up to 1994, but using different measures of immobility.

substantial private school sector, with 45 % and 51 % attending a (costly) private primary or secondary school, respectively, in 2015, and these numbers were already high in the early 2000s (World Bank 2021). Private schools are also found to come with modestly better schooling outcomes for students (Singh 2015). However, even (free) public primary schooling comes with additional costs not covered by the government, such as textbooks or uniforms (Gupta 2019). This explains why India is the only country in the sample where education spending by the parents substantially contributes to the greater success in school by children from better educated parental backgrounds. Still, also in India, children's cognitive ability and aspirations contribute the largest part to the immobility, with 28.9 % and 25.9 %, respectively. Overall, the channels considered in this analysis can explain a large part of the estimated immobility in India.

Peru

Also for Peru, our estimates of educational immobility are lower than previous estimates. In the study of Hertz et al. (2007), in which the considered sample for Peru includes children born up to 1983, the estimated Peruvian immobility is the highest of all 42 countries considered in the sample with an estimated correlation coefficient of 0.66. In comparison to that, the findings of World Bank (2023) (children born up to 1989), already show a lower IGC of 0.51.[33] Our estimation of 0.35 on the more recent (but also smaller) sample is even lower in comparison.[34]

For Peru, the considered pathways can only explain a smaller part of the observed immobility.[35] Cognitive ability and aspirations explain still the largest part, with 15.3 % and 8.9 %, respectively. Child labor is an obstacle for poorer children's school perspectives, explaining 6.9 % of immobility. Despite the still relevant number of child marriages in Peru, this is not a factor in the intergenerational transmission of educational outcomes. An inspection of Table 4 shows the reason for this: although child engagement is an issue in Peru, and those marrying or getting children while still underage have a lower probability of being successful in school, the likelihood of doing so is not influenced by the parents' education.

Vietnam

In Vietnam, the estimation educational immobility in our sample is comparably more similar to previous studies. The estimated values of 0.41 (IGRC) and 0.48 (IGC) are close to

[33] A study by Neidhöfer et al. (2018) estimates the IGRC in Peru to be in the range of 0.4 to 0.48, with a sample ranging up to children born in 1987, suggesting an even lower immobility in that period.

[34] While the study of Munoz (2021) is not directly comparable to our estimations of relative immobility, it estimates Peru to be in the middle of all analyzed Latin American countries in terms of upward mobility.

[35] Also note that Peru has the smallest number of observations in our sample, making the statistical analysis somewhat less authoritative.

those from World Bank (2023) (0.47 and 0.51) and Hertz et al. (2007) (0.58 and 0.4) from earlier generations. After economic reforms in Vietnam in 1986, the country has achieved constant growth rates above 5 % of GDP from 1988 until recently. However, poverty rates are still high and the benefits from growth are distributed unequally. Our results suggest that intergenerational immobility is also still substantial.

More than in any other country in our sample, differences in aspirations between children from varying backgrounds are responsible for the immobility in Vietnam. Explaining 23.4 % of the immobility, this factor constitutes the most important pathway factor. Differences in cognitive ability can 'only' explain 11.4 % here. In Vietnam, where cognitive skills are generally higher and less dispersed (see Table A6), what children want thus appears to be even more important than what they can in order to determine relative outcomes. Child engagement, although overall less of an issue than in other countries, is particularly a concern for children from less educated backgrounds, and is, thus, a relevant contributor to the intergenerational persistence in educational achievements. Child labor also contributes to this (with 2.1 %), but less so than in other countries because it depends only to a small degree on the parents' education level (Table 4).[36] Although private education plays no role in Vietnam and public education is formally free by constitutional pledge, costs for textbooks and uniforms are substantial (Hoang 2013), so that parents' spending on education is also a relevant – although not large – pathway factor (4.1 %) in Vietnam.

Robustness of the Results

The pathway variables considered are most likely not independent from each other. When estimating the relationship between the pathway variables and the children's educational outcome by Equation (4), common variation of the pathways is not taken into account as all variables are jointly included in the estimation. This raises issues of multicollinearity. However, not only are the correlations between the pathway factor variables small, also the variance inflation factors (VIFs) are unproblematic. To further address the issue of multicollinearity, we individually exclude pathway variables from the estimations, which does not affect the results on the relevance of any other pathway factor. Online Appendix E discusses these issues in more detail and presents the respective tests along with the step of estimating the relationship between the pathways and children's education.

As noted above, the estimation of the coefficients may be biased if the error terms $\epsilon_{i,2}$ and ϵ_3 are correlated, such that $\mathbb{E}[\epsilon_{i,2} \cdot \epsilon_3] \neq 0$, for at least one $i \in \mathcal{F}$. The most obvious reason for a correlation of the error terms in the two steps would be the omission of mediating pathway factor variables in Equation (4) which are correlated with the included factors. When multiple mediating variables are included in the estimation, this risk is somewhat reduced (Altonji et al. 2005; Preacher and Hayes 2008) relative to a setting with only one mediating variable (as in, for example, Imai et al. 2010a and Imai et al. 2010b). Altonji et al. (2005, 2011) show that including a large number of observable characteristics also

[36] However, if child labor is pursued, it is equally harmful for success probabilities in school.

controls for unobservable characteristics if the correlation between the variables of interest and the observable characteristics is the same as that between the variables of interest and the unobservable characteristics. In order for this to be the case, the most important variables in determining the outcome must be observable and sufficiently many characteristics not dominant in determining the outcome need to be observed as well. Because the YL survey was designed as to elicit determinants of children's outcomes, it is reasonable to believe that all characteristics that dominate these outcomes are also captured with the data. The estimations can provide an indication whether there are at the same time also sufficient observable characteristics which are not dominant in determining schooling outcomes. Table H1 in Online Appendix H therefore shows the results of estimating Equation (4) when the pathway factors are successively included in the estimation. We start with the one that accounts for the largest part of the immobility observed in Table 3, namely cognitive skills. At some point, adding more pathway factors does neither change the estimated coefficient of the others nor the explained part of the schooling outcome. The results of this exercise are only reported for the sample of all countries for reasons of exposition but they are similar for the individual country samples. This indicates that including a relatively large number of pathway factors, as we do here, reduces the concern about bias stemming from omitted variables that are correlated with the outcome and only one of the factors.

Another potential source of bias could be that parental education influences the *effect* of the pathway factors or control variables on educational outcome (rather than influencing only the *level* of the pathway factors). This would imply that the ζ_3 and ρ_i coefficients are not constant but rather a function of parental education (Heckman et al. 2013). To ensure that this is not the case and that this does not lead to a correlation of the error terms of Equations (3) and (4), we split each country sample by the median in terms of parental education and estimate Equation (4) for each subsample within each country. The comparisons of the $\hat{\zeta}_3$ and $\hat{\rho}_i$ coefficients are reported in Tables H2-H6 in Online Appendix H. They to the most part show no statistically significant difference.[37]

The results of clustering standard errors at the site level are depicted in Table H7 in Online Appendix H. They show that only those results that were marginally significant before lose significance, but all central results do hold even when bootstrapped standard errors are clustered on this very fine-grained level.

Alternatively to the level of the parents' education, other variables can be used to capture the parents' socioeconomic status. On the children's side, this is different, because the children are still too young to have other observable outcomes. Therefore, we elicit the relationship between a wealth index (similar to Landersø and Heckman 2017) or household expenditure and children's education levels and decompose these links. The results of the decomposition with parental wealth and the total expenditures per household member as the measures of parents' socioeconomic status are displayed in Tables H8 and H9. They are very similar to the results based on parental education presented above.

[37] This is in line with Falk et al. (2021), who find that parental background only influences the level, not the productivity, of cognitive and non-cognitive abilities.

Different proceedings to obtain the number of relevant factors and their measures from the survey data lead to slightly different interpretations. However, none of these are at odds with the interpretation of our main results. Online Appendix I shows the relevant groupings when the selection is purely data driven, that is, through an EFA, and presents the findings from the respective decomposition analysis. The results of this exercise confirm that cognitive ability and career focus (which combines measures of underage family engagement and aspirations) tend to be the most important pathway factors for educational immobility in all four countries, whereas spending on education plays only a role in India and Vietnam. The analysis also shows that non-cognitive skill factors have no part in the intergenerational transmission of educational status. Although some differences driven by correlation patterns between the measures are notable, this generally confirms the appropriateness of the intuitive approach presented.[38]

Gender Differences

The effect of parental education on children's education levels is likely to be different for sons and daughters. Therefore, we re-estimate the levels of educational immobility and its pathways for the subsamples of boys and girls in each country. This reduces the size of the respective samples substantially (as far as down to 275 for girls in Peru), but most subsamples are still large enough to conduct the mediation analysis meaningfully. The results are shown in Table 5.

Across all four countries, the point estimates of immobility are larger for girls than for boys, as could be expected from previous studies in developed and developing economies (Corak 2006; Leone 2021).[39] However, this difference is (statistically and economically) only significant in India, which confirms findings by Emran and Shilpi (2015). In Ethiopia, Peru, and Vietnam, the data in our sample cannot confirm a difference between the educational immobility for girls and boys.

[38] Also by applying different forms of an EFA, depending on the selection criterion, different factors can be dropped or their measures merged to one factor. This says nothing about the relationship of the respective group of measures with either the parents' or the children's education. An alternative approach would be to simply include an individual measure as a proxy for measures of interest (such as the hours worked for child labor, the logical skill tests, or the urban/rural divide for infrastructure, which can get dropped otherwise) as a pathway variable. This yields similar results on the absolute and relative importance of the pathway (as it does for infrastructure and the social environment), but for reasons of methodological consistency and the advantages of factor analysis when using information from all available measures, we take the middle ground, using the derived factor scores from an intuitive grouping as pathway factors.

[39] Again, the estimated differences in IGC and IGRC between the two genders are quite similar, mirroring the fact that the variance of daughters' education levels is only slightly smaller than that of sons' education levels, compared to the variance of the respective parents' education levels.

Table 5. *Results by Gender*

	Ethiopia		India		Peru		Vietnam		All	
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
IGRC ($\hat{\beta}$)	0.356***	0.323***	0.323***	0.163***	0.312***	0.284***	0.410***	0.398***	0.352***	0.282***
	(0.045)	(0.052)	(0.037)	(0.035)	(0.056)	(0.052)	(0.040)	(0.040)	(0.020)	(0.022)
IGC	0.424***	0.384***	0.385***	0.194***	0.371***	0.338***	0.488***	0.473***	0.419***	0.336***
	(0.054)	(0.061)	(0.044)	(0.042)	(0.066)	(0.062)	(0.047)	(0.047)	(0.024)	(0.026)
Pathways in percent of $\hat{\beta}$										
Child Labor	0.032	0.080*	0.154***	0.164**	0.055	0.069	0.040**	0.008	0.067***	0.066***
	(0.028)	(0.041)	(0.038)	(0.073)	(0.038)	(0.057)	(0.020)	(0.011)	(0.015)	(0.019)
Infrastructure	0.028	0.102*	0.028	0.044	0.005	0.021	−0.009	0.032	0.012	0.061***
	(0.049)	(0.055)	(0.025)	(0.074)	(0.038)	(0.051)	(0.018)	(0.021)	(0.015)	(0.023)
Education Spending	0.038	−0.031	0.084	0.161**	0.015	0.048	0.047***	0.033	0.015	0.032**
	(0.031)	(0.046)	(0.077)	(0.081)	(0.024)	(0.032)	(0.016)	(0.022)	(0.021)	(0.014)
Child Engagement	0.126***	0.012	0.129***	0.040	0.085	0.004	0.171***	−0.010	0.139***	0.013*
	(0.036)	(0.009)	(0.029)	(0.026)	(0.058)	(0.021)	(0.033)	(0.020)	(0.018)	(0.008)
Parent Attentiveness	0.017	0.009	0.028	0.041	−0.000	0.068	0.022	0.007	0.021**	0.030***
	(0.016)	(0.019)	(0.018)	(0.031)	(0.039)	(0.046)	(0.016)	(0.014)	(0.009)	(0.010)
Social Environment	−0.000	−0.006	−0.001	0.013	0.026	−0.045	0.004	−0.002	−0.000	−0.002
	(0.007)	(0.014)	(0.007)	(0.022)	(0.036)	(0.032)	(0.008)	(0.006)	(0.002)	(0.005)
Health	0.015	0.044	−0.006	0.005	0.004	0.050	0.002	0.025	0.006	0.025**
	(0.015)	(0.036)	(0.015)	(0.041)	(0.021)	(0.041)	(0.018)	(0.016)	(0.008)	(0.013)
Aspirations	0.031	0.068	0.263***	0.239***	0.062	0.108**	0.242***	0.228***	0.169***	0.180***
	(0.027)	(0.048)	(0.051)	(0.081)	(0.046)	(0.047)	(0.051)	(0.048)	(0.023)	(0.025)
Cognitive Ability	0.187***	0.173**	0.206***	0.427***	0.114*	0.184**	0.124***	0.113***	0.183***	0.211***
	(0.056)	(0.071)	(0.047)	(0.121)	(0.069)	(0.074)	(0.038)	(0.034)	(0.024)	(0.029)
Non-Cognitive Ability	−0.000	0.003	−0.001	−0.009	0.020	0.005	0.002	−0.002	0.004	0.009
	(0.006)	(0.017)	(0.012)	(0.026)	(0.025)	(0.026)	(0.009)	(0.008)	(0.004)	(0.006)
Explained component of $\hat{\beta}$	0.475***	0.455***	0.884***	1.124***	0.385***	0.512***	0.645***	0.432***	0.615***	0.624***
	(0.085)	(0.112)	(0.110)	(0.279)	(0.127)	(0.126)	(0.067)	(0.072)	(0.043)	(0.056)
Unexplained component of $\hat{\beta}$	0.525***	0.545***	0.116	−0.124	0.615***	0.488***	0.355***	0.568***	0.385***	0.376***
	(0.085)	(0.112)	(0.110)	(0.279)	(0.127)	(0.126)	(0.067)	(0.072)	(0.043)	(0.056)
Observations	350	402	473	448	275	307	462	444	1560	1601

Notes: This table shows the results from a decomposition approach, as presented in Equation (5). Thereby, the respective coefficients from estimating Equations (3) and (4) are multiplied to elicit the mediating effect of the pathway variables in the transmission of the overall effect of parental education on the children's educational outcome ($\hat{\beta}$). The decomposition approach is conducted for each gender individually, separated by country and all countries combined. The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. The pathway factors mainly represent characteristics of the children between the ages 11 and 18. For detailed information on the pathway factors, see Table 2. The values provided for each pathway presented in the table give the relative share of this pathway factor for $\hat{\beta}$. Bootstrapped standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

Furthermore, most estimated pathway factors are quite similar in importance or non-importance for both genders. There are, however, some exceptions. Most notably, child engagement is only a factor for girls in the three countries in which it is found to be a relevant factor, that is, in Ethiopia, India, and Vietnam.[40] This result is not surprising, given that being married and/or becoming a parent when still a minor is mostly a problem for girls. Our results show that it is an important driver of their worse educational prospects in Ethiopia, India, and Vietnam.

In India, also differences in cognitive ability between boys from different backgrounds play a clearly larger role than those of girls, which could suggest that either boys are more likely to be evaluated according to their cognitive potential, or that girls from better educated backgrounds are less likely to also benefit from their backgrounds along this dimension. The same holds to some extent for Peru, where the difference equally holds for aspirations, and both seem to be a more relevant factor for transmission to the male offspring. Finally, in Vietnam, child labor is an important pathway factor merely for girls.

Pathway Interactions

Above, we noted that the correlations between the individual pathway factors are not as large as to bias the estimations. However, dynamic relationships may still exist between the pathway factors. For example, although the health status is found to have small to no effects on the children's educational outcomes, it may in fact influence the cognitive development of the children, which then determines their success in school (Attanasio et al. 2020; López Boo 2016; Figlio et al. 2014). If early health status was the only source of differences in cognitive skills, our previous estimations would ignore this effect of parental education completely. If there were also other reasons why cognitive abilities differ – besides early health status – the channel would be entirely attributed to cognitive ability, but not to health. However, these types of relationships can explicitly be accounted for in the decomposition analysis. To this end, we conduct a multilevel mediation analysis, similar to Bolt et al. (2021).

For the multilevel mediation analysis, hypothesized relationships between the pathway factors are included by adding the explanatory pathway factor as an independent variable in Equation (3) for the respective explained pathway factor. Thereby, the effect of the common variation of the two pathway factors is attributed to the explanatory pathway, that is, the "previous" factor.[41] From the broad set of pathway factors analyzed, one can assume such a clear one-

[40] Even in Peru, the point estimate on child engagement is positive for girls and of comparable size, but the statistical power is too low to deduct robust interpretation from this.

[41] In the unstructured mediation analysis presented above, common variation of two pathway variables is excluded from the analysis of the relationship between the pathway factors and children's education levels in Equation (4). However, because it is included in the analysis of the relationship between the parents' education and the pathway factors in Equation (3), it is assumed that the common variation with another factor has the same association with children's education as the individual variation of each factor.

directional relationship for some pairs of pathway factors. We explicitly consider the following relationships in the structured multilevel mediation analysis: a) Both health and parent attentiveness could affect cognitive ability as well as non-cognitive skills; b) By affecting the understanding of the own possibilities, cognitive and non-cognitive skills may affect children's aspirations; c) The children's aspirations can affect the likelihood of child labor and that of child engagement because both appear as alternatives to schooling and the choice of either may thus be endogenous.[42]

Thus, we assume, for example, that the impact of childhood health on a child's cognitive ability affects its aspirations, which in turn influences whether the child performs child labor, which then affects the schooling outcome. In this multilevel mediation analysis, this effect is completely attributed to childhood health status. Individual variation in cognitive ability and non-cognitive skills (not jointly moving with the health status) and their indirect effects through aspirations (and through these on child labor and child engagement) would still be attributed to cognitive ability or non-cognitive skills, respectively. Furthermore, we allow that in this structure all previous pathway factors also directly influence later pathway factors. However, not all pathway factors can be assumed to be in such unidirectional dependency relationships with other pathway factors – either because they are likely to be independent characteristics of the children's development situation (infrastructure), or their relationship with other pathway factors is likely to be bidirectional (education spending and social environment). Also in the structured multilevel mediation analysis, we consider these as potential independent pathway factors. Online Appendix J shows a graphical illustration of the relationships and lays out the corresponding system of equations.

Table 6 shows the result of the corresponding decomposition analyses with the aforementioned relationships between pathway factors. In Ethiopia, indeed, the childhood health status slightly increases in importance, once its effects on other pathway factors are accounted for. It now explains some part of the variation that was before attributed to cognitive ability and child labor. However, they are still also important on their own. In India, child labor and child engagement are to some extent explained by differences in cognitive ability. Parent attentiveness can explain a part of what would otherwise be attributed to aspirations in Peru, and becomes a relevant pathway factor, with 11.5 %, when taking the relationships between the two variables, direct and indirect via cognitive ability, into account. For Vietnam, the analysis shows that the outstanding role of aspirations is attenuated somewhat, once the effect of cognitive ability on aspirations is accounted for, which in turn is found to be relatively more important. The small effect of child labor in Vietnam which was observed before is now completely explained by the factors conducive to it. Overall, although we implemented an extensive model of multilevel mediation, the results do not change very much in any of the

[42]The directions of these relationships are also generally consistent with the timing of the respective survey measures. The only exception is non-cognitive skills affecting other factors, as their measures in the baseline regressions (of grit) come from the latest round (when children were 21/22). However, as discussed, one can assume stability of this factor over the children's youth. Thus, the later elicited measures also capture earlier characteristics of the children in this regard.

Table 6. *Results Multilevel Mediation Analysis*

	Ethiopia	India	Peru	Vietnam	All
IGRC ($\hat{\beta}$)	0.339***	0.249***	0.295***	0.405***	0.317***
	(0.034)	(0.025)	(0.038)	(0.028)	(0.015)
IGC	0.404***	0.297***	0.352***	0.482***	0.378***
	(0.040)	(0.030)	(0.045)	(0.033)	(0.018)
Pathways in percent of $\hat{\beta}$					
Child Labor	0.037**	0.066***	0.052*	0.011	0.039***
	(0.016)	(0.022)	(0.027)	(0.007)	(0.008)
Infrastructure	0.069**	0.026	0.003	0.012	0.032**
	(0.035)	(0.027)	(0.030)	(0.013)	(0.013)
Education Spending	0.013	0.125***	0.021	0.041***	0.027**
	(0.023)	(0.041)	(0.020)	(0.013)	(0.012)
Child Engagement	0.069***	0.072***	0.022	0.060***	0.061***
	(0.021)	(0.019)	(0.026)	(0.016)	(0.010)
Parent Attentiveness	0.022	0.069***	0.115***	0.037***	0.055***
	(0.017)	(0.023)	(0.039)	(0.013)	(0.010)
Social Environment	−0.003	0.003	−0.008	0.001	−0.001
	(0.007)	(0.007)	(0.018)	(0.004)	(0.002)
Health	0.057***	−0.013	0.028	0.024*	0.027***
	(0.021)	(0.019)	(0.024)	(0.013)	(0.009)
Aspirations	0.047**	0.230***	0.037	0.164***	0.161***
	(0.023)	(0.039)	(0.032)	(0.028)	(0.017)
Cognitive Ability	0.168***	0.409***	0.145***	0.189***	0.229***
	(0.046)	(0.056)	(0.049)	(0.031)	(0.021)
Non-Cognitive: Grit	−0.003	0.001	0.004	0.001	0.005*
	(0.007)	(0.011)	(0.020)	(0.005)	(0.003)
Explained component of $\hat{\beta}$	0.477***	0.987***	0.416***	0.539***	0.636***
	(0.068)	(0.099)	(0.084)	(0.048)	(0.034)
Unexplained component of β	0.523***	0.013	0.584***	0.461***	0.364***
	(0.068)	(0.099)	(0.084)	(0.048)	(0.034)
Observations	752	921	582	906	3161

Notes: This table shows the results from the decomposition approach based on the structured multilevel mediation analysis, as detailed in Online Appendix J. The decomposition approach is conducted separately for each country individually and all countries combined. The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. The pathway factors mainly represent characteristics of the children between the ages 11 and 18. For detailed information on the pathway factors, see Table 2, and on their considered relationships, see Figure J1 in Online Appendix J. The values provided for each pathway presented in the table give the relative share of this pathway factor of $\hat{\beta}$. Bootstrapped standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

countries compared to the more traceable unstructured mediation analysis presented above.

