Intergenerational Impacts of Secondary Education: Experimental Evidence from Ghana*

Esther Duflo Pascaline Dupas Elizabeth Spelke Mark Walsh

July 15, 2024

Abstract

We provide experimental evidence on the intergenerational impacts of secondary education subsidies in a low-income context, leveraging a randomized controlled trial and 15-year longitudinal follow-up. For young women, receiving a scholarship for secondary school delays childbearing and marriage, and reduces unwanted pregnancies. Female scholarship recipients are more likely to marry a partner with tertiary education and their children have better early childhood development outcomes. In particular, we document a 45% reduction in under-three mortality as well as cognitive development gains of 0.25 standard deviations of test scores once children are of school age. The primary mechanism seems to be that more-educated caregivers have the knowledge and skills to safeguard their children's health and stimulate their cognitive development. In contrast, we find no evidence of a positive impact for the children of male scholarship recipients, who tend to marry less educated partners. Together, these results suggest a key role for maternal education in child outcomes. We also estimate the cost-benefit ratio for secondary school scholarships and find that the impact on child survival alone is sufficient to make them a highly cost-effective investment.

JEL classification: I26, J12, J13, O15
Keywords: child mortality, child development, scholarships, maternal education

*This study is registered in The American Economic Association's registry for randomized controlled trials under RCT ID AEARCTR-0000015. The study protocol was approved by the Ghana Health Service Ethics Review Committee, and the IRBs of MIT, Stanford, and IPA. We thank the Ghana Education Service and IPA Ghana for their collaboration. We are grateful to Stephanie Kabukwor Adjovu, Kwame Akakpo, Isaac Arthur, Flora Bergounhon, Joseph Coffey, Gabriella Fleischman, Erin Grela, Leon Karreman and Antoine de Mascarel for outstanding research assistance. The funding for this study was provided by the British Academy, the JPAL Post-Primary Education Initiative, and USAID-DIV. We thank them, without implicating them, for making this study possible.
Duflo: MIT economics department and Collège de France (eduflo@mit.edu); Dupas: Princeton University economics department (pdupas@princeton.edu); Spelke: Harvard department of psychology (spelke@wjh.harvard.edu); Walsh: Stanford economics department (mwalsh24@stanford.edu).

1 Introduction

Following the widespread adoption of free primary education in low-income countries and the subsequent surges in primary school enrollment rates, policymakers' attention has shifted to secondary school. The United Nations *Sustainable Development Goals* call for "... free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes" (target 4.1). However, the extent to which secondary education should be publicly subsidized is not settled in the academic literature, and is still a very active policy debate in low-income countries. Currently, less than one-third of countries in sub-Saharan Africa offer tuition-free secondary school education. This paper provides experimental evidence on one important aspect of this debate: the extent to which free secondary education has intergenerational impacts. For decades, this claim has been at the heart of the push for girls' education. But despite its prominence in public discourse, rigorous evidence to back it has been scarce.

In Ghana, the setting of this study, debates about whether secondary education should be free have been central to policy discussions over the past 15 years. In 2016, the National Patriotic Party (NPP) won the presidential elections on a promise to make Senior High School (SHS) free for all qualified students and implemented a policy that covered tuition and fees for all Ghanaian students admitted to SHS from the 2017/2018 school year onward. The opposition critiqued the policy as over-committing resources to the education sector and diluting the quality of secondary education.[1] While the free SHS program is popular among Ghanaians, even NPP politicians have raised concerns over the government's ability to fund the program absent increases in tax revenue.[2]

At the heart of the debate is the fact that secondary education is expensive, and free secondary school implies a transfer to households who are sufficiently well off to pay to send their children to secondary school. Offsetting these costs are any benefits of secondary education for all those unable to afford it, as well as the possible externalities to society of a more educated population.

In Duflo et al. (2024), we investigate the labor market impacts of secondary education, and find small impacts on earnings, largely accounted for by hiring in the public sector. In this paper, we investigate one possible source of externality, namely the fact that more educated individuals, especially women, may choose and be able to have fewer children, delay childbearing, and invest

[1] https://www.ghanaweb.com/GhanaHomePage/NewsArchive/Free-SHS-to-go-Mahama-threatens-689275

[2] E.g., "In recent debates over a controversial E-levy (a tax on mobile money transactions), NPP MPs claimed that the free SHS program would have to be discontinued if the E-levy was not enacted." https://www.ghanaweb.com/GhanaHomePage/NewsArchive/Review-Free-SHS-Kwame-Sefa-Kayi-urges-government-1478996

more in the human capital of the children they do have (Becker, 1991). This implies that the benefit of providing free education to one cohort of adolescents would benefit future generations as well.

A large literature finds a correlation between education, lower and delayed fertility, and better outcomes for children (Thomas et al., 1991; Mohanty et al., 2016; Wietzke, 2020; Wodon et al., 2018; Hahn et al., 2018). Moreover, there is a demonstrated causal link between parental inputs in childhood, and cognitive scores and performance in school (Walker et al., 2007; Gertler et al., 2014; Attanasio et al., 2022). It is at least plausible that more educated parents provide more of these inputs (Attanasio et al., 2020).

However, establishing a causal link between education and future family outcomes is difficult: adolescents who receive more education may be different in various ways, which may in turn explain why their own children would be more educated. Countries that invest in the education of one cohort might continue to invest in future cohorts, which means that educational reforms cannot easily be used as natural experiments. To fill this gap, we provide what is, to the best of our knowledge, the first experimental evidence on the impact of secondary education on the timing, quantity, and quality of children, leveraging a randomized controlled trial and a very long longitudinal follow-up.

