Policy Research Working Paper 10309

Cognitive and Socioemotional Skills in Low-Income Countries

Measurement and Associations with Schooling and Earnings

Alice Danon
Jishnu Das
Andreas de Barros
Deon Filmer

WORLD BANK GROUP
Development Economics
Development Research Group
February 2023

Policy Research Working Paper 10309

Abstract

This paper assesses the reliability and validity of cognitive and socioemotional skills measures and investigates the correlation between schooling, skills acquisition, and labor earnings. The primary data from Pakistan incorporates two innovations related to measurement and sampling. On measurement, the paper develops and implements a battery of instruments intended to capture cognitive and socioemotional skills among young adults. On sampling, the paper uses a panel that follows respondents from their original rural locations in 2003 to their residences in 2018, a period over which 38 percent of the respondents left their native villages. In terms of their validity and reliability, our skills measures compare favorably to previous measurement attempts in low- and middle-income countries. The following are documented in the data: (a) more years of schooling are correlated with higher cognitive and socioemotional skills; (b) labor earnings are correlated with cognitive and socioemotional skills as well as years of schooling; and (c) the earnings-skills correlations depend on respondents' migration status. The magnitudes of the correlations between schooling and skills on the one hand and earnings and skills on the other are consistent with a widespread concern that such skills are underproduced in the schooling system.

This paper is a product of the Development Research Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at dfilmer@worldbank.org.

Cognitive and Socioemotional Skills in Low-Income Countries: Measurement and Associations with Schooling and Earnings[1]

Alice Danon (Harvard University)

Jishnu Das (Georgetown University)

Andreas de Barros (MIT)

Deon Filmer (World Bank)

Keywords: Cognitive Skills, Socioemotional Skills, Labor Earnings, Returns to Schooling, Migration

JEL Codes: O1, O12, O15, I2, I25, I26

[1] For helpful comments, we thank Lant Pritchett, participants at the World Bank's conference on linkages between schooling and labor markets, the Methods and Measurement Conference at Northwestern University, and the 2022 RISE conference at Oxford University. Financial support for this research was made available through RISE grant PO13412 and The World Bank's Strategic Research Program Fund. The findings, interpretations, and conclusions expressed in this paper are those of the authors and do not necessarily represent the views of the World Bank, its Executive Directors, or the governments they represent. Please address correspondence to adanon@g.harvard.edu, jishnu.das@georgetown.edu, debarros@mit.edu, and dfilmer@worldbank.org.

INTRODUCTION

Education and migration are two strategies widely believed to improve living standards for populations living in rural areas. A vast literature demonstrates that educated rural households are better able to avail of new technologies, that consumption growth is higher among households who migrate, and that secondary schooling and migration (whether international or within-country) increase earnings.[2] However, two data-related issues have made it harder to assess the returns to specific skills for migrants versus non-migrants and the ability of children to acquire these skills through schooling in the first place. First, despite an extensive literature in the United States that demonstrates the importance of cognitive and socioemotional skills for labor market outcomes (Heckman, 2007), these skills have proven notoriously difficult to measure in low-income countries (Laajaj and Macours, 2021; Valerio et al., 2016). Second, long-term panels with information on skills, schooling, and earnings in low-income countries are extremely rare, although we expect more will become available in the next decade.

We address both of these gaps in this paper. We focus on the measurement of cognitive and socioemotional (SEM) skills and assess their reliability and validity relative to comparable research from the existing literature in low- and middle-income countries. To allow for potential links between migration and skills, we study children who grew up in rural Pakistani villages and were first surveyed in 2003 when they were between the ages of 5 and 15 and then re-surveyed between 2017 and 2018, regardless of where they were living at that time. At this point, 38% had migrated from their native homes, so migration appears as a potentially endogenous response to the skills that respondents have acquired, which in turn may lead to differential returns to these skills. We further corroborate our findings using data from a similar sample from Cambodia, which shares key features with the data from Pakistan.

Our results show that more schooling is associated with higher cognitive skills and, to a smaller extent, greater SEM skills. Labor earnings are correlated with years of schooling, and cognitive and SEM skills conditional on years of schooling, with the size of the associations varying with migration from the

[2] See Foster and Rosenzweig (1996) on the relationship between returns to schooling and technological change in an agrarian economy; Duflo, Dupas, and Kremer (2021) on the experimentally estimated returns to secondary education in Ghana; McKenzie, Stillman, and Gibson (2010) on the returns to international migration between Tonga and New Zealand; Bryan, Chowdhury, and Mobarak (2014) on the experimental returns to seasonal migration in Bangladesh and Beegle, Weerdt, and Dercon (2011) on consumption growth among migrant and non-migrant households in the Kagera region of Tanzania.

village. There is, therefore, some evidence that schools produce the skills that labor markets reward. Yet, as we will show, the results are also consistent with the need for substantially higher school investments that can improve the production of both SEM and cognitive skills.

The two parts of our paper are presented as follows. In the measurement part of our paper, we follow Laajaj and Macours (2021) and document that our measures of SEM skills satisfy several desirable psychometric properties. These include a reasonably high Cronbach's α-statistic, which is a measure of internal consistency, and a factor structure that corresponds to the constructs that are being measured. Exploratory factor analysis identifies five factors; four correspond to the Big Five personality traits, and one corresponds to a standard measure of grit. Nevertheless, there is still room for improvement. For instance, repeated administrations of the tests show a test-retest reliability that is still below ϱ=0.7, compared to a desirable standard of 0.8, and, even though we correct for acquiescence bias or the tendency of respondents to agree with statements, it remains a concern.

Having demonstrated the properties of these measures, in the second part of our paper we then examine associations between these skill measures, years of schooling and labor earnings. In our sample, every year of schooling is associated with a 0.17 standard deviation (sd) increase in English, Mathematics, and Urdu test scores in Pakistan (with very similar results in Cambodia). Nevertheless, in 2018 when the respondents were 24 years old, 52% could not write a complex word in Urdu, 84% could not write a simple sentence in English, and 95% could not solve a simple fraction. This suggests that learning curves are relatively "flat," a point repeatedly made in the literature; in fact, positive selection into years of schooling would suggest that the true annual gain is even smaller (see Hanushek and Kimko (2000) and Pritchett (2013)). At the same time, there is considerable variation in test scores for every year of schooling. For instance, the top 5% of children who have completed Grade 5 (but no more) report test scores higher than the bottom 5% of children who have completed Grade 10 (but no more). It is this variation that allows us to separately estimate the association of labor earnings with cognitive skills and years of schooling.

Socioemotional skills are also positively correlated with years of schooling in our data, but the association is smaller than for cognitive scores (0.03sd per additional year of schooling). This correlation is also likely to be upward biased, as children with higher SEM skills (more grit, more perseverance) may be more likely to continue in school longer. In fact, parallel research by Barrera-

Osorio, de Barros, and Filmer (2018) leverages an experimental design in the same Cambodian sample to demonstrate zero causal impact of additional years of schooling on SEM skills. Interestingly, even though schools appear not central to the production of SEM skills, the mean level of SEM skills is similar to those in rich-country populations. This suggests either that population-level measures do not indicate substantial deficits or that they cannot be used for cross-country comparisons due to population-dependent reference points.

One reason why schools may not be very good at producing these skills is because they focus on *other* unmeasured skills, which are then rewarded in the labor market. Therefore, in order to evaluate this possibility, the third set of results turns to the association of our skills measures with labor force participation and earnings. Labor force participation (LFP) among men who are not enrolled in school or college at the time of the survey in 2018 is 85%, compared to 5% for women, consistent with low female labor force participation rates in South Asia (Field and Vyborny, 2016). Labor force participation increases with years of schooling for women but declines for men. This pattern could reflect different search patterns as men with more years of schooling may be "waiting it out" as they search for a better job, while women in our sample appear to have a limited window during which they can participate in the labor force before marriage. However, conditioning on years of schooling, cognitive skills are not correlated with LFP for either sex and SEM skills are correlated with LFP only for men. Schooling is important for LFP, but specific cognitive and SEM skills less so.

The lack of correlation between cognitive skills and LFP no longer holds once we turn to the correlation between labor earnings. In our data, every year of education is associated with 3.4%-4.1% higher monthly median labor earnings for men and 1.7%-1.9% higher monthly mean earnings for women (the median earning for women is zero at all years of schooling). Restricted to the sample of working adults, the estimates are 5.2%-5.3% for men and 4.8%-5.9% for women.[3] Importantly, and in contrast to labor force participation, labor earnings are correlated with both cognitive and SEM skills, conditional on years of schooling. In our preferred median regressions, a standard deviation

[3] One explanation why the association with labor earnings is smaller for men than the 10%-12% usually found in the literature is that Pakistan, like many other countries, has seen a sharp increase in the return to college education and a decline in the return to primary or secondary education (Montenegro and Patrinos, 2021). In our sample, 9% of the men are still enrolled in education (college), and of these, 80% are not yet in the labor force. If all the returns to education come from those currently enrolled in college, it is too early to pick this up in our sample.

increase in cognitive/SEM skill is associated with \$8.3/\$16.5 higher monthly earnings for men. For women, point estimates are smaller, less stable, and lack statistical precision.

This association between labor earnings and skills for men varies by respondents' migration status.[4] Among the 65% of men still resident in their native village (non-migrants), we find a precisely estimated *zero* association between labor earnings and years of schooling and cognitive skills, but a strong correlation between SEM skills and labor earnings. This result is robust to multiple specifications and sample definitions. Among men who have left the village (migrants), we find a significant correlation between labor earnings and years of schooling. The differences in the association between labor earnings and years of schooling for migrants and non-migrants reflects not just the fact that LFP is higher among migrants, but also that conditional on LFP the association between earnings and years of schooling is also significantly higher for migrants. Finally, we also find a strong correlation between labor earnings and cognitive skills for men in median regressions, but this result is less stable across multiple specifications.

To understand if the lower returns to years of schooling and cognitive skills among non-migrants is a facet of this particular sample, we turn to the second dataset from Cambodia. As in Pakistan, this sample comes from a rural area (poorer than in Pakistan) and includes children who have been tracked and re-surveyed in adulthood. Here, we again find that the correlation with years of schooling is close to zero, and although the results are more imprecise, the association with SEM skills is positive. Thus, for those who have remained in the village, the labor market appears to reward skills that are only weakly correlated with years of schooling; but for those who have left their village, the labor market rewards both the cognitive skills that are produced by schools as well as years of schooling itself. This result is reminiscent of Schultz's (1975) argument that schooling allows individuals to adapt better to changing circumstances (exemplified here by migration).

[4] The high rates of migration mirror Beegle et al.'s (2011) previous study of migration from an initially rural sample in Kagera; one difference in our sample is that 10.5% of the men are working outside the country, with more than 90% in the Arab countries of Saudi Arabia, UAE, Bahrain, Qatar, and Oman. In our sample, female migration is almost entirely due to marriage (of those who have migrated, 90% are married, and only 4% are working), while male migration primarily reflects work opportunities. Migration and occupations for men are linked—those living in the village are either engaged in daily labor, salaried occupation, self-employment (or in their family's business) or in agriculture. Men who have migrated are more likely to be in salaried employment at the expense of other occupations.

Relationship to Literature: In the United States, the returns to schooling arise, in part, from the link between earnings and cognitive skills (Altonji and Pierret, 2001). However, as Bowles and Gintis (1976) and Heckman and Rubinstein (2001) pointed out, individuals with seemingly identical years of schooling and cognitive skills are nevertheless compensated very differently in the labor market. To explain this earnings residual, they argued that schools must produce other skills that are then recognized and rewarded in the labor market. The measurement of, and returns to, these socioemotional skills has been an active area of research for at least the last two decades, with Deming's (2017) recent contribution suggesting that the returns to some SEM skills have increased over time (relative to cognitive skills). Interestingly, to return to Bowles and Gintis's (1976) original program of trying to discern the specific skills valued in the labor market and produced in schools, it is the studies from low-income countries that hold considerable promise. This context still provides the variation in years of schooling that we need to allow for the separate identification of labor market returns to years of education, test scores, and SEM skills.

Unfortunately, despite advances, research that tries to understand the link between labor market outcomes, schooling, and specific skills has proven difficult even before addressing complex identification problems. One set of issues revolves around the measurement of SEM skills in low-income countries. Laajaj and Macours (2021) show, for instance, that the factor structures observed in the United States do not fit the data from low-income countries and that measures of SEM skills suffer from both low reliability and validity concerns. Similarly, when estimating returns to the seven skills measured in the World Bank STEP surveys, Valerio et al. (2016) point to serious reliability concerns.[5] Where SEM skills have been measured successfully, such as the longitudinal Young Lives project in Ethiopia, India, Peru, and Vietnam, the panel of respondents is still too young to enter the labor force. These data are already providing important insights into the importance of SEM skills, but as of yet, they do not include labor market outcomes; only the latest (yet-to-be-released) round of

[5] They find that four coefficients are positive and statistically significant at the 90% level of confidence, three coefficients that are negative and statistically significant and 42 coefficients that are small and imprecise. As the standard errors are not adjusted for multiple-hypotheses testing, even the statistically significant results are likely due to chance. The authors write, "*It must be noted that the noncognitive skills measures are a function of scores on three to five items each. We believe the limited number of items for each (noncognitive skill) scale could be limiting the reliability of these measures and obscuring the true relationship between noncognitive skills and earnings.*" (Valerio et al., 2016: 26).

Young Lives surveys will allow researchers to examine the links between labor earnings, schooling, and skills.[6]

The studies from low- and middle-income countries closest to our work are Glewwe, Huang, and Park (2017) and Glewwe, Song, and Zou (2022), who follow a sample of children in rural China. In 2017, they found no correlation between labor earnings and cognitive or SEM skills. By 2022, earnings were correlated with cognitive skills after controlling for years of schooling, but SEM skills were not. This scholarship complements ours by offering insights from a different context. Yet, the studies also differ from ours in their ability to track students over time: Glewwe, Huang, and Park (2022) have some data on 67% of their original sample and face-to-face interviews on just under 50% (compared to 84.5% and 75.1% in our case). In addition, they focus less explicitly on the (potential) heterogeneity in returns by migration status.

It is in this context that our contribution adds value. We first show that our lengthy tool development process produced measures of SEM skills in low-income countries that satisfied desirable measurement properties. We then document the associations between these skills and labor earnings, highlighting the critical role of migration in a sample that started in rural areas but moved to multiple locations over 15 years. As other longitudinal studies either do not (yet) observe labor market outcomes (e.g., Young Lives) or do not include well-constructed measures of SEM skills, this is only the second long-term research project to establish a pattern of correlations between labor earnings and specific cognitive and SEM skills, in a less-developed country. We hope this descriptive evidence helps resurrect the fundamental question raised by Bowles and Gintis (1976), this time for low-income countries: What are the skills that children learn in schools that are then valued in the labor market?

I. SAMPLE

I.1. SAMPLE SELECTION

The data come from the Learning and Education Achievement in Punjab Schools (LEAPS) project, a longitudinal study of education in Pakistan (Andrabi, Das, and Khwaja 2022). In 2003, the LEAPS project randomly sampled 112 villages from three districts in the province of Punjab (from a list frame

[6] Using the Young Lives data, Mitchell et al. (2020) find that individuals with higher task effectiveness skills are less likely to engage in risky behavior. Singh et al. (2018) show that psychosocial skills are positively correlated with progression through school. Das, Singh, and Yi Chang (2022) show that test scores at age 12 predict years of schooling at age 22 but still leave a substantial role for socioeconomic status.

of villages with at least one private school). The sampled villages were richer and larger than the average village; however, approximately 70% of Punjab's population lived in such villages at the time. As part of the first survey, 1,807 households were surveyed, including information on 5,865 children between the ages of 5 and 15. These households were then revisited four times between 2004 and 2011.

Between 2016 and 2018, we attempted to contact and resurvey the children of the 2003 sample. Over two years of tracking, we completed in-person surveys for 75.1% of the sample. We have information through phone surveys or from a third-party respondent for another 9.4%. We do not have skill measurements for these "indirect" surveys, although we have data on many other outcomes of interest, including years of education, earnings, and migration. Thus, of the 5,865 children, we have at least partial information on 4,956 children, or 84.5% of the original sample, at the individual level. The implied annual attrition rate of (just above) 1% compares favorably to 10-year panels with the highest retention (Outes-Leon and Dercon, 2009).

Appendix A and Table A1 detail the tracking process, the different instruments used, and the types of attrition in the data. Of the 909 individuals on whom we have no data, 43 had died, 186 were living in four villages that fell into a military zone that our team could not access, 395 respondents refused to participate despite multiple attempts, and 285 could not be located. For an additional 550 individuals, the information does not come from face-to-face interviews with the respondent but from third-party surveys or phone calls. We therefore do not have skills measures for these individuals.[7] Respondents with indirect information and phone surveys were more likely to live outside the village at the time of the survey and more likely to be working (Table A2, Panel A). Attritors are less likely to have ever been married and more likely to be living outside their original district and outside the country (Table A2, Panel B). They were also poorer than other households and more likely to have a father living abroad in 2003 (Table A2, Panel D). We account for these differences using a variety of weighting schemes, Heckman selection models, and semi-nonparametric estimation techniques, all described in Appendix B.

[7] They also include 15 respondents who answered the survey in person but whose skills measures cannot be used: 4 respondents did not finish the survey and 11 respondents for whom there was a bug with the test on tablets.

I.2. SAMPLE CHARACTERISTICS

Table 1 summarizes the characteristics of respondents in our sample (Panel A) and compares them to their *parents* in 2003 (Panel B). Both men and women in our sample report over eight years of education, compared to just above three years for the parents; 73% of our sample report that they can read, compared to 37% among the parents. These statistics reflect the well-known and often dramatic improvements in schooling participation over the last two decades (World Bank, 2018).

The changes are just as significant with respect to broader social and occupational regimes. For instance, the age of marriage for women will be at least 22, compared to 19.7 for their mothers. In addition, the share of men working in agriculture has plummeted from 33% among the fathers (in 2003) to 7% among the sons (in 2018). There have also been improvements in living standards: 96% of our sample reports having a toilet on their premises in 2018, and 98% report having access to electricity (compared to 58% and 88%, respectively, in 2003). However, one statistic that remains low and unchanged across generations at 5% is female LFP. Multiple authors have commented on the low and declining LFP among South Asian women, and our data are consistent with these findings from that literature (see Afridi, Dinkelman, and Mahajan (2018) for India, Field and Vyborny (2016), and Subramanian (2020) for Pakistan).

As in other contexts (see Beegle, De Weerdt, and Dercon (2011)), migration has emerged as a critical feature of our respondents' lives. In the sample, 38.7% of men and women no longer reside in their birth village (Figure 1). Of the original sample of men, 10.5% now live outside Pakistan (mostly in Arab countries), 16% live in Pakistan outside their native district, and 65% remain in their native village, with the rest migrating within their native district. The farther men are from their village, the more likely they are to be salaried and the less likely they are to rely on daily wages, their own or family business, or agriculture (Table A3). Earnings are also higher for those who migrate—median/mean monthly earnings for respondents in their original village is USD \$115/\$139, compared to \$173/\$192 for respondents living outside the district and \$337/\$380 for those living outside the country (Figure A1). In the sample of men, more than half (54%) of income stems from the approximately one-third (35%) of men who left their birth village; and about a quarter (27%) comes from those who have left the country. For women, on the other hand, migration is closely tied to marriage and the practice of virilocal residence. Among those who have migrated, 90% are married, in contrast to 31% among those still residing in their original village. Most female migration is within the same district, and only 0.5% (11 women) of the female sample lives outside Pakistan.

II. MEASURING SKILLS

In our 2018 survey, we measured two types of skills that, following the literature, we refer to as cognitive and socioemotional (SEM) skills. Here, we describe the instruments we used to measure each type of skill and the reliability of the respective measures (Table A4 provides an overview of the instruments).

II.1. COGNITIVE SKILLS

Our measurement strategy aimed to capture two sets of cognitive skills. First, we measured English, Mathematics, and Urdu skills that are commonly taught in schools. Second, we measured respondents' proficiency in everyday arithmetic and literacy skills (which may not necessarily be taught in schools; see Banerjee et al., 2022).

Our measures of commonly taught cognitive skills include the same tests used previously in the LEAPS project. These tests are norm rather than criterion-referenced and designed to cover a wide range of topics. Andrabi et al. (2002) and Bau, Das, and Yi Chang (2021) show that the LEAPS tests satisfy the requirements of horizontal and vertical linking (that is, the function that relates a respondent's ability to their likelihood of correctly answering a question is stable across test takers and over time).[8] In the limited number of items where vertical linking does not function well, eliminating the unstable items does not lead to any appreciable difference in the test scores.

As the original LEAPS tests targeted primary-school-age children, we worried that ceiling effects would censor the cognitive skills distributions of the resurvey. Therefore, we worked with an educational organization to design an adaptive test administered on tablets. Each assessment started with a set of simple questions; the difficulty of the following items increased or decreased based on the respondent's performance. The test classified respondents into six levels (Level 1 corresponding

[8] Here, we do not vertically link the scores as we use results from an additional adaptive test as well. Andrabi et al. (2002) discuss test construction and assess the psychometric properties of the original test administered in 2003 sample. Bau, Das, and Yi Chang (2021) further assess vertical linking in the LEAPS test and demonstrate that there was limited differential item functioning between 2003 and 2011. That is, the function relating the latent variable (ability or knowledge) to the likelihood of answering the question correctly remained stable across the years.

to Early Primary and Level 6 to College). Appendix C.1 presents the progress and placement logic of the test.[9]

However, contrary to our expectations, ceiling effects in the LEAPS test were small, with 11.2% of children achieving the maximum in English, 2.7% in Mathematics, and 13.1% in Urdu. There was considerable variation in test scores, with a mean of 51.5%, 54.2%, and 55.2% correct in English, Mathematics, and Urdu, and a standard deviation of 33%, 33.6%, and 34.4%, respectively. In contrast, the adaptive test elicited little meaningful variation: 43%, 64%, and 77% of respondents were classified as Level 1 for Urdu, Mathematics, and English, respectively, and the rest were largely classified as Level 2.

We aggregated all the items from both types of assessments into Urdu, Mathematics, and English scores, using Item Response Theory (IRT) with a two-parameter logistic (2PL) model.[10] Formally, each item characteristic curve is given by the two-parameter logistic:

$$P_j(\theta) = \frac{1}{1 + \exp\{-a_j(\theta - b_j)\}}$$

where $b_j \equiv \theta^* | P_j(\theta^*) = \frac{1}{2}$ is the difficulty parameter, which is the ability level at which the child will answer a given question *j* correctly half the time and $a_j \propto \frac{\partial P_j(\theta)}{\partial \theta}$ at $\theta = b_j$, is the discrimination parameter, which specifies the steepness of the item characteristic curve at the point that the ability of the child is equal to the difficulty of the question (b_j). The joint estimation of θ and these parameters follows the standard maximum-likelihood procedure in IRT.[11] We then assess model fit by comparing actual with predicted item responses based on estimated item parameters and the model assumptions of the 2PL model for each subject (Appendix C.2, Figures A3 to A5 for the LEAPS test, and A6 to A8 for the adaptive test). For most items, the actual and predicted responses match closely, although

[9] The organization that designed the adaptive test developed 324 items ranging from early primary level to college level. The test classified respondents into six groups corresponding to different grades. The mapping between level and grades is as follows: Level 1: Nursery, Grades 1 to 3 (early primary); Level 2: Grades 4 and 5 (late primary); Level 3: Grades 6 to 8 (middle school); Level 4: Grades 9 and 10 (high school); Level 5: Grades 11 and 12 (intermediate); Level 6: College. Like with the LEAPS test, ceiling effects were minimal. In English, Mathematics, and Urdu, 0.3%, 0.2%, and 7% of respondents placed at Level 6. In English and Mathematics, fewer than 3% of respondents were placed in Levels 3 to 6. In Urdu, 12% are placed at Level 3, and 5-7% at each subsequent level.

[10] We exclude items that less than 50 respondents answered as well as items that less than 5% or more than 95% of respondents got the correct answer.

[11] We use Stata's OpenIRT command (see Das and Zajonc, 2010).

there are a few items in the adaptive test where the fit is poor (for instance, items 62, 96, and 107 in Urdu). Re-estimating the model after eliminating these poorly fitting items does not alter the overall estimated score.

One concern with the LEAPS assessment is that it may not adequately reflect functional literacy and numeracy as documented by Banerjee et al. (2022). Therefore, we designed a separate assessment to capture respondents' proficiency in everyday arithmetic and literacy skills. The math questions in this assessment of functional numeracy asked respondents to read an electricity bill, compute the correct amount given arrears (easier) and then recompute the right amount given electricity consumption and non-linear pricing (harder). Another math question assessed competency in a marketplace transaction where respondents purchased multiple items and collected change. To assess functional literacy skills, we asked respondents to read several messages written in Urdu and Roman Urdu (Urdu but using roman language script). We also assessed whether the individual knew how to use a phone by asking them to save a contact on a phone. Appendix C.3 details these items. We argue that someone who can complete these tasks can be considered "functionally" literate; accordingly, we use principal-component factor analysis to aggregate these questions into a single "functional literacy and numeracy" index.

II.2. SOCIOEMOTIONAL SKILLS

Measuring socioemotional (SEM) skills in low-income countries has proven difficult, with evidence of non-classical measurement error for self-reported instruments (Laajaj and Macours, 2021). Consequently, we developed the instruments included in the Pakistan survey through an iterative process. We started with data collection in Cambodia for a related project in 2017 (we will present comparative results from that study in Section III.2.2). As in Pakistan, the sample in Cambodia also consisted of young adults from rural regions. The data collection incorporated a comprehensive assessment of socioemotional skills. However, despite our efforts to mitigate the issues later discussed by Laajaj and Macours (2021), the self-reported scales included in that survey displayed limited internal consistency (see Appendix D).

We built on that experience when designing and implementing our SEM skills assessment in Pakistan in several ways. First, in addition to self-reported measures, we developed bespoke applications on tablets. We assumed these tablet-based measures would be less subject to biases arising from social desirability (the tendency to over-report socially valued attitudes) and acquiescence (the tendency to agree with yes/no questions, regardless of their content). We then conducted a pilot with 403 respondents and (a) included debrief sessions to gauge respondents' understanding of the material and (b) randomly re-surveyed half (201 respondents) two weeks later to assess the reliability of our measures in repeated administrations of the same test.

Following Laajaj and Macours (2021) and the psychometrics literature, we used three main criteria to select the instruments in our survey: face validity, predictive validity, and reliability. The pilot first allowed us to assess the face validity of our tools to ensure that the questions were perceived as measuring the concepts we intended to measure. Following literature from the United States that establishes a link between SEM skills and earnings, we also assessed their predictive validity by calculating the bivariate correlations of each score with years of schooling and earnings (Cunha and Heckman, 2010; Brunello and Schlotter, 2011).

Finally, we computed two reliability estimates: test-retest reliability and internal consistency. Multiple measurements allow us to estimate the test-retest reliability (the correlation of the same measures in repeated test administrations). Under the assumption of classical measurement error, the test-retest correlation estimates the share of the variance of a measure that is explained by the true latent trait we are trying to capture (rather than by measurement error). Specifically, if the measured value X is the true value X^* plus a measurement error ε, $X = X^* + \varepsilon$, then the test-retest correlation is an estimate of the reliability defined as $Relability = \frac{\sigma_{X^*}}{\sigma_X}$. Generally, a value of at least 0.7 is considered to indicate acceptable levels of reliability (and some researchers prefer a value of 0.8 or larger).

To assess internal consistency, we first computed Cronbach's α-statistic, a measure used in psychometrics that indicates the inter-correlation of the items on a scale, commonly interpreted as the extent to which the items of a scale measure the same underlying concept. Cronbach's α-statistic is computed as $\alpha = \frac{K}{K-1}(1 - \frac{\sum_{i=1}^{K} \sigma_{Y_i}^2}{\sigma_X^2})$ where K is the number of items in the scale, σ_X^2 is the variance of the observed total test score, and $\sigma_{Y_i}^2$ is the variance of responses to item i for the current sample of persons. The statistic is a ratio of variances and therefore lies between 0 and 1; a rule of thumb for a measure with high internal consistency is that Cronbach's α should be above 0.7 (and 0.8) (see

Nunnally, 1978). We follow this heuristic with two notes of caution: Cronbach's α may be high due to systematic response biases that lead to a high inter-item correlation, even after correcting for acquiescence bias, and the statistic mechanically increases as the number of items in a scale increases. These results from the pilot and additional details on the tools used to assess the measurement properties of our instruments are reported in Table 2 and Appendix D. Based on these results, we retained two self-reported scales and two task-based scales (administered on tablets).

The first self-reported scale is a 10-item measure of grit—the combination of passion and perseverance for long-term goals—as developed by Duckworth and Quinn (2009).[12] The second self-reported scale measures the "Big Five," a taxonomy of traits that encompasses five dimensions of personality: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. To measure these traits, we used the short 15-item Big Five Inventory (Lang et al., 2011), which consists of three items for each of the five personality traits. All items use a 5-point format ranging from 1 ("Disagree strongly") to 5 ("Agree strongly"). Following Laajaj and Macours (2021), we applied an acquiescence bias correction to account for respondents' tendency to agree with a statement (see Appendix D.1).

The two self-reported scales show reasonably high internal consistency, with Cronbach's α just above 0.7 for the Grit scale and just below 0.7 for the Big Five (Table 2). Cronbach's α's for the Big Five subscales range between 0.53-0.68, even though each scale only comprises three items. As a comparison, in rural Kenya, Laajaj and Macours (2021) found Cronbach's α's for the same constructs ranging from 0.31 to 0.51, with 4 or 5 items per sub-scale. Also, in contrast to Laajaj and Macours (2021), the skills factor structure is closely reproduced in our data (Table 3).[13] Exploratory factor analysis identifies five factors. The Grit scale's items all load on one factor, then the items corresponding to the openness to experience, extraversion, and emotional stability sub-scales of the Big five load on three distinct factors. The only deviation from the original skills structure is that the conscientiousness and agreeableness items load on a common factor. Finally, the acquiescence bias in our sample is 0.26 to 0.31, slightly below the 0.37 reported by Laajaj and Macours (2021) for Colombia. Overall, compared to previous studies, we interpret these results as showing better measurement

[12] The scale can be found here: https://angeladuckworth.com/grit-scale/. There was a mistake in translating the tool from English to Urdu, so only nine items were implemented.
[13] We rotate factor loadings so that each variable loads (mostly) on one factor.

properties for our measures of self-reported SEM skills. Nevertheless, even with this extensive process, Cronbach's α values were just around the acceptable threshold, and the test re-test correlations in the pilot for the chosen items were low. Thus, even if we are measuring the "right" constructs, there is still considerable measurement error.[14]

The second set of SEM skills measures we used was administered on tablets. Contrary to commonly held beliefs among education researchers, but similar to Boon-Falleur et al. (2022), two task-based measurements of grit did not work well, with either a low test-retest correlation of 0.27 or a level of required mathematics ability that was not suitable for our respondents (single- and two-digit addition for a task designed to measure grit by Alan, Boneva, and Ertac, 2019). We dropped these two measures from our final assessment and retained two other tasks.