5. Conclusion

Intergenerational social immobility interferes with the societal goals of equality of opportunities (Black and Devereux 2011) and efficiency (Causa and Johansson 2009). There has been no comprehensive research on the transmission mechanisms of intergenerational social immobility in developing countries. Using the extensive Young Lives dataset, we provide new estimates of educational immobility in four developing countries from different regions of the world and analyze the respective channels of social immobility in Ethiopia, India, Peru, and Vietnam.

Using more recent data on cohorts born in 1995, our estimates of intergenerational educational persistence are lower than in any previous study on cohorts born in the 1980s. This finding coincides with school reforms addressed at supporting poorer children – particularly in India and Ethiopia – and poverty reduction in these countries in the early 2000s. The difference to previous findings is least pronounced in Vietnam, which falls back in terms of the level of relative educational mobility compared to the other three countries. In Vietnam, despite high rates of growth over the last 30 years, broader access to education has only started to become an official policy priority after 2005 and was realized only after 2010 (WENR 2017). Even in our relative recent data, this cannot yet show an effect.

Regarding the channels, we find that educational immobility is most strongly mediated through differences in cognitive skills between children of different backgrounds in all four countries. Children's aspirations are an almost equally important factor in India and Peru, and even more so in Vietnam. For girls, child marriage and/or motherhood, that is, the factor of underage family engagement is an important pathway factor in Ethiopia, India, and Vietnam. Child labor also plays a role in transmitting educational status from parents to children in all four countries, particularly in India, but somewhat less so in Vietnam. Ethiopia, the poorest country in the sample, is also the only country in which the school infrastructure and childhood health status play a role for the link between parents and children in educational outcomes. In India, and to a lesser extent also in Vietnam, the monetary spending on education also determines children's outcomes, contributing to the immobility, which can be explained by the financial requirements for schooling in both countries.

The finding that differences in cognitive skills account for much of the immobility observed is also prevalent in studies of developed countries. In difference to these studies, however, we do not find that available non-cognitive factors play a role. We reported the findings on grit in the main text, but this holds equally for factors of agency/locus of control, pride/self-esteem, inclusion, and trust. This does not imply that non-cognitive skills generally play no role in the transmission of educational attainment across generations because it could be that the part of immobility which remains unexplained in our study is driven by other non-cognitive skills. However, it is surprising that none of the – commonly analyzed – non-cognitive factors measured in this study plays a role. It deserves particular attention for developing countries given the growing interest in the role of non-cognitive skills as transmission channels of social

immobility (see Chowdhury et al. 2022 and Falk et al. 2021).

The rich longitudinal Young Lives data has made it possible for us to incorporate more pathway factors than any previous study, including any on developed countries. Some of the pathway factors analyzed are specific to developing countries, such as underage family engagement and child labor. The number of measurements in the data allowed us to address measurement error through a factor analysis of related measures.

One limitation of our study is that it only captures four developing countries, although the data collection aimed to represent the major developing regions in the world. Causally relating the differences in the developments of both the levels of educational mobility and of the relative importance of its channels to institutional or macroeconomic characteristics would be an interesting avenue for future research. Therefore, having newer data also on other developing countries would increase our understanding of the dynamic processes. Another limitation of this study is that the children were first observed when they were already 8 years old, and, for example, data for cognitive ability and aspirations were obtained only at age 12. This does not allow us to analyze the significance of individual factors or the dynamics between the factors in even earlier childhood. This holds particularly for the pathway of aspirations (of the children and the parents), which are the lesser blurred by the expectation of the actual outcome, the earlier they are elicited. A structured multilevel analysis could thus gain from even earlier observations. Our findings only suggest which factors are most likely to operate from the point on when they were measured, and what their influences are later.

Whereas our paper cannot make causal inferences, it clearly illustrates that the relevant factors in developing countries may differ from the ones in developed countries. Also, the institutional settings are mirrored by the relevant and non-relevant channels of immobility. Furthermore, it provides useful information for policy makers to more effectively target intergenerational social immobility in each of the four developing countries. The analysis can also give an indication beyond the four countries studied, by showing which factors should be considered relevant channels of intergenerational social immobility in developing-country settings in general. In this sense, our study also provides a framework for future, in-depth, causal analyses of particular pathway factors.

References

Acciari, P., A. Polo, and G. L. Violante (2022). "And yet it moves: Intergenerational mobility in Italy." *American Economic Journal: Applied Economics* 14(3), pp. 118–163.

Agupusi, P. (2019). "The Effect of Parents' Education Appreciation on Intergenerational Inequality." *International Journal of Educational Development* 66, pp. 214–222.

Alesina, A., S. Hohmann, S. Michalopoulos, and E. Papaioannou (2021). "Intergenerational mobility in Africa." *Econometrica* 89(1), pp. 1–35.

Altonji, J. G., T. Conley, T. E. Elder, and C. R. Taber (2011). "Methods for Using Selection on Observed Variables to Address Selection on Unobserved Variables (Preliminary and Incomplete)." *IZA Conference Paper*.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005). "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools." *Journal of Political Economy* 113(1), pp. 151–184.

Asher, S., P. Novosad, and C. Rafkin (2024). "Intergenerational Mobility in India: New Measures and Estimates Across Time and Social Groups." *American Economic Journal: Applied Economics* 16(2), pp. 66–98.

Attanasio, O., C. Meghir, and E. Nix (2020). "Human capital development and parental investment in India." *The Review of Economic Studies* 87(6), pp. 2511–2541.

Aurino, E. and F. Burchi (2017). "Children's Multidimensional Health and Medium-Term Cognitive Skills in Low- and Middle-Income Countries." *The European Journal of Development Research* 29(2), pp. 289–311.

Azam, M. and V. Bhatt (2015). "Like Father, Like Son? Intergenerational Educational Mobility in India." *Demography* 52(6), pp. 1929–1959.

Balán, J., H. L. Browning, and E. Jelin (1973). *Men in a Developing Society: Geographic and Social Mobility in Monterrey, Mexico*. University of Texas Press.

Bartlett, M. S. (1951). "The effect of Standardization on a Chi-square Approximation in Factor Analysis." *Biometrika* 38(3/4), pp. 337–344.

Becker, G. S., S. D. Kominers, K. M. Murphy, and J. L. Spenkuch (2018). "A Theory of Intergenerational Mobility." *Journal of Political Economy* 126(S1), S7–S25.

Björklund, A., M. Jäntti, and M. Nybom (2017). "The Contribution of Early-life Versus Labour Market Factors to Intergenerational Income Persistence: A Comparison of the UK and Sweden." *The Economic Journal* 127(605), F71–F94.

Black, S. E. and P. J. Devereux (2011). "Chapter 16 - Recent Developments in Intergenerational Mobility." *Handbook of Labor Economics*. Ed. by D. Card and O. Ashenfelter. Vol. 4. Elsevier, pp. 1487–1541.

Black, S. E., P. J. Devereux, P. Lundborg, and K. Majlesi (2017). "On the Origins of Risk-Taking in Financial Markets." *The Journal of Finance* 72(5), pp. 2229–2278.

Black, S. E., P. J. Devereux, P. Lundborg, and K. Majlesi (2020). "Poor Little Rich Kids? The Role of Nature versus Nurture in Wealth and Other Economic Outcomes and Behaviors." *The Review of Economic Studies* 87(4), pp. 1683–1725.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005). "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education." *The Quarterly Journal of Economics* 120(2), pp. 669–700.

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

Blanden, J., R. Haveman, T. Smeeding, and K. Wilson (2014). "Intergenerational Mobility in the United States and Great Britain: A Comparative Study of Parent–Child Pathways." *Review of Income and Wealth* 60(3), pp. 425–449.

Bolt, U., E. B. French, J. H. MacCuish, and C. ODea (2021). *The Intergenerational Elasticity of Earnings: Exploring the Mechanisms*. IFS Working Paper (No.W21/07). London.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. ter Weel (2008). "The Economics and Psychology of Personality Traits." *Journal of Human Resources* 43(4), pp. 972–1059.

Bossuroy, T. and D. Cogneau (2013). "Social Mobility in Five African Countries." *Review of Income and Wealth* 59, pp. 84–110.

Bowles, S. and H. Gintis (2002). "The Inheritance of Inequality." *Journal of Economic Perspectives* 16(3), pp. 3–30.

Boyden, J. (2018a). *Young Lives: An International Study of Childhood Poverty: Round 2, 2006. [Data Collection]. 3rd Edition.*

— (2018b). *Young Lives: An International Study of Childhood Poverty: Round 3, 2009. [Data Collection]. 3rd Edition.*

— (2018c). *Young Lives: An International Study of Childhood Poverty: Rounds 1-5 Constructed Files, 2002-2016. [Data Collection]. 3rd Edition.*

Cameron, A. C. and D. L. Miller (2015). "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50(2), pp. 317–372.

Case, A., A. Fertig, and C. Paxson (2005). "The Lasting Impact of Childhood Health and Circumstance." *Journal of Health Economics* 24(2), pp. 365–389.

Cattell, R. B. (1966). "The Scree Test For The Number Of Factors." *Multivariate Behavioral Research* 1(2), pp. 245–276.

Causa, O. and A. Johansson (2009). *Intergenerational Social Mobility*. OECD Economics Department Working Papers.

Chetty, R. and N. Hendren (2018a). "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *The Quarterly Journal of Economics* 133(3), pp. 1107–1162.

— (2018b). "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates." *The Quarterly Journal of Economics* 133(3), pp. 1163–1228.

Chetty, R., N. Hendren, and L. F. Katz (2016). "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review* 106(4), pp. 855–902.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *The Quarterly Journal of Economics* 129(4), pp. 1553–1623.

Child, D. (2006). *The Essentials of Factor Analysis*. 3rd Ed. New York: Continuum.

Chowdhury, S., M. Sutter, and K. F. Zimmermann (2022). "Economic preferences across generations and family clusters: A large-scale experiment in a developing country." *Journal of Political Economy* 130(9), pp. 2361–2410.

Conti, G., J. J. Heckman, and R. Pinto (2016). "The Effects of Two Influential Early Childhood Interventions on Health and Healthy Behaviour." *The Economic Journal* 126(596), pp. 28–65.

Corak, M. (2006). "Do Poor Children Become Poor Adults? Lessons From A Cross-Country Comparison Of Generational Earnings Mobility." *Dynamics of Inequality and Poverty*. Ed. by J. Creedy and G. Kalb. Emerald Group Publishing.

— (2013). "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *The Journal of Economic Perspectives* 27(3), pp. 79–102.

Dalton, P. S., S. Ghosal, and A. Mani (2016). "Poverty and Aspirations Failure." *The Economic Journal* 126(590), pp. 165–188.

Darroch, J. E. and S. Singh (2013). "Trends in Contraceptive Need and Use in Developing Countries in 2003, 2008, and 2012: An Analysis of National Surveys." *The Lancet* 381(9879), pp. 1756–1762.

Dercon, S. and A. Singh (2013). "From nutrition to aspirations and self-efficacy: Gender bias over time among children in four countries." *World Development* 45, pp. 31–50.

Doepke, M. and F. Zilibotti (2017). "Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission." *Econometrica* 85(5), pp. 1331–1371.

Duckworth, A., C. Peterson, M. Matthews, and D. Kelly (2007). "Grit: Perseverance and Passion for Long-Term Goals." *Journal of Personality and Social Psychology* 92(6), pp. 1087–1101.

Duflo, E. (2012). "Women Empowerment and Economic Development." *Journal of Economic Literature* 50(4), pp. 1051–1079.

Ejrnæs, M. and C. C. Pörtner (2004). "Birth Order and the Intrahousehold Allocation of Time and Education." *The Review of Economics and Statistics* 86(4), pp. 1008–1019.

Emerson, P. M. and A. P. Souza (2003). "Is There a Child Labor Trap? Intergenerational Persistence of Child Labor in Brazil." *Economic Development and Cultural Change* 51(2), pp. 375–398.

Emran, M. S., W. Greene, and F. Shilpi (2018). "When Measure Matters Coresidency, Truncation Bias, and Intergenerational Mobility in Developing Countries." *Journal of Human Resources* 53(3), pp. 589–607.

Emran, M. S. and F. Shilpi (2011). "Intergenerational Occupational Mobility in Rural Economy Evidence from Nepal and Vietnam." *Journal of Human Resources* 46(2), pp. 427–458.

Emran, M. S. and F. Shilpi (2015). "Gender, geography, and generations: Intergenerational educational mobility in post-reform India." *World Development* 72, pp. 362–380.

Emran, M. S. and F. J. Shilpi (2019). *Economic Approach to Intergenerational Mobility: Measures, Methods, and Challenges in Developing Countries*. WIDER Working Paper (No.No. 2019/98).

Fabrigar, L., D. Wegener, R. MacCallum, and E. Strahan (1999). "Evaluating the Use of Exploratory Factor Analysis in Psychological Research." *Psychological Methods* 4(3), pp. 272–299.

Falk, A., F. Kosse, P. Pinger, H. Schildberg-Hörisch, and T. Deckers (2021). "Socioeconomic status and inequalities in children's IQ and economic preferences." *Journal of Political Economy* 129(9), pp. 2504–2545.

Figlio, D., P. Giuliano, U. Özek, and P. Sapienza (2019). "Long-Term Orientation and Educational Performance." *American Economic Journal: Economic Policy* 11(4), pp. 272–309.

Figlio, D., J. Guryan, K. Karbownik, and J. Roth (2014). "The effects of poor neonatal health on children's cognitive development." *American Economic Review* 104(12), pp. 3921–55.

Gelbach, J. B. (2016). "When Do Covariates Matter? And Which Ones, and How Much?" *Journal of Labor Economics* 34(2), pp. 509–543.

Güell, M., J. V. Rodríguez Mora, and G. Solon (2018). "New Directions in Measuring Intergenerational Mobility: Introduction." *The Economic Journal* 128(612), F335–F339.

Genicot, G. and D. Ray (2017). "Aspirations and Inequality." *Econometrica* 85(2), pp. 489–519.

Goodman, A., I. Koupil, and D. W. Lawson (2012). "Low Fertility Increases Descendant Socioeconomic Position but Reduces Long-Term Fitness in a Modern Post-Industrial Society." *Proceedings of the Royal Society B: Biological Sciences* 279(1746), pp. 4342–4351.

Gorsuch, R. L. (1983). *Factor Analysis*. 2 New edition. Hillsdale, N.J: Taylor & Francis Inc.

Gould, E. D., A. Simhon, and B. A. Weinberg (2020). "Does parental quality matter? Evidence on the transmission of human capital using variation in parental influence from death, divorce, and family size." *Journal of Labor Economics* 38(2), pp. 569–610.

Gupta, P. (2019). *How much it Takes to Educate People in India: Cost of Education revealed in Four Charts*. https://www.financialexpress.com/economy/how-much-it-takes-to-educate-people-in-india-cost-of-education-revealed-in-four-charts/1778220/ (last accessed: 18 May 2022).

Guttman, L. (1954). "Some Necessary Conditions for Common-Factor Analysis." *Psychometrika* 19(2), pp. 149–161.

Haile, G. A. (2018). "Intergenerational Mobility in Socio-Economic Status in Ethiopia." *Journal of International Development* 30(8), pp. 1392–1413.

Heckman, J. J., R. Pinto, and P. Savelyev (2013). "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103(6), pp. 2052–2086.

Hendrickson, A. E. and P. O. White (1964). "Promax: A Quick Method for Rotation to Oblique Simple Structure." *British Journal of Statistical Psychology* 17(1), pp. 65–70.

Hertz, T., T. Jayasundera, P. Piraino, S. Selcuk, N. Smith, and A. Verashchagina (2007). "The Inheritance of Educational Inequality: International Comparisons and Fifty-year Trends." *The BE Journal of Economic Analysis & Policy* 7(2).

Hnatkovska, V., A. Lahiri, and S. B. Paul (2013). "Breaking the Caste Barrier: Intergenerational Mobility in India." *Journal of Human Resources* 48(2), pp. 435–473.

Hoang, L. (2013). *Vietnam: Where Free Education isn't so Free*. https://www.aljazeera.com/features/2013/4/7/vietnam-where-free-education-isnt-so-free (last accessed: 18 May 2022).

Horn, J. L. (1965). "A Rationale and Test for the Number of Factors in Factor Analysis." *Psychometrika* 30(2), pp. 179–185.

Hoxby, C. M. and C. Avery (2013). "The Missing "One-Offs": The Hidden Supply of High-Achieving, Low-Income Students." *Brookings Papers on Economic Activity* 44(1), pp. 1–65.

Hoxby, C. M. and S. Turner (2015). "What High-Achieving Low-Income Students Know about College." *American Economic Review* 105(5), pp. 514–517.

Imai, K., L. Keele, and D. Tingley (2010a). "A General Approach to Causal Mediation Analysis." *Psychological Methods* 15(4), pp. 309–334.

Imai, K., L. Keele, and T. Yamamoto (2010b). "Identification, Inference and Sensitivity Analysis for Causal Mediation Effects." *Statistical Science* 25(1), pp. 51–71.

Jalan, J. and R. Murgai (2008). "Intergenerational Mobility in Education in India." *Indian Statistical Institute, Delhi*.

Jensen, R. (2010). "The (Perceived) Returns to Education and the Demand for Schooling." *The Quarterly Journal of Economics* 125(2), pp. 515–548.

Jones, N. and S. Huttly (2018). *Young Lives: An International Study of Childhood Poverty: Round 1, 2002. [Data Collection]. 6th Edition*.

Kaiser, H. F. (1960). "The Application of Electronic Computers to Factor Analysis." *Educational and Psychological Measurement* 20, pp. 141–151.

— (1961). "A Note on Guttman's Lower Bound for the Number of Common Factors1." *British Journal of Statistical Psychology* 14(1), pp. 1–2.

— (1974). "An Index of Factorial Simplicity." *Psychometrika* 39(1), pp. 31–36.

Kautz, T., J. J. Heckman, R. Diris, B. ter Weel, and L. Borghans (2014). *Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success*. Working Paper (No.20749). National Bureau of Economic Research.

Kazeem, A., L. Jensen, and C. S. Stokes (2010). "School Attendance in Nigeria: Understanding the Impact and Intersection of Gender, Urban-Rural Residence and Socioeconomic Status." *Comparative Education Review* 54(2), pp. 295–319.

Keane, M., S. Krutikova, and T. Neal (2022). "Child work and cognitive development: Results from four low to middle income countries." *Quantitative Economics* 13(2), pp. 425–465.

Kim, J.-O. and C. W. Mueller (1978). *Factor Analysis: Statistical Methods and Practical Issues*. Newbury Park: SAGE.

Kornrich, S. and F. Furstenberg (2013). "Investing in Children: Changes in Parental Spending on Children, 1972–2007." *Demography* 50(1), pp. 1–23.

Kosse, F., T. Deckers, P. Pinger, H. Schildberg-Hörisch, and A. Falk (2020). "The Formation of Prosociality: Causal Evidence on the Role of Social Environment." *Journal of Political Economy* 128(2), pp. 434–467.

Kundu, A. and K. Sen (2023). "Multigenerational mobility among males in India." *Review of Income and Wealth* 69(2), pp. 395–418.

Laajaj, R. and K. Macours (2021). "Measuring skills in developing countries." *Journal of Human Resources* 56(4), pp. 1254–1295.

Lambert, S., M. Ravallion, and D. van de Walle (2014). "Intergenerational Mobility and Interpersonal Inequality in an African Economy." *Journal of Development Economics* 110, pp. 327–344.

Landersø, R. and J. J. Heckman (2017). "The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the US." *Scandinavian Journal of Economics* 119(1), pp. 178–230.

Leone, T. (2018). "The Geography of Intergenerational Mobility: Evidence of Educational Persistence and the "Great Gatsby Curv" in Brazil." *Review of Development Economics*.

Leone, T. (2021). "The Gender Gap in Intergenerational Mobility." *World Development Perspectives* 21, p. 100286.

López Boo, F. (2016). "Socio-economic Status and Early Childhood Cognitive Skills: A Mediation Analysis using the Young Lives Panel." *International Journal of Behavioral Development* 40(6), pp. 500–508.

Maitra, P. and A. Sharma (2009). *Parents and Children: Education across Generations in India*. 5th Annual Conference on Economic Growth and Development, ISI Delhi, Delhi.

Mincer, J. (1974). *Schooling, Experience, and Earnings*. New York: Columbia University Press.

Mohammed, A. R. S. (2019). "Does a Good Father Now Have to Be Rich? Intergenerational Income Mobility in Rural India." *Labour Economics* 60, pp. 99–114.

Munoz, E. (2021). *The Geography of Intergenerational Mobility in Latin America and the Caribbean*. SSRN Working Paper (No.3807350).

Neidhöfer, G., J. Serrano, and L. Gasparini (2018). "Educational inequality and intergenerational mobility in Latin America: A new database." *Journal of Development Economics* 134, pp. 329–349.

Nguyen, M. C. and Q. Wodon (2014). *Estimating the Impact of Child Marriage on Literacy and Education Attainment in Africa*. The World Bank.

Nordman, C. J., L. Sarr, and S. Sharma (2015). *Cognitive, Non-Cognitive Skills and Gender Wage Gaps: Evidence from Linked Employer-Employee Data in Bangladesh*. IZA Discussion Paper (No.9132).

Nutrition Accountability Framework (2021). *Global Nutrition Report*. Retrieved from https://globalnutritionreport.org/resources/nutrition-profiles/.

Pasquier-Doumer, L. and F. Risso Brandon (2015). "Aspiration Failure: A Poverty Trap for Indigenous Children in Peru?" *World Development* 72, pp. 208–223.

Preacher, K. J. and A. F. Hayes (2008). "Asymptotic and Resampling Strategies for Assessing and Comparing Indirect Effects in Multiple Mediator Models." *Behavior Research Methods* 40(3), pp. 879–891.