The trial began in 2008, several years before free secondary education was enacted in Ghana. With philanthropic funding and support from the Ghana Education Service, the NGO IPA awarded four-year secondary school scholarships to 682 adolescents, randomly selected among a study sample of 2,064 rural youth who had gained admission to a public high school but did not immediately enroll because they were not able to pay the fee. In Duflo et al. (2024), we show that adolescents that received a scholarship were 28 percentage points more likely to complete secondary school, compared to those who did not get a scholarship (with results similar for men and women), and received on average 1.33 more years of education.

Since 2013, i.e., right after (potential) graduation from secondary school, we have been regularly following up with the sample to collect data on their occupation, their earnings, and their family formation. In 2017, we began collecting data on the cognitive development of their children at specific milestone ages. We used locally-appropriate tests developed by the Harvard Laboratory for Developmental Studies, based on the best available evidence and practice on how to measure cognitive development in young children. They were also designed to be implemented by regular surveyors, as opposed to trained psychologists (unlike the standard psychometric assessments like the Bailey or MacArthur tests). In the tests, the child plays interactive games that target cognitive abilities that emerge in infancy and remain important through adolescence. To measure parental care-giving behaviors, we complement self-reports with a day-long recording of the child's auditory environment using the Language ENvironment Analysis (LENA^{TM}) system.

Our first set of results is that, for females, receiving a scholarship impacted when they started having children, as well as when and who they married. At our first follow-up in 2013, female scholarships recipients were 7 percentage points (14%) less likely to have had a pregnancy, and had 18% fewer children. This effect on fertility was driven by a 7 percentage point (17%) decrease in unwanted pregnancies. By the end of the follow-up period in 2023, they were as likely to have had at least a child and their number of children had mostly caught up; but the gap reduced slowly over the years (the average female non-recipient has had 1.77 children by 2022). They were also still less likely to be married or cohabiting with a partner. Their current or most recent partner, typically the father of their children, was more educated—in particular, he was more likely to have had tertiary education.

The second result is that children of female scholarship recipients were more likely to survive. Among children born to our female control group respondents, 3.5% died before the age of one and 4.0% before the age of three. These mortality rates fall to 1.7% (p-value=0.028) and 2.2% (p-value=0.065), respectively, among children of female scholarship recipients. This represents a halving of under-one and under-three mortality in our sample.

The third set of results concerns child cognitive development. To avoid bias stemming from the fact that women who received a scholarship started having children later in life, which implies that their children tend to be younger, we collect data on children at specific age milestones: 18 months, two and half years, three and half years, five years, and seven years. We do not find a significant difference in the cognitive scores of children at the lower age ranges, but an advantage emerges over time for children of female scholarship recipients. By five years of age, a child's aggregate score is 0.238 standard deviation higher (p-value=0.005), and by seven it is 0.252 standard deviation higher (p-value=0.035) if their mother received a scholarship. Those are large impacts, found both for average test scores and for most of the cognitive domains tested.

The fourth set of results is that we find none of these impacts for male education subsidies: young men who received a scholarship did not start having children later than those who did not, and their partners were significantly *less* educated than the partners of young men in the comparison group. We find neither mortality impacts nor any positive impact on cognitive scores for children of male scholarship recipients. For child mortality, we are not able to reject equality between the effect for female and male scholarship recipients. However, for cognitive development at 5 and 7 years, we can reject equality, and the point estimates for the treatment effects are in fact negative for children of male recipients.

An important mediator of the impact of the scholarship program on children outcomes is the education of the child's primary caregiver, who is in most instance the child's mother: in the case of female scholarship recipients, the impact is strongly positive (since the scholarship recipient is generally the caregiver); in the case of male scholarship recipients, there is no positive impact since the partners of male scholarship recipients are not more educated than partners of male respondents in the control group. The education of the father is also a potential mediator: partners of female scholarship recipients are more likely to have tertiary education, while male scholarship recipients are not themselves more likely to have attended tertiary school (though they are more likely to have a secondary education).

In turn, parental education (as well as any other direct impact of the scholarship, for example, lower expenses while in school, or greater incentives to stay in school and delay childbearing) could impact child well-being in a number of ways. In our previous work, we do not find large or significant impacts of receiving the scholarship on earnings in the 12 years that follow (Duflo et al., 2024).[3] The main channel for the intergenerational impacts is thus unlikely to be material well-being. We also don't find any difference in formal schooling inputs (time spent in school, age at which they started school) or even educational aspirations, which were very high across the board (81% of the mothers in our sample hope their child will go to university, although only 2% of the mothers themselves went to university).

What seems to be different are inputs that are not costly, but require perhaps more awareness or skills, namely preventive care and time spent interacting with children. Children of mothers who received the scholarship have received more preventive care, and mothers who received a scholarship reported more often playing with their children, doing simple mathematics with them, and singing them songs. These self-reports are confirmed by an objective measure of the interaction between young children and their caregivers. At 18 months, children of mothers who received the scholarship produce more vocalizations and have more conversational turns per minute with an adult in the day-long LENA recording. Since programs training parents to spend time playing and interacting with their children have been shown to impact cognitive development and long-run outcomes (Gertler et al., 2014; Walker et al., 2011; Attanasio et al., 2022), the greater stimulation

[3] A positive effect on the earnings of female scholarship recipients does emerge after 12 years (Duflo et al., 2024), driven by relatively few women who obtain a job in the government sector.

that female scholarship recipients appear to engage in with their young children could explain the children's greater performance on cognitive tests a few years later. Conversely, we find negative effects on the LENA measurements of having a father that has received a scholarship, consistent with the negative impact on cognitive scores that emerges in later years, and possibly owing to the fact that they are less likely to live with their father.