One of the two tasks we retained was the "GoNoGo" task, used to measure impulse control, which had a test-rest correlation of 0.78 (Table 2). The participant is presented with a square on the screen for a brief period. If the square is of any color but black (the "go" stimulus), the participant must touch the screen as quickly as possible. If the square is black (the "no go" stimulus), the respondent must inhibit their response. A total of 72 trials are completed (48 Go and 24 NoGo trials), and our main outcome is the average response time. The second task was the Balloon Analogue Risk Task ("BART"), which measures risk-taking behavior by asking participants to maximize the amount of money they can win from a game. In each trial, respondents are presented with a balloon they can pump. Each pump earned them (real) money but increased the likelihood that the next pump would "pop" the balloon, in which case they lost the accrued cash for that balloon. If they instead chose to stop pumping the balloon, they collected their accrued money and moved to the next trial. The main outcome is the average number of pumps on the balloons that did not explode, and respondents earned, on average, PKR 322 or $3 from the game. While the BART displayed limited reliability (the test-retest correlation was 0.36), it was easy for respondents to understand and provided a measure of risk aversion. As with other skill categories, we aggregate the items from the self-reported scales and the scores from the tasks on tablets using principal-component factor analysis. We also verified that two alternative versions of the index—one in which we only keep the self-reported scales and one in which we further drop items that have poor properties (uniqueness > 0.65) as per the factor analysis—

[14] Interestingly, except for the extraversion subscale of the Big Five, Cronbach's α for the re-test is higher than the one from the test, suggesting that repeated exposure may improve comprehension.

resulted in similar aggregate scales (with correlations of 0.98 and 0.99 with our preferred SEM skills index).

A key lesson from Laajaj and Macours (2021) is that measurement tools developed in high-income countries may exhibit poor reliability and validity in low-income countries. Despite our extensive and iterative approach to building the SEM skills measures, a straightforward response to this challenge remains elusive. For instance, task-based measures may seem attractive *ex-ante*, but they are not necessarily more reliable when education levels are low and with high variability. Instead, our assessment, which builds on Laajaj and Macours's (2021), suggests that more extensive enumerator training, more piloting to aid the selection of tools, and better translations can help mitigate measurement error and response bias (to a limited degree). The final assessments we employ in our survey perform better on multiple measures but still have low levels of test-retest reliability (at least in the pilot). Even if this measurement error is classical, it will remain an important source of attenuation bias when these measures are used as explanatory variables.

III. SCHOOLING, SKILLS, AND LABOR EARNINGS

Having provided evidence for the validity and reliability of our skills measures, we now turn to the associations between schooling, skills, and labor earnings. We are interested in assessing, first, the extent to which schooling is linked to the production of these skills and, second, how cognitive and SEM skills are, in turn, associated with labor force participation and labor earnings, conditional on years of schooling. Throughout, we pay close attention to the role of migration. In documenting these associations, we often aggregate different cognitive and SEM skills into simple averages. This approach simplifies our exploratory analysis and reduces measurement error, but it comes at the cost of ignoring finer distinctions between different skills. We comment on individual skills measures and their correlations with earnings in Section III.2.

III.1. SCHOOLING AND SKILLS

Most of the young adults in our sample are between the ages of 19 and 28 (with an average age of 24). They can count and identify numbers (79%), and a majority can add 3-digit numbers, but only 5% could express the simple fraction $7/3$ as $2\frac{1}{3}$ from among multiple options (Table 4). For the vernacular, Urdu, the majority can write simple words but cannot fill in a blank in a story by selecting the correct word. For English, respondents can match pictures to words but cannot write simple sentences, such as a sentence that uses the word "deep" ("The water is deep" would be graded correctly). As reported previously, our adaptive tests showed that 43%, 64%, and 77% of the young adults had only mastered materials commonly taught in Grades 1 to 3 for Urdu, Mathematics, and English, respectively (Level 1). Only 26%, 34%, and 20% fall in the level corresponding to Grades 4 and 5 (Level 2). These low skill levels imply ceiling effects are limited even though the original LEAPS test targeted children in primary school. Despite the extensive literature pointing to severe learning deficits among students of schools in low-income countries (including Hanushek and Kimko (2000), Pritchett (2013), and The World Bank's World Development Report (2018)), it bears repeating that the learning deficits are so severe that young adults, with an average age of 24 and 8.5 years of schooling, do not top-code on a test designed for third and fourth graders. Even more worryingly, among those young adults with at least some college education, only 4.5% top-coded in the LEAPS test for all three subjects.

Figure 2 then shows the relationship between years of schooling and cognitive skills. Table 5 presents the regression equivalent for all our skills measures, with and without village fixed effects, as villages where the quality of education is higher (perhaps because returns are higher) will both have greater educational attainment and higher cognitive skills. Figure 2 and Table 5 first show that more years of schooling are associated with higher cognitive skills, with stable coefficient estimates of 0.17sd for each additional year of schooling (in specifications with and without village fixed effects). Further, Figure 2 shows considerable variation in cognitive skills for every level of schooling. This finding is a feature of the data we will exploit when examining earnings-skills correlations further below.[15] Figure 2 also suggests that individuals learn less in college (0.28sd increase) compared to earlier schooling, where the moves from primary to middle and from middle to secondary school are each associated with a 0.7sd increase in cognitive skills. This observation is puzzling since gains during the college years reflect a combination of the causal impact of college, the selection into college, and any depreciation in cognitive skills after leaving school. All of these forces should increase the learning gains during college in our data: Selection effects should be stronger for those attending college, and

[15] This result also suggests schooling does not necessarily serve a "selection" or "sorting" function (contradicting Muralidharan (2019), for instance)—educational attainment serves as a poor predictor of cognitive skills at any level of schooling.

children who only completed Grade 5 in our sample would have left school ten years before our survey. Therefore, we may have expected their skills to depreciate.[16]

To investigate this further, we first focused on a sub-sample of children tested multiple times. We find that, for this group, test scores were highest in 2011, when respondents were 17 years old, and then declined by 2018 (Table A5). Surprisingly, although youth who went on to college increased their test scores on every question, there were still basic concepts they did not understand in each of the three subjects. We then looked at depreciation by comparing the cognitive skills index of children who report the same number of schooling years but differ in age—those who are older would have graduated sooner (Table 5). We find that depreciation is small in our data: conditional on years of schooling, the cognitive skills index of a respondent who is a year older (and therefore left school one year earlier compared to another child with the same years of schooling) is only 0.009sd lower. The data suggest that the children who attend college are not necessarily the most selected, and college attendance may not significantly increase cognitive skills (in line with earlier findings by Loyalka et al. (2021) and Bau, Das, and Yi Chang (2021)). Overall, schooling is associated with the development of foundational skills, but acquiring these skills becomes harder in later grades and college, and the children in our sample remain far from mastering more advanced content, even after completing college.

Performance in functional literacy and numeracy was also poor, but again these skills were positively associated with years of schooling (Table 4, Part 1, Panel B, and Table 5). For Mathematics, 80% could read an electricity bill and calculate how much money they owed. However, 50% had difficulty with non-linear pricing in their utility bill, and 36% could not compute the correct change from a market transaction with five items (with five different quantities and prices). The picture is more nuanced for reading (particularly for the English alphabet). Even though 55% could read a complicated text ("*Peace be upon you. How are you and how is everyone at home?*") in Roman Urdu, which is Urdu written in the Roman script in texting language, they could not read the word "dog" in the Roman script. Perhaps the classification of the former as Urdu, not English, allowed respondents to discern the question differently. For Urdu, 73% could read complex text accurately (in the Urdu script). Thus, literacy

[16] Another option is that we are not capturing other skills that colleges impart. However, the skills we are measuring are basic—knowing how to read a paragraph in Urdu or English and how to perform basic arithmetic in Mathematics. It is quite difficult to see what higher-level skills would not require this level of foundational learning.

performance on the test of functional reading skills appears higher than literacy performance measured with the written LEAPS test. Even so, the two measures are strongly correlated (with a coefficient of 0.77), and every year of schooling is associated with a similar increase in functional skills (0.17sd; see Table A6 and Table 5).

Unlike the substantial deficits we document for cognitive skills, the levels of SEM skills in our data compare to those found in richer countries (Table 4, Part 2). Thus, while low levels of cognitive skills in the population are consistent with the narrative that countries with lower GDP are also those with lower human capital, population-level measurements of SEM skills do not support this conclusion. This finding could reflect different reference points, suggesting that these SEM measures are population-dependent; however, it could just be that SEM skills in rural Pakistan are similar to those in wealthier countries. What is unlikely is that schools produce these skills in the first place. At the individual level, every additional year of schooling predicts only a .033sd increase in SEM skills, and even this small correlation potentially reflects reverse causality (Table 5). In fact, Barrera-Osorio, de Barros, and Filmer (2018) demonstrate experimentally that there is zero causal impact of schooling on SEM skills in the Cambodian sample that we discuss later. Depending on how seriously we wish to take the cross-country comparability of SEM skills, it thus seems that this population acquires skills similar to those in high-income countries, but schools are not where these skills are primarily produced.[17]

The acquisition of these different sets of skills varies by gender and age. Women report higher cognitive skills but lower functional and SEM skills.[18] Cognitive skills appear to depreciate slowly with age, while SEM skills appear to increase with age. These patterns suggest that cognitive and SEM skills are different abilities that people bring to the labor market, with different processes of skills acquisition during and after the schooling years. Years of schooling, cognitive and functional skills are all highly

---

[17] We also investigated correlations between years of schooling and different SEM skills, reported in Tables A7 and A16. Using Romano-Wolf p-values to correct for multiple hypotheses testing, we find that conscientiousness, grit, self-control, risk-taking, openness to experience and emotional stability are all positively correlated with more years of schooling. Of these, grit and openness to experience have the largest coefficient estimates, but even here, associations of 0.035 and 0.042 are small in magnitude. Patterns in Cambodia are different—for instance, the coefficient on grit is no longer significant in Cambodia, although that for openness to experience is again the largest in magnitude.

[18] A possibility for this gender differential that we investigated but found limited evidence for was that this was linked to mobility restrictions for women. Specifically, we constructed a mobility index based on the family of birth's discomfort with the respondent traveling outside the village or talking to people they did not know. Including this mobility index as an additional regressor reduced the female deficit in SEM skills from -0.62sd to -.53sd.

correlated, with correlation coefficients ranging between 0.77 and 0.81 but SEM measures capture a different part of the skills set with lower correlation coefficients of 0.14 to 0.18 with cognitive and functional skills respectively (Appendix Table A6). Finally, the inclusion of village fixed-effects (Table 5, even columns) explains little variation in skills measures; it results in virtually no change in the R-squared or the coefficient on years of schooling. This is a remarkable finding as there are large differences in consumption aggregates across these villages, ranging in 2003 from PKR 31,105 at the 10^{th} to PKR 81,718 at the 90^{th} percentile. It suggests that years of schooling are not higher in villages where the association between schooling and skills is stronger.

III.2. LABOR EARNINGS, YEARS OF SCHOOLING, AND SKILLS

Having examined the correlation between years of schooling, cognitive skills, and SEM skills in our sample, we now turn to the correlation between labor earnings and these variables. Given dramatic differences in labor force participation, of 85% among men and 5% among women, we present specifications relating labor market outcomes to schooling and skills separately for men and women. We consider three main outcomes: labor force participation, labor earnings, and migration (with details regarding the measurement of earnings presented in Appendix E). We treat years of schooling, cognitive skills, and SEM skills as conceptually separate. We exclude functional skills because they are highly correlated with cognitive skills; moreover, where the skills profiles differ (women report higher cognitive but lower functional skills), the samples are generally too small to pick up these nuanced differences. Our preferred specification excludes those currently enrolled in school/college and includes age and district fixed effects. In Section III.2.2., we assess the robustness of our findings to different sample definitions, attrition, and the treatment of agricultural income.

III.2.1. Main results

We present three sets of descriptive correlations in Tables 6 to 8. In each table, we begin by estimating

$$y_i = \alpha + \beta s_i + \sum_{j=1}^{N} \gamma_j AgeFE_j + \varepsilon_i \quad (1)$$

, where y_i represents the outcome of interest; s_i is the years of schooling, and $AgeFE_j$ are age fixed effects (one for each age *j*). One estimation includes all the individuals; the remaining estimations exclude those without skills measures. We then include measures of cognitive and socioemotional skills. The respective specification is

$$y_i = \alpha + \beta s_i + \delta \, Cog_i + \gamma \, SEMS_i + \sum_{j=1}^{N} \gamma_j AgeFE_j + \varepsilon_i \quad (2)$$

where Cog_i and $SEMS_i$ refer to the cognitive and socioemotional skills indices. Finally, for men, we look at associations between earnings and skills depending on where the respondent lives in a fully interacted specification (the sample of women working outside the village is too small to investigate this channel)

$$y_i = \alpha + \beta_1 \, s_i + \delta_1 \, Cog_i + \gamma_1 \, SEMS_i + \varphi \, Out_i + \beta_2 \, s_i * Out_i + \delta_2 \, Cog_i * Out_i + \gamma_2 \, SEMS_i * Out_i + \sum_{j=1}^{N} \gamma_j AgeFE_j + \varepsilon_i \quad (3)$$

where Out_i represents an indicator of whether the respondent lives outside the village. In mean regressions, standard errors are clustered at the village level. In the median regression, we report robust standard errors. As is common, labor earnings in our data are highly skewed for both men and women (skew = 13.6 for men and 7.8 for women); therefore, mean regressions may be particularly susceptible to outliers in the earning distribution. We used three different methods—standardized residuals, Cook's distance, and DIF-Beta—to identify potential outliers and assess the sensitivity of our results in specifications for men. Appendix F describes these methods and shows that median and mean regressions yield similar estimates when we remove highly influential observations using either one of these methods. Given these results, in the main tables we report estimates from the median regressions for men, both for labor earnings in the full sample and the sample of men who report working. For women, when we report associations with labor earnings in the full sample, we continue to use mean regressions since the median labor earning is zero. In addition, we report both median and mean regressions for the restricted sample of women who report participating in the labor force.

There are four main results. First, LFP increases with years of schooling for women but declines for men (Table 6). As the specifications control for age fixed-effects, these coefficients do not reflect that respondents with fewer years of schooling will have left school earlier and therefore have been in the labor market longer. Instead, the negative years-of-schooling coefficient likely reflects greater search durations for men, including preparing applications for public sector jobs or waiting for job offers from outside the village and country. In contrast, women in our sample are limited in terms of geographical mobility, and their participation in the labor force plummets with marriage so that women who wish to work have only a limited window to do so (Afridi, Dinkelman and Mahajan, 2018, Field and Vyborny 2016). For women, the size of the coefficient relative to the baseline LFP is striking: Among women with primary schooling or lower, female LFP is 2.4%, but among those with post-secondary education (including those currently enrolled), it rises to 18.8%. This difference in female LFP compares, for instance, to an effect size of 4.9 percentage points in a program that is considered successful in improving women's labor market engagement in a similar context (Bandiera et al. 2020). Importantly, for both men and women, cognitive skills are not correlated with LFP; but for men, a standard deviation increase in SEM skills is associated with a 6-percentage-point increase in LFP (Table 6).

The second result confirms that more years of education are associated with higher earnings for both men and women (Tables 7 and 8). As we are interested in the correlation of our skills measures with total labor earnings, we do not regard this as an estimate of the Mincerian return and include estimates for our entire sample, which capture both the wage and participation effects, as well as associations for the sample of working men and women only. For men, median regressions show that each year of education is associated with $4.1 higher monthly earnings, which translates to 3.4% of the $120 monthly earnings reported in our sample (Table 7).[19] This is smaller than the 10-12% usually found in the literature using similar OLS specifications (Montenegro and Patrinos, 2021) and could reflect that labor earnings are much higher for those attending college in Pakistan. Unfortunately, our sample is too young to observe labor earnings after college completion (and those studying in college report not working).

Even though the correlation between labor earnings and years of schooling does not reflect the experience of those who have studied at the college level, for men, we find that cognitive and SEM skills are highly predictive of labor earnings. Our estimates suggest that, *conditional* on years of schooling, every standard-deviation increase in cognitive skills is associated with an $8.3 increase in median monthly earnings, and a standard-deviation increase in SEM skills is associated with a $16.5

[19] If we consider the mean regressions instead, each year of education is associated with $5.2 higher monthly earnings, which also translates to 3.4% of the $155 mean monthly earnings reported in our sample.

increase in median monthly earnings (Table 7).[20] Further, once we include cognitive and SEM skills as independent variables, the coefficient for years of schooling reduces substantially, confirming that the correlation with schooling captures, in part, the association between earnings and cognitive and socioemotional skills. Finally, we find similar patterns when we restrict our specifications to men who are working, except that the association with SEM skills is smaller (columns 5-8, Table 7).

For women, each additional year of education is associated with a $1.7 monthly increase in earnings, which is 24% of the sample mean of $7.2 (Table 8). Once we focus on working women only (N=111), each additional year of education is associated with an $8.5 monthly increase from a baseline mean of $107, or an 8% increase.[21] The small sample size of working women leads to a high degree of imprecision in the association between earnings and skills however, with coefficients changing signs across mean and median specifications (and statistical insignificance throughout, at conventional confidence levels).

Third, for men, the association between earnings, years of schooling and skills is mediated by the migration status of the respondent (for women, the sample of those working outside the village is too small). One channel that links skills and migration is a higher propensity to migrate for those with higher skills (Appendix Table A8). For men, although years of schooling is not associated with migration, higher levels of cognitive and SEM skills are both associated with a significantly greater likelihood of leaving the village (but not the country).

A second channel through which migration affects the returns to skills is more surprising: our findings suggest the labor market rewards skills differently depending on whether people work within the village or outside. Specifically, for those who chose not to leave the village by the time of the resurvey, there appears to be a precisely estimated *zero* correlation with years of schooling and a strong correlation with SEM skills (Table 7, Column 4). For those who have migrated, the results reverse—the association with SEM skills is statistically insignificant (although more imprecise), and the association with years of schooling is higher, with an additional year of schooling predicting a $4.8

[20] If we consider the mean regressions instead, *conditional* on the years of schooling, every standard deviation increase in cognitive skills is associated with a less precisely measured $6.3 increase in monthly earnings and a standard deviation increase in SEM skills is associated with $14.6 monthly higher earnings.
[21] If we consider the median regressions instead, each year of education is associated with $5.9 higher monthly earnings, which to translates to 9.4% of the $62.5 median monthly earnings reported in the sample of working women.

increase in median monthly earnings. The difference across migrants and non-migrants reflects both the lower labor force participation of non-migrants and differences in the association between years of schooling and earnings conditional on participation. Columns 5-8, Table 7, where we focus only on the sample of men who are working, shows that the association between labor earnings and years of schooling as well as earnings and cognitive skills is significantly lower for non-migrants relative to migrants. Finally, when we discuss the robustness of these findings, we will show that the association of earnings with cognitive skills is more sensitive to whether we look at the mean or the median; the most consistent patterns are that the association between years of schooling, SEM skills and earnings are very different for migrants and non-migrants.

One potential concern is that the interpretation of the regression specification changes if cognitive and SEM skills reflect the acquisition (or depreciation) of skills on the job. To assess this possibility, we computed a "brain" and "brawn" index as well as a teamwork index based on answers to questions about job content and included these as additional regressors in Columns 6 and 8 of Table 7.[22] Differences in the coefficients on skills are consistent both with skills changing on the job and the sorting of skills to jobs, as in a Roy model. Therefore, the resulting coefficient on skills will be a lower bound on the associations between labor earnings and skills. As we would therefore expect, respondents who report that their jobs require more "brain" and less "brawn" also report higher earnings, as do those who report that their jobs require more teamwork. Including these additional job content indices also reduces the coefficient on years of schooling and cognitive skills (but not SEM skills). Still, the main results are robust to the inclusion of these additional variables, both for the full sample and when we estimate the associations between earnings and skills separately for migrants and non-migrants.

Fourth, there has been some discussion in the literature on precisely which cognitive and SEM skills are rewarded in the labor market (Díaz, Arias, and Tudela, 2014; Valerio et al., 2016). For cognitive skills, the association between labor earnings and test scores in Urdu is greater than for mathematics

[22] The brawn index is based on responses to five questions. How often does it happen that you are (1) carrying heavy loads; (2) using dangerous tools; (3) working under the hot sun or under the rain; (4) working with fumes, gases, dust and; (5) operating heavy equipment/machinery. The brain index is based on responses to the following five questions. How often does it happen that you are (1) required to read something; (2) required to write something; (3) required to calculate something; (4) required to operate a computer; (5) required to learn something new. We also created an indicator for jobs in which the individual is required to work in a team often or always.

and English scores. Still, we typically cannot rule out that the estimates are statistically the same (Table A12). The exception is for women, where the association with English is strongest and statistically different from Urdu and mathematics. This result could reflect the preponderance of teaching jobs among women and the demand for English as a subject. For SEM skills, we disaggregated our main measure into six indices corresponding to grit and the five traits of the Big five scale. We then estimated Equation (3) using each measure separately, with results presented in Appendix Tables A13 and A14. Overall, we do not find clear differences in these associations. For men, there is some evidence that the SEM skills with higher correlations are grit, conscientiousness, and emotional stability and that the correlation with extraversion is small. Still, except for the very low coefficient on extraversion, these differences do not point to a particularly strong correlation in the labor market for one particular skill. Again, precision is low for women, although even with the small sample, the correlation with measures of grit remains positive and statistically significant (Table A14).

III.2.2. Robustness checks

We present additional robustness checks in three parts. First, we examine the robustness of our estimates to different specifications, sample definitions, earnings measures, and approaches to account for attrition. Then, we offer a back-of-the-envelope calculation on potential attenuation bias due to measurement error. Finally, we present results from our sample in Cambodia to assess the structure of correlations in a similar study (children born in rural areas who were followed eight years later) but in a different context.

To check the robustness of our estimates to different sample definitions and specifications, we estimated an additional 66 specifications. We investigated whether our estimates were affected by (a) mean or median specifications; (b) the inclusion/exclusion of current students in the sample; (c) the inclusion/exclusion of village fixed effects as a proxy for regional labor market returns, (d) sensitivity to extreme values of earnings, and (e) different ways of accounting for attrition. E*x-ante,* it is difficult to argue one particular strategy is preferable to another in an exploratory exercise such as ours—even including students could be justified if respondents enroll in colleges as an additional activity while searching for jobs (see Jeffrey, 2010). Instead, we opt for a transparent approach and assess whether the correlations we documented previously are consistent across multiple robustness checks.

Rather than present separate tables, we plot all the estimates in specification curves, one each for years of schooling, cognitive skills, and SEM skills for men within and outside the village (Figures 3 to 5). The top panels provide coefficients; the bottom panels provide the linear combination of different restrictions. While most of the checks are self-explanatory, Appendix B discusses how we investigated the sensitivity of our specifications to sample selection using the Heckman two-step selection model and a sequential bivariate semi-nonparametric estimation following Glewwe, Song, and Zou (2022) and De Luca (2008). Even though there are specifications for which the results become imprecise, the general theme of our regressions holds: for those who remained in their original village, SEM skills are strongly correlated with earnings, while years of schooling are not; in turn, for those who left the village, years of schooling are correlated with earnings, while SEM skills are not.

A second exercise aimed to understand the extent of attenuation bias in our estimates for SEM skills. We used the test-retest reliability to estimate measurement error and rescaled the estimates using the standard formula for attenuation bias.[23] Doing so suggests that measurement error-corrected estimates would be \$48.5 instead of the \$16.5 we report in the main results. Our estimated reliability of 0.34 from the average test-retest correlations during the pilot is likely an underestimate of reliability in the final sample as Cronbach's α is 30% higher on average, suggesting lower measurement error in the final data collection. However, even if we adjust for this and estimate reliability to be 0.44, which is now likely an upper bound, the corrected coefficient would be \$37.5.

A third exercise sought to understand whether these patterns are specific to the particular sample from Pakistan. Therefore, we incorporated further data from the second study site in Cambodia. The structure of the sample and the survey is very similar, with children first surveyed in 2008 and then re-contacted and resurveyed in 2016/17, allowing us to look at links between education and skills in an originally rural sample. Unfortunately, the migration status is harder to determine, and the measures of socioemotional skills are less reliable, as discussed below and in Appendix D.

To measure SEM skills, we used the same self-reported Grit and Big Five scales as in Pakistan. Using a four-item scale, we also measured growth mindset—the belief that one can get smarter through hard

[23] For classical measurement error, the observed $\hat{\beta} = \frac{\sigma_{X^*}}{\sigma_{X^*}+\sigma_{\varepsilon}}\beta$ with $\frac{\sigma_{X^*}}{\sigma_{X^*}+\sigma_{\varepsilon}}$ being the reliability, which we approximate by the test-retest correlation.

work and practice. In addition, we asked four questions about locus of control. Internal locus of control measures the degree to which people believe they have power over the outcome of events in their lives, as opposed to external forces beyond their control. Finally, we administered the Strengths and Difficulties Questionnaire (SDQ)—a brief behavioral screening questionnaire aimed at measuring two main constructs: behavioral difficulties and pro-social behavior.

The reliability of, and validity evidence for, these measures are lower than in Pakistan. After correcting for acquiescence bias, Cronbach's α's of the different constructs range between 0.08 and 0.71, with only one of 11 measures passing the 0.7 threshold (Table 2). Moreover, as in Pakistan, the test-retest correlations are low, ranging from 0.14 to 0.43. Finally, the skills factor structure is not reproduced when conducting a factor analysis. Only three factors are retained, and one single factor encompasses a wide range of items that aim to measure different concepts.

Keeping these limitations around reliability and construct validity in mind, Tables A15 to A19 examine the correlation of these different skill measures with years of schooling and labor earnings. As in Pakistan, we document low levels of cognitive skills in these youth populations. For instance, only 34% of respondents correctly answered the following test question: "Three ox carts carry 1,000 kg of rice. How many kg of rice can nine ox carts carry?". We also find strikingly similar associations between years of schooling and cognitive skills. Every additional year of schooling is associated with a 0.17sd increase in cognitive skills in Pakistan and a 0.18sd increase in Cambodia (Table A15). The results from Cambodia also confirm that the correlation between SEM skills and years of schooling is weaker. An additional year of schooling is associated with a 0.05sd increase in SEM skills (and in related experimental work on the value of additional years of schooling, Barrera-Osorio, de Barros, and Filmer (2018) show no causal link between schooling and SEM skills in their sample). Thus, the results replicate across two very different rural regions of the world, with an association between schooling and cognitive skills but weaker associations between schooling and SEM skills.

Turning to the correlation between skills and earnings, we first highlight that the Cambodian sample is very different in its occupational structure, with 84% of respondents reporting agriculture as their primary occupation and LFP rates of 95% for both men and women. The earnings regressions lack precision in many specifications, but if there are any associations between earnings and skills, they are entirely for SEM skills, with zero or even negative associations between earnings and years of schooling or earnings and cognitive skills. In fact, if we only focus on the median regressions, we do not find any correlations between earnings and either skills or schooling for men. In the mean regressions, we find (imprecise) negative associations with cognitive skills and schooling, and positive associations with SEM skills (Table A19). As in Pakistan, the positive correlations with SEM skills are only for men who chose to stay in their original village. We find positive associations with SEM skills in the median and mean specifications for women, although the mean estimates are imprecise (Table A19). Schooling also appears to be more strongly correlated with earnings for women than for men, with a positive (although imprecise) correlation for cognitive skills among women who remained in their village.

IV. DISCUSSION

This discussion highlights three implications of our results and the questions they raise.

IMPLICATION 1: SIZEABLE CORRELATIONS. A one standard-deviation increase in cognitive/SEM skills is associated with an $8/$16.5 increase in median monthly earnings. Table A10 shows that these estimates for median earnings are quite similar to the estimates for mean earnings once we address extreme earnings values in the data. If we were to interpret these effects as causal and, sticking to the lower end of these estimates, we would conclude that a program that can boost test scores by 1sd by the time children leave school will increase earnings by $8 every month or $94 per year. A simple calculation helps contextualize this number in terms of actual school budgets. Specifically, if this is the annual increase in labor earnings for a working life of 40 years, using the World Bank's recommended discount rate of 5% yields an additional lifetime (annually) discounted income of $1700 (Fay et al., 2014). Assuming that the investment will only kick in 10 years after the extra spending, this would still lead to an additional benefit of $1,047 per child. Suppose that children remain in school for the average of eight years we observe in this sample. Then, spread over those eight years, governments should be willing to spend an additional $131 per year, relative to the current average annual spending of $132 per child in 2017-2018 (Shah et al., 2018). For a program that increases SEM skills by a similar amount, governments should be willing to spend twice as much. Yet, any spending close to that amount is regarded as next-to-impossible in a country like Pakistan, and we believe this is at least in part because the benefits of increasing cognitive and socioemotional skills on labor earnings have never been clearly shown.

Of course, there are multiple brave assumptions built into this calculation regarding the depreciation of skills (which we have assumed to be negligible), the demand for skills (which we have assumed to be perfectly elastic), and the extent of omitted variable bias. So, an urgent next step, given the magnitude of these coefficients, is to assess the plausibility of these estimates as reflecting causal links from skills to earnings. If these estimates hold, it could lead to an important reassessment of how much governments should invest in improving cognitive and SEM skills in low-income countries.

IMPLICATION 2: BOWLES AND GINTIS REVISITED. Bowles and Gintis (1976) argued that a central function of schools was to produce workers for the capitalist factory system and instill in them the new skills required to operate assembly lines at the turn of the twentieth century. These required skills not only refer to cognitive skills but also included skills such as punctuality, discipline, and respect for hierarchy—skills that were highly valued on the factory floor. In their formulation, the Mincer returns to years of schooling captured, at least to some extent, the imparting of these skills to students. In current terminology, we may restate their question as assessing the extent to which cognitive and socioemotional skills causally mediate the returns to schooling.

In the United States, the introduction of compulsory secondary education in the early 1900s limited the ability to exploit variation in educational attainment. While comparisons between college and secondary school graduates are common, such an analysis introduces many additional complications arising from the enormous variation in college courses and college quality. In contrast, as in many low-income countries, our data provides considerable variation in the years of schooling, from zero to college completion, along with considerable variation in cognitive skills for each year of schooling. Exploiting this variation shows, just as Bowles and Gintis (1976) predicted, that including cognitive and SEM skills as additional explanatory variables reduces the coefficient on years of schooling (by two-thirds in our case). So, the correlation between earnings and years of schooling captures (in part) the skills we measured.