Psacharopoulos, G. and H. A. Patrinos (2018). "Returns to investment in education: a decennial review of the global literature." *Education Economics* 26(5), pp. 445–458.

Putnick, D. L. and M. H. Bornstein (2015). "Is Child Labor a Barrier to School Enrollment in Low- and Middle-Income Countries?" *International Journal of Educational Development* 41, pp. 112–120.

Raven, J. and J. Raven (2003). "Raven progressive matrices." *Handbook of nonverbal assessment*. Springer, pp. 223–237.

Ray, D. (2016). "Aspirations and the Development Treadmill." *Journal of Human Development and Capabilities* 17(3), pp. 309–323.

Sacerdote, B. (2011). "Chapter 4 - Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" *Handbook of the Economics of Education*. Ed. by E. A. Hanushek, S. Machin, and L. Woessmann. Vol. 3. Elsevier, pp. 249–277.

— (2014). "Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward?" *Annual Review of Economics* 6(1), pp. 253–272.

Sanchez, A., M. Penny, T. Woldehanna, S. Galab, J. Boyden, and L. T. Duc (2018). *Young Lives: An International Study of Childhood Poverty: Round 5, 2016. [Data Collection]*.

Schad, M. (2015). *Intergenerational Income Mobility and Redistributive Policy*. Springer.
Singh, A. (2015). "Private School Effects in Urban and Rural India: Panel Estimates at Primary and Secondary School Ages." *Journal of Development Economics* 113, pp. 16–32.
— (2020). "Learning more with every year: School year productivity and international learning divergence." *Journal of the European Economic Association* 18(4), pp. 1770–1813.
Solon, G. (1999). "Intergenerational Mobility in the Labor Market." *Handbook of Labor Economics*. Ed. by O. C. Ashenfelter and D. Card. Vol. 3. Amsterdam: Elsevier, pp. 1761–1800.
Thompson, B. (2004). *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*. American Psychological Association.
UNICEF (2015a). *Ending Child Marriage: A Profile of Progress in India*.
— (2015b). *Understanding Child Marriage in Viet Nam*.
— (2017). *Ending Child Marriage: A Profile of Progress in Ethiopia*.
— (2018a). *Child Labour | UNICEF*. http://unicef.in/Whatwedo/21/Child-Labour (last accessed: 13 February 2018).
— (2018b). *Impact Evaluation of Alternative Basic Education in Ethiopia*. https://www.unicef.org/ethiopia/reports/impact-evaluation-alternative-basic-education-ethiopia (last accessed: 17 May 2022).
— (2021a). *Child Labour Rises to 160 Million – First Increase in Two Decades*. UNICEF Press Release. https://www.unicef.org/india/press-releases/child- labour-rises-160-million-first-increase-two-decades.
— (2021b). *UNICEF Global Databases*. Retrieved from https://data.unicef.org/topic/child-protection/child-labour/.
United Nations (1989). *OHCHR | Convention on the Rights of the Child*. http://www.ohchr.org/EN/ProfessionalInterest/Pages/CRC.aspx (last accessed: 15 February 2018).
— (2016). *ISCED Mappings*. http://uis.unesco.org/en/isced-mappings (last accessed: 18 March 2017).
UNODC (2017). *World Drug Report 2017*. Retrieved from www.unodc.org/documents/data-and-analysis/dataportal/PublicationReports2017_1a.xlsx.
van der Weide, R., C. Lakner, D. G. Mahler, A. Narayan, and R. Gupta (2024). "Intergenerational mobility around the world: A new database." *Journal of Development Economics* 166. 103167.
Vuri, D. (2008). *The Effect of Availability and Distance to School on Children's Time Allocation in Ghana and Guatemala*. Understanding Children's Work Programme Working Paper.
Wansbeek, T. and E. Meijer (2003). "Measurement Error and Latent Variables." *A Companion to Theoretical Econometrics*. Wiley Online Books, pp. 162–179.
Wantchekon, L., M. Klašnja, and N. Novta (2015). "Education and Human Capital Externalities: Evidence from Colonial Benin." *The Quarterly Journal of Economics* 130(2), pp. 703–757.
WENR (2017). *Education in Vietnam*. https://wenr.wes.org/2017/11/education-in-vietnam (last accessed: 20 May 2022).
Wodon, Q. T. et al. (2017). *Economic Impacts of Child Marriage*. Global Synthesis Report (No.116829). The World Bank.

Woldehanna, T., S. Galab, A. Sanchez, M. Penny, L. T. Duc, and J. Boyden (2018). *Young Lives: An International Study of Childhood Poverty: Round 4, 2013-2014. [Data Collection]. 2nd Edition.*

Woldehanna, T. and A. Gebremedhin (2015). *Is Child Work Detrimental to the Educational Achievement of Children? Results from Young Lives in Ethiopia.* Young Lives.

Woodhead, M., P. Dornan, and H. Murray (2014). "What Inequality Means for Children: Evidence from Young Lives." *International Journal of Children's Rights* 22(3), pp. 467–501.

World Bank (2011). *Perspectives on Poverty in India: Stylized Facts from Survey Data.* World Bank Group.

— (2016). *Basic Profile of Child Marriage in Peru.* Health, Nutrition, and Population (HNP) Knowledge Brief: Child Marriage Series. Washington, D.C. : World Bank Group.

— (2021). *World Development Indicators 2021.* World Bank Group.

— (2023). *Global Database on Intergenerational Mobility.* Retrieved from https://datacatalog.worldbank.org/search/dataset/0050771/Global-Database-on-Intergenerational-Mobility. Washington, D.C.

Young Lives (2017). *A Guide to Young Lives Research.* Retrieved from https://www.younglives.org.uk/ content/guide-young-lives-research.

JENA ECONOMICS RESEARCH PAPERS · # 2025 – 009

ACCOUNTING FOR INTERGENERATIONAL EDUCATIONAL IMMOBILITY IN LOW- AND MIDDLE-INCOME COUNTRIES

Fabian Koenings *Jakob Schwab*

ONLINE APPENDIX

A. Online Appendix A: Summary Statistics

Table A1. *Sampling of Data*

	Ethiopia		India		Peru		Vietnam	
	YL	DHS	YL	DHS	YL	DHS	YL	DHS
Household Characteristics								
Parental Education	3.34	1.77	3.83	5.42	9.10	8.59	7.00	8.04
Household Head's Age	46.70	47.13	42.43	44.67	43.46	45.49	42.12	45.24
Female Household Head	0.25	0.20	0.11	0.12	0.19	0.18	0.13	0.20
Household Size	6.53	6.64	5.19	6.27	5.58	5.59	4.89	5.22
Urban *(Rural)*	0.34	0.10	0.24	0.27	0.78	0.58	0.19	0.16
Girl *(Boy)*	0.47	0.50	0.51	0.48	0.47	0.48	0.51	0.50
Assets in Household								
Radio	0.57	0.34	0.13	0.30	0.70	0.88	0.23	0.41
TV	0.15	0.04	0.47	0.41	0.77	0.68	0.87	0.83
Refrigerator	0.02	0.02	0.06	0.12	0.23	0.30	0.21	0.19
Bicycle	0.04	0.01	0.40	0.59	0.39	0.27	0.88	0.83
Motorcycle/scooter	0.00	0.00	0.12	0.15	0.05	0.04	0.65	0.55
Car	0.01	0.00	0.01	0.02	0.03	0.07	0.01	0.01
Amenities/ Housing								
Drinking Water: Piped into dwelling	0.12	0.00	0.10	0.10	0.78	0.57	0.13	0.18
Floor: Cement	0.13	0.03	0.33	0.30	0.41	0.36	0.47	0.28

Notes: This table shows the mean of some characteristics for the estimation sample from the second observation round of the YL Study and for households with children aged 11-12 from the nationally representative Demographic and Health Surveys-V (2004-6). For comparability with the other tables in the paper, parental education is measured in the first observation round of the YL Study. DHS means are weighted using the survey weights included in the DHS data. With the exception of parental education, which is measured in years, age of household head, and household size, all variables are binary indicators. For binary indicators that are not yes/no indicators, the base category is indicated in parentheses.

Table A2. *Two-Sided t-Tests on Difference between Sample and All Observation Means*

	Mean (All)	Observations (All)	Mean (Sample)	Observations (Sample)	Difference (All - Sample)
Parental Education	5.552	3697	5.592	3161	−0.040
Parental Wealth	0.386	3712	0.389	3154	−0.003
Total Expenditures of HH	8.742	3649	8.752	3161	−0.010
Children's Education	11.311	3488	11.483	3161	−0.171*
Girl *(Boy)*	0.493	3722	0.494	3161	−0.000
Household Head's Age	40.340	3713	40.201	3161	0.139
Urban *(Rural)*	0.358	3722	0.341	3161	0.016

Notes: This table shows the comparison for some variable means between all observations in the dataset and the observations included in the analysis. In Column (5) the difference in means is displayed with the stars indicating the following: * p<0.1, ** p<0.05, *** p<0.01.

Table A3. *Description Factor Measures I*

Measure	Description
Factor 1: Child Labor	
Working Hours	Sum of hours doing household chores, caring for others, the hours spent working in the own farm business and hours spent in a paid activity for a company/person not related to the child on a typical day.
Activity for Money	Did you do anything in the last 12 month for money? (Yes (1) / No (0))
Work Injury	Seriously injured while working in last 4 yrs. (Yes (1) / No (0))
Factor 2: Infrastructure	
Access to Education	Do you have access to education/ schools? (Yes (1)/ No (0))
Type of Living Site	Rural (0)/ Urban (1) living site
Time to School	(Hypothetical) time to get to school in minutes (reversed 'scale')
Factor 3: Education Spending	
Spent on Schooling Fees	Amount spent on payment for schooling fees in last 12 months (in int. dollar cents)
Spent on School Books	Amount spent on school books and stationery in last 12 months (in int. dollar cents)
Private Schooling	Enrolled at private Schooling during second observation period. (Yes (1)/ No (0))
Factor 4: Underage Family	
Child Marriage	Child married while still under age (Yes (1)/ No (0))
Child Parent	Number of children born by child while still under age
Expected Child Marriage	At what age should child get married? Binary variable that is 1 if age is below 18.
Expected Child Parent	At what age should child have a child? Binary variable that is 1 if age is below 18.
Factor 5: Parental Attentiveness	
Without Father	Father not in household/ dead (Yes (0) / No (1))
Number of Children	Number of additional children under age in each household (reversed 'scale')
Attention received	I receive lots of time and attention from my parents. (4 point scale: "Strongly agree" (4), "Agree" (3), "Disagree" (2), "Strongly disagree" (0)) *
Love received	I always feel loved by my parents. (4 point scale: "Strongly agree" (4), "Agree" (3), "Disagree" (2), "Strongly disagree" (0)) *
Conversation	My parents rarely talk to me about the things that matter to me. (4 point scale: "Strongly agree" (0), "Agree" (2), "Disagree" (3), "Strongly disagree" (4)) *
Supported by Parents	My parents never support me in the things i want to do. (4 point scale: "Strongly agree" (0), "Agree" (2), "Disagree" (3), "Strongly disagree" (4)) *
Free Speech	I usually feel able to speak my views and feelings with my parents. (4 point scale: "Strongly agree" (4), "Agree" (3), "Disagree" (2), "Strongly disagree" (0)) *
	* The data for Peru is measured on a 3 point scale. It is adjusted to fit the other countries according to "Yes" (4/1), "More or less" (2.5), "No" (1/4).
Factor 6: Social Environment	
Safe Area	Is the area you live in safe for children? (Yes (1) / No (0))
Friends with Alcohol	How many of your best friends drink alcohol at least once a month? (4 point scale: "All of my friends" (1); "Most of my friends" (2); "A few of my friends" (3); "None of my friends" (4))
Friends with Trouble	Have other young people - tried to get you into trouble with your friends? (4 point scale: "4 or more times" (1); "2-3 times" (2); "Once" (3); "Never" (4))
Friends who Smoke	How many of your friends smoke cigarettes at least once a month? (4 point scale: "All of my friends" (1); "Most of my friends" (2); "A few of my friends" (3); "None of my friends" (4))
Shock-theft	Shock-theft/destruction of housing/consumer goods (Yes (0) / No (1))
Trust in Neighborhood	I feel I can trust my neighbours to look after my house. (3 point scale: "No" (1); "More or less" (2); "Yes" (3))
Safe to go Outside	I think it is safe for my child to go out on the street on his/her own. (3 point scale: "No" (1); "More or less" (2); "Yes" (3))

Notes: This table shows the measures from the survey data by the first six factors that they are associated to as presented in Section 2. Column (2) describes the coding of the respective measures from the survey question and how they enter the factor analysis.

Table A4. *Description Factor Measures II*

Measure	Description
Factor 7: Health	
Weight-for-age	Weight-for-age z-score
Height-for-age	Height-for-age z-score
Thinness Indicator	Low BMI for age (3 point scale: "Severely thin" (0), "Moderately thin" (1), "Not thin" (2))
Factor 8: Aspirations	
Job Aspirations Child	What do you want to be when you grow up? (Occupations coded to a scale of ambition from 1-4 following Pasquier-Doumer and Risso Brandon 2015)
School Aspirations Child	Self-reported grade that children want to complete when finishing school (ISCED classification).
School Aspirations Parents	What level of education would you like child to complete? (ISCED classification)
Factor 9: Cognitive Ability	
PPVT Test Score	Score (0 to 228) of Picture Picture Peabody Vocabulary Test (PPVT) corrected for different mother languages
Math Score	Score of math test (0 to 9) corrected for different countries
Reading Level	Child's reading level (4 point scale)
Writing Level	Child's writing level (3 point scale)
Literacy	Child can read and write a sentence without difficulty (Yes (1)/No (0))
Factor 10: Non-Cognitive Ability: Grit	
Grit 1	New ideas and projects sometimes distract me from previous ones.* (reversed scale)
Grit 2	Setbacks do not discourage me.*
Grit 3	I have been obsessed with a certain idea or project for a short time but later lost interest.* (reversed scale)
Grit 4	I am a hard worker.*
Grit 5	I often set a goal but later choose to pursue a different one.* (reversed scale)
Grit 6	I have difficulty maintaining my focus on projects that take more than a few month to complete.* (reversed scale)
Grit 7	I finish whatever I begin.*
Grit 8	I am diligent.*
	* All GRIT-measures are based on 5 point scale ranging from from "Not like me at all" (1) to "Very much like me" (5)

Notes: This table shows the measures from the survey data by the last four factors that they are associated to as presented in Section 2. Column (2) describes the coding of the respective measures from the survey question and how they enter the factor analysis.

Table A5. *Descriptive Statistics I*

	Mean	SD	Min.	Max.	Observ. Round
Ethiopia ($N = 752$)					
Parental Education	3.34	4.04	0.00	16.00	1
Children's Education	9.61	4.12	0.00	18.00	5
Controls					
Girl *(Boy)*	0.47	0.50	0.00	1.00	1
Household Head's Age	42.63	10.99	18.00	85.00	1
Household Head's Age^2	1938.03	1025.90	324.00	7225.00	1
Birth Order *(First Child)*	3.62	1.85	1.00	6.00	1
Pathways					
Child Labor	0.17	0.83	-0.96	5.02	2
Infrastructure	-0.27	1.09	-4.64	1.58	2
Education Spending	-0.29	0.38	-0.48	6.06	2
Child Engagement	-0.11	0.76	-0.30	7.81	2/4
Parent Attentiveness	-0.45	0.98	-5.03	1.61	1/2
Social Environment	0.24	0.76	-3.20	1.13	1/2/3
Health	-0.26	1.06	-3.85	2.93	1/3
Aspirations	0.03	0.91	-2.76	1.17	2
Cognitive Ability	-0.55	1.18	-3.13	1.09	2
Non-Cognitive Ability: Grit	0.01	0.99	-3.99	3.05	5
India ($N = 921$)					
Parental Education	3.83	4.71	0.00	16.00	1
Children's Education	11.92	4.06	0.00	18.00	5
Controls					
Girl *(Boy)*	0.51	0.50	0.00	1.00	1
Household Head's Age	40.18	10.84	21.00	90.00	1
Household Head's Age^2	1731.78	1050.52	441.00	8100.00	1
Birth Order *(First Child)*	2.45	1.39	1.00	6.00	1
Pathways					
Child Labor	-0.14	1.15	-0.96	4.41	2
Infrastructure	-0.11	0.97	-7.10	1.67	2
Education Spending	0.33	1.47	-0.48	15.14	2
Child Engagement	0.22	1.39	-0.30	12.46	2/4
Parent Attentiveness	0.37	0.88	-3.34	1.69	1/2
Social Environment	0.54	0.74	-4.01	1.13	1/2/3
Health	-0.23	0.95	-3.14	3.46	1/3
Aspirations	-0.34	1.15	-3.43	1.17	2
Cognitive Ability	-0.10	1.06	-3.29	0.86	2
Non-Cognitive Ability: Grit	-0.00	0.99	-3.82	3.23	5
Peru ($N = 582$)					
Parental Education	9.10	4.46	0.00	16.00	1
Children's Education	12.97	3.74	0.00	16.00	5
Controls					
Girl *(Boy)*	0.47	0.50	0.00	1.00	1
Household Head's Age	39.00	10.28	21.00	87.00	1
Household Head's Age^2	1626.33	945.06	441.00	7569.00	1
Birth Order *(First Child)*	2.69	1.68	1.00	6.00	1
Pathways					
Child Labor	0.56	1.09	-0.96	4.21	2
Infrastructure	0.77	0.83	-2.17	1.62	2
Education Spending	0.12	1.16	-0.48	19.04	2
Child Engagement	-0.06	0.79	-0.30	7.81	2/4
Parent Attentiveness	-0.06	1.16	-4.63	1.81	1/2
Social Environment	-0.92	1.08	-4.18	1.13	1/2/3
Health	0.76	0.79	-2.11	3.22	1/3
Aspirations	0.42	0.64	-3.43	1.17	2
Cognitive Ability	0.27	0.75	-3.44	0.92	2
Non-Cognitive Ability: Grit	0.00	1.07	-3.64	3.15	5

Notes: This table shows the descriptive statistics for children's and parents' education levels, control factors, and pathway factors, separately by country and for all countries combined. For categorical variables, the base category is displayed in parentheses.

Table A6. *Descriptive Statistics II*

	Mean	SD	Min.	Max.	Observ. Round
Vietnam ($N = 906$)					
Parental Education	7.00	4.02	0.00	16.00	1
Children's Education	11.63	3.71	0.00	18.00	5
Controls					
Girl *(Boy)*	0.51	0.50	0.00	1.00	1
Household Head's Age	38.98	10.29	23.00	91.00	1
Household Head's Age^2	1625.15	1025.40	529.00	8281.00	1
Birth Order *(First Child)*	2.23	1.36	1.00	6.00	1
Pathways					
Child Labor	-0.35	0.66	-0.96	4.41	2
Infrastructure	-0.16	0.80	-2.87	1.58	2
Education Spending	-0.18	0.28	-0.48	2.72	2
Child Engagement	-0.09	0.75	-0.30	8.13	2/4
Parent Attentiveness	0.04	0.85	-3.59	1.55	1/2
Social Environment	-0.16	0.89	-4.31	1.13	1/2/3
Health	-0.04	0.88	-2.59	3.04	1/3
Aspirations	0.05	0.99	-3.43	1.17	2
Cognitive Ability	0.38	0.62	-3.48	0.75	2
Non-Cognitive Ability: Grit	-0.00	0.98	-4.12	3.06	5
All Countries ($N = 3161$)					
Parental Education	5.59	4.85	0.00	16.00	1
Children's Education	11.48	4.08	0.00	18.00	5
Controls					
Girl *(Boy)*	0.49	0.50	0.00	1.00	1
Household Head's Age	40.20	10.71	18.00	91.00	1
Household Head's Age^2	1730.87	1025.77	324.00	8281.00	1
Birth Order *(First Child)*	2.71	1.65	1.00	6.00	1
Pathways					
Child Labor	0.00	1.00	-0.96	5.02	2
Infrastructure	0.00	1.00	-7.10	1.67	2
Education Spending	0.00	1.00	-0.48	19.04	2
Child Engagement	0.00	1.00	-0.30	12.46	2/4
Parent Attentiveness	0.00	1.00	-5.03	1.81	1/2
Social Environment	0.00	1.00	-4.31	1.13	1/2/3
Health	0.00	1.00	-3.85	3.46	1/3
Aspirations	0.00	1.00	-3.43	1.17	2
Cognitive Ability	0.00	1.00	-3.48	1.09	2
Non-Cognitive Ability: Grit	0.00	1.00	-4.12	3.23	5

Notes: This table shows the descriptive statistics for children's and parents' education levels, control factors, and pathway factors, separately by country and for all countries combined. For categorical variables, the base category is displayed in parentheses.

Table A7. *Sample Restrictions*

	Ethiopia		India		Peru		Vietnam		All	
	Obs.	Missings	Obs.	Missings	Obs.	Missings	Obs.	Missings	Obs.	Missings
Original YL sample size	1000		1008		714		1000		3722	
SES variables										
Parental Education	1000	0	1005	3	698	16	994	6	3697	25
Children's Education	911	89	956	52	656	58	965	35	3488	234
Pathway variables										
Child Labor	979	21	994	14	684	30	990	10	3647	75
Infrastructure	980	20	994	14	685	29	990	10	3649	73
Education Spending	1000	0	1008	0	714	0	1000	0	3722	0
Child Engagement	1000	0	1008	0	714	0	1000	0	3722	0
Parent Attentiveness	999	1	1008	0	714	0	1000	0	3721	1
Social Environment	1000	0	1007	1	714	0	1000	0	3721	1
Health	965	35	1008	0	712	2	1000	0	3685	37
Aspirations	979	21	994	14	685	29	990	10	3648	74
Cognitive Ability	978	22	994	14	684	30	990	10	3646	76
Non-Cognitive: Grit	813	187	921	87	596	118	910	90	3240	482
Control variables										
Gender	1000	0	1008	0	714	0	1000	0	3722	0
Household Head's Age	996	4	1008	0	709	5	1000	0	3713	9
Birth Order	994	6	1008	0	709	5	1000	0	3711	11
Number of observations dropped	248		87		132		94		561	
Final sample size	752		921		582		906		3161	

Notes: This table shows number of observations for each variable that is used in the estimations in the first row for each country and all countries combined. For the pathway factor variables, an observation was included as long as it had at least one measure of a factor available. If information on all (three to eight) measures from one respective factor was missing, the factor variable was assigned missing. The second column for each country or all countries combined shows the missing number of observations for each variable in comparison the orignal YL sample size. The resulting overall number of observations dropped (if at least one of the variables was missing) are reported in the second to last row, and the resulting sample sizes in each country sample are listed in the last row.