Overall, these results strongly support the idea that providing free access to secondary education for young women would ensure that not only they, but also their children, would be more educated and live healthier lives. A cost-benefit analysis suggests that investing in means-tested secondary school scholarships for girls would be a highly socially valuable investment in contexts similar to Ghana's. We estimate that the internal rate of return (IRR) from investing in a females-only means-tested secondary scholarship program is between 27% and 76%. Even if the scholarship cannot be targeted to females only, it would still be a socially-efficient investment, with an IRR between 20% and 51%.

This paper contributes to the literature on the impacts of maternal education on fertility. A quasi-experimental study in Nigeria (Osili and Long, 2008) and an experiment in Kenya (Duflo et al., 2015) find that primary education reduces or, at least, delays fertility. An experiment in Malawi (Baird et al., 2019) finds that conditional cash transfers increase secondary education and reduce fertility among adolescent females in Malawi. Similarly, Ozier (2018) compares the outcomes of Kenyan students just above and below the cutoff score for attending secondary school and estimates lower rates of teen pregnancy among female students above the cutoff (and therefore with a higher likelihood of going to secondary school). Consistent with the findings of Baird et al. (2019) and Ozier (2018), we find a reduction in teen pregnancy due to an increase in secondary education for female students. However, we are able to track fertility over a longer time period, which enables us to learn that the total fertility of women who receive a scholarship eventually catches-up to the total fertility of those who did not, but their children are born later, and are more likely to be wanted. A limitation is that our study only estimates the effect for individuals who qualified for secondary school but could not attend (i.e., well-prepared and low-income).

Existing evidence on the intergenerational impacts of education generally relies on variation induced by policy reforms, such as expansion of education access or compulsory schooling laws. Studies exploiting the expansion of primary education in low-income countries find positive intergenerational effects on educational attainment in Indonesia (Akresh et al., 2023) and economic outcomes in Benin (Wantchekon et al., 2013). The expansion of tertiary education also appears to have positive intergenerational effects. Suhonen and Karhunen (2019) finds an impact of parental tertiary education on a child's educational attainment in Norway, while Currie and Moretti (2003) estimates a positive impact of maternal tertiary education on child health in the United States. In contrast, studies exploiting the introduction of compulsory secondary schooling laws in mid-20th century Europe find small and insignificant effects of parental education on their children's education (Black et al., 2005; Chevalier, 2004). Baird et al. (2019)'s conditional cash transfer experiment in Malawi provides some experimental evidence on the intergenerational impact of secondary education. While they do not measure child cognitive development, they do estimate noisy but positive effects of the conditional cash transfer on height-for-age among children of beneficiaries born within 9 months of the program ending (Baird et al., 2019).[4]

Our study diversifies the evidence base on the intergenerational impacts of education, and our identification strategy and data improves upon the existing evidence. To our knowledge, our study is the first to estimate the intergenerational impacts of secondary education on child mortality and cognitive development and in a low- or middle-income country. In addition, by randomizing secondary school scholarships to individuals, we avoid the potential identification concerns plaguing past studies. Policy reforms that increase educational attainment for a large cohort (such as expanding access or increasing compulsory schooling years) will increase demand for teachers, affect peer quality, and alter the marriage market. These changes could bias the quasi-experimental estimates of education's intergenerational impacts.

Our rich dataset provides not only uniquely rigorous estimates of the intergenerational impact of secondary education on cognitive development, it also shed lights on the mechanisms driving this relationship. The evidence from LENA recordings and parental reports point towards changes in parenting practices, such as playing and conversing with the child, as the likely mechanism. Our finding that cognitive gains appear only once children are of school-age is consistent with recent evidence of complementarity between school-based and home-based investments (Duhon et al., 2024).

Finally, by combining administrative data on secondary school fees with self-reported school expenses and wages during secondary school, we accurately estimate the social cost of secondary school scholarships, a crucial component for cost-effectiveness and cost-benefit analyses.

[4] Baird et al. (2019)'s estimate are noisier than ours since they only have 88 children in the relevant sample (i.e., children born within 9 months of the program ending to mothers who were not attending school at baseline). They do not find effects on children born more than 9 months after the program ended.

2 Setting, Experimental Design and Data

In 2008, Duflo et al. (2024) sampled 2,064 students who had not enrolled in senior high school (SHS) because they could not afford to pay the fees, and initiated a randomized controlled trial in which 682 of these students were selected to receive a scholarship. Below we provide a summary of the important features of the experiment.

2.1 Sampling, randomization and first-generation data

Secondary school admission in Ghana is conditioned on an exam taken at the end of junior high school (JHS). Based on the exam results and their wishes, students who qualify are assigned to a school by a deferred acceptance algorithm.

The scholarship study sampled students who had been offered a spot to start SHS in Fall 2008 but had not yet enrolled (usually due to financial constraints) by the end of the fall quarter (details on the sample construction are provided in Duflo et al. (2024)). The research team administered a baseline survey to the students themselves as well as to one of their guardians, most commonly the mother. After the survey, each student received a basic mobile phone with a SIM card and was assigned a phone number.

The sampled students participated in a lottery where one third of students were randomly assigned to the "treatment group" (offered a scholarship) and two thirds to the "comparison group" (no scholarship), after stratifying by district, senior high school, junior high school, gender and year of junior high school exit exam.[5] The scholarship covered full tuition and fees for a "day" (i.e., non-boarding) student for four years, paid directly to the school. Students who received the scholarship were only responsible for the cost of school materials, transportation to school, and school meals.

Duflo et al. (2024) show that demographic characteristics of study participants were balanced between treatment and control groups. The scholarship-lottery participants were on average 17 years old at the onset of the study and just over 31 at our last follow-up in Spring 2023. Students were from poor families in rural areas. At baseline, over 40% of the students lived in households with no male head and 48% of household heads had only primary education or less, compared to 24% and 35%, respectively, in Ghana as a whole (Ghana Statistical Service, 2013).