Yet, two problems remain. First, schools appear to play only a marginal role in producing the SEM skills we measured, if any. Hence, our results sit uneasily with Bowles and Gintis' preferred explanation that the returns to years of schooling capture schools' ability to socialize and prepare children for factory work in a capitalist system. Second, we can explain at most 21% of the variation in mean labor earnings for men after adjusting for outliers in Table A11 and only 7% for women (Table 8). Thus, even with a comprehensive skills measurement component, a large share of our respondents' earnings reflects other considerations. Another way to see this is to note that the extent to which the coefficient on years of school will decline once we include other skills depends on the covariance of the measured skill with labor earnings and with years of schooling. Using the coefficients for the mean regressions from Table A9, a year of schooling is associated with a \$5.2 increase in labor earnings for men. We also know that a year of schooling is associated with a 0.17sd increase in cognitive and a 0.03sd increase in SEM skills. Multiplying these gains with the associations between earnings and skills suggests that the increase in skills associated with one more year of schooling will increase earnings by \$1.51. This still leaves two-thirds of the correlation between labor earnings and years of schooling as an unexplained residual. These results continue to raise important questions regarding the types of skills that schools impart to their students that are then rewarded in the labor market.

IMPLICATION 3: MIGRATION AND SKILLS. Our final point considers how labor markets reflect skills and relates to a lively debate on migration, labor earnings, and skills. This literature is concerned with the returns to migration. The main challenge is that people who migrate may have systematically different skills so that earnings for migrants reflect both the causal impact of migration and the differences in skills. Although studies have designed clever natural and randomized experiments to get around this problem, we are unaware of studies that directly assess the skills of migrants and non-migrants to understand how these are rewarded in the labor market at different locations. Combining our sampling strategy and skills measurement allows us to present correlational evidence on this question—with several noteworthy findings.

First, we find those with higher cognitive and SEM skills are more likely to migrate (at least within the country, if not outside). Second, accounting for cognitive and SEM skills lowers the difference in labor earnings of migrants versus non-migrants. However, a remarkable—and novel—finding in this sample, which we corroborate with data from Cambodia, is that the association of earnings with SEM skills is positive (and large) *only* for non-migrants, and the correlation with years of schooling is higher for migrants. This result suggests that the returns to different skills vary by the migration status of the respondent. One possible explanation raised by Carranza et al. (2022) in South Africa is that this pattern reflects a lack of information among employers who rely on observable signals instead (years of schooling). They provide job seekers and employers with assessment scores for "non-specialist skills such as communication, numeracy, and grit" and find that both employers and job seekers react positively to the information. The type of information on skills that Carranza et al. (2022) provide to workers and employers may have enormous value for migrants, as years of schooling explains very little of the variation in the SEM skills we measure. A second implication is that, under the current regime, if individuals estimate the returns to years of schooling by looking at non-migrants whose labor earnings are easier to observe, they will underestimate the returns. This point corroborates Jensen's (2010) argument about perceived vs. actual returns to schooling in the Dominican Republic, again demonstrating the potential role for information in affecting schooling decisions.

V. CONCLUSION

We measured cognitive and socioemotional skills in a sample of young adults in rural Pakistan. Fifteen years after the first data collection exercise, we revisited the respondents in their current residence, when they were 24 years old on average. Considerable migration in this population provides insights concerning the link between skills, migration, and labor earnings. Our results suggest that careful design and training can help improve the measurement of socioemotional skills in low-income countries. They also show that cognitive and socioemotional skills are (a) correlated with years of schooling and (b) correlated with labor earnings for men, but the size of these correlations is mediated by migration. For the sample as a whole, the associations between labor earnings and skills (conditional on schooling) suggest that these skills are underproduced in schools.

While this paper makes some progress in addressing several puzzles around the measurement of skills in low-income countries, there is substantial room for further improvements. Our measures of cognitive and SEM skills do not fully explain the association between earnings and years of schooling. It could be that the unexplained variation in labor earnings reflects wedges that lead to inefficient labor markets, as in Carranza et al. (2022). Or it could be that there are genuinely other measures that employers look for and that schools are designed to produce (the "hidden curriculum" in Bowles and Gintis, 1976)—which would require that new contextual measurements be built from the ground up in low-income countries.

References

Afridi, F., Dinkelman, T., Mahajan, K., 2018. Why are fewer married women joining the work force in rural India? A decomposition analysis over two decades. *Journal of Population Economics, 31(3),* 783–818

Alan, S., Boneva, T., Ertac, S., 2019. Ever failed, try again, succeed better: Results from a randomized educational intervention on grit. *The Quarterly Journal of Economics*, *134*(3), 1121-1162.

Altonji, J.G., Pierret, C.R., 2001. Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, *116*(1), 313-350.

Andrabi, T., Das, J., Khwaja, A.I., Farooqi, D., Zajonc, T., 2002. Test feasibility survey Pakistan: Education sector. Available at: https://www.researchgate.net/publication/349377986_Test_Feasibility_Survey_Pakistan

Andrabi, T., Das, J., Khwaja, A.I., Vishwanath, T., Zajonc, T., 2008. Learning and educational achievements in Punjab schools (LEAPS): Insights to inform the education policy debate (Research Report No. 43750), *The World Bank, Washington, D.C.*

Andrabi, T., Das, J., Khwaja, A.I.,2022, The Learning and Educational Achievement in Pakistan Schools (LEAPS) Longitudinal Dataset, 2004-2011, version 11-2022, released November 1st 2022.

Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., Sulaiman, M., 2020. Women's empowerment in action: Evidence from a randomized control trial in Africa. *American Economic Journal: Applied Economics*, *12*(1), 210–259.

Banerjee, A.V.B., Bhattacharjee, S., Chattopadhyay, R., Ganimian, A.J., Duflo, E., Spelke, E.S., 2022. Street smart or school smart? The arithmetic skills of working children in two Indian cities.

Barrera-Osorio, F., de Barros, A., Filmer, D.P., 2018. Long-term impacts of alternative approaches to increase schooling: Experimental evidence from a scholarship program in Cambodia (Working Paper No. WPS8566), *Policy Research Working Paper. The World Bank, Washington, D.C.*

Bau, N., Das, J., Yi Chang, A., 2021. New evidence on learning trajectories in a low-income setting. *International Journal of Educational Development*, *84*, 102430.

Beegle, K., De Weerdt, J., Dercon, S., 2011. Migration and economic mobility in Tanzania: Evidence from a tracking survey. *Review of Economics and Statistics*, *93*(3), 1010-1033.

Boon-Falleur, M., Bouguen, A., Charpentier, A., Algan, Y., Huillery, É., Chevallier, C., 2022. Simple questionnaires outperform behavioral tasks to measure socioemotional skills in students. *Scientific reports*, *12*(1), 1-11.

Bowles, S., Gintis, H., 1976. Schooling in capitalist America: Educational reform and the contradictions of economic life. *Basic Books, New York.*

Brunello, G., Schlotter, M., 2011. Non cognitive skills and personality traits: Labour market relevance and their development in education and training systems (Working paper No. 5743), *IZA Discussion Paper. Institute of Labor Economics (IZA), Bonn, Germany.*

Bryan, G., Chowdhury, S., Mobarak, A.M., 2014. Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica*, *82*(5), 1671–1748.

Carranza, E., Garlick, R., Orkin, K., Rankin, N., 2022. Job search and hiring with limited information about workseekers' skills. *American Economic Review*, *112*(11), 3547–3583

Cunha, F., Heckman, J.J., 2010. Investing in our young people, in: *Reynolds, A.J., Rolnick, A.J., Englund, M.M., Temple, J.A. (Eds.), Childhood Programs and Practices in the First Decade of Life: A Human Capital Integration. Cambridge University Press, New York, p. 381-414.*

Das, J., Singh, A., Yi Chang, A., 2022. Test scores and educational opportunities: Panel evidence from five low- and middle-income countries. *Journal of Public Economics 206*, 104570.

Das, J., Zajonc, T., 2010. India shining and Bharat drowning: Comparing two Indian states to the worldwide distribution in mathematics achievement. *Journal of Development Economics*, *92*(2), 175-187.
De Luca, G., 2008. SNP and SML estimation of univariate and bivariate binary-choice models. *The Stata Journal*, *8*(2), 190-220
Deming, D.J., 2017. The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, *132*(4), 1593-1640.
Díaz, J.-J., Arias, O.S., Tudela, D.V., 2014. Does perseverance pay as much as being smart? The returns to cognitive and non-cognitive skills in urban Peru, in: *9th IZA/World Bank Conference on Employment and Development. IZA - Institute of Labor Economics, Lima, Peru.*
Duckworth, A.L., Peterson, C., Matthews, M.D., Kelly, D.R., 2007. Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*(6), 1087.
Duckworth, A.L., Quinn, P.D., 2009. Development and validation of the short grit scale (Grit–S). *Journal of Personality Assessment*, *91*(2), 166-174.
Duflo, E., Dupas, P., Kremer, M., 2021. The impact of free secondary education: Experimental evidence from Ghana (Working Paper No. 28937), *NBER Working Paper. National Bureau of Economic Research, Cambridge, MA.*
Fay, M., Hallegate, S., Kraay, A., Vogt-Schilb, A., 2014. Discounting costs and benefits in economic analysis of World Bank projects.
Field, E., Vyborny, K., 2016. Female labor force participation in Asia: Pakistan country study. *Background report. Asian Development Bank, Manila.*
Foster, A.D., Rosenzweig, M.R., 1996. Technical change and human-capital returns and investments: Evidence from the green revolution. *American Economic Review*, *86*(4), 931-953.
Glewwe, P., Huang, Q., Park, A., 2017. Cognitive skills, noncognitive skills, and school-to-work transitions in rural China. *Journal of Economic Behavior & Organization*, *134*, 141-164
Glewwe, P., Song, Y., Zou, X., 2022. Labor market outcomes, cognitive skills, and noncognitive skills in rural China. *Journal of Economic Behavior & Organization 193*, 294–311.
Hanushek, E.A., Kimko, D.D., 2000. Schooling, labor-force quality, and the growth of nations. *American Economic Review, 90(5), 1184-1208.*
Hanushek, E.A., Woessmann, L., 2015. Universal basic skills: What countries stand to gain. *OECD Publishing, Paris, France.*
Hanushek, E.A., Woessmann, L., 2008. The role of cognitive skills in economic development. *Journal of Economic Literature*, *46*(3), 607-68.
Heckman, J.J., 2007. The economics, technology, and neuroscience of human capability formation. *Proceedings of the national Academy of Sciences*, *104*(33), 13250-13255.
Heckman, J.J., Rubinstein, Y., 2001. The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review*, *91*(2), 145-149.
Jeffrey, C., 2010. Timepass: Youth, class, and time among unemployed young men in India. *American Ethnologist*, *37*(3), 465-481.
Jensen, R., 2010. The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics*, *125*(2), 515-548.
Laajaj, R., Macours, K., 2021. Measuring skills in developing countries. *Journal of Human resources*, *56*(4), 1254-1295.
Lang, F.R., John, D., Lüdtke, O., Schupp, J., Wagner, G.G., 2011. Short assessment of the Big Five: Robust across survey methods except telephone interviewing. *Behavior Research Methods*, *43*(2), 548-567.
Loyalka, P., Liu, O.L., Li, G., Kardanova, E., Chirikov, I., Hu, S., Yu, N., Ma, L., Guo, F., Beteille, T. and Tognatta, N., 2021. Skill levels and gains in university STEM education in China, India, Russia and the United States. *Nature Human Behaviour, 5(7),* 892-904.

McKenzie, D., Stillman, S., Gibson, J., 2010. How important is selection? Experimental vs. non-experimental measures of the income gains from migration. *Journal of the European Economic Association*, *8*(4), 913-945.

Mitchell, M., Favara, M., Porter, C., Sánchez, A., 2020. Human capital development: New evidence on the production of socioemotional skills (Discussion Paper No. 13804), *Discussion Paper Series. Institute of Labor Economics (IZA), Bonn, Germany.*

Montenegro, C.E., Patrinos, H.A., 2021. A data set of comparable estimates of the private rate of return to schooling in the world, 1970–2014. *International Journal of Manpower ahead-of-print*.

Muralidharan, K., 2019. Reforming the Indian School Education System, in: *What the Economy Needs Now, edited by Abhijit Banerjee, Gita Gopinath, Raghuram Rajan, and Mihir Sharma, Juggernaut*.

Nunnally, J.C., 1978. Psychometric theory, 2nd edition. ed, McGraw-Hill series in *Psychology. McGraw-Hill, New York.*

Outes-Leon, I., Dercon, S., 2009. Survey attrition and attrition bias in Young Lives (Technical Note No. 5), *Young Lives Technical Notes. University of Oxford, Oxford, United Kingdom.*

Pierre, G., Sanchez Puerta, M.L., Valerio, A., Rajadel, T., 2014. STEP skills measurement surveys: Innovative tools for assessing skills (Discussion Paper No. 1421), *Social Protection and Labor Discussion Papers. World Bank, Washington, DC.*

Pritchett, L., 2013. The rebirth of education: Schooling ain't learning. *CGD Books.*

Schultz, T. W., 1975. The value of the ability to deal with disequilibria. *Journal of Economic Literature*, *13*(3), 827-846.

Shah, D., Khatoon, S., Bhatti, M.A., Tatla, A., 2018. Public financing in education sector (Report No. 282), *AEPAM Publications. Academy of Educational Planning and Management, Islamabad, Islamabad, Pakistan.*

Singh, R., Kesarwani, R., Mukherjee, P., 2018. 'I will achieve everything on my own': the association between early psychosocial skills and educational progression through adolescence in India. *Young Lives Working Paper 175.*

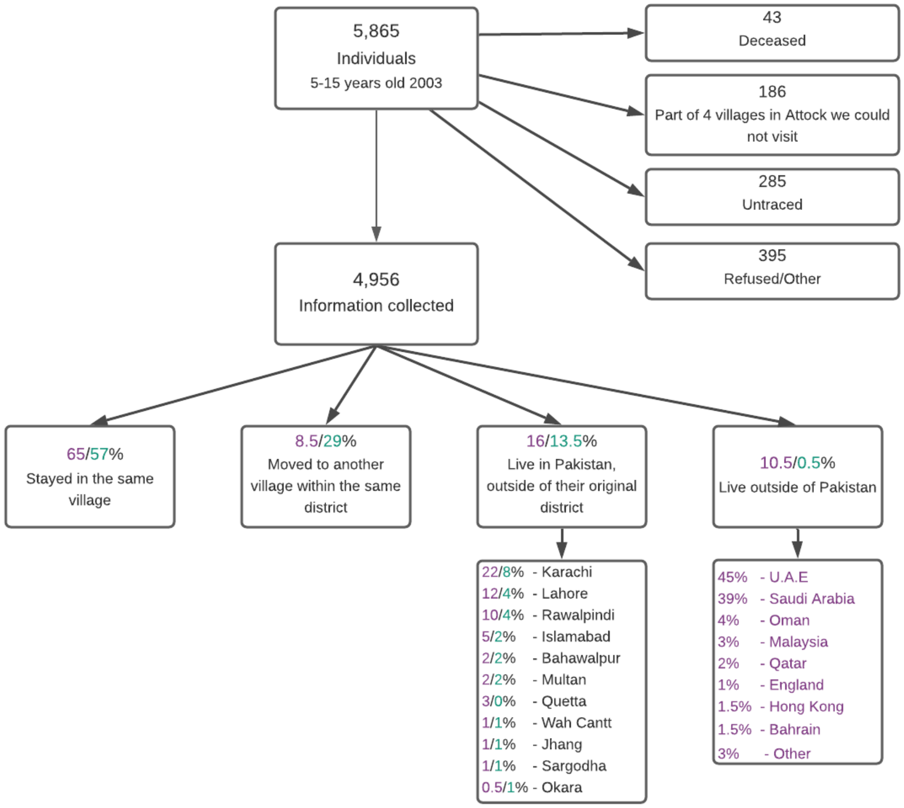
Subramanian, N., 2020. Workplace attributes and women's labor supply decisions: Evidence from a randomized experiment in Pakistan. *Presented at the 2020 IZA/University of Luxembourg Workshop: Matching Workers and Jobs Online - New Developments and Opportunities for Social Science and Practice, IZA - Institute of Labor Economics, Online.*

Valerio, A., Sanchez Puerta, M.L., Tognatta, N., Monroy-Taborda, S., 2016. Are there skills payoffs in low- and middle-income countries? Empirical evidence using STEP data (Working Paper No. 7879), *Policy Research Working Paper. World Bank, Washington, DC.*

World Bank, 2018. World Development Report 2018: Learning to realize education's promise. *World Bank, Washington, DC.*

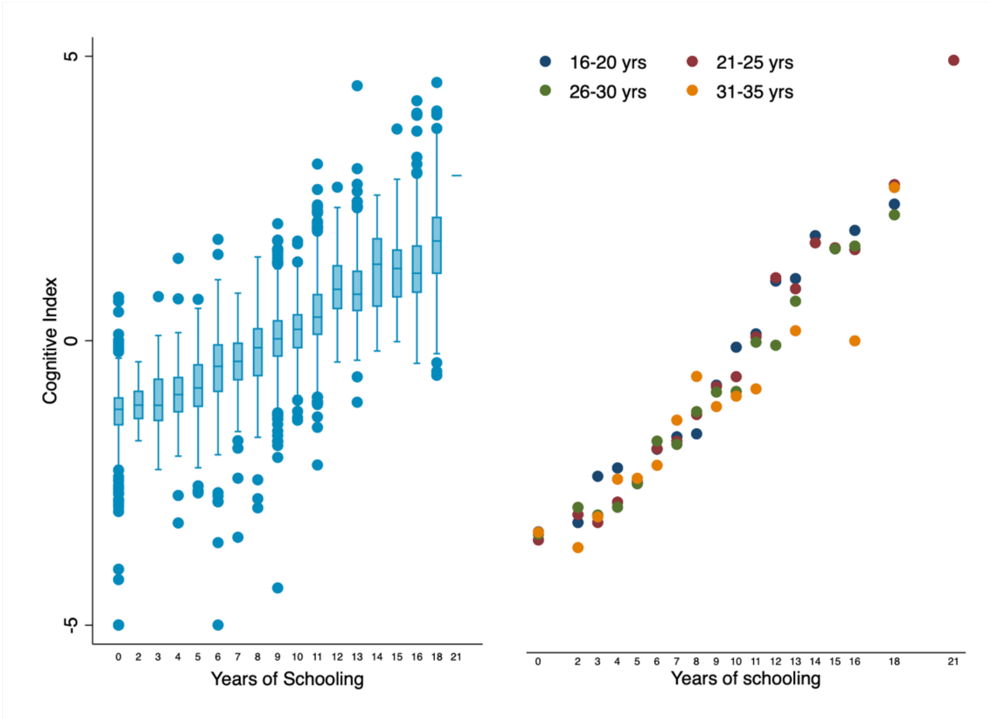
MAIN FIGURES

Figure 1. Migration patterns



Notes. This figure shows where respondents in our sample lived at the time of the follow-up survey in 2018. Numbers for men are in purple and for women in blue.

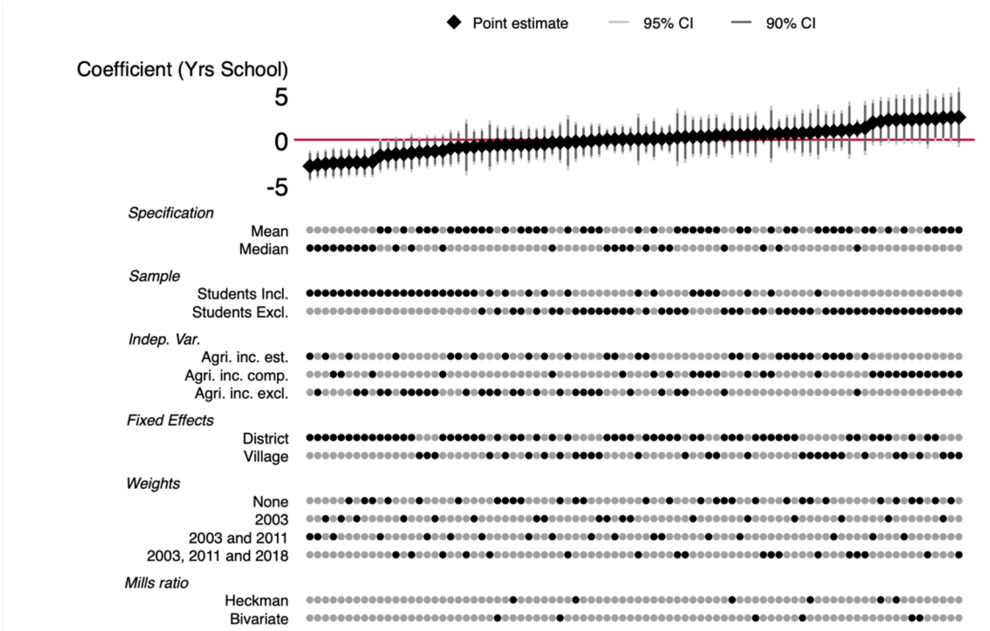
Figure 2. Years of schooling, Age, and Cognitive skills Formation



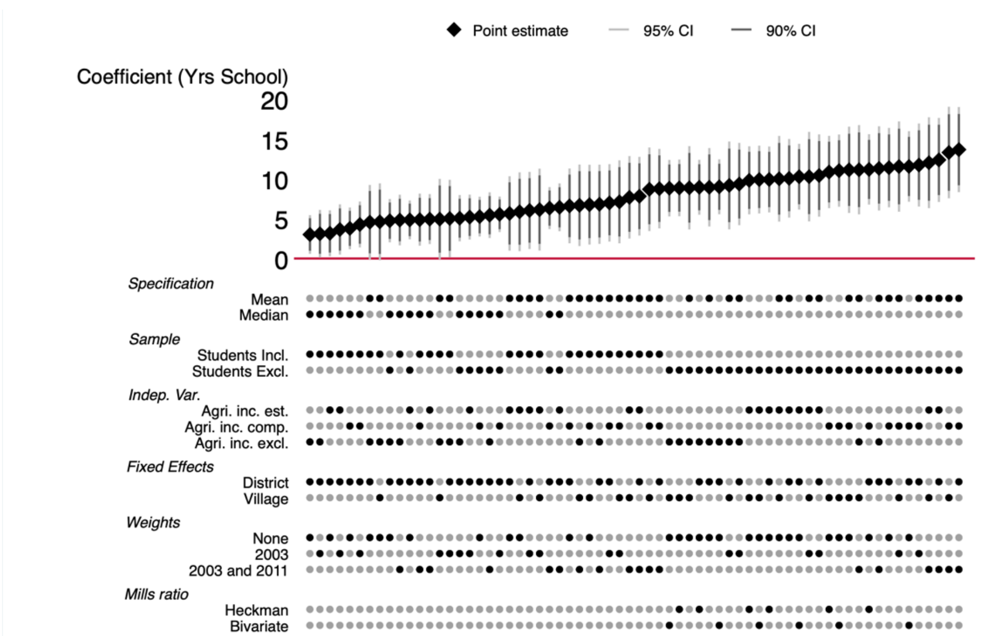
Notes. This graph shows the relationship between schooling and cognitive skills formation. We coded the years of schooling variable using the highest grade enrolled. The Cognitive Index is the mean of the Urdu, Mathematics, and English scores computed using Item Response Theory with a two-parameter logistic (2PL) model. The items included in the model are those from the paper-based test and the computer-adaptive tablet-based test (leaving out items that less than 50 respondents answered and items that less than 5% or more than 95% of respondents solved correctly). The left panel shows, for each year of schooling, the distribution of the cognitive index. The right panel shows, for each year of schooling the average cognitive skills for respondents who are between 16-20 years old, for respondents who are between 21-25 years old, for respondents who are between 26-30 years old, and for respondents who are between 31-35 years old. The sample includes all men and women who have cognitive skills measures in our sample (4,401 respondents).

Figure 3. Specification curves – Years of schooling

Panel A. Men within village



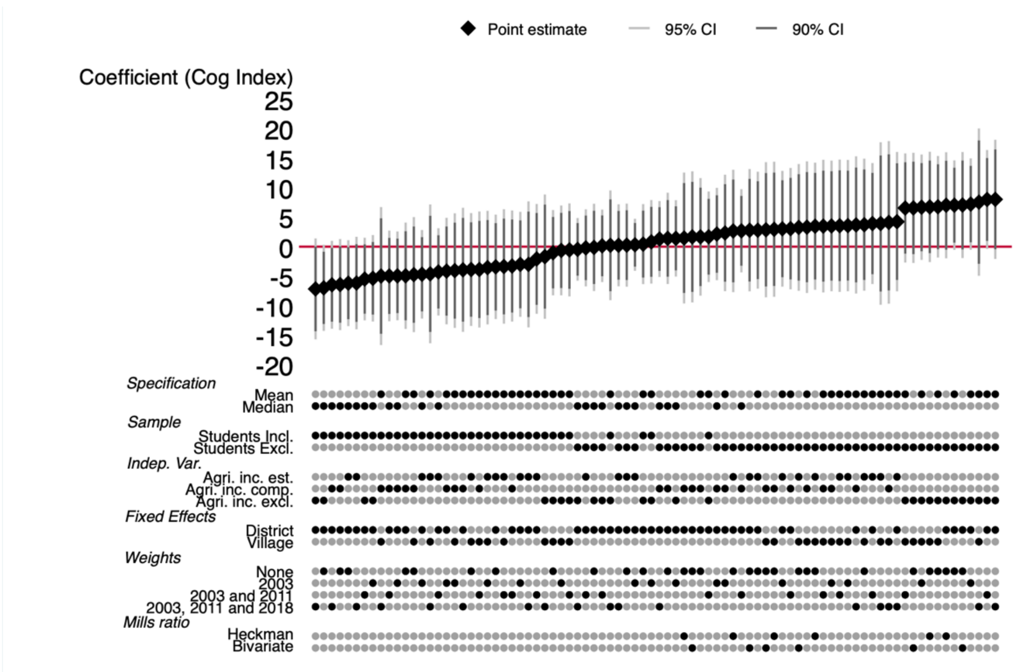
Panel B. Men outside of village



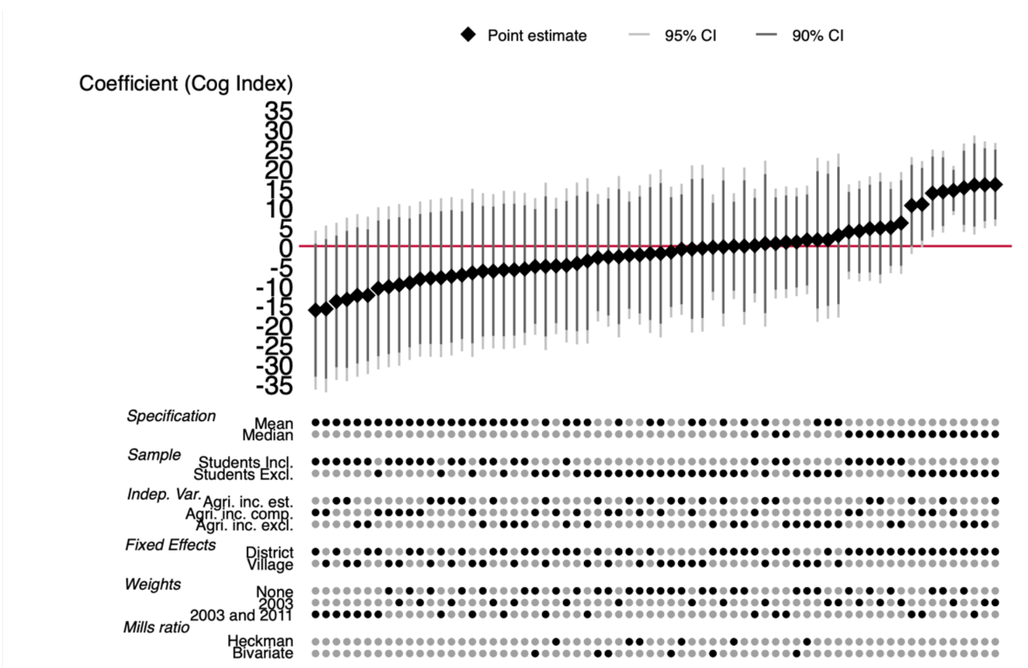
Notes. Each dot in the top panel of the graph depicts the years of schooling coefficient from the earnings returns estimation, using the fully interacted specification for men within the village (Panel A) and men outside the village (Panel B). The dots vertically aligned below indicate the analytical decisions behind those estimates. A total of 66 specifications were estimated. We describe how we compute the weights and the inverse mills ratio in Appendix B.

Figure 4. Specification curves – Cognitive skills

Panel A. Men within village



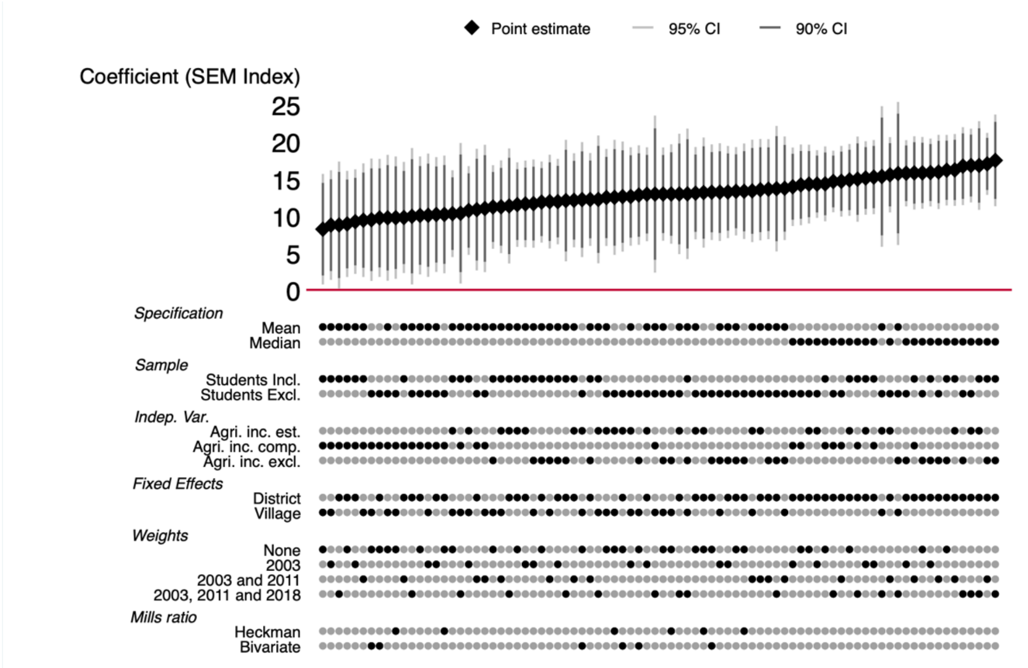
Panel B. Men outside of village



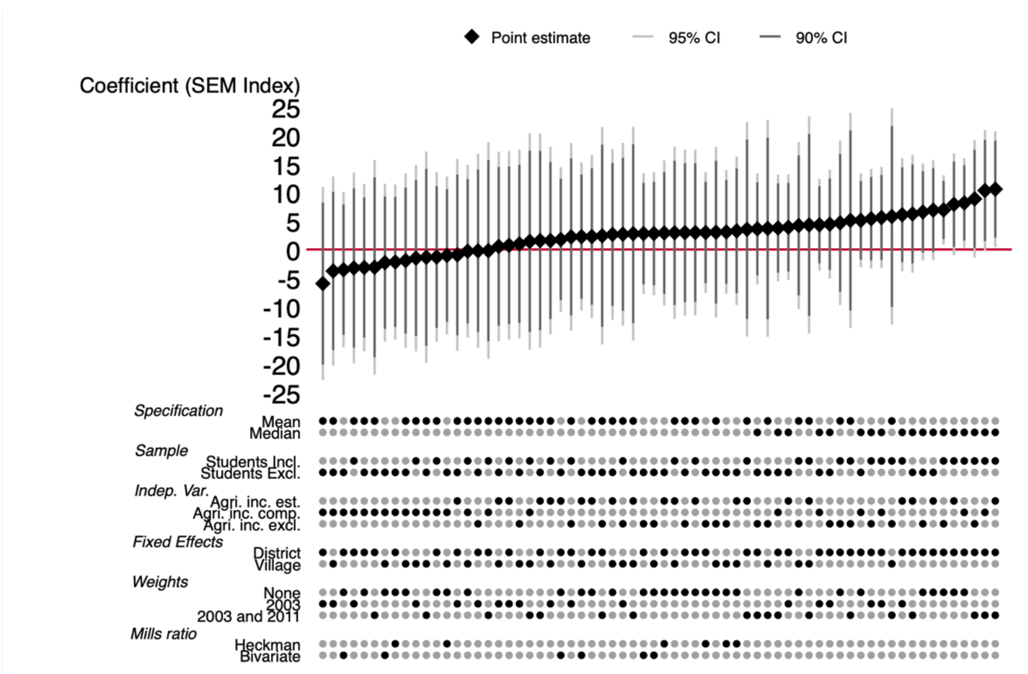
Notes. Each dot in the top panel of the graph depicts the cognitive skills index coefficient from the earnings returns estimation, using the fully interacted specification for men within the village (Panel A) and men outside the village (Panel B). The dots vertically aligned below indicate the analytical decisions behind those estimates. A total of 66 specifications were estimated. We describe how we compute the weights and the inverse mills ratio in Appendix B.