B. Online Appendix B: Decomposition Approach

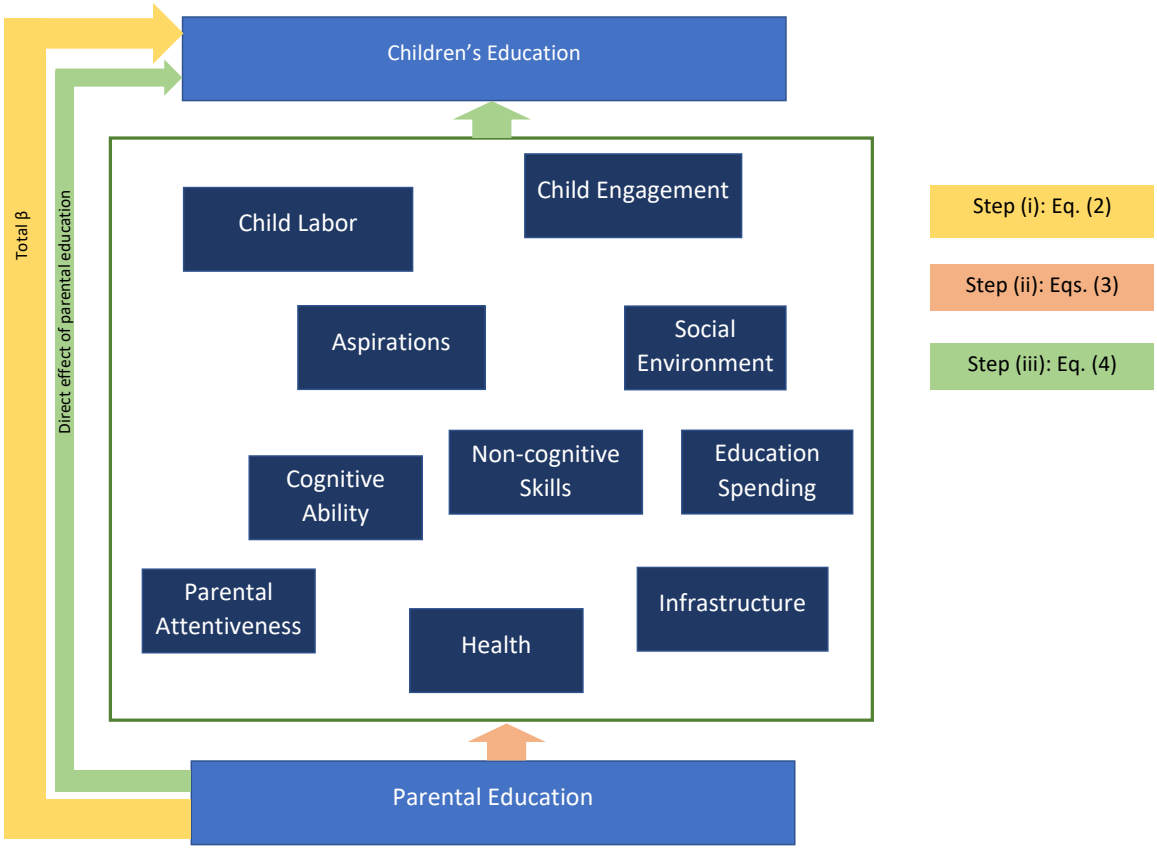

Fig. B1. *Graphical Representation of Decomposition Approach*

Notes: This figure shows the approach to the decomposition approach. The overall effect of the parental background is mediated by the different factors. Arrows pointing to or from a box imply that the relationship holds for all variables within the box. The details of the decomposition approach are presented in Section 3, the derivation of the pathway factor variables from the measures listed in Table 2 is described in Section 2.

C. Online Appendix C: Empirical Strategy for the Sample of all Countries

For the estimations in the sample of all countries combined, we include country dummies in the estimations. The equations (2) to (5) are thus replaced by the following set of equations:

$$Ed_h^C = \beta Ed_h^P + {X_h}'\zeta_1 + \sum_k \alpha_{k,1} d_k + \epsilon_{h,1} \tag{6}$$

$$PW_{i,h} = \lambda_i Ed_h^P + {X_h}'\zeta_{i,2} + \sum_k \alpha_{k,i,2} d_k + \epsilon_{i,h,2} \tag{7}$$

$$Ed_h^C = \sum_i \rho_i PW_{i,h} + \gamma_{Ed^P} Ed_h^P + {X_h}'\zeta_3 + \sum_k \alpha_{k,3} d_k + \epsilon_{h,3} \tag{8}$$

$$\begin{aligned} Ed_h^C =& (\gamma_{Ed^P} + \sum_i \rho_i \lambda_i) Ed_h^P + {X_h}'(\zeta_3 + \sum_i \rho_i \zeta_{i,2}) \\ &+ \sum_k (\alpha_{k,3} + \sum_i \rho_i \alpha_{k,2}) d_k + \sum_i \rho_i \epsilon_{i,h,2} + \epsilon_{h,3} \end{aligned} \tag{9}$$

where the country-specific constants are replaced by the country dummies d_k in order to capture country k-wide effects, such as the overall level of schooling, so that we capture only within-country variation of educational outcomes in our estimations.

D. Online Appendix D: Parental Education on Pathways

In order to play a role in the transmission of socioeconomic status, a pathway factor has to depend on parental background. To find out whether this is the case for the pathways considered, we estimate Equation (3) for each pathway $i \in \mathcal{F}$. We control for the same set of variables X_h (and country fixed effects in the full sample). To facilitate interpretation, we use the plain (non-standardized) number of years of parents' schooling as explanatory variable, analogous to the estimation of the IGRC.[43]

The results of estimating Equation (3) are shown in Tables D1-D3. The coefficients of interest from each estimation are also displayed comprehensively in Table 4 in the main text. Almost each of the hypothesized pathway factors (except for the social environment) is significantly correlated with parental education in all or at least three of the four countries. This also rationalizes our choice of pathways to analyze. Because the pathway variables are standardized, the relative size of the estimated coefficients shows which pathway factors are more strongly correlated with parental schooling.

[43] Using the standardized number of years only re-scales the estimates.

Table D1. *Correlation between Parental Education and Pathways I*

Ethiopia

Panel A Dependent Variable	Child Labor	Infra-structure	Education Spending	Underage Family	Parental Attentiveness
Parental Education	−0.038***	0.091***	0.029***	−0.025***	0.014
	(0.007)	(0.009)	(0.006)	(0.006)	(0.010)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	752	752	752	752	752
Adj. R^2	0.081	0.155	0.420	0.085	0.174

Panel B Dependent Variable	Social Environment	Health	Aspirations	Cognitive Ability	Non-Cognitive: Grit
Parental Education	−0.004	0.037***	0.059***	0.046***	−0.002
	(0.007)	(0.010)	(0.007)	(0.011)	(0.009)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	752	752	752	752	752
Adj. R^2	0.092	0.082	0.065	0.205	0.011

India

Panel A Dependent Variable	Child Labor	Infra-structure	Education Spending	Underage Family	Parental Attentiveness
Parental Education	−0.059***	0.062***	0.136***	−0.051***	0.026***
	(0.007)	(0.007)	(0.013)	(0.008)	(0.006)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	921	921	921	921	921
Adj. R^2	0.090	0.107	0.230	0.142	0.189

Panel B Dependent Variable	Social Environment	Health	Aspirations	Cognitive Ability	Non-Cognitive: Grit
Parental Education	0.012**	0.034***	0.075***	0.065***	0.026***
	(0.005)	(0.007)	(0.007)	(0.006)	(0.007)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	921	921	921	921	921
Adj. R^2	0.364	0.081	0.207	0.109	0.011

Notes: This table shows the results from estimating the relationship between parental education and different pathway factors based on child characteristics as dependent variables, as given by Equation (3). The regression is conducted for each country individually and all countries combined. Parental education is the years of schooling completed by the parents. Pathway factor variables are standardized, such that the coefficients display the increase of one standard deviation of each pathway factor that is associated with one year more of Ed^P. For detailed information on the pathway factors, see Table 2. Robust standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

Table D2. *Correlation between Parental Education and Pathways II*

Peru

Panel A Dependent Variable	Child Labor	Infra-structure	Education Spending	Underage Family	Parental Attentiveness
Parental Education	−0.062***	0.052***	0.055***	−0.011	0.067***
	(0.011)	(0.008)	(0.012)	(0.007)	(0.012)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	582	582	582	582	582
Adj. R^2	0.255	0.506	0.076	0.059	0.058

Panel B Dependent Variable	Social Environment	Health	Aspirations	Cognitive Ability	Non-Cognitive: Grit
Parental Education	−0.034***	0.034***	0.025***	0.046***	0.011
	(0.011)	(0.008)	(0.007)	(0.009)	(0.011)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	582	582	582	582	582
Adj. R^2	0.432	0.521	0.332	0.217	0.014

Vietnam

Panel A Dependent Variable	Child Labor	Infra-structure	Education Spending	Underage Family	Parental Attentiveness
Parental Education	−0.021***	0.040***	0.016***	−0.043***	0.036***
	(0.006)	(0.007)	(0.002)	(0.008)	(0.008)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	906	906	906	906	906
Adj. R^2	0.253	0.156	0.340	0.121	0.025

Panel B Dependent Variable	Social Environment	Health	Aspirations	Cognitive Ability	Non-Cognitive: Grit
Parental Education	0.002	0.038***	0.093***	0.047***	0.021**
	(0.008)	(0.008)	(0.009)	(0.007)	(0.009)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	906	906	906	906	906
Adj. R^2	0.040	0.071	0.211	0.373	0.004

Notes: This table shows the results from estimating the relationship between parental education and different pathway factors based on child characteristics as dependent variables, as given by Equation (3). The regression is conducted for each country individually and all countries combined. Parental education is the years of schooling completed by the parents. Pathway factor variables are standardized, such that the coefficients display the increase of one standard deviation of each pathway factor that is associated with one year more of Ed^P. For detailed information on the pathway factors, see Table 2. Robust standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

Table D3. *Correlation between Parental Education and Pathways III*

	All				
Panel A					
Dependent Variable	Child Labor	Infra-structure	Education Spending	Underage Family	Parental Attentiveness
Parental Education	−0.045***	0.063***	0.070***	−0.037***	0.033***
	(0.004)	(0.004)	(0.006)	(0.004)	(0.004)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	3161	3161	3161	3161	3161
Adj. R^2	0.153	0.215	0.153	0.107	0.112
Panel B					
Dependent Variable	Social Environment	Health	Aspirations	Cognitive Ability	Non-Cognitive: Grit
Parental Education	−0.002	0.037***	0.069***	0.052***	0.016***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	3161	3161	3161	3161	3161
Adj. R^2	0.265	0.175	0.160	0.197	0.003

Notes: This table shows the results from estimating the relationship between parental education and different pathway factors based on child characteristics as dependent variables, as given by Equation (3). The regression is conducted for each country individually and all countries combined. Parental education is the years of schooling completed by the parents. Pathway factor variables are standardized, such that the coefficients display the increase of one standard deviation of each pathway factor that is associated with one year more of Ed^P. For detailed information on the pathway factors, see Table 2. Robust standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

E. Online Appendix E: Pathways on Children's Education

To see whether the potential pathway factors that are all at least in one country (and typically in more) dependent on parental background are also related to children's educational outcomes, we estimate Equation (4) for each country, and all countries combined. We include all respective pathway factors at once, while continuing to control for the same control variables, X_h, as well as country fixed effects in the combined sample. In line with the previous estimations, we use the non-standardized number of years of children's schooling as dependent variable to facilitate interpretation.

Table E1 shows the results from estimating the effect of the pathways on the children's later educational outcomes, as given by Equation (4). The estimated coefficient for parental education, that is, the direct relation between parental education and children's education that is not captured by the considered pathways, is given in the first line. The effects of the pathway variables are shown below. The point estimates for the pathway variables show the effect of one standard deviation of the variable on the expected years of schooling, and are thus also comparable in size. The coefficients of interest are also displayed comprehensively in Table 4 in the main text.

The results show that differences in cognitive ability in all four countries significantly influence outcomes in schooling. Aspirations do so in India, Peru, and Vietnam particularly. Underage marriage and/or parenthood also negatively affects educational achievements in all countries, as does child labor. School infrastructure, children's health status, parent attentiveness, and children's health status only affect schooling outcomes in Ethiopia (parent attentiveness also in India). Non-cognitive skills have a statistically significant effect in Peru and Ethiopia, but it is relatively small compared to other factors.

Since almost all pathway factors are related to parental background in most countries, multicollinearity may be a concern for the results presented here. However, the correlation of the pathway factor variables is not large. Table E2 shows the correlation coefficients (for the entire sample), none of which exceeds 0.31. We furthermore compute the variance inflation factors (VIFs) for the above estimations (Table E3). None of the factors exhibit problematic error correlation with the others in any of the country-level estimations.

Table E1. *Effect of Pathways on Children's Education*

DV: Children's Education	Ethiopia	India	Peru	Vietnam	All
Parental Education	0.177***	0.003	0.172***	0.187***	0.116***
	(0.032)	(0.024)	(0.035)	(0.027)	(0.014)
Pathways					
Child Labor	−0.538***	−0.666***	−0.293**	−0.403***	−0.468***
	(0.170)	(0.108)	(0.124)	(0.152)	(0.067)
Infrastructure	0.260**	0.105	0.016	0.117	0.164***
	(0.122)	(0.100)	(0.168)	(0.122)	(0.062)
Education Spending	0.153	0.229***	0.110	1.024***	0.120**
	(0.247)	(0.072)	(0.096)	(0.300)	(0.056)
Child Engagement	−0.877***	−0.496***	−0.991***	−0.863***	−0.725***
	(0.206)	(0.082)	(0.265)	(0.111)	(0.071)
Parent Attentiveness	0.305***	0.338***	0.158	0.133	0.249***
	(0.112)	(0.116)	(0.113)	(0.110)	(0.057)
Social Environment	0.282**	0.057	0.074	0.139	0.155**
	(0.137)	(0.136)	(0.131)	(0.100)	(0.062)
Health	0.243**	−0.028	0.115	0.136	0.116**
	(0.112)	(0.105)	(0.164)	(0.109)	(0.059)
Aspirations	0.272**	0.862***	1.057***	1.018***	0.843***
	(0.137)	(0.097)	(0.239)	(0.119)	(0.065)
Cognitive Ability	1.380***	1.106***	0.974***	0.982***	1.211***
	(0.108)	(0.106)	(0.258)	(0.167)	(0.067)
Non-Cognitive Ability: Grit	0.215*	−0.038	0.390***	0.023	0.107**
	(0.112)	(0.098)	(0.122)	(0.099)	(0.054)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	752	921	582	906	3161
Adj. R^2	0.907	0.945	0.946	0.950	0.938

Notes: This table shows the results from estimating Equation (4). The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. Pathway factor variables are standardized, such that the coefficients display the increase of Ed^C that is associated with one standard deviation of each pathway factor. For detailed information on the pathway factors, see Table 2. Robust standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

Table E2. *Correlation of Pathway Factors*

	Child Labor	Infra-structure	Education Spending	Underage Family	Parent Attentiven.	Social Environm.	Health	Aspirat.	Cognitive Ability	Non-Cogn.: Grit
Child Labor	1.000									
Infrastructure	−0.076	1.000								
Education Spending	−0.153	0.174	1.000							
Child Engagement	0.162	−0.089	−0.067	1.000						
Parent Attentiveness	−0.177	0.054	0.140	−0.015	1.000					
Social Environment	−0.127	−0.187	0.029	0.029	0.080	1.000				
Health	0.029	0.285	0.112	0.027	0.059	−0.239	1.000			
Aspirations	−0.137	0.220	0.133	−0.202	0.099	−0.109	0.178	1.000		
Cognitive Ability	−0.279	0.239	0.155	−0.123	0.189	−0.113	0.206	0.313	1.000	
Non-Cognitive: Grit	−0.027	0.009	0.054	−0.031	0.030	0.042	0.056	0.039	0.041	1.000

Notes: This table shows the Pearson correlations of the factor scores for the pathway factors analysed in Section 2. For more detailed information on the pathway factors, see Table 2.

Table E3. *Effect of Pathways on Children's Education - VIFs*

DV: Children's Education	Ethiopia		India		Peru		Vietnam		All	
	Regression	VIF	Regression	VIF	Regression	VIF	Regression	VIF	Regression	VIF
Parental Education	0.177***	2.330	0.003	2.530	0.172***	7.910	0.187***	6.174	0.116***	4.168
	(0.032)		(0.024)		(0.035)		(0.027)		(0.014)	
Pathways										
Child Labor	-0.538***	1.240	-0.666***	1.427	-0.293**	1.410	-0.403***	1.454	-0.468***	1.317
	(0.170)		(0.108)		(0.124)		(0.152)		(0.067)	
Infrastructure	0.260**	1.314	0.105	1.188	0.016	2.275	0.117	1.284	0.164***	1.343
	(0.122)		(0.100)		(0.168)		(0.122)		(0.062)	
Education Spending	0.153	1.787	0.229***	1.379	0.110	1.126	1.024***	1.571	0.120**	1.216
	(0.247)		(0.072)		(0.096)		(0.300)		(0.056)	
Child Engagement	-0.877***	1.135	-0.496***	1.274	-0.991***	1.130	-0.863***	1.251	-0.725***	1.179
	(0.206)		(0.082)		(0.265)		(0.111)		(0.071)	
Parent Attentiveness	0.305***	1.273	0.338***	1.308	0.158	1.196	0.133	1.086	0.249***	1.167
	(0.112)		(0.116)		(0.113)		(0.110)		(0.057)	
Social Environment	0.282**	1.157	0.057	1.164	0.074	2.311	0.139	1.137	0.155**	1.271
	(0.137)		(0.136)		(0.131)		(0.100)		(0.062)	
Health	0.243**	1.163	-0.028	1.442	0.115	1.742	0.136	1.477	0.116**	1.313
	(0.112)		(0.105)		(0.164)		(0.109)		(0.059)	
Aspirations	0.272**	1.497	0.862***	1.375	1.057***	1.517	1.018***	1.902	0.843***	1.424
	(0.137)		(0.097)		(0.239)		(0.119)		(0.065)	
Cognitive Ability	1.380***	1.032	1.106***	1.037	0.974***	1.088	0.982***	1.022	1.211***	1.016
	(0.108)		(0.106)		(0.258)		(0.167)		(0.067)	
Non-Cognitive Ability: Grit	0.215*	2.084	-0.038	2.440	0.390***	2.043	0.023	2.278	0.107**	2.156
	(0.112)		(0.098)		(0.122)		(0.099)		(0.054)	
Controls	Yes		Yes		Yes		Yes		Yes	
Observations	752		921		582		906		3161	
Adj. R^2	0.907		0.945		0.946		0.950		0.938	

Notes: The first column of this table shows the results from estimating Equation (4). The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. Pathway factor variables are standardized, such that the coefficients display the increase of Ed^C that is associated with one standard deviation of each pathway factor. For detailed information on the pathway factors, see Table 2. The second column presents the corresponding variance inflation factors (VIF) to the estimation in the first column. Robust standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

Furthermore, multicollinearity would imply that leaving out specific pathway factors changed the results on the other pathway variables. Table E4 reports the results of estimating Equation (4) while excluding one pathway factor at a time in Columns (2) through (11). We report this exercise here for the full sample across all four countries. Column (1) displays the results shown in Table E1 for all countries for comparison. The results for the pathways only marginally change when consecutively excluding one pathway at a time, and this similarity of estimates also holds for the individual country samples. All these exercises confirm that multicollinearity is of no concern with the data at hand.