From 2009-2012, the scholarship study team called the scholarship-lottery participants once a

[5] About 30% of the sample is composed of women who had been admitted in SHS for Fall 2007 but had not enrolled yet by Fall 2008. This group was included to ensure gender balance in the final sample.

year to update their contact information and basic outcomes (education status, fertility, cohabitation). In 2013, a detailed in-person follow-up survey measured schooling, occupation, cognitive skills, labor market expectations, health and fertility, among other topics. The cognitive test was an oral test measuring competencies both in reading and in mathematics, and the ability of the respondent to apply this knowledge to practical situations. Study participants received a phone upgrade at the end of the 2013 survey. In 2015, 2016, 2017, 2019, 2020, 2022, and 2023, 30-minute phone follow-up surveys were conducted to update participants' contact information and outcomes such as tertiary education, fertility, child survival, cohabitation, and labor market activities.

2.2 Impact of the scholarship on education, cognitive skills, and labor market outcomes

Duflo et al. (2024) reports the impact of the scholarship on recipients' education, cognitive skills, and labor market outcomes. We briefly summarize those results here. Winning a scholarship increased the SHS completion rate (the fraction of the entire group—including those that do not enroll—who graduate from SHS) from 39.8% to 67.2% among women (a 69% increase) and from 49.7% to 77.9% among men (a 57% increase) (Table A1). The effect of scholarships on SHS completion is large and statistically significant at the 1% level at all quartiles of the initial test score distribution. Overall, as of 2023, the scholarship had led to an average increase of 1.33 years total years of education.

While this increase is mainly due to more years of secondary education, Duflo et al. (2024) also document significant impacts of the secondary school scholarship on access to tertiary education, but for women only. As of 2019, 12% of women in the comparison group had ever enrolled in tertiary education, and 7.8% had graduated. Treatment increased enrollment rates by 7.4 percentage points and graduation by 4 percentage points. By 2023, the treatment effect on tertiary completion had increased to 10.8 percentage points for women. While average tertiary enrollment was slightly higher among men overall, there was no tertiary education impact of the scholarships for them.

In 2013, scholarship recipients scored 0.16 standard deviations higher on our cognitive tests, with gains found in both math and reading. These gains were experienced across the distribution of test scores, and were higher for females (0.194) than for males (0.113), although the difference is not statistically significant.

In contrast to the clear gains in educational achievements and cognitive skills, the labor market impacts are very mixed and delayed. By 2019, on average, no significant impacts on earnings were observed for either males or females (although the earnings data is quite imprecise). For female scholarship recipients, there was a significantly higher likelihood of having a public sector job, though this concerns a very small share of the sample (10.4% of scholarship recipients vs. 6.3% of the control group). It is only from 2020 onward that labor market gains emerge for women, with earnings 24% higher in 2020 and 30% higher by 2023. There is no discernible labor market returns for males up to the last survey round.

2.3 Child cognitive development test instruments and caregiver surveys

By 2016, many of the scholarship-lottery participants had children of their own, making it possible to assess whether the scholarship affected the cognitive development of recipients' children.

A first task was to develop cognitive tests for a range of children's ages. Most existing batteries of tests to measure early childhood cognitive development were developed and piloted in high-income economies, and were therefore unlikely to be appropriate for Ghanaian children in mostly rural settings.[6] These tests are also expensive, because they need to be administered in controlled conditions by a psychologist. An important contribution of our study is the development of a battery of cognitive tests that can be administered to children (a) by trained surveyors with no psychology degrees, (b) at children's homes, and (c) in low-income contexts.

The psychology Laboratory of Development Studies at Harvard developed these tests, based on research in cognitive science conducted in multiple cultures and with children at diverse economic levels. The tests consist of interactive "games" targeting cognitive abilities, such as language, attention, working memory, executive function, numerical and spatial reasoning, and social cognitive skills including reasoning through beliefs, perception, and emotions. The tests are meant to be engaging and use rules that are easy for children to understand and easy for surveyors to administer (on a laptop computer for children over age five, and using simple concrete materials for the younger children such as pictures, small objects, and cups). Combined, the games lasted 15 to 20 minutes for the youngest age group and 40-50 minutes for the oldest age group (seven year olds). Appendix D provides details on the games for each age group. The tests were developed at Harvard and piloted and validated in Ghana (Coffey and Spelke, 2023).

[6]The Oxford Neurodevelopment Assessment (Ox-NDA), an infant development test adapted for use in low- and middle-income (LMIC) settings, was not available in 2017 (when we began assessments). This test also would not have been appropriate for the older age groups (Fernandes, 2021).

Due to the effect of the scholarship on the timing of fertility (section 4), it would have been inappropriate to survey all the children in one survey round, because the scholarship recipients' children would have been systematically different. First, they would have been younger on average and it is difficult to compare health and cognitive outcomes across ages.[7] Second, even after controlling for age at a given date, systematic differences may persist because scholarship recipients who started childbearing early may be negatively selected.

Instead, we set up an infrastructure that allowed us to administer the tests, in person, to children of scholarship-lottery participants in specific age windows: 14-22 months old (we refer to these as the "1.5 years" group), 39-45 months old ("3.5 years") and 60-69 months old ("5 years"). This approach allows us to measure outcomes for children that would be missed with one survey wave and to compare the outcomes of treatment and control children around the same age.