Figure 5. Specification curves – Socioemotional skills

Panel A. Men within village



Panel B. Men outside of village



Notes. Each dot in the top panel of the graph depicts the socioemotional (SEM) skills index coefficient from the earnings returns estimation, using the fully interacted specification for men within the village (Panel A) and men outside the village (Panel B). The dots vertically aligned below indicate the analytical decisions behind those estimates. A total of 66 specifications were estimated. We describe how we compute the weights and the inverse mills ratio in Appendix B.

MAIN TABLES

Table 1. Summary statistics

	All			Men/Fathers			Women/Mothers			Difference
	Mean	SD	Obs.	Mean	SD	Obs.	Mean	SD	Obs.	Women-Men
Panel A: Individual (2018)										
Age	23.77	3.62	4956	23.8	3.54	2596	23.73	3.7	2360	-0.07
Years of schooling	8.55	4.89	4955	8.92	4.35	2595	8.13	5.39	2360	-0.79***
Respondent can read	0.73	0.44	4402	0.74	0.44	2222	0.73	0.45	2180	-0.01
Ever married	0.45	0.5	4956	0.34	0.47	2596	0.58	0.49	2360	0.25***
Age at first marriage	20.95	3.43	2248	22.37	3.31	874	20.05	3.2	1374	-2.31***
Has children	0.33	0.47	4956	0.22	0.41	2596	0.45	0.5	2360	0.23***
Working (students excl.)	0.47	0.5	4455	0.85	0.36	2351	0.05	0.22	2104	-0.79***
Main work is farming (students excl.)	0.04	0.19	4455	0.07	0.26	2351	0	0	2104	-0.07***
HH has toilets on premises	0.96	0.19	4404	0.97	0.18	2223	0.96	0.2	2181	-0.01
HH has access to electricity	0.98	0.14	4404	0.98	0.13	2223	0.98	0.15	2181	-0.00
Lives in same village as in 2003	0.61	0.49	4956	0.65	0.48	2596	0.57	0.5	2360	-0.08***
Panel B: Parents (2003)										
Parent years of schooling	3.1	4.28	8780	4.74	4.75	4201	1.6	3.12	4579	-3.15***
Parent can read	0.37	0.48	8945	0.53	0.5	4219	0.23	0.42	4726	-0.30***
Parent age at first marriage	22.07	4.82	7522	24.6	4.8	3648	19.69	3.42	3874	-4.92***
Parent main work is farming	0.18	0.39	8964	0.33	0.47	4230	0.05	0.21	4734	-0.29***
Parent has toilets on premises	0.58	0.49	9699	0.58	0.49	4847	0.58	0.49	4852	-0.00
Parent has access to electricity	0.88	0.33	9685	0.88	0.33	4840	0.88	0.33	4845	-0.00

* p<.10 ** p<.05 *** p<.01

Notes. This table provides sample characteristics (overall and by gender). Panel A shows respondent characteristics (for 2018). "Respondent can read" comes from the functional literacy and numeracy section. It indicates that the respondent could read and understand a greeting text message in Urdu script. Only respondents who completed the survey in person filled out this section. "Age at first marriage" excludes those who never married. "HH has toilets on-premises/access to electricity" indicates that the respondent lives in a household with a toilet on-premises/access to electricity. These variables are filled in all in-person surveys. Panel B shows the characteristics of the respondents' parents (for 2003, with the exception of "Parent age at first marriage", which was collected in 2011). The respective sample consists of individuals whose household was surveyed that year and who were living with their parents. The table also shows the differences in the means across men/fathers and women/mothers. Standard errors are clustered at the village level. There are 108 clusters in the sample.

Table 2. Reliability of socioemotional skills measures

Construct	Instrument	Mode	Country	Acquiescence bias correction (1) Alpha test	(2) Alpha re-test	(3) Test-retest	Uncorrected measures (4) Alpha test	(5) Alpha re-test	(6) Test-retest	N	Nb. items
Grit	**Grit scale**	Self-reported	Cambodia	0.39	-	-	0.16	-	-	3282	8
			Pakistan	0.75	-	-	0.68	-	-	4395	9
			Pakistan Pilot	0.57	0.72	0.2	0.47	0.63	0.2	397	10
	Alan & Ertac Grit Task "Additions Game"	Task-based	Pakistan Pilot	-	-	-	-	-	0.42	402	-
	Frustration task	Task-based	Pakistan Pilot	-	-	-	-	-	0.27	402	-
Openness to experience	**Big Five**	Self-reported	Cambodia	0.08	0.22	0.24	0.4	0.35	0.38	3287	3
			Pakistan	0.53	-	-	0.63	-	-	4475	3
			Pakistan Pilot	0.31	0.31	0.29	0.42	0.54	0.27	403	3
Conscientiousness	**Big Five**	Self-reported	Cambodia	0.48	0.39	0.43	0.37	0.43	0.42	3286	3
			Pakistan	0.57	-	-	0.4	-	-	4475	3
			Pakistan Pilot	0.62	0.71	0.13	0.44	0.58	0.09	402	3
Extraversion	**Big Five**	Self-reported	Cambodia	0.41	0.03	0.14	0.2	0.01	0.13	3286	3
			Pakistan	0.68	-	-	0.64	-	-	4475	3
			Pakistan Pilot	0.47	0.44	0.29	0.4	0.35	0.3	403	3
Agreeableness	**Big Five**	Self-reported	Cambodia	0.45	0.44	0.38	0.31	0.47	0.38	3287	3
			Pakistan	0.6	-	-	0.42	-	-	4475	3
			Pakistan Pilot	0.5	0.73	0.13	0.35	0.59	0.11	403	3
Emotional stability	**Big Five**	Self-reported	Cambodia	0.04	0.11	0.23	.	.	0.24	3289	3
			Pakistan	0.6	-	-	0.64	-	-	4475	3
			Pakistan Pilot	0.39	0.51	0.41	0.39	0.47	0.43	402	3
Big Five	**Big Five**	Self-reported	Cambodia	0.52	.	0.36	0.43	0.18	0.32	3279	15
			Pakistan	0.64	-	-	0.56	-	-	4475	15
			Pakistan Pilot	0.62	0.71	0.29	0.53	0.66	0.25	401	15
Locus of control	**Locus of control**	Self-reported	Cambodia	0.58	-	-	0.2	-	-	3287	4
			Pakistan Pilot	0.52	0.61	0.45	0.29	0.31	0.45	403	4
Growth mindset	**Growth mindset**	Self-reported	Cambodia	0.4	-	-	0.4	-	-	3284	4
Behavioral difficulties	**SDQ**	Self-reported	Cambodia	0.71	-	-	0.64	-	-	3280	20
Pro-social behavior	**SDQ**	Self-reported	Cambodia	0.31	-	-	0.63	-	-	3283	5
Impulsiveness	**Barratt Impulsiveness Scale (BIS)**	Self-reported	Pakistan Pilot	0.71	0.77	0.4	0.64	0.71	0.38	397	30
Risk-taking behavior	**Balloon Analogue Risk Task (BART)**	Task-based	Pakistan Pilot	-	-	-	-	-	0.36	402	-

Self-control	**GoNoGo**	Task-based	Pakistan Pilot	-	-	-	-	-	0.78	402	-

Notes. This table reports estimates of the reliability of socioemotional skills measures. "Cambodia" refers to data collected in that country in 2017. A random 13% of respondents were re-surveyed. "Pakistan Pilot" and "Pakistan" refer to data collected in that country in 2018. Half of the pilot respondents were re-surveyed. Appendix D describes each instrument in detail. Columns (1), (2), (4), and (5) provide Cronbach's alpha statistics for the test (columns 1 and 4) and the re-test (columns 2 and 5). Columns (3) and (6) show test-retest correlations. Following Laajaj and Macours (2021), we applied an acquiescence bias correction when applicable (left panel); we also provide uncorrected measures (right panel). Sample sizes differ due to item non-response and different survey modes (in-person vs. phone surveys).

Table 3. Factor analysis of self-reported measures

Skill	Label	Factor	Uniqueness
Grit			
Grit	New ideas and projects sometimes distract me from previous ones	1	0.7149
Grit	Setbacks don’t discourage me. I don’t give up easily	1	0.5782
Grit	I often set a goal but later choose to pursue a different one	1	0.526
Grit	I have difficulty maintaining my focus on project that take more than a few months to complete	1	0.769
Grit	I finish whatever I begin	1	0.6359
Grit	My interests change from year to year.	1	0.6477
Grit	I am diligent. I never give up.	1	0.4094
Grit	I have been obsessed with a certain idea or project for a short time but later lost interest	1	0.545
Grit	I have overcome setback to conquer an important challenge.	1	0.7781
Big Five			
Openness	is original, comes up with new ideas	4	0.4713
Openness	values artistic, aesthetic experiences	4	0.6354
Openness	has an active imagination	4	0.377
Conscientiousness	does a thorough job	2	0.4286
Conscientiousness	tends to be lazy	2	0.703
Conscientiousness	does things efficiently	2	0.599
Extraversion	is talkative	3	0.5105
Extraversion	is outgoing, sociable	3	0.2589
Extraversion	is reserved	3	0.2196
Agreeableness	is sometimes rude to others	2	0.5878
Agreeableness	has a forgiving nature	2	0.4117
Agreeableness	is considerate and kind to almost everyone	2	0.3341
Emotional stability	worries a lot	5	0.69
Emotional stability	gets nervous easily	5	0.2792
Emotional stability	remains calm in tense situations	5	0.3463

Notes. Using the data from Pakistan, this table provides the results from an exploratory factor analysis on self-reported grit and big-five items. All items are in a five-point format, which ranges from 1 ("Disagree strongly") to 5 ("Agree strongly"). Following Laajaj and Macours (2021), we applied an acquiescence bias correction (see Appendix D). We perform a principal factor estimation and retain five factors. We rotate the factor loadings so that each item mainly loads on one factor. For each item, we indicate the main factor the item loads on. We also provide the item’s uniqueness, which is the item’s percentage of variance that is not explained by the common factors.

Table 4. Skills of young adults in Pakistan (Part 1)

Subject	Difficulty level	Question	Percent correct	N
Panel A. Cognitive Test				
Mathematics	70% or more get	Tick box next number that matches the number of objects	79 %	4139
Mathematics	around 50% get	678+923	56 %	4139
Mathematics	30% or less get	7/3=__	5 %	4139
Urdu	70% or more get	Match picture: Book	78 %	4139
Urdu	around 50% get	Join letters and write word: m-a-l-k	48 %	4139
Urdu	30% or less get	Fill blank in the story by selecting the correct word	29 %	4139
English	70% or more get	Match picture: Banana	77 %	4139
English	around 50% get	Missing letter to match picture: Flag	53 %	4139
English	30% or less get	Use word in sentence: deep	16 %	4139
Panel B. Functional Literacy and Numeracy Assessment				
Numeracy	70% or more get	Read elec bill. How much money do you need to pay for the month of November?	80 %	4402
Numeracy	around 50% get	Multiplication per bracket - Elec bill. How much money do you have to pay?	49 %	4402
Numeracy	30% or less get	Multiplication: Kv* cost per Kv - Elec bill. How much money do you have to pay	16 %	4402
Literacy	70% or more get	Respondent was able to understand the greeting message in Urdu	73 %	4402
Literacy	around 50% get	Respondent was able to understand the greeting message in Roman Urdu	55 %	4402
Literacy	70% or more get	Respondent was able to save contact on mobile phone	69 %	4402

Notes. This table shows select questions for the three subjects tested on paper (Mathematics, Urdu, and English; Panel A), and for the functional numeracy and literacy skills assessment (Panel B). For each category, we show a question that at least 70% of respondents solved correctly, a question that approximately 50% solved correctly, and a question that less than 30% of respondents solved correctly. All functional literacy questions were correctly solved by at least 30% of the respondents. For each question, we indicate the proportion of respondents who answered the question correctly and the number of respondents. Sample sizes differ due to item non-response.

Table 4. Skills of young adults in Pakistan (Part 2)

Skill	Sample	Mean	SD	Min	Max	N	Source
Panel A. LEAPS sample							
Grit	LEAPS	3.45	0.71	1	5	4395	LEAPS Data
Openness to experience	LEAPS	3.42	1.06	1	5	4475	LEAPS Data
Conscientiousness	LEAPS	4.17	0.72	1	5	4475	LEAPS Data
Extraversion	LEAPS	3.31	1.09	1	5	4475	LEAPS Data
Agreeableness	LEAPS	4.33	0.69	1	5	4475	LEAPS Data
Emotional stability	LEAPS	2.28	1.08	1	5	4475	LEAPS Data
Panel B. Other samples							
Grit	Adults aged 25 and older (US)	3.65	0.73	1	5	1545	Duckworth et al. (2007)
Grit	West point cadets 2010 (US)	3.75	0.54	1	5	1308	Duckworth et al. (2007)
Grit	Adults between 15-64 years old	2.72	0.6	1	4	3843	STEP Data Kenya
Grit	Adults between 15-64 years old	2.98	0.61	1	4	3978	STEP Data Macedonia
Openness to experience	Adults between 15-64 years old	3	0.56	1	4	3844	STEP Data Kenya
Openness to experience	Adults between 15-64 years old	3.28	0.55	1	4	3979	STEP Data Macedonia
Conscientiousness	Adults between 15-64 years old	3.22	0.52	1.33	4	3844	STEP Data Kenya
Conscientiousness	Adults between 15-64 years old	3.05	0.5	1	4	3979	STEP Data Macedonia
Extraversion	Adults between 15-64 years old	2.85	0.59	1	4	3845	STEP Data Kenya
Extraversion	Adults between 15-64 years old	3.02	0.61	1	4	3979	STEP Data Macedonia
Agreeableness	Adults between 15-64 years old	2.86	0.57	1	4	3843	STEP Data Kenya
Agreeableness	Adults between 15-64 years old	3.28	0.59	1	4	3978	STEP Data Macedonia
Emotional stability	Adults between 15-64 years old	2.69	0.5	1	4	3843	STEP Data Kenya
Emotional stability	Adults between 15-64 years old	2.09	0.66	1	4	3979	STEP Data Macedonia

Notes. This table shows descriptive statistics for the grit and big-five measures. Panel A reports on the LEAPS sample. Panel B shows results from other samples reported on in the literature. For the STEP data, we computed the results using the STEP surveys conducted by the World Bank, in 2013. Sample sizes differ due to item non-response.

Table 5. Relationship between schooling and skills formation

	(1) Cognitive skills	(2) Cognitive skills	(3) Functional Lit. and Num.	(4) Functional Lit. and Num.	(5) SEM skills	(6) SEM skills
Years of schooling	0.169***	0.168***	0.165***	0.165***	0.033***	0.033***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Respondent age	-0.009***	-0.009***	0.001	0.001	0.007**	0.008*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Respondent is female	0.124***	0.129***	-0.156***	-0.156***	-0.621***	-0.616***
	(0.022)	(0.022)	(0.020)	(0.020)	(0.033)	(0.035)
Mother highest grade	0.011**	0.012**	0.001	-0.001	-0.001	-0.001
	(0.005)	(0.005)	(0.004)	(0.005)	(0.006)	(0.006)
Father highest grade	0.010***	0.010***	-0.001	-0.000	-0.003	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
HH SES in 2003	0.015***	0.011*	-0.001	-0.004	0.007	0.010
	(0.005)	(0.006)	(0.006)	(0.006)	(0.008)	(0.009)
Constant	-1.330***	-1.367***	-1.324***	-1.379***	-0.214*	-0.056
	(0.076)	(0.074)	(0.087)	(0.083)	(0.113)	(0.107)
Observations	4,399	4,399	4,402	4,402	4,395	4,395
Adjusted R-squared	0.659	0.665	0.668	0.677	0.153	0.171
Village FE	No	Yes	No	Yes	No	Yes
District FE	Yes	No	Yes	No	Yes	No
Number of clusters	108	108	108	108	108	108

* p<.10 ** p<.05 *** p<.01

Notes. The dependent variable in columns (1) and (2) is the mean of the Urdu, Mathematics, and English scores computed using Item Response Theory with a two-parameter logistic (2PL) model. Items come from the paper-based and tablet-based, computer-adaptive tests (leaving out items that less than 50 respondents answered and items that less than 5% or more than 95% of respondents solved correctly). The dependent variable in columns (3) and (4) is a functional literacy and numeracy index computed using principal-component factor analysis on 17 real-life literacy and numeracy questions. The dependent variable in columns (5) and (6) is a socioemotional (SEM) skills index computed using principal-component factor analysis on the big-five items, grit items, BART, and GoNoGo scores. The sample for each regression consists of respondents who completed the direct version of the questionnaire, in-person. Sample sizes differ due to item non-response. Regressions in columns (2), (4), and (6) include village fixed effects; regressions in columns (1), (3), and (5) include district fixed effects. Standard errors are in parentheses (clustered at the village level).

Table 6. Schooling, skills and labor force participation

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Working	Working	Working	Working	Working	Working	Working	Working
Years of schooling	-0.022***	-0.024***	-0.017***	-0.0076**	0.012***	0.012***	0.0097***	0.0099***
	(0.0018)	(0.0020)	(0.0031)	(0.0032)	(0.0012)	(0.0013)	(0.0021)	(0.0022)
Cognitive skills			-0.046***	-0.016			0.016	0.013
			(0.013)	(0.014)			(0.011)	(0.012)
SEM skills			0.054***	0.056***			0.0016	0.00068
			(0.0097)	(0.0098)			(0.0062)	(0.0059)
Constant	0.17***	0.20***	0.18***	1.05***	-0.058***	-0.062***	-0.048***	-0.070***
	(0.024)	(0.028)	(0.033)	(0.032)	(0.0062)	(0.0067)	(0.012)	(0.022)
Observations	2595	2217	2217	1978	2360	2174	2174	1925
Adjusted R-squared	0.14	0.15	0.16	0.060	0.060	0.060	0.060	0.060
Mean dependent	0.79	0.76	0.76	0.83	0.06	0.07	0.07	0.06
Sample	All	Skills measures	Skills measures	Skills measures, no students	All	Skills measures	Skills measures	Skills measures, no students
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* p<.10 ** p<.05 *** p<.01

Notes. This table reports estimates of the relationship between schooling, skills, and labor force participation for men and women in Pakistan. All the columns report mean regressions estimates. "Working" is an indicator variable equal to 1 if the respondent is currently employed. The cognitive skills index is the mean of the Urdu, Mathematics, and English scores computed using Item Response Theory with a two-parameter logistic (2PL) model. The socioemotional (SEM) skills index is computed using principal-component factor analysis. Column (1) includes all men in the sample. Among them, one has missing information for years of schooling. Column (5) includes all women in the sample. For columns (2), (3), (6), and (7), the sample includes men/women who answered the direct version of the questionnaire. Columns (4) and (8) exclude any currently enrolled individuals. All regressions include age and district fixed effects. Standard errors are in parentheses (clustered at the village level).

Table 7. Schooling, skills, migration, and earnings for men in Pakistan

	All				Working men			
	(1) Monthly earnings	(2) Monthly earnings	(3) Monthly earnings	(4) Monthly earnings	(5) Monthly earnings	(6) Monthly earnings	(7) Monthly earnings	(8) Monthly earnings
Years of schooling (a1)	4.12***	3.37***	1.10*	0.063	2.89***	2.76***	1.97***	2.09***
	(0.50)	(0.46)	(0.60)	(0.84)	(0.62)	(0.60)	(0.57)	(0.61)
Cognitive skills (a2)			8.31***	0.26	9.50***	3.47*	7.71***	3.31*
			(2.77)	(3.59)	(2.35)	(1.78)	(2.04)	(1.94)
SEM skills (a3)			16.5***	14.3***	10.3***	6.00***	11.6***	7.88***
			(1.86)	(2.71)	(1.98)	(2.02)	(1.81)	(1.64)
Out Village (a4)				32.9**		35.4***		33.5***
				(13.2)		(13.4)		(12.2)
Interaction YrsSchooling and Out Village (b1)				4.78***		2.36*		1.97
				(1.44)		(1.42)		(1.33)
Interaction Cog and Out Village (b2)				13.6**		11.6**		10.3*
				(6.35)		(5.45)		(6.04)
Interaction SEM and Out Village (b3)				-7.37		-4.49		-4.02
				(5.17)		(4.69)		(4.16)
Brain index (job content)							9.20***	6.95***
							(2.21)	(2.12)
Brawn index (job content)							-7.32***	-5.45***
							(1.81)	(1.86)
Team work (job content)							23.1***	11.6***
Observations	2340	1978	1978	1978	1636	1636	1636	1636
Pseudo R-squared	0.034	0.030	0.043	0.092	0.055	0.090	0.068	0.096
Median Dependent	120.19	115.38	115.38	115.38	134.62	134.62	134.62	134.62
Sample	All	Has skills measures	Has skills measures	Has skills measures	All	All	All	All
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
a1 + b1=0				4.84***		5.12***		4.06***
				(1.24)		(1.29)		(1.24)
a2 + b2=0				13.89***		15.07***		13.62**
				(5.33)		(5.17)		(5.78)
a3 + b3=0				6.89		1.51		3.86
				(4.45)		(4.27)		(3.90)

* p<.10 ** p<.05 *** p<.01

Notes. This table reports estimates of the relationship between schooling, skills, migration, and earnings for men in Pakistan. All the columns report median regressions estimates. The dependent variable is raw monthly earnings. The cognitive skills index is the mean of the Urdu, Mathematics, and English scores computed using Item Response Theory with a two-parameter logistic (2PL) model. The socioemotional (SEM) skills index is computed using principal-component factor analysis. The out of village variable indicates whether the respondent lives outside the village where we initially surveyed his household. The brain index indicates how much the individual has to engage into intellectual activities as part of their job. The brawn index indicates how much physical and difficult work the individual's job requires. The teamwork variable is an indicator for jobs in which the individual is required to often or always work in a team. Column (1) refers to all men in the sample who are not currently enrolled. The sample for columns (2) to (4) consists of those who answered the direct version of the questionnaire, in-person. The sample for columns (5) to (8) includes all working men. All regressions include age and district fixed effects. Robust standard errors are in parentheses.

Table 8. Schooling, skills and earnings for women in Pakistan

	All			Working women					
	Mean regressions			Median regressions			Mean regressions		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings	Monthly earnings	Monthly earnings	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)
Years of schooling	1.73***	1.91***	1.35***	5.92***	4.81***	5.25**	8.47***	8.51***	8.14*
	(0.28)	(0.31)	(0.43)	(1.47)	(1.60)	(2.31)	(1.96)	(1.95)	(4.12)
Cognitive skills			3.13*			-3.22			1.74
			(1.83)			(9.75)			(18.2)
SEM skills			0.80			-7.45			3.51
			(0.91)			(7.99)			(7.99)
Constant	-11.9***	-13.1***	-8.22*	1.48	19.2	9.56	-28.6	-27.6	-23.2
	(2.37)	(2.49)	(4.28)	(37.4)	(43.6)	(48.4)	(31.0)	(31.8)	(47.3)
Observations	2104	1925	1925	112	111	111	112	111	111
R-squared	0.060	0.060	0.070	0.22	0.22	0.22	0.26	0.26	0.25
Median/Mean dependent	7.22	7.84	7.84	62.50	57.69	57.69	107.33	107.43	107.43
Sample	All	Has skills measures	Has skills measures	All	Has skills measures	Has skills measures	All	Has skills measures	Has skills measures
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* p<.10 ** p<.05 *** p<.01

Notes. This table reports estimates of the relationships between schooling, skills, and earnings for women in Pakistan. The dependent variable "Monthly earnings" in columns (4) to (6) is the raw monthly earnings while the dependent variable "Monthly earnings (TC)" in columns (1) to (3) and (7) to (8) is top coded at 100,000 PKR per month (961.5 USD). The cognitive skills index is the mean of the Urdu, Mathematics, and English scores computed using Item Response Theory with a two-parameter logistic (2PL) model. The items included in the model are those from the paper test and the computer adaptive test on tablet (excluding items that less than 50 respondents answered and those that less than 5% or more than 95% of respondents got it right). The socioemotional skills index is computed using principal factor analysis on the Big-Five items, Grit items, BART, and GoNoGo scores. The sample for column (1) is all women in the sample who are not currently enrolled, that is 2,104 women. For columns (2) and (3), the sample is only those who answered the direct version of the questionnaire, in-person (they have skills measures): 1,925 women. The sample for columns (4) and (7) is all women in the sample who are working, that is 112 women. For columns (5), (6), (8), and (9), the sample is only those who answered the direct version of the questionnaire, in-person: 111 women. All the regressions include age and district fixed effects. Robust standard errors are in parentheses. The R-squared shown is the pseudo R-squared for median regressions and the adjusted R-squared for mean regressions.

APPENDICES

Appendix A. Tracking and Survey Instruments

The sample for this long-term tracking exercise consisted of all individuals between 5 and 15 years old who were part of the 1,807 households sampled for the first LEAPS round in 2003 (5,865 respondents). We had last attempted to survey these respondents in 2011 before we started a tagging and tracking exercise in 2016. The 2011 survey was conducted at the household level, and we collected information about individual respondents through household rosters. In 2016, we were interested in tracking individuals.

The first step in the tagging and tracking process was to go back to the latest addresses we had for these respondents' households. We could not locate the households of 285 respondents (Table A1). 79% of those we could not find belonged to households that had already attrited by 2011.

For the households we could find, we implemented the following protocol. If the target individual had migrated, we would collect their address and contact information. We attempted to survey all the respondents in person—by visiting them at their new address (for those who had migrated within the country) or by waiting for them to visit their relatives (for those who had migrated abroad). For any individual we were able to meet in person, we would administer five different instruments:

- Questionnaire
- Cognitive skills assessment (on paper)
- Adaptive Cognitive skills assessment (tablet-based)
- Socioemotional skills assessment (on paper)
- Socioemotional skills assessment (tablet-based)

We administered a shorter survey over the phone to individuals we could not meet in person. The phone-based survey was similar to the questionnaire administered in person but shorter. We could not conduct any cognitive or socioemotional skills assessment over the phone.[24] We are therefore missing skills information for all the individuals surveyed over the phone (Table A1).

Finally, we could not survey some respondents in person or over the phone. These respondents mainly fell into two categories: women who got married and were living with their family-in-law and our team was not authorized to visit (30%), and men working in neighboring Arab countries for whom we did not have contact information (30%). Whenever possible, we collected information about these respondents from another person in their original household (parents or siblings, in 80% of the cases). We called this survey mode the "indirect" version of the survey as someone else was *indirectly* giving information about the respondent of interest. This questionnaire was short and

[24] Crawfurd et al. (2021) report differential item functioning across in-person and phone surveys for a cognitive assessment in Sierra Leone. Moreover, the adaptive test and part of the socioemotional skills assessment had to be implemented on tablets.

designed to collect the main variables of interest that a close relative would know about (such as educational attainment, employment status, marital status, and the number of children). Table A1 summarizes the number of respondents for each survey type and the corresponding type of information we have.

Appendix B. Attrition and Sample Selection

To account for attrition, we apply inverse-probability weights (IPW) following Wooldridge (2010). We predict the probability that the respondent has skills measures rather than the probability of having any data (see Appendix A and Table A1 for more details). We follow this approach as our main results relate skills to labor market outcomes, effectively treating individuals with missing skills measures as attritors in these regressions.

We implement two alternative approaches to modeling this probability. We start with a basic model that uses only correlates of attrition from the first round of data collection of the LEAPS project in 2003. We then add variables indicating if the respondent had already attrited in 2011 and 2016 (when we started the tracking process). We describe these two approaches in turn.

Model 1. Probability of having full data (including skills measures) as a function of variables collected in 2003

We start with a list of 35 variables from 2003 and select variables that are predictive of having skills measures using Lasso. This procedure leads us to keep 31 variables, including the respondent's sex, age, housing characteristics (toilet on-premises, electricity access, etc.), parental education and occupation, whether the household head could read and write, and the language of the interview. We then use a probit model to predict the probability that we have full data for the respondent.

Model 2. Probability of having full data (including skills measures) as a function of variables collected in 2003 and indicators of attrition in 2011 and 2016

We use the same specification as in Model 1 but add two variables: a dummy indicating that the individual was already an attritor in 2011 and one indicating that the individual was an attritor in 2016, at the early stage of the tagging and tracking process.

Then, we used two additional methods to assess the robustness of our main results for sample selection:

- First, we applied the Heckman two-step selection model. For the selection equation, we used as instruments an indicator for whether the mother was born in the village, the number of boys in the family, whether the father was living in the household in 2003, whether the male interview was conducted in another language than Punjabi in 2003, and the number of schools in the village in 2003. We also control for the household size and wealth in 2003, whether the family owns their house in 2003, the quality of home construction (whether the house uses permanent materials like concrete) in 2003 and the individual's age in 2003.