Table E4. *Effect of Pathways on Children's Education (excluding pathway factors one by one)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Parental Education	0.116***	0.122***	0.123***	0.122***	0.130***	0.120***	0.115***	0.118***	0.151***	0.137***	0.117***
	(0.014)	(0.015)	(0.014)	(0.014)	(0.015)	(0.014)	(0.015)	(0.014)	(0.015)	(0.015)	(0.014)
Pathways											
Child Labor	−0.468***		−0.483***	−0.477***	−0.562***	−0.489***	−0.474***	−0.466***	−0.522***	−0.672***	−0.468***
	(0.067)		(0.066)	(0.066)	(0.067)	(0.067)	(0.067)	(0.067)	(0.070)	(0.069)	(0.067)
Infrastructure	0.164***	0.207***		0.169***	0.179***	0.157**	0.159**	0.177***	0.196***	0.268***	0.161***
	(0.062)	(0.062)		(0.062)	(0.064)	(0.062)	(0.062)	(0.062)	(0.063)	(0.067)	(0.062)
Education Spending	0.120**	0.156***	0.127**		0.145**	0.124**	0.122**	0.125**	0.160***	0.155***	0.124**
	(0.056)	(0.057)	(0.057)		(0.057)	(0.056)	(0.056)	(0.056)	(0.060)	(0.060)	(0.056)
Child Engagement	−0.725***	−0.779***	−0.728***	−0.728***		−0.725***	−0.728***	−0.716***	−0.805***	−0.770***	−0.726***
	(0.071)	(0.070)	(0.072)	(0.071)		(0.071)	(0.071)	(0.070)	(0.074)	(0.076)	(0.071)
Parent Attentiveness	0.249***	0.284***	0.246***	0.251***	0.251***		0.256***	0.253***	0.307***	0.328***	0.251***
	(0.057)	(0.057)	(0.057)	(0.057)	(0.058)		(0.056)	(0.057)	(0.058)	(0.061)	(0.057)
Social Environment	0.155**	0.176***	0.150**	0.157**	0.173***	0.168***		0.146**	0.176***	0.138**	0.162***
	(0.062)	(0.063)	(0.062)	(0.062)	(0.063)	(0.062)		(0.062)	(0.064)	(0.065)	(0.062)
Health	0.116**	0.108*	0.134**	0.121**	0.057	0.124**	0.104*		0.144**	0.204***	0.122**
	(0.059)	(0.060)	(0.059)	(0.059)	(0.059)	(0.059)	(0.059)		(0.061)	(0.063)	(0.059)
Aspirations	0.843***	0.873***	0.849***	0.849***	0.920***	0.862***	0.847***	0.847***		1.073***	0.844***
	(0.065)	(0.065)	(0.064)	(0.065)	(0.065)	(0.065)	(0.065)	(0.065)		(0.066)	(0.065)
Cognitive Ability	1.211***	1.296***	1.226***	1.215***	1.244***	1.230***	1.209***	1.220***	1.384***		1.213***
	(0.067)	(0.066)	(0.068)	(0.067)	(0.068)	(0.068)	(0.067)	(0.067)	(0.066)		(0.067)
Non-Cognitive: Grit	0.107**	0.108**	0.104*	0.110**	0.114**	0.110**	0.114**	0.113**	0.111**	0.127**	
	(0.054)	(0.054)	(0.054)	(0.054)	(0.055)	(0.054)	(0.054)	(0.054)	(0.055)	(0.058)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3161	3161	3161	3161	3161	3161	3161	3161	3161	3161	3161
Adj. R^2	0.938	0.937	0.938	0.938	0.935	0.937	0.938	0.938	0.934	0.931	0.938

Notes: The first column of this table shows the results from estimating Equation (4). The subsequent columns show the results from estimating Equation (4) while dropping one pathway in each column. The analyses include all countries. The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. The pathway factors mainly represent characteristics of the children between ages 11 and 18, and are standardized, such that the coefficients display the increase of Ed^C that is associated with one standard deviation of each pathway factor. For detailed information on the pathway factors, see Table 2. Robust standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

F. Online Appendix F: Alternative Measures of Cognitive Ability

The measures of cognitive ability that were used in the main text include measures of math, writing, and reading ability at age 11/12. Although these are often even the only measures of cognitive ability used in the literature on drivers of intergenerational immobility (Blanden et al. 2007; Bolt et al. 2021), this may be problematic: Given that acquired math, reading, and writing abilities are a part of schooling success up to that point, they may actually capture parts of the outcome variable, as later schooling success is likely to be correlated with success in school in younger years. However, the YL data allows for two ways of looking into this in more detail. In both cases, we test whether the result holds if other measures of cognitive ability are used. In doing so, it can be excluded that the acquired math, reading, and writing abilities are what drives the central result of the outstanding importance of cognitive ability as a mediator in the process of the intergenerational transmission of schooling outcomes.

First, we use only the PPVT measure as the pathway variable for cognitive skills. By doing so, we do not exploit the opportunities of factor analysis when generating the pathway variables, but at the same time exclude the common variation with skills acquired at school. As in the main analysis, the PPVT test scores are de-meaned by the respective language and then standardized. The results of this approach are depicted in Table F1.

The results remain basically unchanged compared to those reported in the main text. The importance of the factor minimally declines in all four countries. However, it is unclear whether this is due to an actually smaller role of the factor or due to the presence of measurement error. The largely unchanged result is somewhat surprising because the PPVT test score is the measure that loads the weakest (although still significantly) on the factor of cognitive ability as constructed in the main text (see Table 2).

Besides the PPVT score, the YL data also includes another measure of cognitive ability: The Raven's Progressive Matrices (RPM) test (Raven and Raven 2003), which was conducted when the children were 7 or 8 years old. The RPM test aims to measure fluid intelligence – that is, the ability to process information and to reason, based only minimally on prior learning – and can already be well conducted with children. It would, thus, constitute a good independent measure of cognitive ability. The disadvantage of this measure is that the RPM test was not conducted with all children in the sample and to different extents in the four countries. Therefore, when we use the RPM test scores as the pathway variable for cognitive ability, the sample size reduces somewhat in India and Peru, and greatly in Ethiopia and Vietnam, to 174 and 186 children, respectively. Hence, the findings for these two countries are not very authoritative.[44] We nonetheless report the full results of this exercise in Table F2.[45]

[44] Also bootstrapping of standard errors becomes problematic in the two smaller samples but still runs with around 950 (Ethiopia) and 980 (Vietnam) replications.

[45] It is worth noting that in the estimation results in Table F2, the explained component of $\hat{\beta}$ exceeds 100 % in Ethiopia and India. In the main text, this component is already large in India, and in the results reported in Table F1 for India, it is exactly 100 %. The explained component can be close to or exceed 1, when the direct effect is small or negative, that is, if given all pathway variables, the parental background has no additional or even a negative effect on the children's outcome.

Table F1. *Decompositions Using PPVT as Cognitive Ability Indicator*

	Ethiopia	India	Peru	Vietnam	All
IGRC ($\hat{\beta}$)	0.339***	0.250***	0.296***	0.405***	0.318***
	(0.034)	(0.026)	(0.038)	(0.029)	(0.015)
IGC	0.404***	0.297***	0.352***	0.482***	0.378***
	(0.041)	(0.031)	(0.045)	(0.034)	(0.018)
Pathways in percent of $\hat{\beta}$					
Child Labor	0.059***	0.154***	0.059*	0.020**	0.065***
	(0.022)	(0.032)	(0.031)	(0.010)	(0.011)
Infrastructure	0.067**	0.024	−0.014	0.008	0.027**
	(0.034)	(0.027)	(0.030)	(0.012)	(0.012)
Education Spending	0.011	0.122***	0.015	0.041***	0.026**
	(0.023)	(0.044)	(0.017)	(0.013)	(0.012)
Child Engagement	0.065***	0.101***	0.039	0.091***	0.083***
	(0.019)	(0.021)	(0.025)	(0.017)	(0.010)
Parent Attentiveness	0.012	0.034**	0.041	0.011	0.024***
	(0.010)	(0.015)	(0.028)	(0.010)	(0.007)
Social Environment	−0.003	0.002	−0.018	0.001	−0.001
	(0.007)	(0.007)	(0.017)	(0.004)	(0.002)
Health	0.028*	−0.007	0.007	0.011	0.012*
	(0.015)	(0.016)	(0.020)	(0.011)	(0.007)
Aspirations	0.044*	0.252***	0.083***	0.222***	0.174***
	(0.025)	(0.040)	(0.031)	(0.033)	(0.017)
Cognitive Ability: PPVT	0.185***	0.269***	0.114**	0.084***	0.182***
	(0.041)	(0.045)	(0.049)	(0.023)	(0.018)
Non-Cognitive: Grit	0.019	0.048	0.164***	0.085***	0.068***
	(0.024)	(0.032)	(0.061)	(0.029)	(0.017)
Explained component of $\hat{\beta}$	0.486***	1.000***	0.490***	0.574***	0.661***
	(0.069)	(0.097)	(0.086)	(0.050)	(0.035)
Unexplained component of $\hat{\beta}$	0.514***	0.000	0.510***	0.426***	0.339***
	(0.069)	(0.097)	(0.086)	(0.050)	(0.035)
Observations	752	921	582	906	3161

Notes: This table shows the results from a decomposition approach, as presented in Equation (5), but using only the PPVT test score variable as pathway variable for cognitive ability. Thereby, the respective coefficients from estimating Equations (3) and (4) are multiplied to elicit the mediating effect of the pathway variables in the transmission of the overall effect of parental education on the children's educational outcome ($\hat{\beta}$). The decomposition approach is conducted separately for each country individually and all countries combined. The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. The pathway factors mainly represent characteristics of the children between the ages 11 and 18. For detailed information on the pathway factors, see Table 2. The PPVT test scores are de-meaned by language. The values provided for each pathway presented in the table give the relative share of this pathway factor for $\hat{\beta}$. Bootstrapped standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

Table F2. *Decompositions using Raven (RPM) Test Scores as Cognitive Ability Indicator*

	Ethiopia	India	Peru	Vietnam	All
IGRC ($\hat{\beta}$)	0.113*	0.248***	0.295***	0.461***	0.273***
	(0.059)	(0.025)	(0.037)	(0.057)	(0.019)
IGC	0.135*	0.295***	0.351***	0.548***	0.324***
	(0.070)	(0.029)	(0.044)	(0.068)	(0.022)
Pathways in percent of $\hat{\beta}$					
Child Labor	0.018	0.162***	0.063**	0.014	0.099***
	(1.768)	(0.034)	(0.031)	(0.019)	(0.018)
Infrastructure	−0.015	0.026	−0.001	0.009	0.010
	(0.914)	(0.026)	(0.031)	(0.020)	(0.016)
Education Spending	0.384	0.122***	0.017	0.029	0.059***
	(5.038)	(0.042)	(0.018)	(0.028)	(0.021)
Child Engagement	0.024	0.103***	0.039	0.044	0.082***
	(0.292)	(0.020)	(0.024)	(0.035)	(0.013)
Parent Attentiveness	0.197	0.035**	0.044	−0.014	0.042***
	(4.191)	(0.014)	(0.030)	(0.022)	(0.012)
Social Environment	−0.030	0.002	−0.013	0.000	−0.001
	(1.141)	(0.008)	(0.017)	(0.010)	(0.002)
Health	0.052	−0.006	0.009	0.019	0.004
	(0.678)	(0.015)	(0.019)	(0.028)	(0.010)
Aspirations	0.512	0.259***	0.088***	0.236***	0.222***
	(6.197)	(0.040)	(0.032)	(0.080)	(0.025)
Cognitive Ability: Raven (RPM)	−0.009	0.281***	0.145***	0.098*	0.208***
	(7.247)	(0.045)	(0.050)	(0.051)	(0.026)
Non-Cognitive: Grit	0.104	0.034	0.084***	0.015	0.050***
	(0.708)	(0.023)	(0.032)	(0.021)	(0.015)
Explained component of $\hat{\beta}$	1.238	1.019***	0.475***	0.450***	0.775***
	(12.569)	(0.101)	(0.086)	(0.105)	(0.055)
Unexplained component of $\hat{\beta}$	−0.238	−0.019	0.525***	0.550***	0.225***
	(12.569)	(0.101)	(0.086)	(0.105)	(0.055)
Observations	174	916	580	186	1856

Notes: This table shows the results from a decomposition approach, as presented in Equation (5), but using only the RPM test score variable as pathway variable for cognitive ability. Thereby, the respective coefficients from estimating Equations (3) and (4) are multiplied to elicit the mediating effect of the pathway variables in the transmission of the overall effect of parental education on the children's educational outcome ($\hat{\beta}$). The decomposition approach is conducted separately for each country individually and all countries combined. The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. The pathway factors mainly represent characteristics of the children between the ages 11 and 18. For detailed information on the pathway factors, see Table 2, for a description of the Raven (RPM) test score, see the text. The values provided for each pathway presented in the table give the relative share of this pathway factor for $\hat{\beta}$. Bootstrapped standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

For India and Peru, with still sizable samples, the findings on the role of cognitive ability (and all other pathways) are still very similar to those reported in the main text, making cognitive ability as measured by the RPM test the most important pathway factor. Even in Vietnam, the results are very similar to those with different, more broadly available measures for cognitive ability. Its importance in the transmission of educational status between parents and their children is still found to be at roughly 10 % and even weakly significant in this much smaller sample. Only in Ethiopia, we cannot confirm any relationship in the reduced sample, not for cognitive ability and not for any other pathway factor. Overall, we still take the unchanged results on the role of cognitive ability in the observable countries as suggestive evidence that it does not make much of a difference whether we take measures of crystallized or fluid intelligence as pathway variables for cognitive ability.

Taken together, the two robustness tests suggest that, in line with the literature, it appears reasonable to use a set of measures including math, reading, and writing ability and interpret it as "cognitive ability". The central results hold in any case, even with other measures less related to school performance.

G. Online Appendix G: Additional Measures of Non-Cognitive Ability

The YL data contains several measures of non-cognitive skills of the children surveyed. The measures are grouped along five dimensions: *agency*, *trust*, *pride*, *inclusion*, and *grit*. The questions to elicit agency are similar to what is often referred to in the literature as the locus of control, that is, how much the children feel that their own actions affect their personal situation. Trust refers to how much a child trusts others (not, how trustworthy he or she is). Pride is close to the concept of self-esteem. Inclusion captures how comfortable a child feels in its broader social environment. Grit is the persistency and determination to follow their own goals. The measures of agency, trust, pride, and inclusion were elicited in the second round, when the children were 11 or 12 years old. Only grit was elicited in the fifth round at age 21 or 22, mirroring the relevance that this particular skill was found to have for socioeconomic outcomes in developed countries (Duckworth et al. 2007).

In order to summarize survey measures of non-cognitive skills, previous studies employing the YL data in different contexts used simple indices of the respective items measuring agency, trust, pride, and inclusion (see Dercon and Singh 2013). We began by grouping the measurement items accordingly in line with the surveys, but then employed factor analysis to create our own measurement system from the pre-defined groups instead of using the "naïve" means of the different factor measures as the factor variable.

Laajaj and Macours (2021) point out that survey-based measures of non-cognitive skills in developing countries are problematic, most importantly because some questions developed for developed countries may be poorly understood in other contexts and when respondents have lower education levels. This can result in larger measurement error in the questions and increased acquiescence bias (stemming from respondents' tendency to affirm statements in surveys, irrespective of their content and whether they are positively or negatively framed). Also, interviewer effects may be more pronounced when interviewers have to bridge a cultural and potentially language induced gap between the survey designer and the respondent. Ideally, surveys should be designed to minimize these problems, for example, by balancing the number of negatively and positively framed questions within each subset of measures, and by adjusting questions according to the test-retest stability. Since the data is from an intensive survey program conducted over more than 15 years, adjusting the survey questions is no option for our study. Instead, we tested the data for signs of typical measurement errors and applied ex-post corrections, following recommendations by Laajaj and Macours (2021). The data provides enough information for us to do that: When factors are measured by positively and negatively framed questions, these are numerically quite well-balanced. Also, although interviewer identification is not part of the original YL dataset, it was made available for this study by YL.

We correct for the influence of interviewers by subtracting interviewer fixed effects for each item. To correct for acquiescence bias, we first invert the reverse-coded items. We then take the difference between the respective average answers of initially positively and initially negatively framed items within each subset of factor measures for each individual, and divide it by two. This gives the acquiescence score. We then subtract each respondent's acquiescence score from the initially positively coded items and add it to the previously reverse-coded items. We standardize the resulting item measures to facilitate interpretation. After these corrections, we apply the factor analysis to create our own scale tailored to the data, as described in Section 2. Those measurement items that are not strongly correlated with others of the same factor obtain lower factor loadings. This applies

especially for the items apparently badly understood. These, then, play no big role for the resulting measurement system of the factors.[46] The individual measures and respective factor loadings of the non-cognitive factors not presented in the main text are displayed in Table G1.

[46] This procedure alone generally addresses much of the measurement error in the data.

Table G1. *Factor Analysis: Pathway Factors and Measures III - Additional Non-Cognitive Ability Factors*

Measure	Description	Observ. Round (Child's Age)	Loading
Factor: Trust			
Trust 1	Most people in my neighbourhood are basically honest. (4 point scale: "Strongly Agree (4), "Agree" (3), "Disagree" (2), "Strongly Disagree" (1))	round 2 (11/12)	.56170596
Trust 2	Most people in my neighbourhood can be trusted. (4 point scale: "Strongly Agree (4), "Agree" (3), "Disagree" (2), "Strongly Disagree" (1))	round 2 (11/12)	.56966085
Trust 3	I believe the government does what is right for people like me. (4 point scale: "Strongly Agree (4), "Agree" (3), "Disagree" (2), "Strongly Disagree" (1))	round 2 (11/12)	.0824734
Trust 4	I feel safe when I go out of the house on my own. (4 point scale: "Strongly Agree (4), "Agree" (3), "Disagree" (2), "Strongly Disagree" (1))	round 2 (11/12)	.1072991
Factor: Agency			
Agency 1	If I try hard I can improve my situation in life. (4 point scale: "Strongly Agree (4), "Agree" (3), "Disagree" (2), "Strongly Disagree" (1))	round 2 (11/12)	.4500641
Agency 2	Other people in my family make all the decisions about how I spend my time. (4 point scale: "Strongly Agree (1), "Agree" (2), "Disagree" (3), "Strongly Disagree" (4))	round 2 (11/12)	.1071642
Agency 3	I like to make plans for my future studies and work. (4 point scale: "Strongly Agree (4), "Agree" (3), "Disagree" (2), "Strongly Disagree" (1))	round 2 (11/12)	.4356309
Agency 4	If I study hard I will be rewarded with a better job in the future. (4 point scale: "Strongly Agree (4), "Agree" (3), "Disagree" (2), "Strongly Disagree" (1))	round 2 (11/12)	.4024711
Factor: Pride			
Pride 1	I feel proud to show my friends where I live. (4 point scale: "Strongly Agree (4), "Agree" (3), "Disagree" (2), "Strongly Disagree" (1))	round 2 (11/12)	.3780161
Pride 2	I am ashamed of my clothes. (4 point scale: "Strongly Agree (1), "Agree" (2), "Disagree" (3), "Strongly Disagree" (4))	round 2 (11/12)	.5609666
Pride 3	I feel proud of the job done by the head of household. (4 point scale: "Strongly Agree (4), "Agree" (3), "Disagree" (2), "Strongly Disagree" (1))	round 2 (11/12)	.3792417
Pride 4	I am often embarrassed because I do not have the right supplies for school. (4 point scale: "Strongly Agree (1), "Agree" (2), "Disagree" (3), "Strongly Disagree" (4))	round 2 (11/12)	.378107
Pride 5	I am proud of my achievements at school. (4 point scale: "Strongly Agree (4), "Agree" (3), "Disagree" (2), "Strongly Disagree" (1))	round 2 (11/12)	.3310045
Pride 7	I am ashamed of my shoes. (4 point scale: "Strongly Agree (1), "Agree" (2), "Disagree" (3), "Strongly Disagree" (4))	round 2 (11/12)	.5308839
Pride 8	I am worried that I don't have the correct uniform. (4 point scale: "Strongly Agree (1), "Agree" (2), "Disagree" (3), "Strongly Disagree" (4))	round 2 (11/12)	.3337205
Factor: Inclusion			
Inclusion 1	At the shops I am treated with fairness. (4 point scale: "Strongly Agree (4), "Agree" (3), "Disagree" (2), "Strongly Disagree" (1))	round 2 (11/12)	.3217927
Inclusion 2	Adults in my street treat me worse than other children of my age. (4 point scale: "Strongly Agree (1), "Agree" (2), "Disagree" (3), "Strongly Disagree" (4))	round 2 (11/12)	.3176727
Inclusion 3	Other children in my class treat me with respect. (4 point scale: "Strongly Agree (4), "Agree" (3), "Disagree" (2), "Strongly Disagree" (1))	round 2 (11/12)	.3996843
Inclusion 4	Other pupils in my class tease me. (4 point scale: "Strongly Agree (4), "Agree" (3), "Disagree" (2), "Strongly Disagree" (1))	round 2 (11/12)	.3266642
Inclusion 5	My teachers treat me worse than other children. (4 point scale: "Strongly Agree (1), "Agree" (2), "Disagree" (3), "Strongly Disagree" (4))	round 2 (11/12)	.207126

Notes: This table shows the measures from the survey data by the non-cognitive factors they are associated to in the surveys. Column (2) describes the coding of the respective measures from the survey question. Column (3) lists the observation round and respective age of the children when the information was elicited. The factor loading of each measure are depicted in Column (4). The data for Peru in this table is measured on a 3 point scale. It is adjusted to fit the other countries according to "Yes" (4/1), "More or less" (2.5), "No" (1/4). Data on the measures "I have no choice about the work I do." (Agency 5), "I am embarrassed by the work I have to do." (Pride 6) and "The job I do makes me feel proud." (Pride 9) is only available for a small subset of the sample ($\sim$ 1400 respondents) and are, hence, not included in the analysis.

The resulting factor variables of the different non-cognitive skills are not strongly correlated with each other. Table G2 shows the respective correlation coefficients, none exceeding 0.36. This implies that the different dimensions should not be merged into one larger factor of non-cognitive skills. In order to not overload the presentations in the main text with different measures of non-cognitive skills, we chose to display the results for grit only, which is methodologically the most advanced measure in the surveys.

Table G2. *Correlation of Non-Cognitive Ability Factors*

	Grit	Trust	Agency	Pride	Inclusion
Grit	1.000				
Trust	0.034	1.000			
Agency	0.039	−0.037	1.000		
Pride	0.041	0.178	0.364	1.000	
Inclusion	0.042	0.226	0.210	0.266	1.000

Notes: This table shows the Pearson correlations of the factor scores for the non-cognitive ability measures available. For more detailed information on the non-cognitive ability measures, see Table G1.

However, the role of all other non-cognitive factors may be interesting. Table G3 shows the results of the decomposition in percentage terms when the additional non-cognitive skills are included all together, for each country and all countries combined. Table G4 shows the results when the non-cognitive skill factors are included one-by-one for the sample of all countries, Column (1) replicating the results described in the main text. Neither non-cognitive factor plays a significant role in the transmission of relative educational attainment in either of the four developing countries. This is because they are partially not strongly enough related to parental background, but mostly because they do not significantly predict children's educational outcomes (intermediate steps not shown here but are available upon request). All are relevant in one step, which suggests that the measurement system successfully captures underlying factors that either depend on parental background or relate to educational outcomes, but do not contribute to the overall link between parental background and children's educational outcomes.