Beginning in June 2017, we administered the child cognitive tests and caregiver surveys whenever children in the sample entered the targeted age windows. To be eligible, the child had to be a biological child of an initial scholarship-lottery participant. A surveyor contacted the primary caregiver of the eligible child to arrange an interview and the child's cognitive test.[8] The caregiver interview covered respondent demographics, respondent education, respondent health, indicators of household socioeconomic status, caregiver beliefs, child health, child health care, child education, cognitive stimulation of the child by household members, child time use, and, for children in the age range 14-22 months, infant language development. After we received approval from the Ghana Health Service in late 2017, we also began measuring the height and weight of children who were 24 months old or more.[9]

We began administering a test for 84-96 month olds ("7 years") in May 2019, and added a 30-36 month old test ("2.5 years") in July 2021. Starting in January 2018, we permitted the field team to survey children slightly above the maximum age for an age window if, due to time constraints among the field team, the child had not yet been surveyed for that age window.[10]

In total, we conducted 15 child measurement rounds from 2017-2022. Before each round, a

[7]For example, there is an age gradient in height-for-age z-scores that makes it difficult to compare, e.g., a 6 month-old to an 18 month-old (Aiyar and Cummins, 2021).

[8]The primary caregiver was defined as the person 'making the day-to-day decisions about the child's life.'

[9]We did not measure the height of children under 24 months old because our measurement tool (a stadiometer) required the child to stand up straight without assistance from an adult.

[10]Starting in January 2018, the surveyors were permitted to survey children up to 25 months old using the 14-22 month old instrument, children up to 55 months old using the 39-45 month old instrument and children up to 83 month olds using the 60-69 month old instrument. The surveyors were permitted to survey children up to 99 month olds using the 84-96 month old instrument and up to 39 month olds using the 30-36 month old instrument. We control for child age in months in the analysis.

member of the Laboratory of Development Studies met with the enumerators to review videotapes of selected field sessions and discuss ways to improve measurement quality. Then, the enumerators gathered updated location information for the eligible children through a short phone survey with their primary caregivers. Using this location information, local research staff assigned sets of respondents to enumerator teams based on geographic proximity. Finally, enumerator teams tracked and conducted measurements with the respondents over a 3-4 month period. We completed two child measurement rounds in 2017, three in 2018, three in 2019, and one in 2020, before we had to pause fieldwork from March-October 2020 due to the COVID 19 pandemic. We resumed field work in October 2020, completing one round. In 2021, we hired additional surveyors to make up for the missed workdays in 2020. We completed three rounds with this larger team in 2021, three in 2022, and two in 2023 (Figure A.1 shows the number of surveys by year). Across all measurement rounds, we administered 3,853 tests to 1,920 unique children.

2.4 LENA recordings

To complement caregiver-reported information about the day-to-day environment around young children in the study, for the "1.5 years" age group, we gathered day-long recordings of the auditory environment using a recording device called LENA (Language ENvironment Analysis) starting in February 2020. If the caregiver consented to the LENA procedures (80% of caregivers did), in the day that followed the cognitive tests measurements and caregiver survey, the child would wear a specially-designed shirt with an attached recording device for at least 8 hours. The LENA device uses speech recognition software to process the sounds around the child into count-based metrics such as adult word count, adult-child conversational turns, and child vocalizations (see Appendix E for more details). Because LENA devices had never been used in Ghana, an environment where it is common for toddlers to spend hours on their mother's back, we first validated the device's accuracy by asking a few individuals (from outside the study sample) to record their activity for a few hours while their child was wearing the device. While a number of pilot studies in the United States have used the LENA device to estimate impacts of parenting interventions (see for example Leung et al. (2020), or used LENA to provide feedback to parents (Suskind et al., 2013), we are not aware of any prior RCT having used the LENA device to measure outcomes of interest in a low-income context. Based on the measurement success of this study, a subset of us used the LENA device in another study of determinants of early childhood development in Ghana (Dupas et al., 2024).

3 Empirical specifications

The analyses follow a pre-analysis plan filed on the AEA registry for social experiments.[11] To evaluate the impact of the scholarship, we run intent-to-treat regressions at the scholarship-lottery participant (indexed by j) or the child (indexed by i) level. Since gender differences in the impact of scholarships were a core question of interest at the onset, the randomization of scholarships was stratified by gender, and we study effects separately by gender of the scholarship-eligible individual. As specified in our pre-analysis plan, we adjust for multiple hypothesis testing among our pre-specified primary outcomes, namely, child survival and child cognitive development.

3.1 Fertility and family formation

To study first-generation impacts of the scholarship on fertility and family formation, we run regressions at the scholarship-lottery participant level, straightforwardly regressing outcomes of interest on an indicator for treatment status, and controlling for region fixed effects and JHS exit exam score as in Duflo et al. (2024).

3.2 Child mortality

To study impacts on child mortality, our sample consists of children of scholarship-lottery participants. Table A4 shows that, mechanically given the impact of the scholarship on fertility, children of female scholarship recipients are younger than children of females in the control group (4.71 months younger; p=.067). We therefore focus on survival to age 1 (or 3) rather than being alive at the time of follow-up. This means we limit the sample to children who, based on their date of birth, had, or would have, reached age 1 (or 3) by the time their scholarship-lottery participant parent was last surveyed.

We run the following regression:

$$Y_{ij} = \alpha_{ij} + \beta_1 T_j + \beta_0 X_{ij} + \epsilon_i \tag{1}$$

Where Y_{ij} is the outcome (*child survived to age 1 (or 3)*) for child i of scholarship-lottery

[11]The scholarship study started before the AEA RCT registry existed; it was registered immediately upon the creation of the registry in 2013. At the time, we had not anticipated being able to follow-up with the children of the initial study participants. We registered a pre-analysis plan for the intergenerational impact study in February 2022, after the Spelke lab had investigated construct validity blind to treatment status: https://www.socialscienceregistry.org/trials/15.

participant j; T_j is an indicator that the scholarship-lottery participant (the child's parent) was randomly selected to receive a scholarship; and X_{ij} is a set of control variables. Because some scholarship-lottery participants had more than one child, we cluster the standard errors at the scholarship-lottery participant level, i.e., at the biological mother or father level depending on the scholarship-lottery participant's gender.