- Second, we also conducted a sequential bivariate semi-nonparametric estimation for the correction term: we first predict attrition from the sample, and then whether the individual is included in our wage regression. To predict sample attrition, we used an indicator for whether the mother was born in the village, the number of boys in the family, whether the father was living in the household in 2003, and whether the male interview was conducted in another language than Punjabi in 2003. To predict inclusion in our regression analysis sample, we use the number of schools in the village in 2003. We use the Stata package developed by De Luca (2008) to carry out the sequential bivariate semi-nonparametric estimation for the sample attrition and selection.

Appendix C. Measurement of Cognitive Skills

C.1. Adaptive Cognitive Test

To capture the skills of our diverse pool of respondents, we worked with an organization to design an adaptive tablet-based test. The organization we partnered with developed 324 items ranging from early primary level to college level. The test classified respondents into six levels that correspond to different grades. The mapping between levels and grades is as follows:

- Level 1: Nursery, Grades 1 to 3 (early primary)
- Level 2: Grades 4 and 5 (late primary)
- Level 3: Grades 6 to 8 (middle school)
- Level 4: Grades 9 and 10 (high school)
- Level 5: Grades 11 and 12 (intermediate)
- Level 6: College

The items of the tests were designed to capture the following learning domains: (1) mastery over concepts and definitions (e.g., "what is a pronoun?"), (2) application mastery (e.g., "add 2+2"), and (3) evaluation mastery (e.g., "two boys meet two girls, one boy leaves; how many children are left?"). As we expected many respondents to have been out of school for a long time, items were designed to test general mastery (as opposed to specific terms or formulae). All the test items were multiple-choice questions with four possible answer choices and one correct answer.

The logic of the test was as follows.

- Everyone started at the same level – Level 2 for Urdu and Mathematics and Level 1 for English – and answered a batch of 6 questions.
- If the respondent got 5-6 questions right, they moved to the next higher level (or, if at Level 6, remained at Level 6).
- If the respondent got 3-4 questions right, they stayed at the same level.

- If the respondent got 0-2 questions right, they moved to the next lower level (or, if at Level 1, remained at Level 1).

Then, the placement logic of the test was as follows.

- The first time that a respondent completed three batches of 6 questions at any level
    - If the level was Level 1, and the last score was 0-2 questions right, the respondent was placed at Level 1.
    - If the level was Level 1-6, and the last score was 3-4 questions right, the respondent was placed at that level.
    - If the level was Level 6, and the last score was 5-6 questions right, the respondent was placed at Level 6.
    - If the level was Level 1-5, and the last score was 5-6, the respondents was moved to the next higher level, and the test continued.
    - If the level was Level 2-6, and the last score was 0-2, the respondents was moved to the next lower level, and the test continued.
- The first time a respondent completed three batches of questions at any level and any time after that at the next higher level scored 0-2 questions right, the respondent was placed at the level where the respondent had completed three batches of questions.

To complete the placement, the minimum batches of questions were three (and the minimum number of questions was 18). To complete the placement, the maximum batches of questions was 17 (and the maximum number of questions was 102). The test took 15 minutes to complete on average. Examples of progress and placement logic are provided below.

Example 1.

Batch	Level	# of items correct
1	2	3
2	2	4
3	2	4

The respondent answered 18 questions and was placed at Level 2 (the respondent completed three batches of questions at Level 2, and the last score was 3-4 questions right).

Example 2.

Batch	Level	# of items correct

1	2	5
2	3	2
3	2	4
4	2	5
5	3	2

The respondent answered 30 questions and was placed at Level 2 (the respondent completed three batches of questions at Level 2, and the first time after that they got to Level 3, they only got two questions right).

Example 3.

Batch	Level	# of items correct
1	2	2
2	1	6
3	2	0
4	1	4
5	1	4

The respondent answered 30 questions and was placed at Level 1 (the respondent completed three batches of questions at Level 1, and the last score was 3-4 questions right).

C.2. Item Response Theory

We aggregate items from the tablet-based adaptive test and the paper-based test for each subject into Urdu, Mathematics, and English scores using Item Response Theory (IRT). Item Response Theory is a set of mathematical models that describe the relationship between an individual's latent "trait" θ, and their manifestations (performance on a test). They establish a link between that latent trait, the properties of items on a test, and how individuals respond to these items.
In the two-parameter IRT model, the likelihood of answering a question correctly is determined by the ability of the respondent, θ, and two items parameters – difficulty (labeled a) and discrimination (labeled b). In this model, the probability that a respondent answers a given question *j* correctly is modeled as:

$$P_j(\theta) = \frac{1}{1 + \exp\left(-a_j(\theta_j - b_j)\right)}$$

The difficulty parameter, a, represents the ability level at which 50% of respondents get the item right. For example, for an item with a difficulty parameter of 1, a respondent with an ability level of 1 standard deviation over the mean has a 50% chance of getting this item right. The discrimination parameter, a, captures how quickly the likelihood of success changes with respect to ability.
The Item Characteristic Curve (ICC) depicts the likelihood of a correct answer, $P_j(\theta)$, as a function of θ. The higher the individual's ability, the higher the probability of a correct response. We plot the Item Characteristic Curve (solid line) and the actual pattern responses against 40 quantiles of θ for the paper test items for each of the three subjects in Figures A3 to A5. We observe a tight fit between the predicted responses based on the ICC and the actual responses in the data for all the items of the LEAPS test. Then, we produce the same plots for 25 randomly selected items from the adaptive test, for each of the three subjects, in Figures A6 to A8. For the adaptive test, the fit between the predicted and actual responses varies depending on the subject and specific items.

C.3. Functional Literacy and Numeracy Assessment

We also designed an assessment to capture proficiency in everyday arithmetic and literacy skills. The assessment was divided into three sections.

The first part asked the respondent to read an electricity bill and answer the following questions:

- How much money do you need to pay for the month of November?
- How much money will you pay if you pay the amount after the due date?
- Imagine the following: The meter reading on your electricity bill for the month of March is 2500 kV units. Each kV unit costs Rs. 2. Please note that you have to pay a late fee of Rs.500 if you miss the due date. How much money do you have to pay for the month of March to cover your electricity bill if you pay before the due date? After the due date?
- Now, imagine a scenario where the electricity bill is charged as follows. The meter reading is 2500 kV units. The first 500 kV will be charged at Rs. 10; any kV units more than 500 will be charged at Rs. 20, as shown in the table below:

kV	Cost per unit (in Rs.)
0-500	10
501 and above	20

  Please note that you must pay a late fee of Rs. 500 if you miss the due date. How much money do you have to pay for the month of March to cover your electricity bill if you pay before the due date? After the due date?

The second part of the assessment asked respondents to read text messages written in Urdu and Roman Urdu (Urdu but using roman language script). We asked them to read one greeting message, one conversation, one advertisement text, two messages from a school, and four emergency text messages (in Urdu and Roman Urdu). The messages in Roman Urdu were:

Greeting messages

Peace be upon you. How are you and how is everyone at home?
[Urdu: Asalam O alikum kya haal hai app ka aur ghar mein sub ka kya haal hai?]

Conversation

Person 1: How are you? [Urdu: Kya haal hai?]
Person 2: I am fine. [Urdu: Mein theek hoon]
Person 1: Anything new? [Urdu: Koi nayi tazi?]
Person 2: Exams are going on at school. [Urdu: Papers chal rahay hain school mein.]
Person 1: To obtain something and to be successful, you will often face problems in life, but success comes to those who work hard and do not get scared. Work hard, well done. [Urdu: Kuch hasil kar ney aur maqaam bananay kay liye zindagi mein mushkilaat aati hain aur kamyabi aun ko milti hai jo dat jatay hain aur dartnay nahi. Mehnat karo Shabaash.]
Person 2: Thank you. [Urdu: Shukariya.]

Advertisement

Great news! Before the 30th of this month, recharge your balance and get 1000 minutes and 2000 texts absolutely free. To get more information, dial 1212.

[Urdu: Shandar Khabar! Iss mahinay ki 30 tareekh se phelay apnay balance ko recharge ki jiye aur payi 1000 minutes aur 2000 sms bilkul muft. Mazeed malomaat kay liye 1212 dial karien.]

School Text messages

1. Your child has not come to school today. Is everything ok? [Urdu: Apka bacha aj school nahi aya. Sub khariat hai?]
2. Your child was absent today again, and today was his exam in Urdu. Is everything ok? [Urdu: Apkay bachay ne aj phir chuti kar li hai aur aj aus ka urdu ka imtehaan thaa. Kya sub khairiat hai?]

Emergency text messages

1. I am going to be home late today. Please don't worry. [Urdu: Mein aaj ghar per dair se aaon ga. Pareshaan mat hona.]
2. Friend, I am stopped outside the village and my motorcycle has broken down. Can you pick me up? [Urdu: Yaar mein gaon se bahir ruka huwa hoon aur meri motorcycle kharab hogayi hai. Kya mujhay lenay aasaktay ho?]
3. Tomorrow there will be no water in the houses of this village from morning till evening, so please make your arrangements beforehand. [Urdu: Kal gaon kay gharon mein subah se sham tak paani nahi aye ga tou isliye phelay se intezaam kar lain.]

4. Brother, we ran out of flour, so on your way back, can you pick up some flour because we need to make rotis for dinner. [Urdu: Bhai ata khatam hogaya hai wapisi per aata letay ana khanay kay liye roti banani hai.]

We then asked similar messages in the Urdu script.

Finally, the last part mimicked a market transaction. We asked:

Imagine you go to a shop and you have Rs. 300 with you. You get the following items from the shopkeeper and give him three notes of Rs. 100 each.

1 kg rice	Rs. 30
1 kg potatoes	Rs. 20
1 kg sugar	Rs. 20
Surf	Rs. 25
Cooking oil	Rs. 100

How much money should the shopkeeper return to you if you purchase all the items at the same shop?

Appendix D. Measurement of Socioemotional Skills

The design of the socioemotional skills assessment resulted from an iterative process that started with data collection in Cambodia for a related project in 2017. The data from this project suffers from some shortcomings that we detail in section D.3. below. Nevertheless, the data are useful for two reasons. First, we used what we learned from Cambodia when designing and piloting our instruments in Pakistan. In particular, this is what motivated us to conduct a pilot that (1) was large-scale, (2) took place far ahead of the data collection, and (3) included a test-retest component. Second, we can check whether the general pattern of results is consistent across Cambodia and Pakistan.

We start by describing our methods to assess the reliability of, and validity evidence for, our instruments. We then describe our process to select the instruments we kept in the full data collection.

D.1. Acquiescence bias correction

We assessed the quality of our measures by evaluating their reliability and validity evidence. Before assessing these two aspects, we followed Laajaj and Macours (2021) and corrected the self-reported items for acquiescence bias (the tendency to agree rather than disagree with questions). We apply the following procedure[25]:

1. Compute the average score on reverse-coded items and the average score on non-reversed-coded items.
2. Take the average of the two averages obtained in the first step.
3. Subtract the scale mid-point (for instance 3, if the possible answers follow a 5-point format that ranges from 1 ("Disagree strongly") to 5 ("Agree strongly")). This step gives an estimate of acquiescence bias.
4. Subtract the acquiescence bias score obtained in the third step from every non-reverse-coded item and add it to every reverse-coded item.

D.2. Measures of reliability and validity

D.2.1. Reliability

A measure is highly reliable if it produces similar results under consistent conditions. We provide two types of reliability estimates: internal consistency and test-retest reliability.

Internal consistency

Internal consistency is the extent to which all the items in a scale reliably measure the same attributes, or the interrelatedness of scale items. To assess internal consistency, we compute the Cronbach's alpha statistic (Cronbach, 1951). Cronbach's alpha is computed as

$$\alpha = \frac{K}{K-1}\left(1 - \frac{\sum_{i=1}^{K}\sigma_{Y_i}^2}{\sigma_X^2}\right)$$

, where K is the number of items in the scale, σ_X^2 is the variance of the observed total test score, and $\sigma_{Y_i}^2$ is the variance of responses to item i for the current sample of persons. It measures how correlated the items of a scale are and is also a direct function of the number of items in the scale.
The statistic is a ratio of variances and therefore lies between 0 and 1. A Cronbach's above 0.7 indicates acceptable internal consistency, but a Cronbach's alpha above 0.9 may indicate item redundancy (Oxford Mind & Behaviour team[26]).

Cronbach's alpha provides an assessment of the reliability of a scale as well as its construct validity. If there is a lot of classical measurement error (low reliability), the items of a scale will be poorly correlated, and Cronbach's alpha will be low. However, even with no measurement error, if the items do not measure the same underlying construct, their correlation will also be low, leading to a small Cronbach's alpha. Finally, non-classical measurement error can lead to an artificially high Cronbach's alpha. If all the items suffer from systematic response bias, they may be highly correlated, resulting in

[25] This procedure is equivalent to the procedure described by Laajaj and Macours (2021).
[26] https://mbrg.bsg.ox.ac.uk/method/measuring-non-cognitive-skills-psychometric-validation-scales

a high Cronbach's alpha. Although we corrected for acquiescence bias before estimating Cronbach's alpha for the different scales, this remains a potential limitation.

Test-retest reliability

The test-retest reliability measures how correlated the responses of the same individuals to the same instrument are at two different points in time. Suppose an instrument measures the true ability we intend to measure. Then, this ability should not vary over a short period (usually between two weeks to one month), and the two measures should be highly correlated. On the other hand, if the measure mostly captures (classical) measurement error, the test-retest correlation will be low.
The test-retest correlation is a measure of reliability. Under classical measurement error, the observed value of the variable X is equal to the true value of X plus a purely random component. We can write the measured value of X as the sum of the true value X^*, plus a measurement error ε:

$$X = X^* + \varepsilon$$

Where $E(\varepsilon) = 0$ and $cov(\varepsilon, X) = 0$.
Then, the variance of X is equal to:

$$\sigma_X = \sigma_{X^*} + \sigma_\varepsilon$$

The test-retest correlation is a measure of reliability, defined as the share of the variance of X driven by the true variance of the variable, as opposed to measurement error:

$$Relability = \frac{\sigma_{X^*}}{\sigma_X} = \frac{\sigma_X - \sigma_\varepsilon}{\sigma_X} = 1 - \frac{\sigma_\varepsilon}{\sigma_X}$$

A test-retest correlation above 0.7 is generally considered high. If measurement error is non-classical, errors could be correlated over time. In this case, we would overestimate the reliability of a measure. For instance, this would be the case if the answer patterns suffer from systematic acquiescence or social desirability bias.

D.2.2. Validity

An instrument is said to be valid in a specific context if it measures what it is supposed to measure. We investigate three aspects of validity: face validity, predictive validity, and content validity.

Face validity

Face validity ensures that the questions asked are perceived as measuring the concepts the instrument intends to measure. In other words, when we ask respondents a question aimed at assessing their emotional stability, they should subjectively perceive it as such. We assessed respondents' understanding and perception of the questions through debriefing sessions during the pilot.

Predictive validity

Predictive validity ensures that the measures are correlated with the variables we would expect, according to theory or existing empirical evidence. For instance, in theory, we would expect grit to be correlated with educational attainment and fewer career changes (Duckworth et al., 2007). Similarly, locus of control should be positively correlated with desirable labor market outcomes. People with a stronger internal locus of control perform more complex activities and have better job performance (Judge and Bono, 2001).

Content validity

Content validity refers to the extent to which the items of a test represent all facets of a given construct. For instance, the content of an instrument aiming at measuring the Big Five personality traits is said to be valid if it captures the five dimensions of personality: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. To assess content validity, we rely on exploratory factor analysis. Exploratory analysis is used to analyze patterns of correlations between the items variables to infer their relationship to an unknown variable—here, an index of socioemotional skills. To determine the number of factors, we use two criteria commonly used in the literature: the Kaiser criterion (1958), where one keeps only the factors with eigenvalues[27] higher than 1 and the scree plot criterion (Cattell, 1966) where one only keeps the factors up until the line (which plots the eigenvalues) becomes flatter. We perform exploratory factor analysis on all the items from the self-reported scales, corrected for acquiescence bias. In Table 3, we show the results from the factor analysis for the full data collection. For each item, we indicate the main factor the item loads on. Items measuring the same skill/construct are expected to load on the same factor. We also show the uniqueness of the item, which is the percentage of variance for the item that is not explained by the common factors. A high uniqueness could indicate either measurement error, or that the item is measuring something different than the remaining items.

D.3. Selection process

Here, we describe the iterative process that led us to select the two self-reported scales and the two tablet-based tasks included in our data collection.

D.3.1. Cambodia project

In Cambodia, the study took place in the context of a randomized control trial aimed at measuring the long-run impacts of primary-school scholarships. In 2008, the Cambodian government offered scholarships to students as they were beginning the fourth grade of primary school. The new study tracked and attempted to survey a subsample of 3,825 children from three poor and remote

[27] The higher the eigenvalue the higher the percentage of the total variation in the variables that is explained by the factor.

Cambodian provinces that constituted the original experiment population. These children were in grades 3 and 4 in 2008. Data collection for the baseline took place from December 2008 to January 2009, and data collection for the new round of follow-up surveys took place from December 2016 to May 2017. The subsample targeted those children that had not attrited in a previous round of tracking in 2011. The team was able to survey 3,294 respondents, or 86.1% of the target subsample. However, overall attrition rates were higher than in Pakistan and differential across experimental groups (e.g., 24.2% among grade-4 students in the control group and 32.7% among grade-4 students in the treatment group). Another shortcoming of the data collection in Cambodia is that we did not systematically collect the current location of respondents (this information is missing for 18% of the surveyed sample). The survey content was very similar to that of Pakistan. In particular, the team implemented both a cognitive and socioemotional skills assessment and collected data on labor market outcomes and family formation.

The cognitive assessment was a computer-adaptive math test in which respondents answered ten questions from a larger pool of 23 items. These items are aggregated in a cognitive skills index with a two-parameter Item Response Theory model.

To measure socioemotional skills, the team used the same Grit and “Big Five” self-reported scales as in Pakistan. We also measured growth mindset—the belief that one can get smarter through hard work and practice—using a 4-item scale. We also measured respondents’ locus of control—the degree to which people believe that they have power over the outcome of events in their lives, as opposed to external forces beyond their control—using a four-item scale. Finally, we administered the Strengths and Difficulties Questionnaire (SDQ), which is a brief behavioral screening questionnaire. All items are answered using a 5-point answer format that ranges from 1 (“Disagree strongly”) to 5 (“Agree strongly”). Cronbach’s alphas for these different measures, before and after acquiescence bias corrections, are displayed in Table 2. Only the SDQ passes the 0.7 threshold. The Cronbach’s alphas for the other measures range from 0.04 to 0.58. Moreover, we randomly re-surveyed 13% of randomly selected respondents within the same week, using a subset of items. Thus, we can compute the test-retest correlations for the Big-Five constructs. These are low, with an average of 0.3 (Table 2), reflecting a large amount of measurement error. Moreover, when conducting exploratory factor analysis, the skills factor structure is not reproduced. Only three factors are retained, and one single factor encompasses many items that aim to measure different concepts.

D.3.2. Pakistan pilot

Given the results from Laajaj and Macours (2021) and those from the data collection in Cambodia, we decided to conduct a large pilot before our data collection in Pakistan. We conducted the pilot in the district of Okara, between February and March 2018. Interested participants were invited to a central location, resulting in a total sample size of 403 pilot respondents. Then, two weeks after the completion of the first round, we tracked and re-surveyed a random half of them (201 respondents).

The survey was again completed with this group, which allowed us to compute the test-retest reliability for our instruments.

On top of the instruments included in the final data collection and described in Section II.2. of the paper, we included two additional self-reported scales and tasks on tablets. The first self-reported scale was the same locus of control scale as in Cambodia. The second was the Barratt Impulsiveness Scale (BIS), which is a 30-item scale aimed at measuring impulsiveness. Then, we also included two tasks aimed at measuring grit. The first was a "frustration task." It consists of a split-screen interface with the option to either complete a difficult mirror-tracing task or play some games. It lasts five minutes and the outcome is the proportion of time spent on the tracing task rather than playing the games. Then, we also used the "Alan and Ertac Grit task" (Alan, Boneva, and Ertac, 2019). In this task, respondents are presented with a grid that contains different numbers where the goal is to find pairs of numbers that add up to 100. There is one easy game and one difficult game, the latter of which provides a higher reward. At the end of each round, individuals receive feedback, and they get to choose which type of task they want to do in the following round. The outcome is the probability of choosing the difficult game in all rounds. The Barratt Impulsiveness Scale showed satisfactory reliability measures. Its Cronbach's alpha was 0.71, and the test-retest correlation was 0.4 (which is low but among the highest in the Pakistan pilot). However, it took a long time to administer this instrument; therefore, we preferred to keep the Grit and Big Five scales that were also used in Cambodia and found more reliable during the Pakistan pilot. We dropped the frustration task because respondents had difficulty understanding it and since the screen size was not well suited for it. Finally, we found the Alan and Ertac Grit task to measure a combination of cognitive skill and grit rather than grit itself. Therefore, we decided against keeping it.

Appendix E. Earnings Measurements and Distributions

In this section, we describe how we collected information on respondents' earnings in the different versions of the survey (i.e., in-person, phone, and indirect) and how we aggregated these responses.

We started by collecting respondents' employment status for their two main activities. For each activity, respondents could fall into five employment categories: daily wages, salaried, self-employed, family business, and agriculture. For salaried respondents, we collected their monthly wages. For respondents with daily wages, who were self-employed or working in a family business, we collected monthly earnings for the past two months (to account for income volatility) and averaged the two measures. Agricultural earnings were collected in multiple ways. For respondents who answered the survey in-person, we collected detailed information on output quantities, prices, and input costs. Using these data, we calculated a measure of "computed agricultural earnings." For all respondents, we checked if they had been engaged in agriculture for more than four weeks during the last year with the goal of selling their production, and if so, how much money they earned from it. We divide their answer by 12 and call this variable the monthly "estimated agriculture earnings." Moreover, to accurately capture the earnings of women who may be informally working in their village, we also asked if they were engaged in tutoring, sewing, or any other activities against payment during the last month.

We compute three versions of the earnings variable. For each version, we aggregate the monthly earnings from the different activities in which the respondent was engaged. For respondents who completed the in-person version of the questionnaire, we also added the extra income earned by women. The three versions are:

- Version 1: We use the "estimated agriculture earnings" for everyone, including respondents who answered the direct version of the questionnaire and are doing agriculture as their main activity. We replace the earnings of respondents who do not earn any money with zero.
- Version 2: We use the "computed agriculture earnings" for respondents who completed the in-person survey and are doing agriculture as their main activity. For respondents who completed the indirect or phone-based version of the questionnaire, as well as for other respondents who have "estimated" agriculture earnings but are not engaged in agriculture as their main activity, we use the "estimated agriculture earnings." We replace the earnings of people who do not earn any money with zero.
- Version 3: We do not include agriculture earnings in this version, as there are a lot of outliers in these earnings. We replace the wage of people who do not earn any money with zero.

We create a capped version of each variable, with a cap of 100,000 PKR (approximately 961 USD). We convert the raw and the capped variables to USD. Figure A2 shows the distribution of the raw and capped earnings for male respondents. We use the capped earnings for all mean regressions and the raw earnings for all median regressions of the paper.

Appendix F. Mean vs. Median Regressions and Influential Observations

Following a referee's suggestions, we searched for observations influencing the mean regression results and then investigated whether dropping these observations helped reconcile the mean and median regression estimates from Table 7. We used three different methods to identify these observations. The methods are:

- Standardized residuals. We drop observations with residuals that are two or more standard deviations away from the expected value.
- Cook's distance. Cook's distance measures the effect of deleting an observation with large residuals (outliers) and/or high leverage. Cook's distance is calculated by removing the i_{th} data point from the model and recalculating the regression. We drop observations with a Cook's distance higher than $\frac{4}{N}$, where $N$is the number of observations.
- Dif beta. Dif beta measures how much an observation influences a parameter estimate, say b_j. It is calculated by removing an observation, say i, recalculating b_j , say $b_{j_{\{-i\}}}$, taking the difference in betas and standardizing it. We drop observations with a dif beta above $\frac{2.}{\sqrt{N}}$.

When we drop these observations, the mean and median regression results are similar and close to the initial median regression results (see Tables A10 and A11). We show the mean and median regressions from Table 7 and the mean and median regressions when dropping observations according to the three criteria above. Our preferred criterion is the Cook's distance, which shows that:

a) The coefficient on years of schooling is small and not significant in either the mean or median regression. This result is closer to the outcome of the median regression using the whole sample.
b) The coefficient on cognitive skills is highly significant for both the median and mean regressions. This finding is also in line with the outcomes of the median regression using the full sample.
c) The coefficient on the interaction term for cognitive skills and "Out village" is positive and significant at the 5% level in the median regression and positive but not statistically significant at conventional levels and smaller in the mean regression.

Given these results, our preferred estimates use the median regressions for men. Note that for women, we need to continue reporting the mean regressions for the full sample, as the median earnings are 0.

Appendix – References

Alan, S., Boneva, T., Ertac, S., 2019. Ever failed, try again, succeed better: Results from a randomized educational intervention on grit. *The Quarterly Journal of Economics*, *134*(3), 1121-1162.

Cattell, R.B., 1966. The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), pp.245-276.

Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), pp.297-334.

De Luca, G., 2008. SNP and SML estimation of univariate and bivariate binary-choice models. *The Stata Journal*, *8*(2), 190-220.

Duckworth, A.L., Peterson, C., Matthews, M.D., Kelly, D.R., 2007. Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*(6), 1087.

Judge, T.A. and Bono, J.E., 2001. Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: A meta-analysis. *Journal of applied Psychology*, *86*(1), p.80.

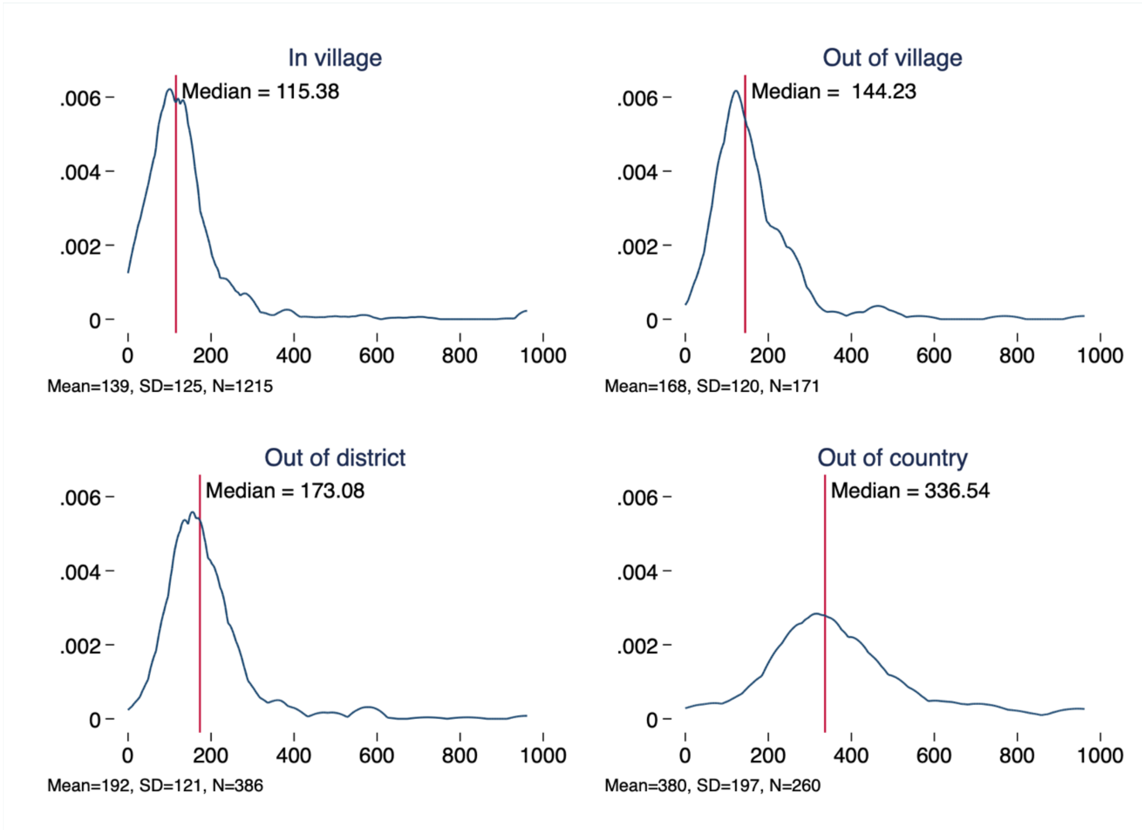
Kaiser, H.F., 1958. The varimax criterio n for analytic rotation in factor analysis. *Psychometrika*, *23*(3), pp.187-200.

Laajaj, R., Macours, K., 2021. Measuring skills in developing countries. *Journal of Human resources*, *56*(4), 1254-1295.

Wooldridge, J. M.,2010. Econometric analysis of cross section and panel data. *MIT press*.