Table G3. *Results with all Non-Cognitive Factors*

	Ethiopia	India	Peru	Vietnam	All
IGRC ($\hat{\beta}$)	0.339***	0.250***	0.296***	0.405***	0.318***
	(0.034)	(0.025)	(0.036)	(0.028)	(0.016)
IGC	0.404***	0.297***	0.352***	0.482***	0.378***
	(0.041)	(0.030)	(0.042)	(0.033)	(0.019)
Pathways in percent of $\hat{\beta}$					
Child Labor	0.063***	0.159***	0.058**	0.020**	0.067***
	(0.023)	(0.032)	(0.030)	(0.010)	(0.011)
Infrastructure	0.067**	0.028	0.005	0.013	0.032**
	(0.034)	(0.027)	(0.031)	(0.013)	(0.013)
Education Spending	0.008	0.124***	0.020	0.041***	0.026**
	(0.023)	(0.042)	(0.019)	(0.013)	(0.012)
Child Engagement	0.065***	0.103***	0.038*	0.089***	0.084***
	(0.019)	(0.021)	(0.023)	(0.017)	(0.010)
Parent Attentiveness	0.012	0.039**	0.037	0.006	0.025***
	(0.010)	(0.016)	(0.027)	(0.011)	(0.007)
Social Environment	−0.003	0.003	−0.010	0.001	−0.001
	(0.007)	(0.008)	(0.017)	(0.004)	(0.002)
Health	0.027*	−0.005	0.013	0.012	0.013*
	(0.015)	(0.015)	(0.021)	(0.011)	(0.007)
Aspirations	0.046*	0.262***	0.087***	0.222***	0.181***
	(0.026)	(0.040)	(0.031)	(0.033)	(0.017)
Cognitive Ability	0.187***	0.292***	0.153***	0.114***	0.199***
	(0.041)	(0.046)	(0.052)	(0.023)	(0.018)
Non-Cognitive: Grit	−0.001	−0.003	0.014	0.001	0.005*
	(0.006)	(0.011)	(0.017)	(0.006)	(0.003)
Non-Cognitive: Trust	−0.004	−0.003	0.001	−0.014*	−0.001
	(0.006)	(0.006)	(0.008)	(0.008)	(0.002)
Non-Cognitive: Agency	0.001	−0.026	0.024	0.005	−0.002
	(0.007)	(0.016)	(0.036)	(0.011)	(0.007)
Non-Cognitive: Pride	0.008	0.011	−0.025	0.023	0.010
	(0.012)	(0.015)	(0.033)	(0.014)	(0.009)
Non-Cognitive: Inclusion	−0.001	0.012	0.003	0.007	0.002
	(0.005)	(0.010)	(0.010)	(0.008)	(0.003)
Explained component of $\hat{\beta}$	0.475***	0.996***	0.417***	0.540***	0.641***
	(0.066)	(0.098)	(0.083)	(0.050)	(0.035)
Unexplained component of $\hat{\beta}$	0.525***	0.004	0.583***	0.460***	0.359***
	(0.066)	(0.098)	(0.083)	(0.050)	(0.035)
Observations	752	921	582	906	3161

Notes: This table shows the results from a decomposition approach, as laid out in Section 3, using all non-cognitive ability factors, that are available, at the same time. The results are displayed for each country separately (Columns (1)-(4)) and all countries combined (Column (5)). For each estimation, the respective coefficients from estimating Equations (3) and (4) are multiplied to elicit the mediating effect of the pathway variables in the transmission of the overall effect of parental education on the children's educational outcome ($\hat{\beta}$). The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. The pathway factors mainly represent characteristics of the children between ages 11 and 18. For detailed information on the pathway factors, see Tables 2 and G1. The values provided for each pathway presented in the table give the relative share of this pathway factor for $\hat{\beta}$ when including all different factors for non-cognitive ability at the same time. Bootstrapped standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

Table G4. *Results with Varying Non-Cognitive Ability Factors*

	Grit	Trust	Agency	Pride	Inclusion
IGRC ($\hat{\beta}$)	0.318***	0.318***	0.318***	0.318***	0.318***
	(0.014)	(0.016)	(0.015)	(0.015)	(0.015)
IGC	0.378***	0.378***	0.378***	0.378***	0.378***
	(0.017)	(0.018)	(0.018)	(0.018)	(0.018)
Pathways in percent of $\hat{\beta}$					
Child Labor	0.067***	0.067***	0.067***	0.067***	0.067***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Infrastructure	0.032**	0.031***	0.032**	0.032**	0.032***
	(0.013)	(0.012)	(0.012)	(0.013)	(0.012)
Education Spending	0.027**	0.027**	0.027**	0.027**	0.027**
	(0.013)	(0.012)	(0.013)	(0.012)	(0.012)
Child Engagement	0.084***	0.084***	0.084***	0.084***	0.084***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Parent Attentiveness	0.026***	0.027***	0.025***	0.024***	0.026***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Social Environment	−0.001	−0.001	−0.001	−0.001	−0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Health	0.013*	0.014**	0.014**	0.014**	0.014**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Aspirations	0.183***	0.182***	0.183***	0.182***	0.183***
	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)
Cognitive Ability	0.200***	0.200***	0.200***	0.200***	0.200***
	(0.019)	(0.018)	(0.018)	(0.018)	(0.019)
Varying Non-Cognitive Ability	0.005	−0.000	0.002	0.007	0.001
	(0.003)	(0.002)	(0.007)	(0.008)	(0.003)
Explained component of $\hat{\beta}$	0.636***	0.632***	0.634***	0.637***	0.633***
	(0.033)	(0.034)	(0.034)	(0.035)	(0.035)
Unexplained component of $\hat{\beta}$	0.364***	0.368***	0.366***	0.363***	0.367***
	(0.033)	(0.034)	(0.034)	(0.035)	(0.035)
Observations	3161	3161	3161	3161	3161

Notes: This table shows the results from a decomposition approach, as laid out in Section 3, separately for the usage of different factors for non-cognitive ability. Thereby, for each estimation, the respective coefficients from estimating Equations (3) and (4), are multiplied to elicit the mediating effect of the pathway variables in the transmission of the overall effect of parental education on the children's educational outcome ($\hat{\beta}$). The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. The pathway factors mainly represent characteristics of the children between ages 11 and 18. For detailed information on the pathway factors, see Tables 2 and G1. The values provided for each pathway presented in the table give the relative share of this pathway factor for $\hat{\beta}$ when including varying non-cognitive ability factors. Bootstrapped standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

H. Online Appendix H: Additional Decompositions and Tables

Table H1. *Effect of Pathways on Children's Education (including pathway factors one by one)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Parental Education	0.229***	0.172***	0.153***	0.140***	0.131***	0.124***	0.119***	0.116***	0.115***	0.116***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.015)	(0.014)
Cognitive Ability	1.684***	1.427***	1.371***	1.265***	1.247***	1.242***	1.220***	1.211***	1.209***	1.211***
	(0.065)	(0.067)	(0.067)	(0.068)	(0.068)	(0.068)	(0.067)	(0.067)	(0.067)	(0.067)
Aspirations		1.024***	0.929***	0.886***	0.879***	0.872***	0.851***	0.848***	0.847***	0.843***
		(0.067)	(0.066)	(0.065)	(0.065)	(0.065)	(0.064)	(0.065)	(0.065)	(0.065)
Child Engagement			−0.792***	−0.727***	−0.725***	−0.721***	−0.721***	−0.729***	−0.728***	−0.725***
			(0.070)	(0.071)	(0.070)	(0.070)	(0.070)	(0.071)	(0.071)	(0.071)
Child Labor				−0.522***	−0.506***	−0.495***	−0.472***	−0.475***	−0.474***	−0.468***
				(0.067)	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)
Infrastructure					0.170***	0.163***	0.169***	0.156**	0.159**	0.164***
					(0.062)	(0.062)	(0.062)	(0.063)	(0.062)	(0.062)
Education Spending						0.135**	0.130**	0.126**	0.122**	0.120**
						(0.056)	(0.056)	(0.056)	(0.056)	(0.056)
Parent Attentiveness							0.261***	0.258***	0.256***	0.249***
							(0.056)	(0.056)	(0.056)	(0.057)
Health								0.111*	0.104*	0.116**
								(0.059)	(0.059)	(0.059)
Non-Cognitive: Grit									0.114**	0.107**
									(0.054)	(0.054)
Social Environment										0.155**
										(0.062)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3161	3161	3161	3161	3161	3161	3161	3161	3161	3161
Adj. R^2	0.926	0.932	0.936	0.937	0.937	0.937	0.938	0.938	0.938	0.938

Notes: This table shows the results from estimating Equation (4) while stepwise including one pathway factor after another according to their importance in the decomposition (Table 3. The analyses include all countries. The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. The pathway factors mainly represent characteristics of the children between ages 11 and 18, and are standardized, such that the coefficients display the increase of Ed^C that is associated with one standard deviation of each pathway factor. For detailed information on the pathway factors, see Table 2. Robust standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

Table H2. *Test of Independence of Coefficients from Parental Background – Ethiopia*

	Below Median	Above Median	Test on Difference
Parental Education	0.635**	0.143**	
	(0.282)	(0.057)	
Controls			
Girl	1.369***	1.114***	0.272
	(0.351)	(0.358)	*0.602*
Household Head's Age	0.099	0.140	0.073
	(0.096)	(0.121)	*0.788*
Household Head's Age^2	−0.001	−0.002	0.244
	(0.001)	(0.001)	*0.621*
Birth Order *(Base: First Child)*			
Second Child	−0.503	0.539	1.962
	(0.644)	(0.522)	*0.161*
Third Child	−0.167	−0.263	0.014
	(0.620)	(0.587)	*0.907*
Fourth Child	−0.625	−0.010	0.451
	(0.650)	(0.682)	*0.502*
Fifth Child	0.759	0.511	0.074
	(0.671)	(0.618)	*0.785*
Sixth Child	−0.342	1.128*	3.092
	(0.578)	(0.621)	*0.079**
Ethiopia	7.285***	6.541***	0.048
	(2.253)	(2.521)	*0.827*
Pathways			
Child Labor	−0.657***	−0.412	0.486
	(0.201)	(0.250)	*0.486*
Infrastructure	0.416**	0.080	1.834
	(0.164)	(0.181)	*0.176*
Education Spending	0.885	0.203	0.451
	(1.122)	(0.341)	*0.502*
Child Engagement	−0.776***	−1.390***	2.676
	(0.183)	(0.397)	*0.102*
Parent Attentiveness	0.402**	0.240	0.508
	(0.189)	(0.160)	*0.476*
Social Environment	0.369*	0.222	0.299
	(0.215)	(0.233)	*0.585*
Health	0.211	0.244	0.022
	(0.173)	(0.155)	*0.881*
Aspirations	0.084	0.439**	1.752
	(0.189)	(0.210)	*0.186*
Cognitive Ability	1.408***	1.416***	0.001
	(0.152)	(0.164)	*0.973*
Non-Cognitive: Grit	0.072	0.289*	0.944
	(0.176)	(0.167)	*0.331*

Notes: This table shows the results from estimating Equation (4) for the Ethiopian subsamples below and above the median of parents' completed years of schooling, respectively, in Column (1) and (2). The dependent variable is the years of schooling completed by the child at age 21–22. The pathway factors mainly represent characteristics of the children between ages 11 and 18, and are standardized, such that the coefficients display the increase of Ed^C that is associated with one standard deviation of each pathway factor. For detailed information on the pathway factors, see Table 2. Standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01). The third column presents the result of a "Chow test" on the difference on the coefficients in the two estimations, reporting χ^2 values and *p-values*.

Table H3. *Test of Independence of Coefficients from Parental Background – India*

	Below Median	Above Median	Test on Difference
Parental Education	0.000	0.030	
	(.)	(0.046)	
Controls			
Girl	−0.184	0.098	0.428
	(0.315)	(0.293)	*0.513*
Household Head's Age	−0.095	−0.071	0.037
	(0.090)	(0.088)	*0.848*
Household Head's Age^2	0.001	0.001	0.000
	(0.001)	(0.001)	*0.995*
Birth Order *(Base: First Child)*			
Second Child	−0.289	−0.310	0.002
	(0.394)	(0.329)	*0.968*
Third Child	−1.067**	−0.194	2.492
	(0.432)	(0.399)	*0.114*
Fourth Child	−0.467	−0.907	0.272
	(0.538)	(0.653)	*0.602*
Fifth Child	−1.084*	0.097	0.880
	(0.597)	(1.349)	*0.348*
Sixth Child	−0.989	0.106	1.297
	(0.632)	(0.772)	*0.255*
India	14.924***	13.488***	0.271
	(2.002)	(1.968)	*0.603*
Pathways			
Child Labor	−0.699***	−0.522***	0.619
	(0.126)	(0.182)	*0.432*
Infrastructure	0.181	0.009	0.640
	(0.170)	(0.146)	*0.424*
Education Spending	0.437**	0.170**	3.063
	(0.181)	(0.082)	*0.080**
Child Engagement	−0.401***	−0.933***	8.699
	(0.092)	(0.167)	*0.003****
Parent Attentiveness	0.207	0.535***	1.995
	(0.159)	(0.181)	*0.158*
Social Environment	0.148	−0.104	0.872
	(0.183)	(0.217)	*0.351*
Health	−0.018	−0.022	0.000
	(0.159)	(0.148)	*0.986*
Aspirations	0.862***	0.817***	0.056
	(0.136)	(0.146)	*0.813*
Cognitive Ability	1.103***	1.053***	0.051
	(0.140)	(0.181)	*0.821*
Non-Cognitive: Grit	0.001	−0.070	0.130
	(0.150)	(0.135)	*0.718*

Notes: This table shows the results from estimating Equation (4) for the Indian subsamples below and above the median of parents' completed years of schooling, respectively, in Column (1) and (2). The dependent variable is the years of schooling completed by the child at age 21–22. The pathway factors mainly represent characteristics of the children between ages 11 and 18, and are standardized, such that the coefficients display the increase of Ed^C that is associated with one standard deviation of each pathway factor. For detailed information on the pathway factors, see Table 2. For detailed information on the pathway factors, see Table 2. Standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01). The third column presents the result of a "Chow test" on the difference on the coefficients in the two estimations, reporting χ^2 values and *p-values*. There is no coefficient for below the median as all have zero years of education.

Table H4. *Test of Independence of Coefficients from Parental Background – Peru*

	Below Median	Above Median	Test on Difference
Parental Education	0.103	0.172***	
	(0.084)	(0.065)	
Controls			
Girl	0.429	0.788**	0.466
	(0.443)	(0.320)	*0.495*
Household Head's Age	−0.117	0.055	0.984
	(0.135)	(0.114)	*0.321*
Household Head's Age^2	0.001	−0.001	1.413
	(0.001)	(0.001)	*0.235*
Second Child	1.393**	−0.040	4.141
	(0.658)	(0.375)	*0.042***
Third Child	0.551	−0.366	1.090
	(0.767)	(0.484)	*0.296*
Fourth Child	0.585	−0.111	0.336
	(0.855)	(0.635)	*0.562*
Fifth Child	0.883	0.846	0.001
	(0.913)	(0.771)	*0.971*
Sixth Child	1.575*	0.803	0.687
	(0.833)	(1.077)	*0.407*
Peru	12.641***	10.322***	0.396
	(2.929)	(2.450)	*0.529*
Pathways			
Child Labor	−0.262	−0.503***	1.034
	(0.186)	(0.170)	*0.309*
Infrastructure	−0.285	0.506**	6.602
	(0.248)	(0.250)	*0.010***
Education Spending	0.728	0.013	6.068
	(0.453)	(0.102)	*0.014***
Child Engagement	−0.855***	−1.289***	0.901
	(0.260)	(0.225)	*0.342*
Parent Attentiveness	0.067	0.225	0.519
	(0.193)	(0.147)	*0.471*
Social Environment	−0.082	0.150	0.759
	(0.216)	(0.143)	*0.384*
Health	0.137	0.087	0.025
	(0.283)	(0.219)	*0.874*
Aspirations	1.185***	0.480	2.122
	(0.304)	(0.355)	*0.145*
Cognitive Ability	1.228***	−0.048	10.457
	(0.272)	(0.345)	*0.001****
Non-Cognitive: Grit	0.224	0.424***	0.631
	(0.222)	(0.140)	*0.427*

Notes: This table shows the results from estimating Equation (4) for the Peruvian subsamples below and above the median of parents' completed years of schooling, respectively, in Column (1) and (2). The dependent variable is the years of schooling completed by the child at age 21–22. The pathway factors mainly represent characteristics of the children between ages 11 and 18, and are standardized, such that the coefficients display the increase of Ed^C that is associated with one standard deviation of each pathway factor. For detailed information on the pathway factors, see Table 2. For detailed information on the pathway factors, see Table 2. Standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01). The third column presents the result of a "Chow test" on the difference on the coefficients in the two estimations, reporting χ^2 values and *p-values*.

Table H5. *Test of Independence of Coefficients from Parental Background – Vietnam*

	Below Median	Above Median	Test on Difference
Parental Education	0.098*	0.428***	
	(0.051)	(0.075)	
Controls			
Girl	1.311***	0.700***	2.620
	(0.275)	(0.266)	*0.105*
Household Head's Age	−0.124	−0.002	0.897
	(0.099)	(0.092)	*0.344*
Household Head's Age^2	0.001	−0.000	0.734
	(0.001)	(0.001)	*0.391*
Second Child	0.520	0.049	1.038
	(0.357)	(0.307)	*0.308*
Third Child	−0.304	−0.328	0.001
	(0.439)	(0.418)	*0.969*
Fourth Child	0.225	−1.008*	2.701
	(0.547)	(0.596)	*0.100*
Fifth Child	0.035	0.280	0.052
	(0.636)	(0.869)	*0.820*
Sixth Child	−0.832	−0.419	0.056
	(0.684)	(1.269)	*0.814*
Vietnam	12.735***	6.995***	3.984
	(2.182)	(2.104)	*0.046***
Pathways			
Child Labor	−0.381**	−0.339	0.014
	(0.182)	(0.260)	*0.906*
Infrastructure	0.017	0.149	0.288
	(0.186)	(0.173)	*0.592*
Education Spending	1.778***	0.442	2.947
	(0.616)	(0.414)	*0.086**
Child Engagement	−0.895***	−1.312***	3.596
	(0.162)	(0.277)	*0.058**
Parent Attentiveness	−0.029	0.353**	3.120
	(0.155)	(0.159)	*0.077**
Social Environment	0.138	0.168	0.023
	(0.139)	(0.160)	*0.879*
Health	0.190	0.138	0.060
	(0.158)	(0.157)	*0.807*
Aspirations	1.034***	0.966***	0.076
	(0.139)	(0.199)	*0.783*
Cognitive Ability	1.000***	1.239***	0.219
	(0.197)	(0.410)	*0.639*
Non-Cognitive: Grit	0.096	−0.019	0.357
	(0.134)	(0.133)	*0.550*

Notes: This table shows the results from estimating Equation (4) for the Vietnamese subsamples below and above the median of parents' completed years of schooling, respectively, in Column (1) and (2). The dependent variable is the years of schooling completed by the child at age 21–22. The pathway factors mainly represent characteristics of the children between ages 11 and 18, and are standardized, such that the coefficients display the increase of Ed^C that is associated with one standard deviation of each pathway factor. For detailed information on the pathway factors, see Table 2. For detailed information on the pathway factors, see Table 2. Standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01). The third column presents the result of a "Chow test" on the difference on the coefficients in the two estimations, reporting χ^2 values and *p-values*.

Table H6. *Test of Independence of Coefficients from Parental Background*

	Below Median	Above Median	Test on Difference
Parental Education	0.102***	0.215***	
	(0.036)	(0.034)	
Controls			
Girl	0.594***	0.813***	0.961
	(0.160)	(0.156)	*0.327*
Household Head's Age	−0.025	−0.010	0.046
	(0.047)	(0.050)	*0.830*
Household Head's Age^2	0.000	0.000	0.051
	(0.000)	(0.001)	*0.821*
Second Child	0.066	0.208	0.253
	(0.232)	(0.185)	*0.615*
Third Child	−0.508**	−0.060	1.711
	(0.250)	(0.241)	*0.191*
Fourth Child	−0.372	−0.331	0.008
	(0.293)	(0.337)	*0.929*
Fifth Child	0.180	0.095	0.025
	(0.311)	(0.420)	*0.874*
Sixth Child	−0.217	0.639	2.287
	(0.287)	(0.474)	*0.130*
Ethiopia	10.717***	8.785***	1.535
	(1.053)	(1.126)	*0.215*
India	12.443***	10.123***	2.309
	(1.053)	(1.113)	*0.129*
Peru	12.209***	9.973***	2.028
	(1.072)	(1.108)	*0.154*
Vietnam	10.042***	9.337***	0.211
	(1.058)	(1.104)	*0.646*
Pathways			
Child Labor	−0.553***	−0.331***	2.514
	(0.078)	(0.106)	*0.113*
Infrastructure	0.208**	0.046	1.704
	(0.082)	(0.097)	*0.192*
Education Spending	0.470***	0.098	8.664
	(0.142)	(0.064)	*0.003***
Child Engagement	−0.608***	−1.193***	12.244
	(0.068)	(0.129)	*0.000***
Parent Attentiveness	0.258***	0.246***	0.010
	(0.082)	(0.083)	*0.918*
Social Environment	0.136	0.145*	0.005
	(0.091)	(0.088)	*0.942*
Health	0.090	0.149*	0.258
	(0.083)	(0.090)	*0.611*
Aspirations	0.758***	0.941***	1.678
	(0.077)	(0.108)	*0.195*
Cognitive Ability	1.258***	1.015***	2.391
	(0.077)	(0.125)	*0.122*
Non-Cognitive: Grit	0.095	0.099	0.002
	(0.078)	(0.074)	*0.968*

Notes: This table shows the results from estimating Equation (4) for the subsamples below and above the median of parents' completed years of schooling, respectively, in Column (1) and (2). The dependent variable is the years of schooling completed by the child at age 21–22. The pathway factors mainly represent characteristics of the children between ages 11 and 18, and are standardized, such that the coefficients display the increase of Ed^C that is associated with one standard deviation of each pathway factor. For detailed information on the pathway factors, see Table 2. For detailed information on the pathway factors, see Table 2. Standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01). The third column presents the result of a "Chow test" on the difference on the coefficients in the two estimations, reporting χ^2 values and *p-values*.