Besides the scholarship-lottery participant level controls mentioned above (region of birth and junior high school finishing exam scores), we control for the child's birth order and birth year.[12] If we did not control for birth order, our results could be driven by differential composition effects, since children of scholarship-recipients are more likely to be first-born children (Table A4), and survival rates may depend on birth order. If we did not control for year of birth, our results could be driven by general time trends in our outcomes of interest. For example, since infant mortality is falling over time in Ghana, secondary education scholarships would lower infant mortality for the treatment group even if secondary education only delayed fertility and did not affect survival rate conditional on birth cohort. While this (positive) "cohort" effect of education subsidies is important and valuable in itself, we choose to control for it in order to focus our estimation on the effects that would be at play even in an environment with no trend in mortality. We control for the time trend with either a linear control in year of birth or year of birth fixed effects (the results are unchanged).

3.3 Child cognitive development

Our sample for child cognitive development outcomes consists of the children of scholarship-lottery participants ever eligible for at least one of our cognitive tests, i.e., between 14 months and seven years old *at some point* between Fall 2017 and Spring 2023.

To estimate impacts on cognitive development, we use Equation 1, but alter the outcome variable and the set of controls. The outcome variable in these regressions is the child's score on the age-appropriate cognitive test. Since we administer different tests to 1.5 year olds, 2.5 year olds, 3 year olds, 5 year olds, and 7 year olds, we separately estimate effects for each of these age groups. To control for updates in enumerators' training across measurement rounds, we include round fixed effects. Since the children are in narrow age groups and rounds only last 2-4 months, round fixed effects also effectively control for year of birth. Additionally, we include a linear control for age in

[12]We cannot control for child gender because we are missing gender for 40% of the deceased child sample. This is due to our survey design. To minimize the psychological cost to respondents being asked to recall children who were born alive but passed away, we only asked for the child's birth date or the age when the child passed away and skipped all other questions.

months at the time of measurement, to increase estimate precision.[13] Because the children of female scholarship recipients in the cognitive games sample are more likely to be first-borns (Table A4), we control for the child's birth order. We also control for child gender, as well as the same scholarship-lottery participant level variables as in the survival analysis.

Since the caregiver survey administered alongside the child measurements was the same every round, we have multiple caregiver-reported outcomes for a given child who was measured at different age windows. For this reason, we run regressions at the child–age-window level and include age-window fixed effects. Otherwise, analysis of these outcomes is identical to the specification described in this subsection.

3.4 Threats to validity

The most important threat to the validity of our estimates is sampling bias. In this section, we discuss sampling bias concerns for each of our main outcomes.

Given high survey rates in our follow-up surveys, there is little risk of sampling bias for the fertility and family formation outcomes. 95% of the sample could be surveyed after 11 years (2019) and there was no differential attrition between the treatment and the control (Table A2). Survey rates dropped somewhat in the 2022 (85%) and 2023 (80%) follow-ups. In 2023, attrition among male participants was higher in the control group (22% vs. 15%, p-value=0.005), so we focus on 2019 and 2022 outcomes when possible.

For the child mortality outcomes, we ensure low attrition rates by relying on data obtained through surveys conducted with first-generation respondents almost yearly between 2009 and 2023. While scholarship-lottery participant-level attrition in these surveys was minimal (up to 2019), child-level attrition in any single survey round may be non-trivial as parents may neglect to mention children who passed away years ago. For example, when asked "Did you (or your partner) ever give birth to a biological child who was born alive but did not survive?", 22 of the 49 respondents who responded *yes* in 2017 responded *no* in 2019. After being specifically asked about the deceased child they listed previously, 21 of the 22 agreed that they had a child who passed away. This highlights the benefit of having conducted surveys almost yearly. Drawing upon all follow-up surveys from 2009-2023, we have survival status for 95% (94%) of all children (ever mentioned by a scholarship-lottery participant) who had, or would have, turned at least 1 (3) by the last follow-up survey in

[13] For reference, Figure A.2 plots cognitive development score against child age in months by age window among the control group.

Spring 2023 (Table A2).

A limitation of our child mortality data is that we miss children who would have turned 1 or 3 after we stopped surveying in Spring 2023. Of the children ever mentioned by a scholarship-lottery participant, 8% had not turned 1 and 25% had not turned 3 by Spring 2023. There may also be a sample selection bias among children ever born. Since Ghana has a total fertility rate of ≈3.6 per woman (United Nations, 2022), it is likely that our sample will have more children over the course of their lives. Of particular concern is the subset of participants who do not yet have children but may have children in the future. By Spring 2023 (when scholarship-lottery participants were 31 years old on average), 75% of women and 52% of men reported having at least one child (Table A2). While these proportions are not significantly different between treatment and control groups, the characteristics of respondents who had at least one child by Spring 2023 differ by treatment status among men (but not women), as shown in panel B of Table A3. In particular, among men who had a child by Spring 2023, scholarship recipients are 9.3 percentage points (47%; p-value=0.004) more likely to come from a household with no male head at baseline compared to those in the control group. This result suggests that male scholarship recipients for whom we can study offspring survival may be negatively selected (in terms of baseline SES) relative to control men. To understand whether this could be driving the results, we perform robustness checks with entropy balancing following Hainmueller (2012).[14]

For the cognitive tests, we had relatively high tracking rates. Table A2 shows that tracking rates ranged from 78-94% for eligible children. Some children were never eligible for the cognitive tests. These children were either already past seven years when we began measurements or did not reach 14 months before the cessation of our measurements. Having children who were too old when assessments began accounts for the largest share of never-eligible parents (5% of all female scholarship-lottery participants and 2% all of males fall into this category).[15] Overall, we have measures of cognitive development for at least one eligible child for 64% of female scholarship-lottery participants (86% of those who had a child) and 41% of male scholarship-lottery participants (79% of those who had a child) (Table A2). As with child survival, we use entropy balancing to address the fact that men with children in the cognitive games sample have slightly different characteristics between treatment and control group (Table A3, Panel B.)