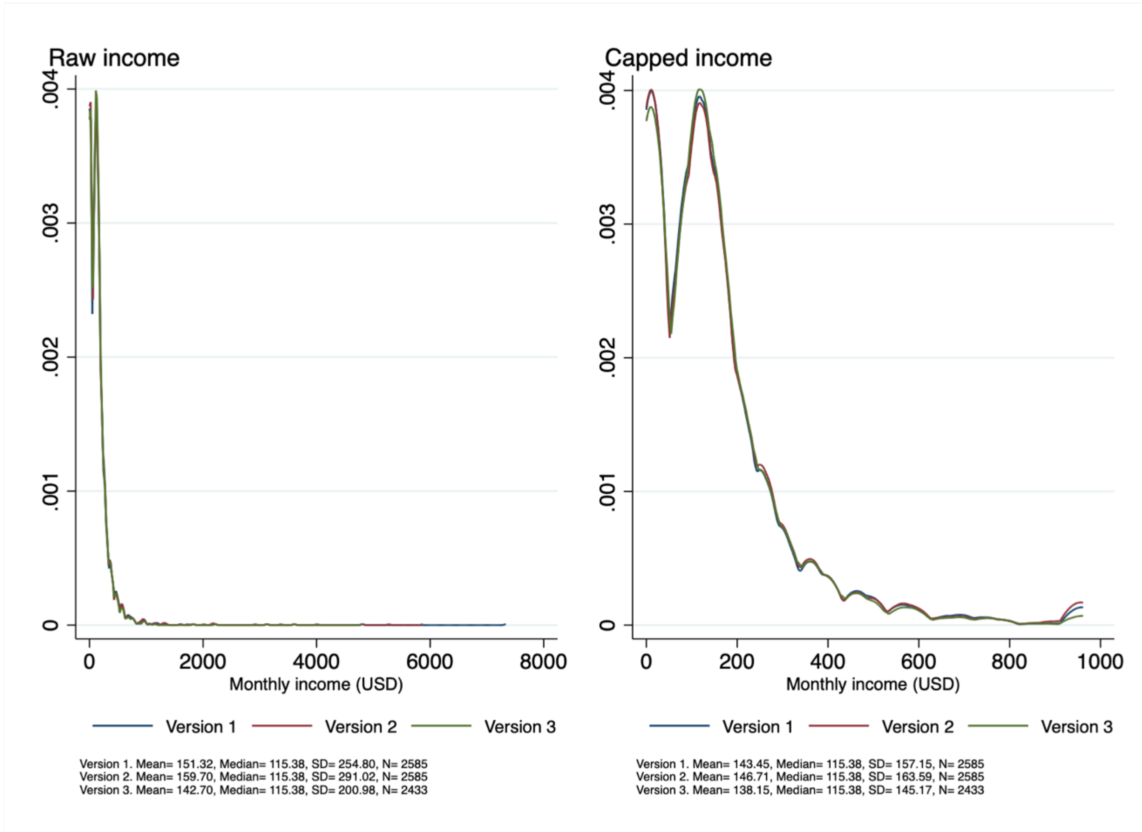
APPENDICES – FIGURES

Figure A1. Monthly earnings of men by location



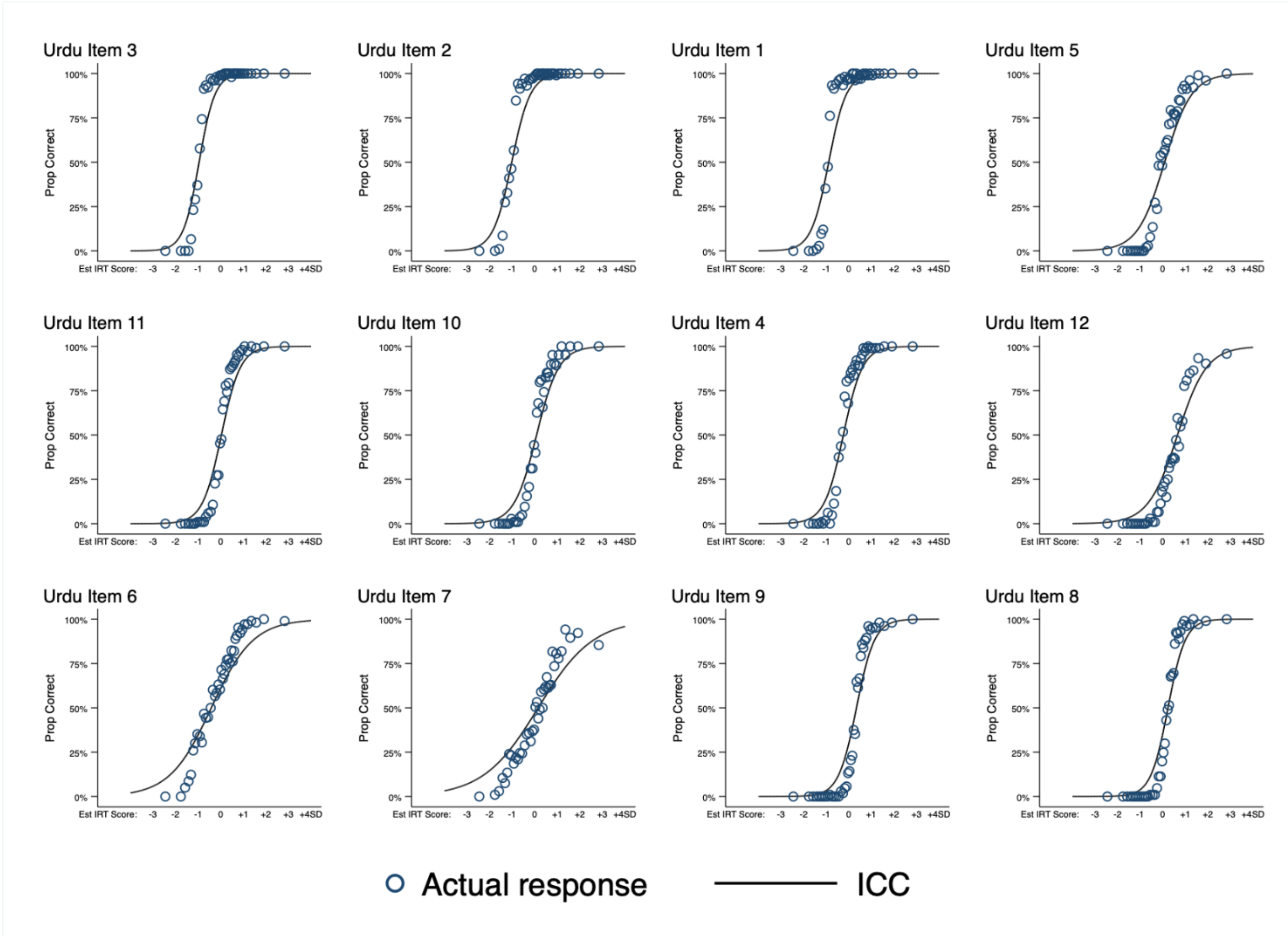
Notes. This figure shows the distribution of monthly earnings for male respondents in our sample depending on where they lived in 2018. Red lines indicate the median. The sample excludes those not working.

Figure A2. Distribution of monthly earnings among men



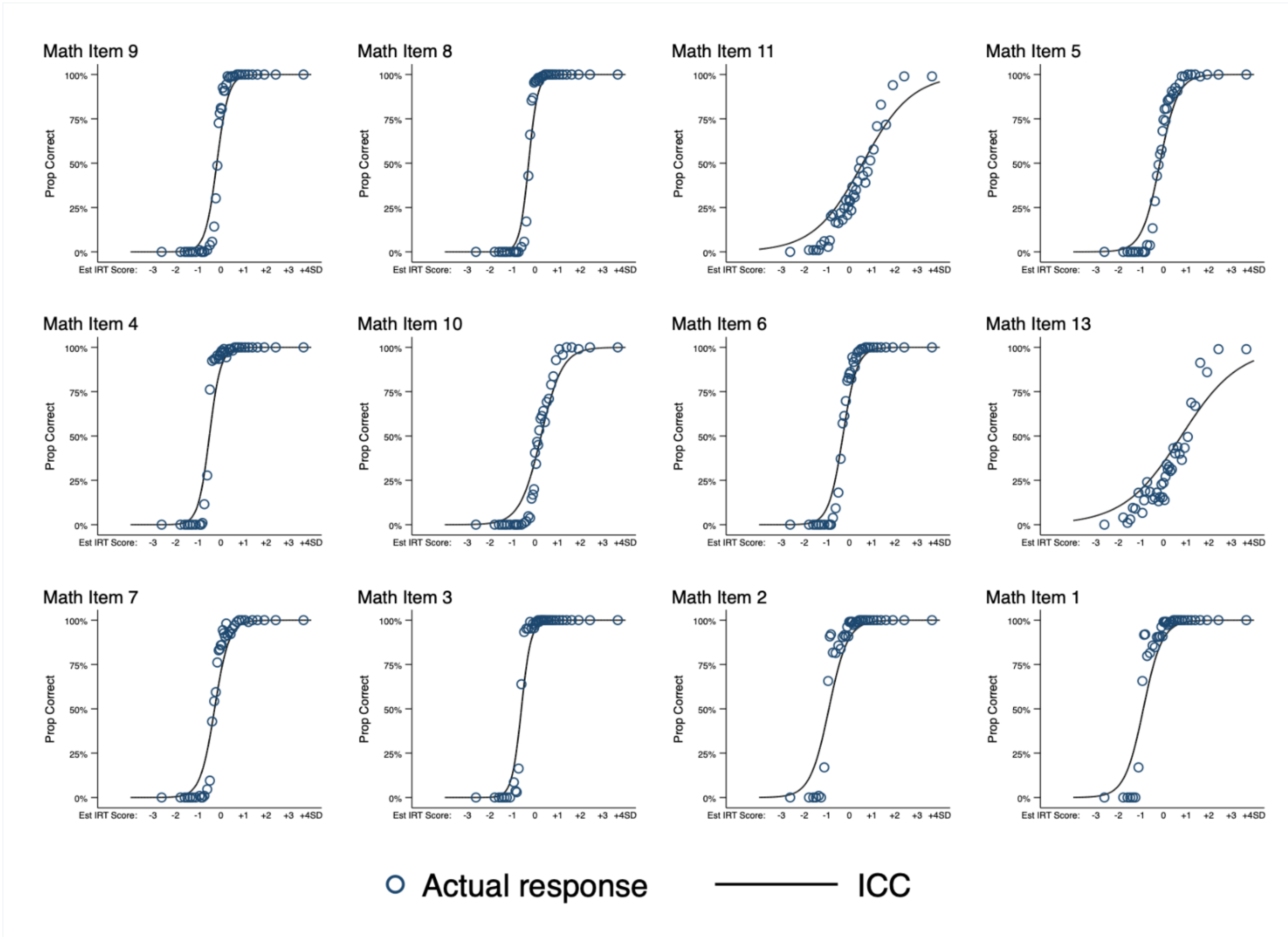
Notes. This figure shows the distribution of monthly earnings for men in our sample, using different ways of aggregating respondents' earnings.

Figure A3. Item Characteristics Curves and actual response patterns—paper-based test, Urdu



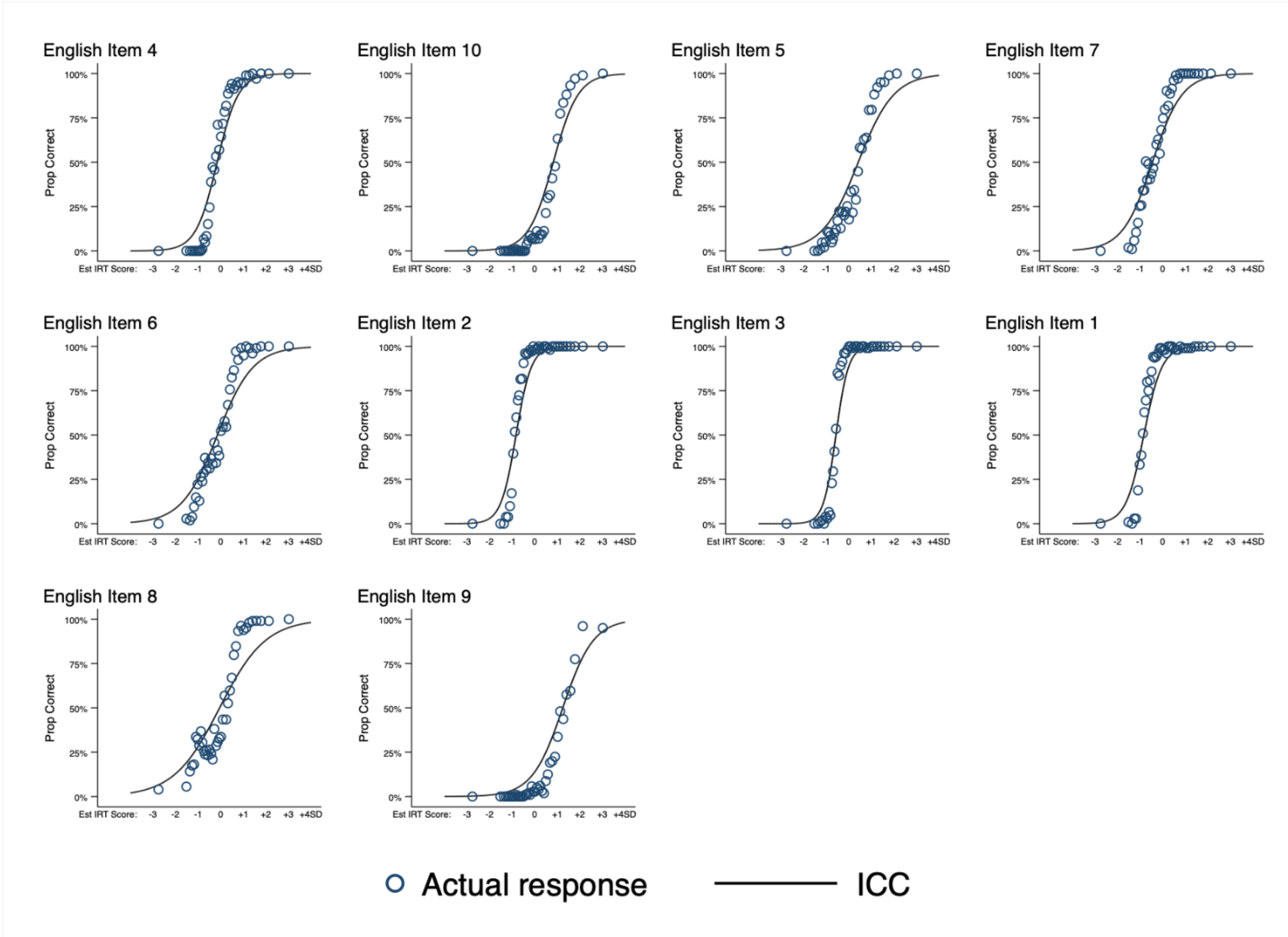
Notes. We use a two-parameter Item Response Theory (IRT) to model the likelihood of answering a question correctly. In this model, the probability of getting a question right is determined by the ability of the respondent, θ, and two item parameters—difficulty and discrimination. The solid line in each graph is the Item Characteristic Curve (ICC), which represents the expected patterns of responses for each θ. The actual pattern of responses against 40 quantiles of θ is then plotted against it.

Figure A4. Item Characteristics Curves and actual response patterns—paper-based test, Mathematics



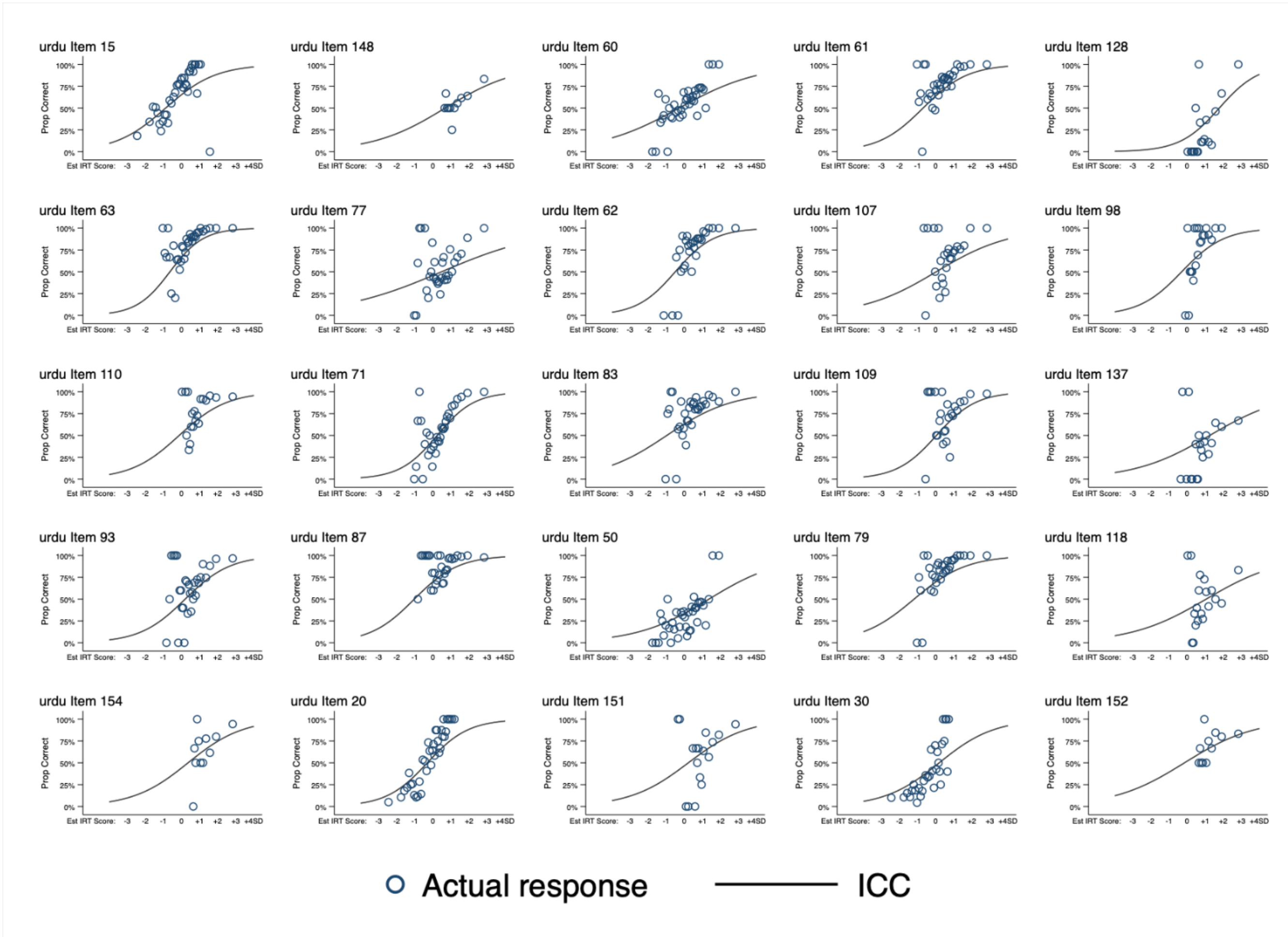
Notes. We use a two-parameter Item Response Theory (IRT) to model the likelihood of answering a question correctly. In this model, the probability of getting a question right is determined by the ability of the respondent, θ, and two item parameters—difficulty and discrimination. The solid line in each graph is the Item Characteristic Curve (ICC), which represents the expected patterns of responses for each θ. The actual pattern of responses against 40 quantiles of θ is then plotted against it.

<u>Figure A5</u>. Item Characteristics Curves and actual response patterns—paper-based test, English



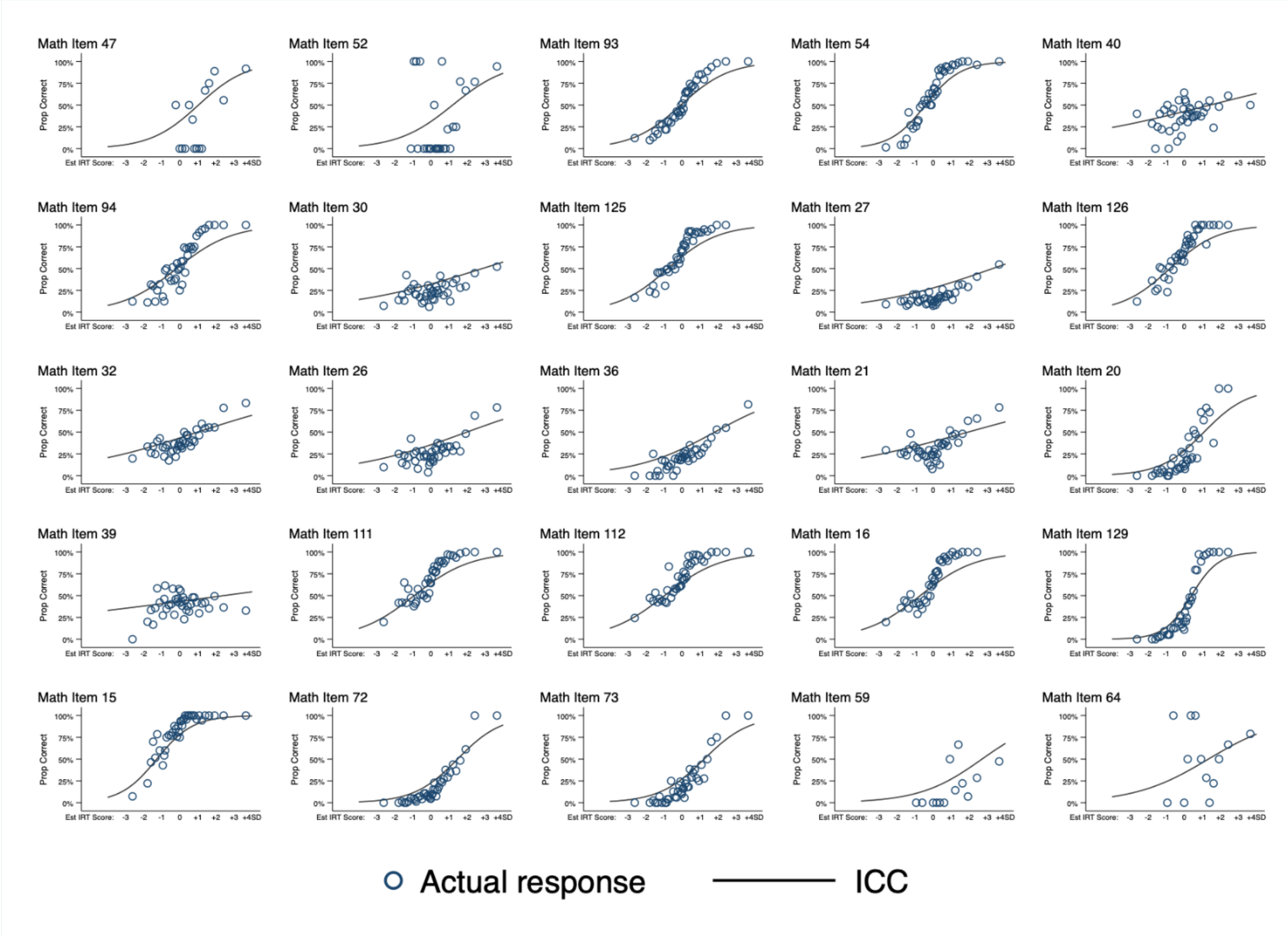
Notes. We use a two-parameter Item Response Theory (IRT) to model the likelihood of answering a question correctly. In this model, the probability of getting a question right is determined by the ability of the respondent, θ, and two item parameters—difficulty and discrimination. The solid line in each graph is the Item Characteristic Curve (ICC), which represents the expected patterns of responses for each θ. The actual pattern of responses against 40 quantiles of θ is then plotted against it.

Figure A6. Item Characteristics Curves and actual response patterns—adaptive test, Urdu



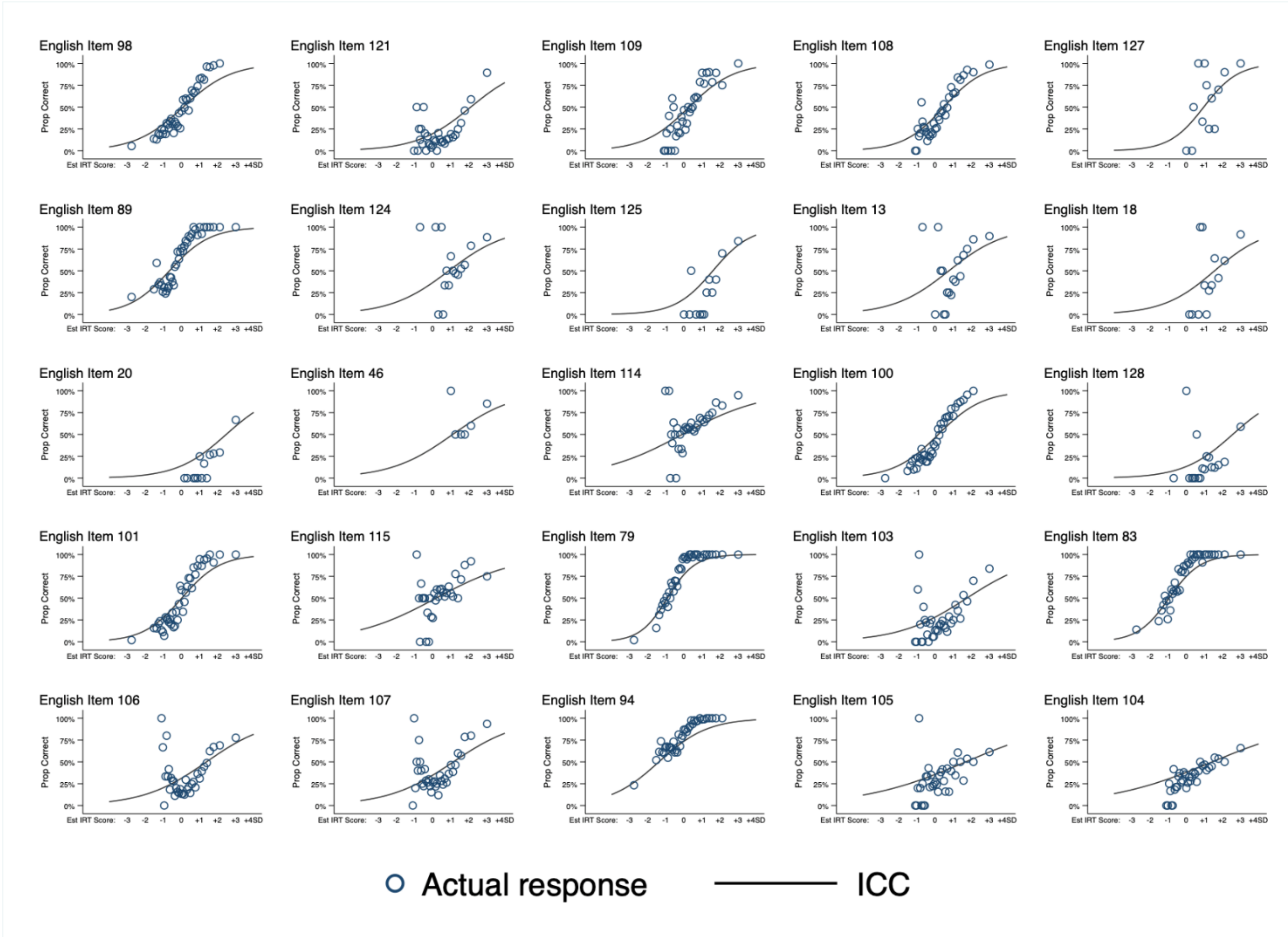
Notes. We use a two-parameter Item Response Theory (IRT) to model the likelihood of answering a question correctly. In this model, the probability of getting a question right is determined by the ability of the respondent, θ, and two item parameters—difficulty and discrimination. The solid line in each graph is the Item Characteristic Curve (ICC), which represents the expected patterns of responses for each θ. The actual pattern of responses against 40 quantiles of θ is then plotted against it.

Figure A7. Item Characteristics Curves and actual response patterns—adaptive test, Mathematics



Notes. We use a two-parameter Item Response Theory (IRT) to model the likelihood of answering a question correctly. In this model, the probability of getting a question right is determined by the ability of the respondent, θ, and two item parameters—difficulty and discrimination. The solid line in each graph is the Item Characteristic Curve (ICC), which represents the expected patterns of responses for each θ. The actual pattern of responses against 40 quantiles of θ is then plotted against it.

Figure A8. Item Characteristics Curves and actual response patterns—adaptive test, English



Notes. We use a two-parameter Item Response Theory (IRT) to model the likelihood of answering a question correctly. In this model, the probability of getting a question right is determined by the ability of the respondent, θ, and two item parameters—difficulty and discrimination. The solid line in each graph is the Item Characteristic Curve (ICC), which represents the expected patterns of responses for each θ. The actual pattern of responses against 40 quantiles of θ is then plotted against it.

APPENDICES – TABLES

Table A1. Definition of attrition

	In-person	Phone	Indirect	Died	4 Attock villages	No info	Refused/Other	Total
No attrition	4391	0	0	0	0	0	0	4391
Skills measures missing	15	79	471	0	0	0	0	565
Full attrition	0	0	0	43	186	285	395	909
Total	4406	79	471	43	186	285	395	5865

Notes. Respondents with "No Attrition" are respondents for whom we have the complete questionnaire, cognitive skills measures (either from the test on tablet or on paper), and socioemotional skills measures. Respondents with "skills measures missing" are respondents for whom we either do not have the cognitive skills measures (the paper-based test and tablet-based test missing) or the socioemotional skills measures. These respondents are excluded from the regressions in which we include skills measures. For the in-person version of the survey, four respondents did not finish the survey, 11 respondents for whom the paper-based test was forgotten, and respondents for whom there was a bug with the test on tablets. Respondents with "Full Attrition" are respondents for whom we could not collect questionnaires. Forty-three of them had died, 186 were living in four villages we could not visit in Attock because they fell under military control, 285 were part of households we could not locate, and 395 refused.

Table A2. Analysis of differential attrition (Panel A)

	N	(1) **Full data** *Mean [SD]*	N	(2) **No skills measures** *Mean [SD]*	N	(3) **Full attrition** *Mean [SD]*	t-test *Diff* *(1)-(2)*	t-test *Diff* *(1)-(3)*	*Total obs.*
Location - 2018									
Original village	4391	0.67	565	0.17	395	NA	0.50***	NA	4956
		[0.47]		[0.38]			(0.02)		
Within district	4391	0.19	565	0.15	395	NA	0.04**	NA	4956
		[0.39]		[0.36]			(0.02)		
Within country	4391	0.13	565	0.29	395	NA	-0.16***	NA	4956
		[0.34]		[0.46]			(0.02)		
Outside country	4391	0.01	565	0.39	395	NA	-0.37***	NA	4956
		[0.12]		[0.49]			(0.02)		
Individual - 2018									
Age	4391	23.62	565	24.96	395	NA	-1.34***	NA	4956
		[3.64]		[3.23]			(0.15)		
Years of schooling	4391	8.56	564	8.40	395	NA	0.16	NA	4955
		[4.96]		[4.33]			(0.20)		
Ever married	4391	0.45	565	0.48	395	NA	-0.04*	NA	4956
		[0.50]		[0.50]			(0.02)		
Age at first marriage	1974	20.89	274	21.40	395	NA	-0.51**	NA	2248
		[3.39]		[3.69]			(0.25)		
Has children	4391	0.33	565	0.35	395	NA	-0.02	NA	4956
		[0.47]		[0.48]			(0.02)		
Working	4391	0.42	565	0.64	395	NA	-0.22***	NA	4956
		[0.49]		[0.48]			(0.03)		
Main work is farming	4391	0.04	565	0.01	395	NA	0.03***	NA	4956
		[0.19]		[0.08]			(0.00)		
HH has toilets on premises	4391	0.96	13	0.85	395	NA	0.12	NA	4404
		[0.19]		[0.38]			(0.09)		
HH has access to electricity	4391	0.98	13	0.92	395	NA	0.06	NA	4404
		[0.14]		[0.28]			(0.07)		

* p<.10 ** p<.05 *** p<.01

Notes. Columns (1) to (3) display the means for the group that answered all parts of the survey, the group surveyed that did not complete the cognitive and socioemotional assessments (either because they were surveyed over the phone or information was collected indirectly), and the group who attrited fully. Standard deviations are shown in brackets. Column (4) is the difference between the mean of respondents with complete data and those who do not have skills measures. Column (5) is the difference between the mean of respondents with complete data and the mean of attritors. Panel A shows differences in variables measured in this latest round of data collection, in 2018, so we cannot report data on full attritors. Panel B reports data measured during the tracking exercise of 2016. Panel C reports data measured in 2003 at the individual level. Panel D reports data measured in 2003 at the household level. Differences in means are computed by OLS regressions. All standard errors in parentheses are clustered at the village level.