Table H7. *Results taking Site-level Clusters into Account*

	Ethiopia	India	Peru	Vietnam	All
IGRC ($\hat{\beta}$)	0.339***	0.250***	0.296***	0.405***	0.318***
	(0.065)	(0.027)	(0.030)	(0.029)	(0.022)
IGC	0.404***	0.297***	0.352***	0.482***	0.378***
	(0.077)	(0.032)	(0.035)	(0.034)	(0.026)
Pathways in percent of $\hat{\beta}$					
Child Labor	0.060**	0.157***	0.061	0.021*	0.067***
	(0.024)	(0.039)	(0.038)	(0.012)	(0.016)
Infrastructure	0.069	0.026	0.003	0.012	0.032*
	(0.053)	(0.028)	(0.037)	(0.020)	(0.017)
Education Spending	0.013	0.124***	0.020	0.041*	0.027**
	(0.032)	(0.044)	(0.019)	(0.022)	(0.013)
Child Engagement	0.065***	0.102***	0.038*	0.091***	0.084***
	(0.018)	(0.023)	(0.022)	(0.025)	(0.013)
Parent Attentiveness	0.012	0.035*	0.036	0.012	0.026***
	(0.012)	(0.018)	(0.024)	(0.012)	(0.008)
Social Environment	−0.003	0.003	−0.008	0.001	−0.001
	(0.011)	(0.005)	(0.016)	(0.004)	(0.002)
Health	0.027	−0.004	0.013	0.013	0.013
	(0.026)	(0.014)	(0.023)	(0.010)	(0.009)
Aspirations	0.047	0.259***	0.089***	0.234***	0.183***
	(0.030)	(0.050)	(0.032)	(0.036)	(0.025)
Cognitive Ability	0.188**	0.289***	0.153*	0.114***	0.200***
	(0.077)	(0.028)	(0.079)	(0.022)	(0.023)
Non-Cognitive: Grit	−0.001	−0.004	0.014	0.001	0.005
	(0.004)	(0.011)	(0.012)	(0.005)	(0.003)
Explained component of $\hat{\beta}$	0.477***	0.987***	0.419***	0.539***	0.636***
	(0.104)	(0.096)	(0.106)	(0.072)	(0.053)
Unexplained component of $\hat{\beta}$	0.523***	0.013	0.581***	0.461***	0.364***
	(0.104)	(0.096)	(0.106)	(0.072)	(0.053)
Observations	752	921	582	906	3161

Notes: This table shows the results from a decomposition approach, as presented in Equation (5). Thereby, the respective coefficients from estimating Equations (3) and (4), as given in Tables D1-D3 and E1, are multiplied to elicit the mediating effect of the pathway variables in the transmission of the overall effect of parental education on the children's educational outcome ($\hat{\beta}$). The decomposition approach is conducted separately for each country individually and all countries combined. The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. The pathway factors mainly represent characteristics of the children between the ages 11 and 18. For detailed information on the pathway factors, see Table 2. The values provided for each pathway presented in the table give the relative share of this pathway factor for $\hat{\beta}$. Bootstrapped standard errors taking the 20 sentinel sites in each country into account are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

Table H8. *Results with Parental Wealth*

	Ethiopia	India	Peru	Vietnam	All
IGRC ($\hat{\beta}$)	11.135***	5.679***	5.664***	7.362***	7.282***
	(0.686)	(0.576)	(0.656)	(0.534)	(0.307)
IGC	0.616***	0.314***	0.313***	0.407***	0.403***
	(0.038)	(0.032)	(0.036)	(0.030)	(0.017)
Pathways in percent of $\hat{\beta}$					
Child Labor	0.058**	0.181***	0.072**	0.028**	0.073***
	(0.024)	(0.036)	(0.035)	(0.013)	(0.012)
Infrastructure	−0.009	0.023	−0.013	−0.038	0.003
	(0.042)	(0.043)	(0.053)	(0.038)	(0.020)
Education Spending	−0.001	0.112***	0.026	0.052***	0.023**
	(0.016)	(0.042)	(0.024)	(0.017)	(0.011)
Child Engagement	0.038***	0.083***	0.070***	0.096***	0.072***
	(0.013)	(0.021)	(0.026)	(0.018)	(0.009)
Parent Attentiveness	0.020	0.021*	0.045	0.010	0.025***
	(0.012)	(0.011)	(0.031)	(0.008)	(0.007)
Social Environment	0.004	0.001	−0.012	0.001	−0.002
	(0.004)	(0.004)	(0.022)	(0.004)	(0.002)
Health	0.020	−0.004	0.003	0.018	0.013
	(0.015)	(0.014)	(0.037)	(0.011)	(0.008)
Aspirations	0.037**	0.201***	0.132***	0.255***	0.163***
	(0.017)	(0.035)	(0.040)	(0.039)	(0.015)
Cognitive Ability	0.273***	0.270***	0.198***	0.112***	0.220***
	(0.034)	(0.042)	(0.059)	(0.023)	(0.018)
Non-Cognitive: Grit	−0.002	−0.004	0.009	−0.002	0.002
	(0.005)	(0.010)	(0.016)	(0.004)	(0.002)
Explained component of $\hat{\beta}$	0.438***	0.884***	0.530***	0.532***	0.592***
	(0.063)	(0.093)	(0.108)	(0.067)	(0.036)
Unexplained component of $\hat{\beta}$	0.562***	0.116	0.470***	0.468***	0.408***
	(0.063)	(0.093)	(0.108)	(0.067)	(0.036)
Observations	751	921	577	905	3154

Notes: This table shows the results from a decomposition approach, as presented in Equation (5) while the wealth index is used as a proxy for the socioeconomic status of the parents. The wealth index, which is frequently used in studies on developing countries, is a relative measure ranging from 0 to 1 that is constructed on the basis of three main components: indices for housing quality, consumer durables, and access to services. For the estimation, the respective coefficients from estimating Equations (3) and (4) are multiplied to elicit the mediating effect of the pathway variables in the transmission of the overall effect of parental wealth on the children's educational outcome ($\hat{\beta}$). The dependent variable is the years of schooling completed by the child at age 21–22. The parental wealth was retrieved when the child was 7 or 8 years old. The pathway factors mainly represent characteristics of the children between the ages 11 and 18. For detailed information on the pathway factors, see Table 2. The values provided for each pathway presented in the table give the relative share of this pathway factor for $\hat{\beta}$. Bootstrapped standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

Table H9. *Results with Total Expenditures*

	Ethiopia	India	Peru	Vietnam	All
IGRC ($\hat{\beta}$)	1.759***	1.521***	1.603***	2.212***	1.804***
	(0.248)	(0.222)	(0.222)	(0.214)	(0.112)
IGC	0.280***	0.242***	0.255***	0.353***	0.287***
	(0.040)	(0.035)	(0.035)	(0.034)	(0.018)
Pathways in percent of $\hat{\beta}$					
Child Labor	0.063**	0.146***	0.071**	0.044**	0.074***
	(0.027)	(0.037)	(0.033)	(0.020)	(0.013)
Infrastructure	0.086**	0.029	0.006	0.034	0.049***
	(0.034)	(0.030)	(0.044)	(0.028)	(0.015)
Education Spending	0.035	0.165**	0.027	0.074***	0.045**
	(0.031)	(0.066)	(0.037)	(0.021)	(0.018)
Child Engagement	0.004	0.050**	0.033	0.107***	0.051***
	(0.025)	(0.023)	(0.032)	(0.026)	(0.012)
Parent Attentiveness	0.043*	0.023*	0.050	0.025	0.040***
	(0.023)	(0.014)	(0.036)	(0.018)	(0.011)
Social Environment	−0.008	0.002	−0.007	−0.000	−0.005
	(0.009)	(0.006)	(0.020)	(0.005)	(0.003)
Health	0.052*	−0.006	0.015	0.030	0.022**
	(0.030)	(0.020)	(0.036)	(0.022)	(0.011)
Aspirations	0.075**	0.306***	0.133***	0.334***	0.225***
	(0.036)	(0.058)	(0.042)	(0.050)	(0.023)
Cognitive Ability	0.268***	0.216***	0.197***	0.157***	0.218***
	(0.056)	(0.044)	(0.063)	(0.034)	(0.022)
Non-Cognitive: Grit	0.004	−0.004	0.018	0.001	0.005
	(0.009)	(0.011)	(0.017)	(0.004)	(0.003)
Explained component of $\hat{\beta}$	0.622***	0.926***	0.544***	0.803***	0.723***
	(0.105)	(0.128)	(0.106)	(0.090)	(0.050)
Unexplained component of $\hat{\beta}$	0.378***	0.074	0.456***	0.197**	0.277***
	(0.105)	(0.128)	(0.106)	(0.090)	(0.050)
Observations	752	921	582	906	3161

Notes: This table shows the results from a decomposition approach, as presented in Equation (5) while the natural logarithm of monthly total expenditures per household member in PPP-adjusted dollar cents is used as a proxy for the socioeconomic status of the parents. Thereby, the respective coefficients from estimating Equations (3) and (4) are multiplied to elicit the mediating effect of the pathway variables in the transmission of the overall effect of total expenditures per household member on the children's educational outcome ($\hat{\beta}$). The dependent variable is the years of schooling completed by the child at age 21–22. The total expenditures per household member were retrieved when the child was 11 or 12 years old. The pathway factors mainly represent characteristics of the children between the ages 11 and 18. For detailed information on the pathway factors, see Table 2. The values provided for each pathway presented in the table give the relative share of this pathway factor for $\hat{\beta}$. Bootstrapped standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01)..

I. Online Appendix I: Exploratory Factor Analyses

In this section, we introduce the results from the decomposition analysis when the factors are identified by an EFA. For a more detailed formal discussion, see Section H in the online appendix of Heckman et al. (2013). Intuitively, the EFA procedure aims to identify the set of underlying common factors $\mathcal{F}$ and the respective measures $j \in \mathcal{M}^i$ for each factor $i \in \mathcal{F}$ from correlations between all variables that can serve as potential measures. We follow standard EFA procedure (see, Child 2006, Thompson 2004) in obtaining the meaningful sets and the respective ψ_j^i given by Equation (1). All 46 measures used in the main text are considered as potential measures. Measures are retained if they load strongly on one and only one factor. The factors are likely to be correlated. To obtain the factor loadings, we apply an oblique rotation, which allows for the factors to be correlated. For the rotation criterion, we apply the promax criterion with power three throughout (Hendrickson and White 1964). Other oblique rotation methods produce similar results (see Fabrigar et al. 1999). In an EFA, determining the number of relevant factors to be retained is key. It should be the minimum number that can meaningfully explain much of the common variation of most measures. There are several selection criteria to determine this, each of which can yield quite different results. Below we present the results when applying the two methods most regularly used in the literature, the scree test (Cattell 1966) and the parallel analysis test (Horn 1965). In both cases as many factors as there are measures are determined to capture all common variation of the measures by factors. Then the contribution of each factor is assessed against the respective criterion, and only those that fulfill it are retained.

Scree Test

For the scree test, all the eigenvalues of the respective factors are plotted in descending order. The eigenvalues indicate the variation of all measures explained by each factor. The scree plot for our data is shown in Figure I1. A kink in the plot indicates that one additional factor does not add much to explaining the overall variation of the measures. We observe a kink before Factor 5, implying that five factors should be potentially retained.[47] All measures that do not load strongly on one of the six factors or not on one in particular are then dropped iteratively, repeating the factor analysis with the remaining measures. As a result, the 28 measures listed in Table I1 – grouped by the underlying factors they load on – are retained. The KMO statistic of the remaining set of measures is well above 0.7, and the Bartlett (1951)-Test of sphericity is significant at the 1 %-level, indicating that the data is well suited for a factor analysis. The respective factor loadings are shown next to each measure in Table I1. For comparison, the last column lists the factor each measure was associated with in the main analysis. We see that all measures for parental spending on education do load on one factor, which we thus name equivalently. Measures associated with the factors of underage family engagement and aspirations before jointly load on one factor, which we thus interpret as the focus on one's career. Previous measures of infrastructure, the social environment and childhood health load on one factor. It is slightly dominated by health measures (also in terms of factor loadings), and concerns of safety and the rural/urban distinction may also relate to the health status, which is why we interpret this factor as health. We see that most measures associated with cognitive ability

[47] Also taking more factors to potentially be retained, such as 8, does not change the final outcome or groupings after the next steps.

before jointly load strongly on another factor. Although it also comprises measures for the number of additional children in the household and the hours worked, these load less strongly, and we thus again name this factor cognitive ability. Almost all measures of grit load on one common factor, which we thus name accordingly.

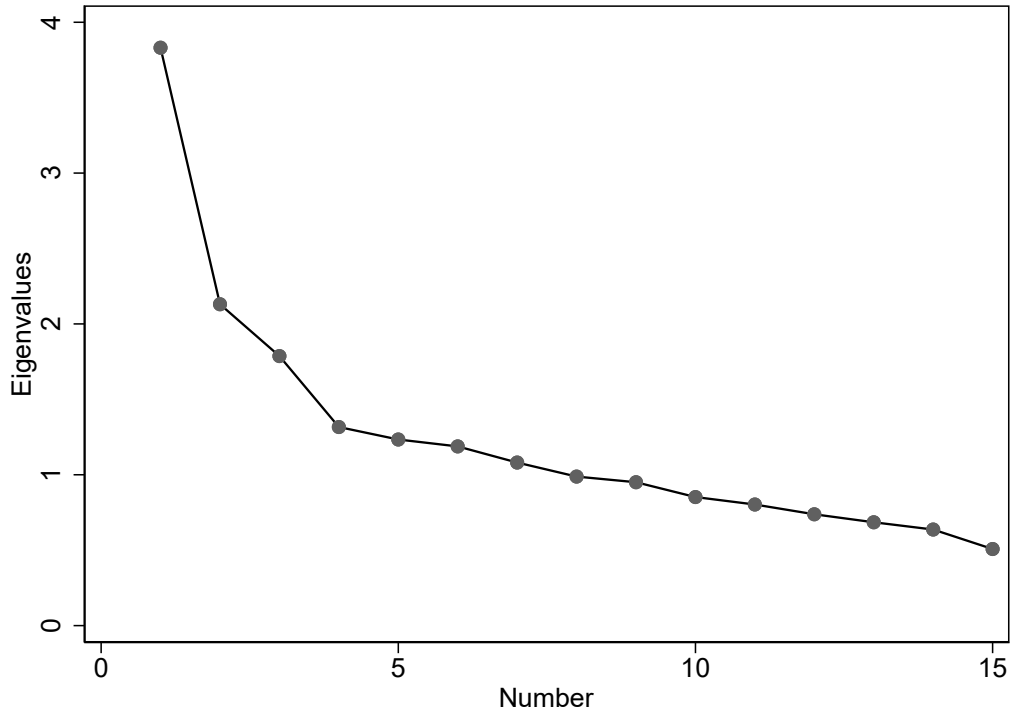

Fig. I1. *Screeplot of Eigenvalues*

Notes: This figure shows the eigenvalues of the first 15 factors derived from an exploratory factor analysis on the measures presented in Table 2, in descending order.

The results of the decomposition analysis employing the five factors identified from the EFA with the scree test as retention criterion are shown in Table I2. They show that cognitive ability identified through the measures at hand contributes the most to the observed immobility in Ethiopia, India, and Peru. The broader factor of career focus also explains substantial shares of the immobility, and even the most in Vietnam, just as in the intuitive grouping. Equivalently, parental spending on education contributes significantly to the transmission of socioeconomic status in India, and to some extent in Vietnam. Children's health status is a relevant factor only in Ethiopia, and grit is in none of the countries. These results are very much in line with the results presented in the main text, except that the factors derived by an EFA do not allow for differentiating between the role played by different aspects of career focus (which is expressed in measures of early aspirations by children and parents as well as becoming a spouse and/or a parent when underage) and between those of cognitive ability and child labor, which are also captured in one factor.

Table I1. *Pathway Factors - Scree Test*

Measure	Loading	Factor in Main Analysis
Factor 1: Education Spending		
Spent on Schooling Fees	.7125171	Education Spending
Spent on School Books	.5584029	Education Spending
Private Schooling	.5337841	Education Spending
Factor 2: Career Focus		
Child Marriage	.6030985	Underage Family
Child Parent	.5752947	Underage Family
Expected Child Marriage	.4150853	Underage Family
Expected Child Parent	.3272172	Underage Family
Job Aspirations Child	-.2338879	Aspirations
School Aspirations Parents	-.3266726	Aspirations
Factor 3: Health		
Type of Living Site	.4189597	Infrastructure
Trust in Neighborhood	-.1884576	Social Environment
Safe to Go Outside	-.3474156	Social Environment
Weight-for-age	.8498806	Health
Height-for-age	.5876808	Health
Thinness Indicator	.3781074	Health
Factor 4: Cognitive Ability		
Working Hours	-.3910649	Child Labor
Number of Children	.2867198	Parental Attentiveness
Math Score	.5085645	Cognitive Ability
Reading Level	.7084015	Cognitive Ability
Writing Level	.8854833	Cognitive Ability
Literacy	.8500995	Cognitive Ability
Factor 5: Non-Cognitive: Grit		
Grit 1	.2708071	Non-Cognitve: Grit
Grit 3	.4078536	Non-Cognitve: Grit
Grit 4	.5128819	Non-Cognitve: Grit
Grit 5	.3989184	Non-Cognitve: Grit
Grit 6	.4080568	Non-Cognitve: Grit
Grit 7	.5689889	Non-Cognitve: Grit
Grit 8	.5780654	Non-Cognitve: Grit

Notes: This table shows the factors derived from an exploratory factor analysis with the scree-test as criterion for determination of the number of factors. For each factor, the measures associated with the factor are listed. Column (2) depicts the respective factor loadings. Column (3) shows the factor that each measure was associated with in the analysis in the main text.

Table I2. *Results – Scree Test*

	Ethiopia	India	Peru	Vietnam	All
IGRC ($\hat{\beta}$)	0.339***	0.250***	0.296***	0.405***	0.318***
	(0.033)	(0.026)	(0.037)	(0.029)	(0.015)
IGC	0.404***	0.297***	0.352***	0.482***	0.378***
	(0.040)	(0.031)	(0.043)	(0.034)	(0.018)
Pathways in percent of $\hat{\beta}$					
Education Spending	0.025	0.187***	0.029	0.053***	0.045***
	(0.024)	(0.047)	(0.021)	(0.015)	(0.013)
Career Focus	0.120***	0.248***	0.083***	0.203***	0.183***
	(0.026)	(0.036)	(0.030)	(0.027)	(0.016)
Health	0.049**	−0.005	0.020	0.021*	0.026***
	(0.021)	(0.021)	(0.026)	(0.011)	(0.009)
Cognitive Ability	0.192***	0.374***	0.158***	0.098***	0.210***
	(0.047)	(0.051)	(0.048)	(0.023)	(0.019)
Non-Cognitive: Grit	−0.001	0.000	0.015	0.000	0.005*
	(0.006)	(0.012)	(0.018)	(0.006)	(0.003)
Explained component of $\hat{\beta}$	0.385***	0.804***	0.305***	0.375***	0.470***
	(0.058)	(0.093)	(0.068)	(0.040)	(0.030)
Unexplained component of $\hat{\beta}$	0.615***	0.196**	0.695***	0.625***	0.530***
	(0.058)	(0.093)	(0.068)	(0.040)	(0.030)
Observations	752	921	582	906	3161

Notes: This table shows the results from a decomposition approach, as presented in Equation (5). Thereby, the respective coefficients from estimating Equations (3) and (4) are multiplied to elicit the mediating effect of the pathway variables in the transmission of the overall effect of parental education on the children's educational outcome ($\hat{\beta}$). The decomposition approach is conducted separately for each country individually and all countries combined. The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. The pathway factors mainly represent characteristics of the children between the ages 11 and 18, and are derived from the available measures using a Scree Test. For detailed information on the pathway factors, see Table I1. The values provided for each pathway presented in the table give the relative share of this pathway factor for $\hat{\beta}$. Bootstrapped standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

Horn's Parallel Analysis

Horn's parallel analysis is an extension of the Kaiser-Guttman criterion (Kaiser 1960; Kaiser 1961; Guttman 1954) for the retention of factors. This criterion would retain all factors with an eigenvalue greater than zero, as they contribute to explaining common variation. It typically yields too many factors to be meaningfully interpreted. Horn (1965)'s parallel analysis corrects for the fact that some common variation is likely to arise from noise in the data. Applied to the 46 measures from the survey, the parallel analysis still recommends to retain 27 factors. For most of these factors, however, not enough measures load strongly enough or not only on the respective factor, so only six factors are retained that can be meaningfully interpreted, captured by 20 measures overall. The KMO statistic of the sample is again above 0.7 and the Bartlett (1951)-test of sphericity yields a significance above 1 %. The measures retained are listed in Table I3, next to the respective factor loadings and the factor they were associated with in the main analysis. Slightly different concepts are dropped in comparison to the scree test but the factors identified are still very much in line with the most important ones obtained through intuitive association. Child labor here gets retained as an individual factor. Measures of parents' spending on education constitute another common factor, as do those of parental attentiveness. The social environment forms another individual factor. Children's and parents' expressions of aspirations are also singled out by this procedure. Of the previous measures of cognitive ability, only those of literacy are retained to measure a common factor. Of the initial factors from the intuitive grouping, only the measures of infrastructure, health, child engagement and non-cognitive skills are not identified as to have enough measures to commonly load on common factors, and those for cognitive ability are limited to include those of literacy.

Table I4 depicts the results from the decomposition analysis with the six factors obtained from an EFA applying the Horn (1965) criterion. Literacy alone accounts for smaller shares of the overall association between parental and offspring's education if measures of mathematical-logical thinking are not considered, but is still an important factor in Ethiopia and India in particular. The ascribed role of aspirations is then even larger, except for in Ethiopia, where it was anyway found to be of lesser importance. The same somewhat increased measured importance applies to child labor, in proportion to its measured roles discussed above. Education spending is also with this approach mostly a factor to the intergenerational transmission of educational achievements in India. Parental attentiveness plays no big role in any country. Although the information content derived through including lesser factors and measures than in the intuitive approach, these results give some support to the central interpretations from the main analysis.

These analyses confirm that the results derived using an intuitive approach of identifying factors from the given set of measures are not at odds with those from the more data driven approach of an EFA. They also show that the former allows for a slightly more nuanced interpretation and the inclusion of potentially important factors which have less but still meaningful measures in the data. Comparing the results also suggests that the fewer factors are included, the greater is the risk of omitted variables. However, the main results remain valid for all countries irrespective of the specific approach to dimensionality reduction.