[14]It is possible that men in the treatment group are more likely to be *aware* of, or to *recognize*, children they may have had outside of wedlock. While this is a fascinating question, we are unable to investigate it since we don't know the "true" fertility rate among men.

[15]These respondents had children prior to May 2012 (when they were 21 years old on average) and none later.

4 Results

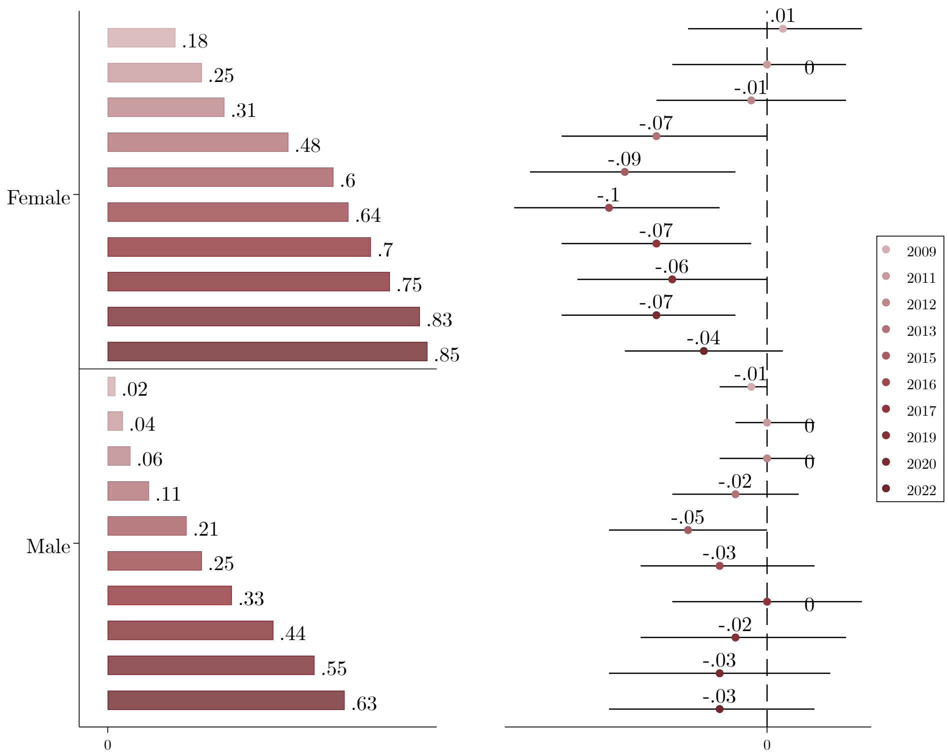
In this section, we start by presenting the impact of the scholarship on family formation and fertility choices. We then show the impact on child survival for all children born to scholarship-lottery participants and the impact on child cognitive development for those with children who completed our cognitive development measures. We show that female scholarship recipients (who were, on average, 27 percentage points more likely to complete secondary school) delay fertility and marriage relative to the control group. When these women have children, their children have lower child mortality and, by age 5 and 7, have significantly higher cognitive scores than children of non-recipients–but we see no such effects for children of male scholarship recipients. Finally, we present evidence on possible channels explaining these striking results, including parent/caregiver characteristics and behavior.

4.1 Fertility and family formation

Table 1 presents the impacts of the scholarship on fertility and marriage. Scholarships lead to delays in childbearing onset and reduce unwanted pregnancies for women. By 2013, women in the scholarship arm were 6.9 percentage points less likely (p-value=0.039) to have ever been pregnant (on a base of 48.3% in the control group). Because the great majority of first pregnancies are reported to be unwanted, the fertility decline is almost exclusively a decline in unplanned, out-of-wedlock pregnancies (column 2). As shown in Figure 1, the delay in childbearing onset is sustained over many years. By 2019, female scholarship recipients are still 6 percentage points less likely to have started childbearing than non-recipients and had fewer children (-0.152 fewer children, p-value 0.065) (column 3 of Table 1). These results are consistent with those of an earlier randomized experiment that reduced the cost of access to upper primary school in Kenya and found that the onset of childbearing was also delayed, with no-catch up in the three years following school exit (Duflo et al., 2015). They are also consistent with estimates based on natural experiments, such as the discontinuity created by admissions cutoff for secondary school in Kenya (Ozier, 2018) or the introduction of free primary school in Uganda (Keats, 2018). Eventually, fertility does catch up for female scholarship recipients. By 2022, the difference in the number of children ever had is not statistically significant.

The finding that the gap in childbearing between treatment and comparison groups persists once the majority of scholarship recipients are out of school suggests that the mechanism is not an

Figure 1: Impact on childbearing onset—-ever pregnant/had a pregnant partner, by year



Notes: Data from 2013 in-person follow-up and yearly phone surveys. The outcome shown is "Ever pregnant" (for females) and "Ever had a pregnant partner" (for males). Left half of graph shows means in comparison group; right half shows estimated treatment effects of scholarship and 95% confidence intervals.