Table A2. Analysis of differential attrition (Panel B)

	N	(1) **Full data** *Mean [SD]*	*N*	(2) **No skills measures** *Mean [SD]*	*N*	(3) **Full attrition** *Mean [SD]*	t-test *Diff (1)-(2)*	t-test *Diff (1)-(3)*	*Total obs.*
Location - 2016									
Original village	4375	0.74	562	0.42	552	0.71	0.32***	0.03	5489
		[0.44]		[0.49]		[0.45]	(0.02)	(0.03)	
Within district	4375	0.15	562	0.12	552	0.11	0.02	0.04***	5489
		[0.35]		[0.33]		[0.31]	(0.02)	(0.02)	
Within country	4375	0.09	562	0.21	552	0.13	-0.12***	-0.04**	5489
		[0.29]		[0.41]		[0.34]	(0.02)	(0.02)	
Outside country	4375	0.02	562	0.25	552	0.05	-0.22***	-0.03**	5489
		[0.16]		[0.43]		[0.22]	(0.02)	(0.01)	
Individual - 2016									
Respondent is female	4337	0.49	531	0.33	339	0.54	0.16***	-0.05*	5207
		[0.50]		[0.47]		[0.50]	(0.02)	(0.03)	
Age	4337	22.36	531	23.83	339	22.73	-1.48***	-0.38**	5207
		[3.72]		[3.17]		[3.46]	(0.15)	(0.18)	
Ever married	4337	0.38	531	0.40	339	0.29	-0.03	0.09***	5207
		[0.48]		[0.49]		[0.45]	(0.02)	(0.03)	
Enrolled	4173	0.15	381	0.03	299	0.19	0.12***	-0.03	4853
		[0.36]		[0.17]		[0.39]	(0.01)	(0.02)	

* p<.10 ** p<.05 *** p<.01

Notes. Columns (1) to (3) display the means for the group that answered all parts of the survey, the group surveyed that did not complete the cognitive and socioemotional assessments (either because they were surveyed over the phone or information was collected indirectly), and the group who attrited fully. Standard deviations are shown in brackets. Column (4) is the difference between the mean of respondents with complete data and those who do not have skills measures. Column (5) is the difference between the mean of respondents with complete data and the mean of attritors. Panel A shows differences in variables measured in this latest round of data collection, in 2018, so we cannot report data on full attritors. Panel B reports data measured during the tracking exercise of 2016. Panel C reports data measured in 2003 at the individual level. Panel D reports data measured in 2003 at the household level. Differences in means are computed by OLS regressions. All standard errors in parentheses are clustered at the village level.
.

Table A2. Analysis of differential attrition (Panel C)

	N	(1) Full data *Mean [SD]*	*N*	(2) No skills measures *Mean [SD]*	*N*	(3) Full attrition *Mean [SD]*	t-test *Diff* *(1)-(2)*	t-test *Diff* *(1)-(3)*	*Total obs.*
Individual - 2003									
Respondent age	4391	9.79	565	10.64	909	10.13	-0.85***	-0.34***	5865
		[2.99]		[2.78]		[2.85]	(0.13)	(0.10)	
Respondent is female	4391	0.49	565	0.33	909	0.50	0.16***	-0.01	5865
		[0.50]		[0.47]		[0.50]	(0.02)	(0.02)	
Highest grade completed	4267	2.10	552	2.76	879	2.22	-0.67***	-0.13	5698
		[2.22]		[2.30]		[2.30]	(0.12)	(0.10)	
Respondent with disability	4236	0.01	550	0.03	863	0.05	-0.01	-0.03***	5649
		[0.12]		[0.16]		[0.22]	(0.01)	(0.01)	
How good is respondent health, max is 16	4308	15.62	557	15.63	881	15.60	-0.01	0.02	5746
		[1.26]		[1.40]		[1.47]	(0.08)	(0.07)	
How intelligent is respondent, max is 5	4279	3.19	554	3.19	884	3.10	-0.00	0.09**	5717
		[0.69]		[0.74]		[0.71]	(0.03)	(0.03)	
How hardworking is respondent, max is 5	4279	3.06	554	2.98	884	2.99	0.08**	0.07*	5717
		[0.76]		[0.81]		[0.79]	(0.04)	(0.03)	
Was tested in 2003	4391	0.15	565	0.18	909	0.15	-0.02	0.00	5865
		[0.36]		[0.38]		[0.36]	(0.02)	(0.02)	
English IRT Score (2003)	670	-0.55	99	-0.88	138	-0.44	0.33**	-0.11	907
		[1.16]		[1.31]		[1.02]	(0.16)	(0.13)	
Math IRT Score (2003)	670	-0.31	99	-0.44	138	-0.41	0.13	0.11	907
		[0.96]		[1.04]		[1.07]	(0.13)	(0.11)	
Urdu IRT Score (2003)	670	-0.49	99	-0.77	138	-0.49	0.28**	0.01	907
		[1.18]		[1.24]		[1.13]	(0.12)	(0.12)	
Mean IRT Score (2003)	670	-0.45	99	-0.69	138	-0.45	0.25**	0.00	907
		[0.93]		[1.00]		[0.94]	(0.12)	(0.10)	

* p<.10 ** p<.05 *** p<.01

Notes. Columns (1) to (3) display the means for the group that answered all parts of the survey, the group surveyed that did not complete the cognitive and socioemotional assessments (either because they were surveyed over the phone or information was collected indirectly), and the group who attrited fully. Standard deviations are shown in brackets. Column (4) is the difference between the mean of respondents with complete data and those who do not have skills measures. Column (5) is the difference between the mean of respondents with complete data and the mean of attritors. Panel A shows differences in variables measured in this latest round of data collection, in 2018, so we cannot report data on full attritors. Panel B reports data measured during the tracking exercise of 2016. Panel C reports data measured in 2003 at the individual level. Panel D reports data measured in 2003 at the household level. Differences in means are computed by OLS regressions. All standard errors in parentheses are clustered at the village level.

<u>Table A2</u>. Analysis of differential attrition (Panel D)

		(1)		(2)		(3)	t-test	t-test	
		Full data		**No skills measures**		**Full attrition**	***Diff***	***Diff***	***Total obs.***
	N	***Mean [SD]***	***N***	***Mean [SD]***	***N***	***Mean [SD]***	***(1)-(2)***	***(1)-(3)***	
Household - 2003									
HH SES in 2003	4382	-0.12	564	-0.02	905	-0.38	-0.09	0.27*	5851
		[1.99]		[1.90]		[1.99]	(0.12)	(0.14)	
Family owns house it is living in	4391	0.95	565	0.96	909	0.90	-0.01	0.05**	5865
		[0.21]		[0.18]		[0.30]	(0.01)	(0.02)	
Number of rooms house	4391	2.50	565	2.53	909	2.45	-0.03	0.04	5865
		[1.36]		[1.36]		[1.36]	(0.08)	(0.11)	
Type of house is permanent	4391	0.68	562	0.67	909	0.59	0.01	0.09***	5862
		[0.46]		[0.47]		[0.49]	(0.02)	(0.03)	
HH has toilets on premises	4390	0.58	565	0.60	909	0.52	-0.02	0.06	5864
		[0.49]		[0.49]		[0.50]	(0.03)	(0.04)	
HH has hard roof	4391	0.55	565	0.51	909	0.53	0.03	0.01	5865
		[0.50]		[0.50]		[0.50]	(0.03)	(0.04)	
HH has access to electricity	4383	0.88	565	0.88	905	0.84	-0.01	0.04	5853
		[0.33]		[0.32]		[0.37]	(0.02)	(0.04)	
Relative HH wealth compared to rest of village, max is 4	4391	3.37	565	3.36	909	3.42	0.02	-0.04	5865
		[0.89]		[0.87]		[1.01]	(0.05)	(0.07)	
HH size	4391	8.86	565	8.57	909	8.24	0.28	0.62**	5865
		[4.07]		[3.15]		[2.77]	(0.22)	(0.28)	
HH religion is not Islam	4381	0.02	565	0.01	909	0.03	0.01**	-0.01	5855
		[0.15]		[0.09]		[0.17]	(0.01)	(0.01)	
Male interview language is not Punjabi	4383	0.26	561	0.17	906	0.21	0.10***	0.06*	5850
		[0.44]		[0.37]		[0.41]	(0.02)	(0.03)	
Father is not living in HH	4293	0.12	555	0.16	884	0.21	-0.03*	-0.08***	5732
		[0.33]		[0.36]		[0.40]	(0.02)	(0.02)	
Mother is not living in HH	4298	0.02	555	0.03	884	0.02	-0.01	-0.00	5737
		[0.15]		[0.17]		[0.16]	(0.01)	(0.01)	
Parents - 2003									
Parent years of schooling	7791	3.13	989	2.86	1538	2.84	0.27	0.29	10318
		[4.31]		[4.05]		[4.08]	(0.18)	(0.23)	
Parent can read	7944	0.38	1001	0.35	1561	0.36	0.02	0.02	10506
		[0.48]		[0.48]		[0.48]	(0.02)	(0.02)	
Parent number of children	7386	5.37	937	5.50	1421	5.29	-0.13	0.08	9744

		[1.82]		[1.82]		[1.65]	(0.11)	(0.12)	
Parent is working	7958	0.51	1006	0.50	1564	0.48	0.02	0.03*	10528
		[0.50]		[0.50]		[0.50]	(0.01)	(0.02)	
Parent main work is farming	7958	0.18	1006	0.18	1564	0.13	0.00	0.05***	10528
		[0.39]		[0.39]		[0.34]	(0.01)	(0.02)	

* p<.10 ** p<.05 *** p<.01

Notes. Columns (1) to (3) display the means for the group that answered all parts of the survey, the group surveyed that did not complete the cognitive and socioemotional assessments (either because they were surveyed over the phone or information was collected indirectly), and the group who attrited fully. Standard deviations are shown in brackets. Column (4) is the difference between the mean of respondents with complete data and those who do not have skills measures. Column (5) is the difference between the mean of respondents with complete data and the mean of attritors. Panel A shows differences in variables measured in this latest round of data collection, in 2018, so we cannot report data on full attritors. Panel B reports data measured during the tracking exercise of 2016. Panel C reports data measured in 2003 at the individual level. Panel D reports data measured in 2003 at the household level. Differences in means are computed by OLS regressions. All standard errors in parentheses are clustered at the village level.

Table A3. Employment category and respondent location for men

		Respondent location			
		In village	**Out of village**	**Out of district**	**Out of country**
Employment	**Daily wage**	30%	23%	14%	15%
	Salaried	31%	51%	76%	81%
	Self-employed or family business	26%	24%	9%	4%
	Agriculture	13%	2%	2%	0%
		100%	100%	100%	100%

Notes. This table shows the share of respondents in each employment category depending on where they currently live. Daily wage refers to someone working for an employer that pays a wage daily. Salaried refers to someone working for an employer that pays a wage monthly. It can be either in the formal or the informal sector as long as the individual receives a wage. Self-employed or family business refers to someone working for themselves or a family member (outside agriculture and livestock). Agriculture refers to someone working in agriculture and livestock for themselves or their family. If the respondent is doing agriculture for someone else for a monthly wage, they are categorized as Salaried (and not Agriculture). The sample includes all men currently working, including those who are simultaneously enrolled (2,043 men).

Table A4. Overview of instruments

Index	Construct	Instrument	Mode
Cognitive	Urdu	LEAPS test	Paper-based
Cognitive	Mathematics	LEAPS test	Paper-based
Cognitive	English	LEAPS test	Paper-based
Cognitive	Urdu	Adaptive test designed for study	Tablet
Cognitive	Mathematics	Adaptive test designed for study	Tablet
Cognitive	English	Adaptive test designed for study	Tablet
Functional	Literacy	Read and interpret electricity bill	Administered by enumerator
Functional	Numeracy	Read text messages in Urdu, roman Urdu and English	Administered by enumerator
Socioemotional	Grit	Grit scale	Self-reported instrument, administered by enumerator
Socioemotional	Big Five	Big Five scale	Self-reported instrument, administered by enumerator
Socioemotional	Self-control	GoNoGo task	Task-based (tablet)
Socioemotional	Risk-taking behavior	Balloon Analogue Risk Task (BART)	Task-based (tablet)

Notes. In this table, we present the instruments used to measure the three different skills we capture in our survey: cognitive skills, functional literacy and numeracy, and socioemotional skills. To measure cognitive skills, we first used a paper-based test with 12 items for Urdu, 13 for Mathematics, and 10 for English. Then, we also used an adaptive test administered on tablets designed especially for this study. We also designed 17 questions that tested functional literacy and numeracy, which are expected to be helpful in the respondents' everyday lives. Examples include understanding how much is due from an electricity bill or reading text messages. The socioemotional skills instruments included self-reported scales and tasks administered on tablets. We use the term "self-reported" as the respondents answered the items, but an enumerator was reading the question out loud, given that some respondents were illiterate.

Table A5. Learning over time in Pakistan

Subject	What is the question	% correct 2003	% correct 2011	% correct 2018	% correct 2018 (college)	N
Panel A. Respondents with test scores in 2003, 2011, and 2018						
Mathematics	Tick box next number that matches the number of objects	49 %	89 %	94 %	99 %	467
Mathematics	678+923	56 %	83 %	77 %	94 %	467
Mathematics	7/3=__	3 %	16 %	7 %	18 %	467
English	Match picture: Banana	63 %	94 %	91 %	99 %	467
English	Missing letter to match picture: Flag	27 %	76 %	70 %	96 %	467
English	Use word in sentence: deep	1 %	29 %	22 %	63 %	467
Urdu	Match picture: Book	74 %	98 %	95 %	98 %	467
Urdu	Join letters and write word: m-a-l-k	36 %	76 %	64 %	83 %	467
Urdu	Fill blank in the story by selecting the correct word 3	NA	NA	NA	NA	0
Panel B. Respondents with test scores in 2011 and 2018						
Mathematics	Tick box next number that matches the number of objects	NA	82 %	85 %	100 %	1643
Mathematics	678+923	NA	71 %	62 %	93 %	1643
Mathematics	7/3=__	NA	13 %	5 %	20 %	1643
English	Match picture: Banana	NA	88 %	82 %	99 %	1643
English	Missing letter to match picture: Flag	NA	69 %	60 %	97 %	1643
English	Use word in sentence: deep	NA	21 %	18 %	61 %	1643
Urdu	Match picture: Banana	NA	90 %	86 %	100 %	1643
Urdu	Join letters and write word: m-a-l-k	NA	66 %	53 %	85 %	1643
Urdu	Fill blank in the story by selecting the correct word 3	NA	44 %	36 %	74%	1643

Notes. This table shows a sample of questions for the three subjects tested on paper: Mathematics, Urdu, and English. We show the same sample questions as in Table 4 (Part 1). For each question, we indicate the proportion of respondents who answered the question correctly in 2003, 2011, and 2018, as well as the number of respondents. Panel A is restricted to respondents who were tested in 2003, 2011, and 2018, while Panel B is restricted to respondents who were tested in 2001 and 2018. The number of respondents for "% correct 2018 (college)" is restricted to the sub-sample of respondents who went to college (82 respondents in Panel A and 234 respondents in Panel B).

Table A6. Correlations between skills measures and education

	Schooling	Cognitive	Socioemotional	Functional lit. and num.
Years of schooling	1			
Cognitive skills index	0.8	1		
Socioemotional skills index	0.15	0.14	1	
Functional literacy and numeracy index	0.81	0.77	0.18	1

Notes. This table shows the bivariate correlations between years of schooling, the cognitive skills index, the socioemotional skills index, and the functional literacy and numeracy index. The cognitive skills index is the mean of the Urdu, Mathematics, and English scores computed using Item Response Theory with a two-parameter logistic (2PL) model. The items included in the model are those from the paper-based test and the tablet-based computer-adaptive test (leaving out items that less than 50 respondents answered and items that less than 5% or more than 95% of respondents solved correctly). The socioemotional skills index is computed using principal-component factor analysis on the Big Five items, Grit items, BART, and GoNoGo scores. The dependent variable in columns (5) and (6) is a functional literacy and numeracy index computed using principal-component factor analysis on 17 functional literacy and numeracy questions.

Table A7. Relationship between socioemotional skills and schooling in Pakistan

	(1) Conscientiousness	(2) Grit	(3) Self-control (GoNoGo)	(4) Risk-taking (BART)	(5) Openness to experience	(6) Extraversion	(7) Agreeableness	(8) Emotional stability
Years of schooling	0.019***	0.035***	-0.050***	0.020***	0.042***	0.002	0.002	0.009**
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Respondent age	0.006	0.009**	0.026***	0.006	-0.013***	-0.007	0.011***	-0.014***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Respondent is female	-0.264***	-0.454***	0.428***	0.016	-0.093***	0.070**	-0.219***	-0.733***
	(0.034)	(0.032)	(0.037)	(0.037)	(0.032)	(0.034)	(0.030)	(0.034)
Mother highest grade	-0.010	0.007	-0.011*	0.011	0.014**	0.002	-0.009	0.001
	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Father highest grade	0.002	-0.005	0.004	-0.005	0.002	0.008**	-0.008*	0.008*
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
HH SES in 2003	0.002	0.001	-0.008	0.003	-0.002	-0.001	0.003	0.026***
	(0.009)	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.009)
Constant	5.331***	4.883***	6.233***	1.509***	3.427***	2.799***	5.705***	2.838***
	(0.121)	(0.114)	(0.113)	(0.114)	(0.114)	(0.120)	(0.102)	(0.118)
Observations	4,475	4,395	4,378	4,377	4,475	4,475	4,475	4,475
Romano-Wolf p-values - Years of schooling	0.004	0.004	0.004	0.004	0.004	0.693	0.693	0.004
Adjusted R-squared	0.088	0.106	0.166	0.094	0.123	0.008	0.059	0.168
Village FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	No	No	No	No	No	No	No	No
Number of clusters	108	108	108	108	108	108	108	108
Test Yrs of school = Yrs of school (1) - Chi2 (p-value)	NA	13.83 (0.00)	132.44 (0.00)	0.05 (0.83)	20.56 (0.00)	9.87 (0.00)	17.47 (0.00)	3.13 (0.08)
Test Yrs of school = Yrs of school (2) - Chi2 (p-value)	13.83 (0.00)	NA	207.92 (0.00)	7.08 (0.01)	1.68 (0.20)	44.65 (0.00)	57.14 (0.00)	29.73 (0.00)
Test Yrs of school = Yrs of school (3) - Chi2 (p-value)	132.44 (0.00)	207.92 (0.00)	NA	115.85 (0.00)	217.49 (0.00)	78.02 (0.00)	83.18 (0.00)	87.06 (0.00)
Test Yrs of school = Yrs of school (4) - Chi2 (p-value)	0.05 (0.83)	7.08 (0.01)	115.85 (0.00)	NA	14.45 (0.00)	10.34 (0.00)	11.27 (0.00)	3.94 (0.05)
Test Yrs of school = Yrs of school (5) - Chi2 (p-value)	20.56 (0.00)	1.68 (0.20)	217.49 (0.00)	14.45 (0.00)	NA	55.23 (0.00)	54.18 (0.00)	44.72 (0.00)
Test Yrs of school = Yrs of school (6) - Chi2 (p-value)	9.87 (0.00)	44.65 (0.00)	78.02 (0.00)	10.34 (0.00)	55.23 (0.00)	NA	0.00 (0.97)	2.37 (0.12)
Test Yrs of school = Yrs of school (7) - Chi2 (p-value)	17.47 (0.00)	57.14 (0.00)	83.18 (0.00)	11.27 (0.00)	54.18 (0.00)	0.00 (0.97)	NA	2.03 (0.15)
Test Yrs of school = Yrs of school (8) - Chi2 (p-value)	3.13 (0.08)	29.73 (0.00)	87.06 (0.00)	3.94 (0.05)	44.72 (0.00)	2.37 (0.12)	2.03 (0.15)	NA

* p<.10 ** p<.05 *** p<.01

Notes. This table reports estimates of the relationships between socioemotional skills and schooling in Pakistan. The sample for columns (1), (5), (6), (7), and (8) includes all respondents who answered the direct version of the questionnaire, whether in person or over the phone. The sample for columns (2), (3), and (4) is restricted to respondents who answered the questionnaire in person. All regressions control for village fixed effects. Standard errors are in parentheses (clustered at the village level).

Table A8. Relationship between migration, schooling, and skills in Pakistan

	Men						Women		
	(1) Out of village	(2) Out of village	(3) Out of village	(4) Out of Pakistan	(5) Out of Pakistan	(6) Out of Pakistan	(7) Out of village	(8) Out of village	(9) Out of village
Years of schooling	0.006**	0.007**	-0.002	-0.000	0.001	0.002	-0.002	-0.002	0.003
	(0.003)	(0.003)	(0.004)	(0.002)	(0.001)	(0.001)	(0.003)	(0.003)	(0.004)
Cognitive skills			0.039**			-0.005			-0.031
			(0.015)			(0.005)			(0.020)
SEM skills			0.046***			-0.000			-0.009
			(0.010)			(0.003)			(0.012)
Mother highest grade	0.006	0.004	0.003	0.003	0.001	0.001	0.003	0.003	0.003
	(0.004)	(0.004)	(0.004)	(0.003)	(0.001)	(0.001)	(0.004)	(0.005)	(0.005)
Father highest grade	-0.002	0.000	-0.000	-0.003*	-0.001	-0.001	-0.002	-0.000	0.000
	(0.003)	(0.003)	(0.003)	(0.002)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
HH SES in 2003	-0.005	-0.010*	-0.010	0.006*	0.002*	0.002*	-0.014**	-0.016***	-0.015***
	(0.006)	(0.006)	(0.006)	(0.003)	(0.001)	(0.001)	(0.006)	(0.006)	(0.005)
Constant	0.007	0.003	0.106**	0.017	0.019	0.011	0.007	-0.001	-0.033
	(0.040)	(0.040)	(0.045)	(0.025)	(0.011)	(0.014)	(0.019)	(0.019)	(0.027)
Observations	2,595	2,217	2,217	2,595	2,217	2,217	2,360	2,174	2,174
Adjusted R-squared	0.026	0.023	0.035	0.025	0.006	0.005	0.128	0.129	0.130
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All	Has skills measures	Has skills measures	All	Has skills measures	Has skills measures	Has skills measures	Has skills measures	Has skills measures
Number of clusters	108	108	108	108	108	108	108	108	108
Mean dependent	0.348	0.267	0.267	0.104	0.0289	0.0289	0.430	0.395	0.395

* p<.10 ** p<.05 *** p<.01

Notes. The dependent variable "Out of village" in columns (1) to (3) and (7) to (9) indicates whether the respondent lives outside the village where we initially surveyed their household (their natal village most of the time). The dependent variable "Out of Pakistan" in columns (4) to (6) indicates whether the respondent lives outside of Pakistan. Out of the 2,360 women in the sample, only 11 reported residing outside of Pakistan. Therefore, we do not run these regressions for the sample of women. The cognitive skills index is the mean of the Urdu, Mathematics, and English scores computed using Item Response Theory with a two-parameter logistic (2PL) model. The items included in the model are those from the paper-based test and the tablet-based computer-adaptive test (leaving out items that less than 50 respondents answered and items that less than 5% or more than 95% of respondents solved correctly). The socioemotional (SEM) skills index is computed using principal-component factor analysis on the Big-Five items, Grit items, BART, and GoNoGo scores. The sample for columns (1) and (4) is all men surveyed. The sample for columns (2), (3), (5), and (6) is all men who answered the direct version of the questionnaire (for whom we have skills measures). The sample for column (7) is all women. The sample for columns (8) and (9) is all women who answered the direct version of the questionnaire (for whom we have skills measures). Standard errors are in parentheses (clustered at the village level). All the regressions include age and district fixed effects.

Table A9. Schooling, skills, migration, and earnings for men in Pakistan—Mean regression estimates

	(1) Monthly earnings (TC)	(2) Monthly earnings (TC)	(3) Monthly earnings (TC)	(4) Monthly earnings (TC)
Years of schooling (a1)	5.22***	4.55***	2.88**	0.69
	(0.93)	(0.86)	(1.15)	(1.06)
Cognitive skills (a2)			6.28	2.79
			(4.74)	(5.00)
SEM skills (a3)			14.6***	13.0***
			(3.28)	(3.14)
Out Village (a4)				8.35
				(22.8)
Interaction YrsSchooling and Out Village (b1)				9.36***
				(2.51)
Interaction Cog and Out Village (b2)				-4.71
				(11.5)
Interaction SEM and Out Village (b3)				-10.8
				(7.68)
Constant	27.1***	33.7***	56.3***	46.1***
	(8.44)	(7.99)	(9.71)	(10.0)
Observations	2340	1978	1978	1978
R-squared	0.060	0.040	0.050	0.13
Mean dependent	154.70	134.71	134.71	134.71
Sample	All	Has skills measures	Has skills measures	Has skills measures
District FE	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes
a1 + b1=0				10.05***
				(2.49)
a2 + b2=0				-1.92
				(10.22)
a3 + b3=0				2.27
				(7.19)

* p<.10 ** p<.05 *** p<.01

Notes. This table reports estimates of the relationships between schooling, skills, migration, and earnings for men in Pakistan. All the columns report median regressions estimates. The dependent variable is top-coded at 100,000 PKR per month (961.5 USD). The cognitive skills index is the mean of the Urdu, Mathematics, and English scores computed using Item Response Theory with a two-parameter logistic (2PL) model. The socioemotional (SEM) skills index is computed using principal-component factor analysis. The Out of village variable is a dummy variable taking the value 1 if the respondent lives outside the village where we originally surveyed their household. The sample for column (1) is all men in the sample who are not currently enrolled. For the rest of the columns, the sample is only those who answered the direct version of the questionnaire, in person (they have skills measures). All the regressions include age and district fixed effects. Standard errors are in parentheses (clustered at the village level)

Table A10. Reconciling mean and median regression results – first specification

	Paper regression		High Cook's distance		High standardized residuals		High dfbeta - cog		High dfbeta - school	
	Median	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*
	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Years of schooling (a1)	1.10*	2.88**	0.39	0.64	0.55	0.61	0.86	1.69*	1.08*	1.20
	(0.60)	(1.15)	(0.58)	(0.76)	(0.52)	(0.77)	(0.62)	(0.93)	(0.63)	(0.79)
Cognitive skills (a2)	8.31***	6.28	8.13***	8.53***	7.87***	8.60***	9.00***	10.6***	7.81***	10.3***
	(2.77)	(4.74)	(2.65)	(2.92)	(2.49)	(2.95)	(2.82)	(3.00)	(2.73)	(3.24)
SEM skills (a3)	16.5***	14.6***	17.1***	14.8***	17.3***	15.4***	17.2***	16.2***	17.3***	17.0***
	(1.86)	(3.28)	(1.80)	(2.09)	(1.66)	(2.11)	(1.89)	(2.75)	(1.95)	(2.69)
Constant	42.1	56.3***	24.8	23.6	41.5	54.2***	24.5	62.5	18.8	60.8
	(62.9)	(9.71)	(23.7)	(18.3)	(88.9)	(7.10)	(43.7)	(37.7)	(43.1)	(37.2)
Observations	1978	1978	1901	1901	1901	1901	1899	1899	1883	1883
R-squared	0.047	0.050	0.057	0.080	0.058	0.080	0.056	0.070	0.059	0.070

* p<.10 ** p<.05 *** p<.01

Notes. Columns (1) and (2) show mean and median regressions that correspond to columns (3) in Table 7 and Table A9. Columns (3) and (4) drop observations with Cook's distance>4/N. Columns (5) and (6) drop observations with standardized residuals>2. Columns (7) and (8) drop observations with dfbeta>2/sqrt(N) for the cognitive skills coefficient. Columns (9) and (10) drop observations with dfbeta>2/sqrt(N) for the years of schooling coefficient.

Table A11. Reconciling mean and median regression results – second specification

	Paper regression		**High Cook's distance**		**High standardized residuals**	
	Median	*Mean*	*Median*	*Mean*	*Median*	*Mean*
	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)
	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling (a1)	0.063	0.69	-0.11	-0.42	-0.21	-0.29
	(0.84)	(1.06)	(0.80)	(0.82)	(0.76)	(0.84)
Out Village (a4)	32.9**	8.35	33.1**	23.4	31.0**	12.3
	(13.2)	(22.8)	(13.9)	(14.4)	(13.2)	(15.3)
Cognitive skills (a2)	0.26	2.79	-0.48	4.15	-0.048	4.27
	(3.59)	(5.00)	(3.36)	(3.25)	(3.39)	(3.44)
SEM skills (a3)	14.3***	13.0***	14.6***	12.8***	15.0***	12.2***
	(2.71)	(3.14)	(2.51)	(2.24)	(2.31)	(2.11)
Interaction YrsSchooling and Out Village (b1)	4.78***	9.36***	4.53***	6.16***	4.75***	7.35***
	(1.44)	(2.51)	(1.47)	(1.72)	(1.39)	(1.80)
Interaction Cog and Out Village (b2)	13.6**	-4.71	**15.6*** **	**5.51**	14.3***	-1.00
	(6.35)	(11.5)	(5.17)	(6.37)	(4.62)	(6.87)
Interaction SEM and Out Village (b3)	-7.37	-10.8	-5.56	-9.53*	-5.99	-5.91
	(5.17)	(7.68)	(4.71)	(5.11)	(4.52)	(5.15)
Constant	36.3	46.1***	25.7	29.4	37.0	43.5***
	(46.7)	(10.0)	(16.9)	(18.1)	(50.0)	(7.42)
Observations	1978	1978	1895	1895	1902	1902
R-squared	0.10	0.13	0.12	0.21	0.12	0.21

(continued)

Table A11. Reconciling mean and median regression results – second specification (continued)

	High dfbeta - out village		High dfbeta - school and out		High dfbeta - cog and out		High dfbeta - noncog and out	
	Median Monthly earnings (TC)	*Mean* Monthly earnings (TC)	*Median* Monthly earnings (TC)	*Mean* Monthly earnings (TC)	*Median* Monthly earnings (TC)	*Mean* Monthly earnings (TC)	*Median* Monthly earnings (TC)	*Mean* Monthly earnings (TC)
	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Years of schooling (a1)	-0.22	0.33	0.12	0.69	-0.050	-0.23	0.19	-0.0051
	(0.86)	(1.02)	(0.90)	(1.04)	(0.83)	(0.92)	(0.78)	(0.94)
Out Village (a4)	32.5**	8.01	28.5*	16.5	31.7*	8.27	34.8***	1.36
	(15.7)	(15.0)	(16.1)	(15.5)	(16.7)	(15.4)	(11.0)	(16.4)
Cognitive skills (a2)	0.47	2.11	-0.38	2.53	-0.23	3.60	-1.19	3.30
	(3.85)	(4.33)	(3.95)	(4.36)	(3.34)	(3.76)	(3.54)	(4.11)
SEM skills (a3)	14.2***	13.7***	14.4***	12.6***	14.7***	13.7***	14.5***	11.6***
	(2.77)	(2.89)	(2.83)	(3.02)	(2.73)	(2.86)	(2.51)	(2.44)
Interaction YrsSchooling and Out Village (b1)	4.86***	7.96***	4.87***	5.97***	4.71**	7.76***	4.47***	8.29***
	(1.73)	(1.90)	(1.79)	(1.57)	(1.87)	(1.86)	(1.12)	(1.87)
Interaction Cog and Out Village (b2)	12.6*	-0.21	13.2*	3.85	**15.6***	**4.37**	16.7***	0.19
	(6.87)	(6.90)	(7.22)	(7.52)	(8.21)	(6.48)	(3.92)	(7.82)
Interaction SEM and Out Village (b3)	-7.41*	-12.0*	-5.27	-8.80	-7.34	-13.2**	-7.08	-6.10
	(4.01)	(7.19)	(5.25)	(6.35)	(4.75)	(6.37)	(5.03)	(4.54)
Constant	38.5	50.2***	35.2	47.7***	37.1	50.1***	34.7	41.7***
	(37.9)	(10.0)	(56.2)	(9.46)	(41.0)	(8.61)	(60.7)	(8.25)
Observations	1894	1894	1892	1892	1896	1896	1888	1888
R-squared	0.11	0.13	0.11	0.11	0.11	0.13	0.11	0.14

* p<.10 ** p<.05 *** p<.01

Notes. Columns (1) and (2) show mean and median regressions that correspond to columns (4) in Table 7 and Table A9. Columns (3) and (4) drop observations with Cook's distance>4/N. Columns (5) and (6) drop observations with standardized residuals>2. Columns (7) and (8) drop observations with dfbeta>2/sqrt(N) for the cognitive skills coefficient. Columns (9) and (10) drop observations with dfbeta>2/sqrt(N) for the years of schooling coefficient. Columns (7) and (8) drop observations with dfbeta>2/sqrt(N) for the "out village" coefficient. Columns (9) and (10) drop observations with dfbeta>2/sqrt(N) for the "interaction yrsschooling and out village" coefficient. Columns (11) and (12) drop observations with dfbeta>2/sqrt(N) for the "interaction cog and out village coefficient." Columns (13) and (14) drop observations with dfbeta>2/sqrt(N) for the interaction NonCog and Out Village coefficient.