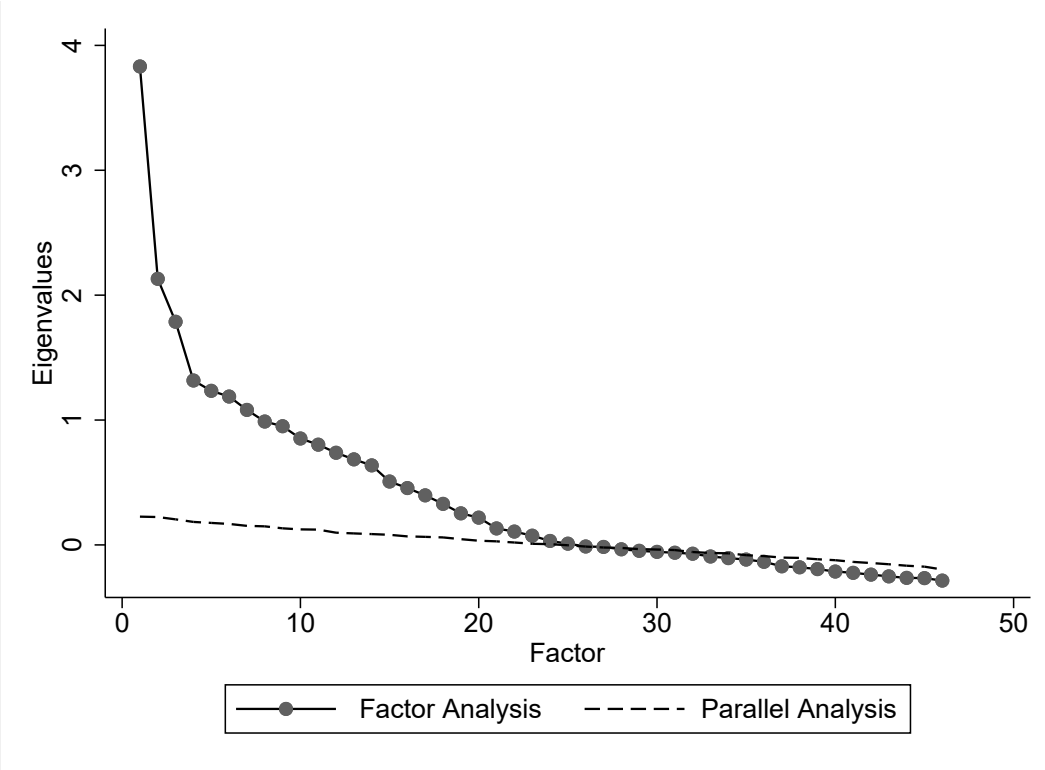

Fig. I3. *Plot Parallel Analysis*

Notes: This figure shows the eigenvalues of all potential 46 factors derived from an exploratory factor analysis on the measures presented in Tables2, in descending order, against the expected eigenvalues that would arise from a purely random but orthogonal distribution of the measures.

Table I3. *Pathway Factors - Parallel Analysis*

Measure	Loading	Factor in Main Analysis
Factor 1: Child Labor		
Activity for Money	.3760703	Child Labor
Working Hours	.5405388	Child Labor
Number of Children	-.3512614	Parental Attentiveness
Factor 2: Education Spending		
Spent on Schooling Fees	.7125171	Education Spending
Spent on School Books	.5584029	Education Spending
Private Schooling	.5337841	Education Spending
Factor 3: Parental Attentiveness		
Attention received	.440274	Parental Attentiveness
Love received	.4796468	Parental Attentiveness
Conversation	.3756971	Parental Attentiveness
Supported by Parents	.4331997	Parental Attentiveness
Free Speech	.4125525	Parental Attentiveness
Factor 4: Social Environment		
Friends with Alcohol	.520359	Social Environment
Friends with Trouble	.2963123	Social Environment
Friends who Smoke	.5287578	Social Environment
Factor 5: Aspirations		
School Aspirations Child	.5789526	Aspirations
Job Aspirations Child	.3690699	Aspirations
School Aspirations Parents	.5713992	Aspirations
Factor 6: Literacy		
Reading Level	.6873109	Cognitive Ability
Writing Level	.8836308	Cognitive Ability
Literacy	.8817762	Cognitive Ability

Notes: This table shows the factors derived from an exploratory factor analysis applying Horn's parallel analysis to determine the number of factors. For each factor, the measures associated with the factor are listed. Column (2) depicts the respective factor loadings. Column (3) shows the factor that each measure was associated with in the analysis in the main text.

Table I4. *Results – Parallel Analysis*

	Ethiopia	India	Peru	Vietnam	All
IGRC ($\hat{\beta}$)	0.339***	0.250***	0.296***	0.405***	0.318***
	(0.035)	(0.025)	(0.036)	(0.028)	(0.015)
IGC	0.404***	0.297***	0.352***	0.482***	0.378***
	(0.042)	(0.030)	(0.043)	(0.033)	(0.018)
Pathways in percent of $\hat{\beta}$					
Child Labor	0.131***	0.236***	0.105**	0.049***	0.126***
	(0.030)	(0.038)	(0.041)	(0.017)	(0.015)
Education Spending	0.016	0.140***	0.033*	0.053***	0.040***
	(0.024)	(0.043)	(0.019)	(0.015)	(0.013)
Parental Attentiveness	0.010	0.034**	0.040	0.010	0.022***
	(0.010)	(0.015)	(0.027)	(0.010)	(0.007)
Social Environment	0.012	0.006	−0.013	0.002	0.003
	(0.010)	(0.008)	(0.013)	(0.004)	(0.003)
Aspirations	0.061**	0.292***	0.115***	0.265***	0.219***
	(0.026)	(0.044)	(0.036)	(0.034)	(0.019)
Literacy	0.119***	0.199***	0.068**	0.049***	0.118***
	(0.038)	(0.036)	(0.034)	(0.017)	(0.014)
Explained component of $\hat{\beta}$	0.349***	0.908***	0.347***	0.427***	0.528***
	(0.058)	(0.096)	(0.066)	(0.046)	(0.032)
Unexplained component of $\hat{\beta}$	0.651***	0.092	0.653***	0.573***	0.472***
	(0.058)	(0.096)	(0.066)	(0.046)	(0.032)
Observations	752	921	582	906	3161

Notes: This table shows the results from a decomposition approach, as presented in Equation (5). Thereby, the respective coefficients from estimating Equations (3) and (4) are multiplied to elicit the mediating effect of the pathway variables in the transmission of the overall effect of parental education on the children's educational outcome ($\hat{\beta}$). The decomposition approach is conducted separately for each country individually and all countries combined. The dependent variable is the years of schooling completed by the child at age 21–22. Parental education is the years of schooling completed by the parents. The pathway factors mainly represent characteristics of the children between the ages 11 and 18, and are derived from the available measures using a Parallel Analysis. For detailed information on the pathway factors, see Table I3. The values provided for each pathway presented in the table give the relative share of this pathway factor for $\hat{\beta}$. Bootstrapped standard errors are reported in parentheses (* p<0.1, ** p<0.05, *** p<0.01).

J. Online Appendix J: Multilevel Mediation Analysis

In this section, we lay out the multilevel mediation analysis approach. This allows for mediating relationships between the pathway factors. The approach follows the method by Bolt et al. (2021). In the multilevel mediation analysis, we impose a structure of unidirectional relationships of influence between factor variables. This structure is theory-driven and displays relationships found in the literature (such as between early health status and cognitive development, Attanasio et al. 2020; López Boo 2016; Figlio et al. 2014) and obvious potential relationships. The timing of the factor measurements can further inform possible relationships, but this is limited by the fact that measured characteristics can also be stable over time. The structure of the multilevel mediation analysis is displayed in Figure J1. Arrows from below to above indicate relationships of influence, where factors further up the chains of influence ("later" factors) are also assumed to potentially be directly influenced by factors further below ("previous" factors). Arrows between boxes around groups of variables indicate that each variable in the box has a (potential) directed relationship of influence with each variable in the respective other box. For the pathway factors of infrastructure, education spending, and social environment, we cannot assume a clear unidirectional relationship of influence with other pathway factors. They are still included as potential pathway factors on their own. Parental education is still assumed to potentially influence all variables individually (as represented by the box around all pathway variables).

For the estimation of the model of multilevel mediation, we replace the set of Equations (3) by the

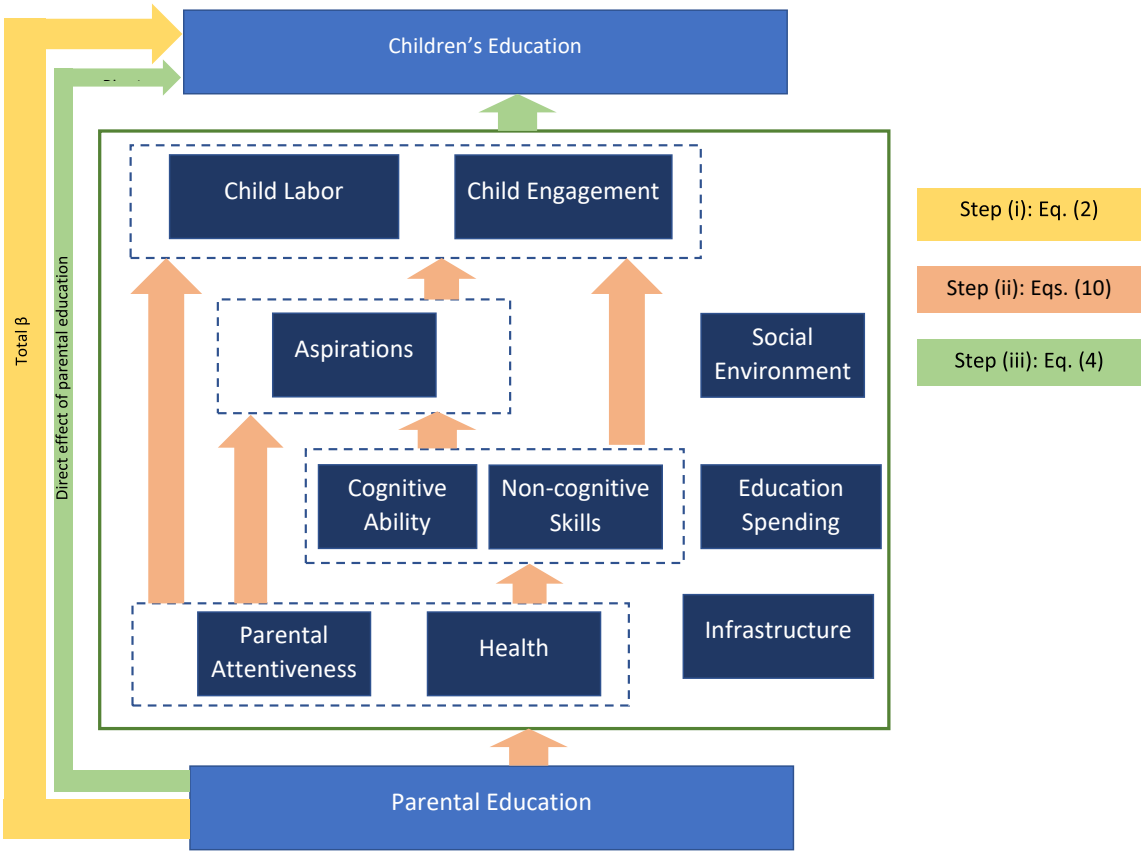

Fig. J1. *Graphical Representation of Multilevel Mediation Analysis*

Notes: This figure shows the interrelations between the pathways that we consider in a multilevel mediation analysis approach, as detailed in Equations (10), which replace Equation (3) in the approach laid out in Section 3. Transparent items display the steps are unchanged compared to the unstructured decomposition analysis. The derivation of the pathway factor variables from the measures listed in Table 2 is described in Section 2.

following set of equations for the respective pathways $PW_{i,h}$ with $i \in \mathcal{F}$:

$$
\begin{aligned}
&Health_h = \alpha_{H,2} + \lambda_H Ed_h^P + X_h{}'\zeta_{H,2} + \epsilon_{H,h,2} \\
&ParentAttentiveness_h = \alpha_{PA,2} + \lambda_{PA} Ed_h^P + + X_h{}'\zeta_{PA,2} + \epsilon_{PA,h,2} \\
&CognitiveAbility_h = \tilde{\alpha}_{CA,2} + \tilde{\lambda}_{CA} Ed_h^P + \delta_{H,CA} Health_h + \delta_{PA,CA} ParentAttentiveness_h \\
&+ X_h{}'\tilde{\zeta}_{CA,2} + \tilde{\epsilon}_{CA,h,2} \\
&NonCognitive_h = \tilde{\alpha}_{NC,2} + \tilde{\lambda}_{NC} Ed_h^P + \delta_{H,NC} Health_h + \delta_{PA,NC} ParentAttentiveness_h \\
&+ X_h{}'\tilde{\zeta}_{NC,2} X_h + \tilde{\epsilon}_{NC,h,2} \\
&Aspirations_h = \tilde{\alpha}_{A,2} + \tilde{\lambda}_A Ed_h^P + \delta_{CA,A} CognitiveAbility_h + \delta_{NC,A} NonCognitive_h \\
&+ \delta_{H,A} Health_h + \delta_{PA,A} ParentAttentiveness_h + X_h{}'\tilde{\zeta}_{A,2} + \tilde{\epsilon}_{A,h,2} \\
&ChildLabor_h = \tilde{\alpha}_{CL,2} + \tilde{\lambda}_{CL} Ed_h^P + \delta_{A,CL} Aspirations_h + \delta_{CA,CL} CognitiveAbility_h \\
&+ \delta_{NC,CL} NonCognitive_h + \delta_{H,CL} Health_h + \delta_{PA,CL} ParentAttentiveness_h + X_h{}'\tilde{\zeta}_{CL,2} + \tilde{\epsilon}_{CL,h,2} \\
&ChildEngagement_h = \tilde{\alpha}_{CE,2} + \tilde{\lambda}_{CE} Ed_h^P + \delta_{A,CE} Aspirations_h + \delta_{CA,CE} CognitiveAbility_h \\
&+ \delta_{NC,CE} NonCognitive_h + \delta_{H,CE} Health_h + \delta_{PA,CE} ParentAttentiveness_h + X_h{}'\tilde{\zeta}_{CE,2} + \tilde{\epsilon}_{CE,h,2} \\
&Infrastructure = \alpha_{I,2} + \lambda_I Ed_h^P + X_h{}'\zeta_{I,2} + \epsilon_{I,h,2} \\
&EducationSpending = \alpha_{ES,2} + \lambda_{ES} Ed_h^P + X_h{}'\zeta_{ES,2} + \epsilon_{ES,h,2} \\
&SocialEnvironment = \alpha_{SE,2} + \lambda_{SE} Ed_h^P + X_h{}'\zeta_{SE,2} + \epsilon_{SE,h,2}
\end{aligned}
\tag{10}
$$

where a "˜" denotes that the respective variables are different estimates than those in the main text, and the $\delta_{A,B}$ denote the relationship between pathway factors A and B.

Equation (4) is estimated as in the unstructured single level mediation analysis:

$$
Ed_h^C = \alpha_3 + \sum_i \rho_i PW_{i,h} + \gamma_{Ed^P} Ed_h^P + X_h{}'\zeta_3 + \epsilon_{h,3} \tag{4}
$$

Inserting Equations (10) (starting from bottom) into Equation (4) solves for the multilevel mediation analysis, but all estimations, those of Equations (4) and (10), are pursued individually. The components of β mediated by each pathway, denoted β_i, for $i \in \mathcal{F}$, are then given by the following terms, respectively:

$$\beta_H = \lambda_H \Bigg[\rho_H + \Big(\delta_{H,CA} \Big(\rho_{CA} + \delta_{CA,A} (\rho_A + \delta_{A,CL}\rho_{CL} + \delta_{A,CE}\rho_{CE}) \Big) \Big) + \Big(\delta_{H,NC} \Big(\rho_{NC} + \delta_{NC,A} (\rho_A + \delta_{A,CL}\rho_{CL} + \delta_{A,CE}\rho_{CE}) \Big) \Big) + \delta_{H,A} \Big(\rho_A + \Big(\delta_{A,CL}\rho_{CL} + \delta_{A,CE}\rho_{CE} \Big) \Big) + \delta_{H,CL}\rho_{CL} + \delta_{H,CE}\rho_{CE} \Bigg]$$

$$\beta_{PA} = \lambda_{PA} \Bigg[\rho_{PA} + \Big(\delta_{PA,CA} \Big(\rho_{CA} + \delta_{CA,A} (\rho_A + \delta_{A,CL}\rho_{CL} + \delta_{A,CE}\rho_{CE}) \Big) \Big) + \Big(\delta_{PA,NC} \Big(\rho_{NC} + \delta_{NC,A} (\rho_A + \delta_{A,CL}\rho_{CL} + \delta_{A,CE}\rho_{CE}) \Big) \Big) + \delta_{PA,A} \Big(\rho_A + \Big(\delta_{A,CL}\rho_{CL} + \delta_{A,CE}\rho_{CE} \Big) \Big) + \delta_{PA,CL}\rho_{CL} + \delta_{PA,CE}\rho_{CE} \Bigg]$$

$$\beta_{CA} = \tilde{\lambda}_{CA} \Bigg(\rho_{CA} + \Big(\delta_{CA,A} \Big(\rho_A + \delta_{A,CL}\rho_{CL} + \delta_{A,CE}\rho_{CE} \Big) \Big) + \delta_{CA,CL}\rho_{CL} + \delta_{CA,CE}\rho_{CE} \Bigg)$$

$$\beta_{NC} = \tilde{\lambda}_{NC}\Bigg(\rho_{NC} + \Big(\delta_{NC,A}\Big(\rho_A + \delta_{A,CL}\rho_{CL} + \delta_{A,CE}\rho_{CE}\Big)\Big) + \delta_{NC,CL}\rho_{CL} + \delta_{NC,CE}\rho_{CE}\Bigg)$$

$$\beta_A = \tilde{\lambda}_A\Big(\rho_A + +\delta_{A,CL}\rho_{CL} + \delta_{A,CE}\rho_{CE}\Big)$$

$$\beta_{CL} = \tilde{\lambda}_{CL}\rho_{CL} \tag{11}$$

$$\beta_{CE} = \tilde{\lambda}_{CE}\rho_{CE}$$

$$\beta_I = \lambda_I\rho_I$$

$$\beta_{ES} = \lambda_{ES}\rho_{ES}$$

$$\beta_{SE} = \lambda_{SE}\rho_{SE}$$

By the structure of the multilevel mediation analysis, the estimated β_i's for factors that are influenced by other factors (that is, "later" factors in Figure J1) are smaller or equal to the estimates of the components of β for the same factor in the unstructured mediation analysis (if – as expected – the correlation between the influencing and the influenced factors is positive, or negative if one has a negative correlation with parents' and children's schooling). In turn, the estimated β_i's for factors that influence other factors (that is, "previous" factors in Figure J1) are larger or equal to the estimates of the components of β for the same factor in the unstructured mediation analysis (if the direction of the relationships is as expected). Indirect effects mediated through multiple pathway factors are hence attributed to the initial factor in this line.

Take, for example, the component of β that is mediated by cognitive ability according to the multilevel mediation analysis from Equation (11), $\beta_{CA} = \tilde{\lambda}_{CA}(\rho_{CA} + (\delta_{CA,A}(\rho_A + \delta_{A,CL}\rho_{CL} + \delta_{A,CE}\rho_{CE})) + \delta_{CA,CL}\rho_{CL} + \delta_{CA,CE}\rho_{CE})$. The following describes the individual terms to illustrate the method:

$\tilde{\lambda}_{CA}\rho_{CA}$ is the correlation between children's education and parental education that is 'directly' mediated by cognitive ability, namely the correlation of cognitive ability with parental education when controlling for the influence of health and parent attentiveness ($\tilde{\lambda}_{CA}$ from Equation (10)) multiplied with the correlation of cognitive ability with children's education when controlling for all other pathway factors and the direct effect of parental education (ρ_{CA} from Equation (4)).

$\tilde{\lambda}_{CA}\delta_{CA,A}\rho_A$ is the component that is mediated by cognitive ability through its correlation with aspirations (assuming the influence of cognitive ability on aspirations) and the 'direct' correlation of aspirations with children's education.

$\tilde{\lambda}_{CA}\delta_{CA,A}(\delta_{A,CL}\rho_{CL} + \delta_{A,CE}\rho_{CE})$ are the components that are mediated by cognitive ability through

its assumed effect on (measured by the correlation with) aspirations, and through that on/with child labor and child engagement, respectively.

$\tilde{\lambda}_{CA}(\delta_{CA,CL}\rho_{CL} + \delta_{CA,CE}\rho_{CE})$ are the components that are mediated by cognitive ability via its effect on/correlation with child labor and child engagement, which are not correlated with/mediated by aspirations.

All of these components summed up constitute the pathway of cognitive ability in this robustness test because we assume that common variation between cognitive ability and the respective "later" pathways (aspirations, child labor, and child engagement) can be attributed to cognitive ability. In turn, cognitive ability only depends on parental education apart from the relationship between parental education and cognitive ability that can be explained by health and parental attentiveness ($\tilde{\lambda}_{CA}$), instead of the entire correlation (λ_{CA}).

Similarly, the correlation between parental education and children's education that is mediated by child labor according to the unstructured single level mediation analysis ($\lambda_{CL}\rho_{CL}$) is, in the structured multilevel mediation analysis, attributed to child labor as a pathway factor ($\tilde{\lambda}_{CL}\rho_{CL}$) only to that extent as it is related to the parents' education independent of (that is, controlling for) the "previous" factors, which we assume to influence it.

The results of the structured multilevel mediation analysis are presented in Table 6 in the main text. The results are similar to those when all pathway factors are considered on an equal footing only through their direct interdependencies with parental and children's education in the unstructured single level mediation analysis.

IMPRESSUM
Jena Economics Research Papers
ISSN 1864-7057
Friedrich Schiller University Jena
Faculty of Economics and Business Administration
Carl-Zeiss-Str. 3
D-07743 Jena, Germany

Email: office.jerp@uni-jena.de
Editor: Silke Übelmesser
Website: www.wiwi.uni-jena.de/en/jerp