Table A12. Cognitive skills subjects and earnings in Pakistan

	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
	Cognitive skills: English	Cognitive skills; Mathematics	Cognitive skills: Urdu	Cognitive skills: English	Cognitive skills; Mathematics	Cognitive skills: Urdu
	Monthly earnings (TC)			Monthly earnings (TC)		
Years of schooling	2.79**	3.72***	2.44**	1.14***	1.63***	1.71***
	(1.10)	(0.97)	(1.22)	(0.37)	(0.39)	(0.44)
Cognitive Index	6.70	1.25	8.47*	4.38**	1.49	0.94
	(4.08)	(3.41)	(4.82)	(1.73)	(1.46)	(1.56)
SEM Skills Index	14.72***	14.37***	14.24***	0.78	0.85	0.90
	(3.27)	(3.28)	(3.27)	(0.90)	(0.91)	(0.91)
Constant	56.45***	49.78***	61.95***	-6.10	-10.83***	-11.75***
	(9.12)	(9.00)	(10.74)	(4.19)	(3.71)	(4.12)
Observations	1,978	1,977	1,977	1,925	1,925	1,925
Adjusted R-squared	0.05	0.05	0.05	0.07	0.06	0.06
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
District FE	No	No	No	No	No	No
N_Clusters	108	108	108	108	108	108

* p<.10 ** p<.05 *** p<.01

Test of coefficients equality – Men sample

	English=Mathematics	**English=Urdu**	**Mathematics=Urdu**
Schooling	1.52 (0.22)	0.26 (0.61)	3.30 (0.07)
Cognitive skills	1.86 (0.17)	0.24 (0.62)	3.41 (0.06)
SEM skills	1.67 (0.20)	3.00 (0.08)	1.03 (0.31)

Test of coefficients equality – Women sample

	English=Mathematics	**English=Urdu**	**Mathematics=Urdu**
Schooling	4.06 (0.04)	3.97 (0.05)	0.12 (0.73)
Cognitive skills	3.75 (0.05)	3.97 (0.05)	0.16 (0.69)
SEM skills	0.74 (0.39)	2.48 (0.12)	0.47 (0.49)

Notes. This table reports (mean) estimates of the relationships between schooling, skills, and earnings in Pakistan. The dependent variable "monthly earnings (TC)" is top-coded at 100,000 PKR per month (961.5 USD). The cognitive skills index is computed using Item Response Theory with a two-parameter logistic (2PL) model on the English/Mathematics/Urdu items only in columns (1) and (4)/(2) and (5)/(3) and (6), respectively. The socioemotional (SEM) skills index is computed using principal-component factor analysis on the Big-Five items, Grit items, BART, and GoNoGo scores. The sample for columns (1) to (3)/(4) to (6) are all men/women who are not currently enrolled and answered the direct version of the questionnaire, in-person. All the regressions include age and district fixed effects. Standard errors are in parentheses (clustered at the village level).

Table A13. Socioemotional skills factors and earnings for men in Pakistan

	Conscientiousness (1)				Openness to experience (2)				Agreeableness (3)			
	Monthly earnings				*Monthly earnings*				*Monthly earnings*			
	Mean		*Median*		*Mean*		*Median*		*Mean*		*Median*	
Years of schooling (a1)	3.14***	0.98	1.71***	0.15	3.27***	1.09	1.85***	0.78	3.39***	1.27	2.26***	1.36*
	(1.15)	(1.05)	(0.64)	(0.76)	(1.15)	(1.02)	(0.65)	(0.74)	(1.16)	(1.04)	(0.70)	(0.78)
Cognitive skills (a2)	6.37	2.79	7.90***	0.57	5.63	0.87	7.81***	-0.29	6.26	2.05	6.40**	-1.71
	(4.76)	(5.03)	(2.76)	(3.71)	(4.79)	(4.98)	(2.99)	(2.47)	(4.78)	(4.99)	(3.03)	(3.33)
SEM skills (a3)	11.8***	8.71***	11.3***	7.66***	6.69**	12.0***	5.36***	7.17***	4.56	7.85**	3.22	5.66**
	(3.07)	(3.10)	(2.07)	(2.58)	(3.23)	(3.37)	(1.75)	(2.51)	(3.17)	(3.04)	(2.45)	(2.75)
Out Village (a4)		7.08		22.0*		11.9		38.9**		17.1		41.3***
		(23.6)		(13.2)		(21.8)		(15.7)		(22.5)		(15.0)
Interaction YrsSchooling and Out Village (b1)		9.15***		5.57***		8.92***		4.30**		8.53***		4.07***
		(2.57)		(1.49)		(2.49)		(1.71)		(2.52)		(1.57)
Interaction Cog and Out Village (b2)		-5.20		10.4		-0.36		16.2***		-2.71		16.3***
		(11.7)		(6.55)		(11.4)		(6.29)		(11.4)		(5.92)
Interaction SEM and Out Village (b3)		-3.22		3.82		-22.4***		-8.95*		-20.5***		-5.37
		(7.77)		(5.22)		(7.81)		(4.98)		(6.84)		(5.76)
Constant	53.9***	41.9***	35.6	28.9	44.3***	35.8***	26.7	22.1	50.3***	47.2***	26.0	26.8
	(10.2)	(10.4)	(39.6)	(37.6)	(9.71)	(9.45)	(74.9)	(51.3)	(11.1)	(11.4)	(57.3)	(45.5)
Observations	1978	1978	1978	1978	1978	1978	1978	1978	1978	1978	1978	1978
R-squared	0.050	0.12	0.036	0.088	0.040	0.13	0.032	0.086	0.040	0.13	0.032	0.086
Median/Mean dependent	134.7	134.7	115.4	115.4	134.7	134.7	115.4	115.4	134.7	134.7	115.4	115.4
a1+b1=0		10.13***		5.73***		10.01***		5.08***		9.80***		5.43***
		(2.54)		(1.34)		(2.45)		(1.58)		(2.48)		(1.40)
a2+b2=0		-2.41		10.98**		0.51		15.94***		-0.66		14.57***
		(10.37)		(5.47)		(10.10)		(5.86)		(10.03)		(5.01)
a3+b3=0		5.49		11.48**		-10.46		-1.77		-12.61*		0.29
		(7.37)		(4.65)		(6.88)		(4.36)		(6.36)		(5.11)

(continued)

Table A13 (continued)

	Extraversion (4)				**Emotional stability** (5)				**Grit** (6)			
	Monthly earnings				*Monthly earnings*				*Monthly earnings*			
	Mean		*Median*		*Mean*		*Median*		*Mean*		*Median*	
Years of schooling (a1)	3.39***	1.20	2.21***	0.81	3.08***	0.90	1.89***	1.08	3.05**	0.95	1.65***	0.55
	(1.16)	(1.05)	(0.65)	(0.79)	(1.18)	(1.04)	(0.49)	(0.78)	(1.16)	(1.09)	(0.63)	(0.83)
Cognitive skills (a2)	6.35	2.18	6.97**	-0.73	5.84	2.43	5.68**	-1.10	6.60	2.65	7.24***	-0.078
	(4.79)	(5.02)	(2.71)	(3.83)	(4.75)	(5.00)	(2.33)	(3.63)	(4.72)	(5.00)	(2.79)	(3.74)
SEM skills (a3)	-0.32	4.55	0.48	4.60**	11.5***	9.98***	7.98***	5.26**	10.6***	6.90	14.1***	12.5***
	(3.12)	(2.93)	(2.11)	(2.28)	(4.06)	(3.60)	(1.84)	(2.52)	(3.87)	(4.22)	(2.00)	(2.66)
Out Village (a4)		8.93		37.0**		7.13		36.1**		6.65		31.9***
		(21.8)		(16.2)		(21.6)		(16.1)		(22.1)		(11.8)
Interaction YrsSchooling and Out Village (b1)		8.97***		4.48***		9.21***		4.28**		9.14***		4.86***
		(2.50)		(1.74)		(2.52)		(1.70)		(2.54)		(1.32)
Interaction Cog and Out Village (b2)		-3.70		16.0**		-5.66		14.2**		-4.66		13.3***
		(11.4)		(6.67)		(11.2)		(6.39)		(11.3)		(4.25)
Interaction SEM and Out Village (b3)		-8.54		-5.11		-1.93		-1.00		-1.07		-7.99
		(7.28)		(4.89)		(8.28)		(5.33)		(8.17)		(5.23)
Constant	41.8***	33.7***	22.7	20.6	43.9***	34.5***	22.2	19.2	45.6***	34.8***	31.9	24.8
	(9.76)	(9.51)	(67.6)	(36.5)	(9.56)	(9.40)	(69.2)	(62.0)	(9.57)	(9.64)	(72.8)	(45.7)
Observations	1978	1978	1978	1978	1978	1978	1978	1978	1978	1978	1978	1978
R-squared	0.040	0.12	0.031	0.085	0.050	0.13	0.034	0.086	0.050	0.12	0.041	0.089
Median/Mean dependent	134.7	134.7	115.4	115.4	134.7	134.7	115.4	115.4	134.7	134.7	115.4	115.4
a1+b1=0		10.17***		5.29***		10.10***		5.35***		10.09***		5.40***
		(2.47)		(1.58)		(2.52)		(1.55)		(2.49)		(1.09)
a2+b2=0		-1.52		15.31***		-3.24		13.06**		-2.01		13.27***
		(10.12)		(5.58)		(9.81)		(5.37)		(9.99)		(2.41)
a3+b3=0		-3.99		-0.51		8.04		4.25		5.83		4.53
		(6.95)		(4.33)		(8.17)		(4.72)		(6.69)		(4.52)

* p<.10 ** p<.05 *** p<.01

Notes. This table reports estimates of the relationships between schooling, skills, migration, and earnings for men in Pakistan. The dependent variable "monthly earnings" is the raw monthly earnings for median regressions; they are top-coded at 100,000 PKR per month (961.5 USD) for mean regressions. The cognitive index is the mean of the Urdu, Mathematics, and English scores computed using Item Response Theory with a two-parameter logistic (2PL) model. The items included in the model are those from the paper test and the tablet-based computer-adaptive test (leaving out items that less than 50 respondents answered and items that less than 5% or more than 95% of respondents solved correctly). The socioemotional (SEM) skills index is computed using principal-component factor analysis on the Grit items for the Grit column. The SEM index is computed using principal-component factor analysis on the openness to experience, agreeableness, extraversion, and emotional stability sub-scales of the Big Five for the respective columns. "Out of village" indicates whether the respondent lives outside the village where we initially surveyed their household (their natal village most of the time). The sample is all men who answered the direct version of the questionnaire in person. All the regressions include age and district fixed effects. Robust standard errors are in parentheses. The R-squared shown is the pseudo R-squared for median regressions and the adjusted R-squared for mean regressions.

Table A14. Socioemotional skills factors and earnings for Women in Pakistan

	Conscientiousness (1)			Openness to experience (2)			Agreeableness (3)		
	Monthly earnings			*Monthly earnings*			*Monthly earnings*		
	All	*Working women*		*All*	*Working women*		*All*	*Working women*	
	Mean	*Mean*	*Median*	*Mean*	*Mean*	*Median*	*Mean*	*Mean*	*Median*
Years of schooling (a1)	1.36***	8.10*	6.63**	1.32***	7.74*	4.92*	1.39***	8.17**	5.75**
	(0.43)	(4.14)	(3.06)	(0.43)	(4.28)	(2.84)	(0.43)	(4.09)	(2.83)
Cognitive skills (a2)	3.22*	2.17	-6.60	3.14*	1.53	-0.96	3.05*	1.97	-3.51
	(1.81)	(18.5)	(13.2)	(1.79)	(18.1)	(13.1)	(1.81)	(18.1)	(11.2)
Socioemotional skills (a3)	-0.024	-5.46	-0.94	1.30	8.71	1.84	-1.92**	1.81	1.93
	(0.78)	(9.96)	(8.65)	(0.79)	(8.57)	(7.07)	(0.83)	(8.67)	(7.92)
Constant	-8.00*	-20.6	-11.0	-8.35*	-19.5	2.52	-6.81	-22.7	7.62
	(4.23)	(47.6)	(40.1)	(4.50)	(49.2)	(49.6)	(4.12)	(49.7)	(48.6)
Observations	1927	111	111	1927	111	111	1927	111	111
R-squared	0.070	0.25	0.22	0.070	0.25	0.22	0.070	0.25	0.22
Median/Mean dependent	7.84	107.43	57.69	7.84	107.43	57.69	7.84	107.43	57.69

	Extraversion (4)			Emotional stability (5)			Grit (6)		
	Monthly earnings			*Monthly earnings*			*Monthly earnings*		
	All	*Working women*		*All*	*Working women*		*All*	*Working women*	
	Mean	*Mean*	*Median*	*Mean*	*Mean*	*Median*	*Mean*	*Mean*	*Median*
Years of schooling (a1)	1.36***	8.19**	6.76***	1.35***	7.55*	6.44***	1.36***	8.15*	6.20**
	(0.43)	(4.04)	(2.27)	(0.42)	(4.11)	(2.43)	(0.43)	(4.14)	(2.43)
Cognitive skills (a2)	3.26*	1.90	-4.63	3.24*	4.33	-3.40	2.98	1.75	-6.24
	(1.81)	(18.1)	(10.5)	(1.80)	(18.5)	(11.2)	(1.84)	(18.5)	(10.5)
Socioemotional skills (a3)	-0.38	0.49	-1.82	1.24	13.8	3.41	1.40*	2.52	-8.92
	(0.72)	(7.28)	(8.28)	(1.34)	(11.9)	(12.8)	(0.76)	(8.99)	(6.54)
Constant	-7.90*	-24.8	-9.61	-8.09*	-0.76	-8.19	-7.83*	-24.1	-23.5
	(4.15)	(46.0)	(48.0)	(4.72)	(46.0)	(51.3)	(4.10)	(47.3)	(49.2)
Observations	1927	111	111	1927	111	111	1925	111	111
R-squared	0.070	0.25	0.23	0.070	0.26	0.22	0.070	0.25	0.23
Median/Mean dependent	7.84	107.43	57.69	7.84	107.43	57.69	7.84	107.43	57.69

* p<.10 ** p<.05 *** p<.01

Notes. This table reports estimates of the relationships between schooling, skills, and earnings for women in Pakistan. The dependent variable "monthly earnings" is the raw monthly earnings for median regressions; they are top-coded at 100,000 PKR per month (961.5 USD) for mean regressions. The cognitive index is the mean of the Urdu, Mathematics, and

English scores computed using Item Response Theory with a two-parameter logistic (2PL) model. The items included in the model are those from the paper-based test and the tablet-based computer-adaptive test (excluding items that less than 50 respondents answered and those that less than 5% or more than 95% of respondents got it right). The socioemotional (SEM) skills index is computed using principal-component factor analysis on the Grit items for the Grit column. The SEM skills index is computed using principal-component factor analysis on the openness to experience, agreeableness, extraversion, and emotional stability sub-scales of the Big Five for the respective columns. The sample for the first column for each factor is all women who answered the direct version of the questionnaire, in person. The sample for the rest of the columns is all women who answered the direct version of the questionnaire, in person and are working. All the regressions include age and district fixed effects. Robust standard errors are in parentheses. The R-squared shown is the pseudo R-squared for median regressions and the adjusted R-squared for mean regressions.

Table A15. Relationship between schooling and skills formation in Cambodia

	(1) Cognitive skills	(2) Cognitive skills	(3) Socioemotional skills	(4) Socioemotional skills
Years of schooling	0.191***	0.183***	0.060***	0.051***
	(0.012)	(0.014)	(0.011)	(0.013)
Respondent Age	-0.097***	-0.102***	0.002	0.007
	(0.009)	(0.010)	(0.008)	(0.010)
Respondent is female	-0.280***	-0.242***	-0.106***	-0.091**
	(0.030)	(0.033)	(0.034)	(0.039)
Household head can read (2011)	0.016	0.015	0.336***	0.338***
	(0.108)	(0.127)	(0.105)	(0.122)
Household head can write (2011)	0.128	0.095	-0.269**	-0.285**
	(0.109)	(0.129)	(0.107)	(0.122)
HH SES in 2008	0.038	0.022	0.006	-0.027
	(0.024)	(0.027)	(0.021)	(0.026)
Constant	0.208	0.349	-1.739***	-0.297
	(0.199)	(0.222)	(0.183)	(0.219)
Observations	3,285	3,285	3,264	3,264
Adjusted R-squared	0.175	0.222	0.014	0.038
Sample	All	All	All	All
Village FE	No	Yes	No	Yes
Province FE	Yes	No	Yes	No
Number of clusters	423	423	423	423

* p<.10 ** p<.05 *** p<.01

Notes. The dependent variable in columns (1) and (2) is the Mathematics score, computed using Item Response Theory with a two-parameter logistic (2PL) model. The items included in the model are those from the tablet-based computer-adaptive test. The dependent variable in columns (3) and (4) is a socioemotional skills index computed using principal-component factor analysis on 15 Big Five items and 8 Grit items. The sample for this long-term follow-up is 3,294 respondents. Among them, some did not complete the computer-adaptive test, and some answered "Don't know" to the socioemotional skills questions. Therefore, the sample is 3,285 respondents for the cognitive index and 3,264 respondents for the socioemotional skills index. Regressions in even columns control for village fixed effects, while regressions in odd columns control for province fixed effects. Standard errors are in parentheses (clustered at the village level).

Table A16. Relationship between socioemotional skills and schooling in Cambodia

	(1) Conscientiousness	(2) Grit	(3) Openness to experience	(4) Extraversion	(5) Agreeableness	(6) Emotional stability
Years of schooling	0.044***	0.000	0.044***	0.027*	0.035**	0.017
	(0.014)	(0.014)	(0.013)	(0.014)	(0.015)	(0.014)
Respondent age	0.018*	0.012	-0.012	-0.004	0.009	-0.018*
	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)	(0.010)
Respondent is female	-0.151***	-0.094**	-0.067	0.067*	0.063	-0.025
	(0.040)	(0.044)	(0.043)	(0.039)	(0.041)	(0.043)
Household head can read (2011)	0.193	0.236**	0.197*	0.195*	0.264*	0.029
	(0.126)	(0.110)	(0.118)	(0.118)	(0.137)	(0.104)
Household head can write (2011)	-0.125	-0.255**	-0.088	-0.180	-0.264*	0.025
	(0.128)	(0.111)	(0.121)	(0.116)	(0.136)	(0.103)
HH SES in 2008	-0.027	0.018	-0.005	0.004	-0.041	0.003
	(0.026)	(0.028)	(0.028)	(0.029)	(0.029)	(0.034)
Constant	4.904***	7.022***	5.119***	3.897***	5.755***	4.994***
	(0.230)	(0.229)	(0.228)	(0.247)	(0.225)	(0.221)
Observations	3,278	3,274	3,279	3,278	3,279	3,281
Romano-Wolf p-values - Years of schooling	0.004	1.000	0.004	0.024	0.004	0.147
Adjusted R-squared	0.027	0.031	0.027	0.019	0.007	0.002
Sample	All	All	All	All	All	All
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	No	No	No	No	No	No
Test Yrs of school = Yrs of school (1) - Chi2 (p-value)	NA	6.28 (0.01)	0.00 (1.00)	1.03 (0.31)	0.31 (0.58)	1.97 (0.16)
Test Yrs of school = Yrs of school (2) - Chi2 (p-value)	6.28 (0.01)	NA	6.23 (0.01)	2.11 (0.15)	3.44 (0.06)	0.84 (0.36)
Test Yrs of school = Yrs of school (3) - Chi2 (p-value)	0.00 (1.00)	6.23 (0.01)	NA	1.11 (0.29)	0.27 (0.60)	2.31 (0.13)
Test Yrs of school = Yrs of school (4) - Chi2 (p-value)	1.03 (0.31)	2.11 (0.15)	1.11 (0.29)	NA	0.26 (0.61)	0.23 (0.63)
Test Yrs of school = Yrs of school (5) - Chi2 (p-value)	0.31 (0.58)	3.44 (0.06)	0.27 (0.60)	0.26 (0.61)	NA	0.79 (0.37)
Test Yrs of school = Yrs of school (6) - Chi2 (p-value)	1.97 (0.16)	0.84 (0.36)	2.31 (0.13)	0.23 (0.63)	0.79 (0.37)	NA

* p<.10 ** p<.05 *** p<.01

Notes. This table reports estimates of the relationships between socioemotional skills and schooling in Cambodia. All regressions control for village fixed effects. Standard errors are in parentheses (clustered at the village level).

Table A17. Relationship between migration, schooling and skills in Cambodia

	All		Men		Women	
	(1)	(2)	(3)	(4)	(5)	(6)
	Out of village	Out of village	Out of village	Out of village	Out of village	Out of village
Years of schooling	-0.002	-0.003	-0.015*	-0.014	0.005	0.004
	(0.005)	(0.005)	(0.009)	(0.010)	(0.008)	(0.008)
Cognitive skills		0.006		-0.002		0.005
		(0.007)		(0.013)		(0.009)
Socioemotional skills		0.002		-0.006		0.004
		(0.005)		(0.010)		(0.009)
Household head can read (2011)	-0.025	-0.026	-0.075	-0.072	0.033	0.033
	(0.032)	(0.032)	(0.052)	(0.053)	(0.053)	(0.053)
Household head can write (2011)	0.002	0.003	0.070	0.068	-0.056	-0.057
	(0.030)	(0.030)	(0.052)	(0.052)	(0.053)	(0.053)
HH SES in 2008	-0.008	-0.008	0.000	-0.000	-0.014	-0.014
	(0.009)	(0.009)	(0.016)	(0.016)	(0.012)	(0.012)
Respondent is female	0.001	0.002				
	(0.012)	(0.012)				
Constant	0.124***	0.138***	0.162**	0.136	0.199***	0.211***
	(0.035)	(0.040)	(0.073)	(0.091)	(0.053)	(0.055)
Observations	2,688	2,688	1,249	1,249	1,439	1,439
Adjusted R-squared	0.066	0.066	0.084	0.083	0.055	0.054
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	All	All	All
Number of clusters	378	378	309	309	324	324
Mean dependent	0.0763	0.0763	0.0817	0.0817	0.0716	0.0716

* p<.10 ** p<.05 *** p<.01

Notes. The dependent variable "out of village" is a dummy variable that indicates whether the respondent lives outside the village where we originally surveyed their household. The cognitive skills index is the Mathematics score, computed using Item Response Theory with a two-parameter logistic (2PL) model. The items included in the model are those from the tablet-based computer-adaptive test. The socioemotional skills index is computed using principal-component factor analysis on 15 Big Five items and 8 Grit items. The sample for this long-term follow-up is 3,294 respondents. We only have information on where the respondent currently lives for 2,706 respondents (1259 men and 1,447 women). Among them, some respondents did not complete the tablet-based computer-adaptive test or answered "Don't know" to the socioemotional skills questions, leading to a sample of 2,688 respondents for regressions that include these skills measures (1,249 men and 1,439 women). All regressions control for village and age fixed effects. Standard errors are in parentheses (clustered at the village level).

Table A18. Schooling, skills, migration, and earnings for men in Cambodia

	Median regressions				Mean regressions			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Monthly earnings	Monthly earnings	Monthly earnings	Monthly earnings	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)
Years of schooling (a1)	-3.74	-1.24	-1.05	-2.33	-13.0	-15.5	-13.8	-15.5
	(2.82)	(4.52)	(4.68)	(4.59)	(10.9)	(12.3)	(12.5)	(12.7)
Cognitive skills (a2)			-0.62	1.35			-22.5	-14.3
			(5.56)	(6.24)			(17.4)	(17.1)
SEM skills (a3)			4.03	2.03			34.3**	38.1**
			(4.35)	(4.69)			(16.8)	(17.7)
Out Village (a4)				-22.2				-50.1
				(63.4)				(452.3)
Interaction YrsSchooling and Out Village (b1)				1.79				16.1
				(11.0)				(67.6)
Interaction SEM and Out Village (b2)				1.36				-50.4
				(32.6)				(56.3)
Interaction Cog and Out Village (b3)				-17.4				-87.3
				(21.1)				(71.9)
Constant	44.4	72.2	31.9	38.7	100.7	115.6	187.1	219.1*
	(670756.7)	(658360.5)	(169809.0)	(167865.0)	(87.7)	(98.6)	(127.5)	(127.6)
Observations	1451	1173	1173	1173	1451	1173	1173	1173
R-squared	0	0	0	0	0.0090	-0.022	-0.017	-0.016
Median/Mean dependent	160.9	160.9	160.9	160.9	296.3	292.1	292.1	292.1
Village FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
a1+b1=0				-0.54				0.61
				(11.18)				(65.33)
a2+b2=0				-16.03				-101.61
				(19.24)				(71.30)
a3+b3=0				3.39				-12.31
				(32.22)				(52.84)

* p<.10 ** p<.05 *** p<.01

Notes. This table reports estimates of the relationships between schooling, skills, migration, and earnings for men in Cambodia. The dependent variable "monthly earnings" in columns (1) to (4) is the raw monthly earnings; the dependent variable "monthly earnings (TC)" in columns (5) to (8) is top-coded at 2,000 USD per month. The cognitive index is the Mathematics score, computed using Item Response Theory with a two-parameter logistic (2PL) model. The items included in the model are those from the tablet-based computer-adaptive test. The socioemotional (SEM) skills index is computed using principal-component factor analysis on 15 Big Five items and 8 Grit items. The "out of village" variable is a dummy variable that indicates whether the respondent lives outside the village where we originally surveyed their household. The sample for columns (1) and (5) are all men in the sample who are working and are not currently enrolled (1,451 men). For the rest of the columns, the sample is only those who have skills and location measures (1,173 men). Robust standard errors are shown in parentheses. All the regressions include age and village fixed effects. The R-squared shown is the pseudo R-squared for median regressions and the adjusted R-squared for mean regressions.

Table A19. Schooling, skills, migration, and earnings for women in Cambodia

	Median regressions				Mean regressions			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Monthly earnings	Monthly earnings	Monthly earnings	Monthly earnings	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)	Monthly earnings (TC)
Years of schooling (a1)	2.27	5.00***	3.65*	1.99	7.57	9.21	7.39	8.35
	(1.75)	(1.72)	(2.11)	(2.43)	(9.11)	(10.3)	(10.7)	(10.0)
Cognitive skills (a2)			2.23	2.43			6.26	8.92
			(3.34)	(3.34)			(13.3)	(13.6)
SEM skills (a3)			7.44***	7.31**			18.5	19.6
			(2.52)	(2.87)			(13.2)	(13.3)
Out Village (a4)				-32.7				26.9
				(78.6)				(371.6)
Interaction YrsSchooling and Out Village (b1)				4.02				-5.46
				(12.2)				(53.4)
Interaction SEM and Out Village (b2)				14.1				-19.8
				(14.8)				(28.2)
Interaction Cog and Out Village (b3)				13.7				-38.0
				(24.4)				(38.7)
Constant	163.7	141.1	167.1	215.1	-243.1***	-244.5***	-223.7***	-235.6***
	(602.3)	(576.4)	(472.3)	(1248.2)	(56.8)	(60.7)	(66.5)	(65.6)
Observations	1642	1354	1354	1354	1642	1354	1354	1354
R-squared	0	0	0	0	-0.0054	-0.051	-0.050	-0.053
Median/Mean dependent	123.1	123.8	123.8	123.8	208.7	205.0	205.0	205.0
Village FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
a1+b1=0				6.01				2.90
				(12.10)				(52.82)
a2+b2=0				16.16				-29.09
				(25.02)				(38.10)
a3+b3=0				21.43				-0.17
				(14.49)				(29.03)

* p<.10 ** p<.05 *** p<.01

Notes. This table reports estimates of the relationships between schooling, skills, migration, and earnings for women in Cambodia. The dependent variable “monthly earnings” in columns (1) to (4) is the raw monthly earnings; the dependent variable “monthly earnings (TC)” in columns (5) to (8) is top-coded at 2,000 USD per month. The cognitive index is the Mathematics score, computed using Item Response Theory with a two-parameter logistic (2PL) model. The items included in the model are those from the tablet-based computer-adaptive test. The socioemotional (SEM) skills index is computed using principal-component factor analysis on 15 Big Five items and 8 Grit items. The “out of village” variable is a dummy variable that indicates whether the respondent lives outside the village where we originally surveyed their household. The sample for columns (1) and (5) is all women in the sample who are working and are not currently enrolled (1,642 women). For the rest of the columns, the sample is only those who have skills and location measures (1,354 women).

Robust standard errors are shown in parentheses. All the regressions include age and village fixed effects. The R-squared shown is the pseudo R-squared for median regressions and the adjusted R-squared for mean regressions.