"incarceration effect", preventing fertility for a few years while in school (Black et al., 2008). Our rich data helps shed light on the relative importance of the alternative mechanisms most discussed in the literature. These include (1) an increase in the opportunity cost of bearing and raising children (Becker, 1991); (2) the ability to control fertility due to better decoding of information (Rosenzweig and Schultz, 1989); (3) changes in the type or preferences of the partner, and in the bargaining power of each partner; and (4) a decrease in the cost of investing in each child's quality (education and health), which in turns affects the demand for the quantity of children (Becker, 1991). These channels can of course operate conjointly.

In Duflo et al. (2024), we find that, consistent with channel (1), female scholarship recipients are more likely to have regular salaried employment than female non-recipients, which presumably increases the opportunity cost of a child. We also document increases in learning and cognitive scores for both men and women, which could facilitate channel (2).

Here we document patterns consistent with channel (3). First, fertility changes coincide with changes in cohabiting behavior. By 2016 (age 25 on average), female scholarship recipients were 12.1 percentage points (24% of the control mean) less likely to report having ever lived with a partner (Table 1, column 4). As of 2019, they are 6.2 percentage points (p-value=0.067) less likely to be married or cohabiting (compared to a base of 47.5% in the control group). Conditional on having a partner, they are more likely to have a partner that completed tertiary education (p-value=0.071; column 8).

In contrast, we see few changes in fertility and marriage behavior for male scholarship recipients, although it is worth noting that men marry/cohabit later and that parenthood is likely measured with much more error for them. Since many pregnancies are out of wedlock and not all of them lead to marriages, it is possible that male respondents under-report births they may have been responsible for. One clear impact on male scholarship recipients is that they are more likely to still be living with their parents (+ 7.8 percentage points, or 30% of the control mean, in 2019), which is not the case for female scholarship recipients.

In the rest of the paper, we show evidence that is consistent with either channel (4) (reduced costs of investing in children quality) or a direct impact of the "wantedness" of children on their quality. Indeed, children of scholarship recipients are healthier, and they have higher cognitive achievement.

4.2 Child survival

In Table 2, we present the results on child survival. The unit of observation in this table is the child. In columns 1 and 2, we control for birth linearly. Among children of female respondents, we find a 1.8 percentage point increase in the probability of surviving until age 1 (51% decrease in mortality, p-value=0.028) and an increase of 1.8 percentage points in survival-to-age-3 probability (a 44% decrease in mortality, p-value=0.065).

For children of male respondents, the estimates are smaller and noisier. We estimate an insignificant 1.4 percentage point increase in the probability to survive until age 1 (47% decrease in mortality, p-value=0.161) and an insignificant increase of 0.9 percentage points in survival-to-age-3 probability (a 31% decrease in mortality, p-value=0.549). These estimates are not significantly different from the estimates for female respondents.

Varying the covariates included does not substantively affect our estimates. In columns 3 and 4 of Table 2, we use birth year fixed effects rather than controlling for birth year linearly. The results only change slightly. For children of female respondents, the effect on survived-to-1 is slightly lower (1.7 percentage point; p-value=0.032). For children of male respondents, the survived-to-3 estimate is slightly higher (0.9 pp; p=.477). In Table A5, we drop birth order fixed effects (columns 1 and 2), add a control for mother's age at birth (columns 3 and 4), or drop all controls for birth year (columns 5 and 6). In all cases, the results barely change from those presented in columns 1 and 2 of Table 2.

To understand whether differences in the characteristics of the scholarship recipients and non-recipients who had children drives our effects, we use entropy balancing (Table A6). The results for children of female recipients do not change, which we expected since there is no imbalance between the recipients who had children and non-recipients who had children. The estimates also do not change for children of male respondents, suggesting that the imbalances estimated in Table A3 are not driving the results.

We adjust for multiple hypothesis testing using the step-down procedure proposed by Romano and Wolf (2005). The p-value on the survived-to-1 effect increases from 0.028 to 0.119 for children of female scholarship-lottery participants, while the p-value on the survived-to-one effect for the children of male participants increases from 0.161 to 0.562 (column 1 of Table A8).

In general, the child mortality results for children of male scholarship-lottery participants are more sensitive to perturbations of the data than our other results. This sensitivity is driven by the fact that the sample size is smaller since fewer male participants report having children, and few children pass away prior to the age of 1 or 3. Focusing on survival-to-1 among the children of male respondents, only 26/1,016 passed away before age 1 (20 in the control group; 6 in the treatment group). Among the children of female respondents, 50/2,716 passed away before age 1 (41 in the control group; 9 in the treatment group). Nonetheless, the stark decline observed in our sample is remarkable.

4.3 Child cognitive development

In Table 3, we present results on child cognitive development; once again, the unit of observation is the child. We estimate a child's cognitive development by age window using item response theory (IRT). For each measure, we estimate a one-parameter logistic model on the relevant cognitive games questions.[16] The model assigns a difficulty-level to each question and then a latent trait

[16]Specifically, we estimate the model on a set of binary variables indicating whether the child was correct or incorrect on a given trial.

to each individual which measures their ability to respond correctly to the questions. We use the standardized latent trait assigned to a child as a measure of the child's cognitive ability (we will call this their IRT score). Consistent with the pre-analysis plan, non-responses by the child are dropped from our analysis since these were often caused by distractions arising in the field (e.g., other children distracting the child) or equipment failures. In Table C1, we show that our results are robust to scoring these questions as incorrect responses.

For children of female scholarship recipients, the estimated treatment effects are insignificant and slightly negative for 18 month olds (-0.066 standard deviations (SDs); p-value=0.489) and 2.5 year olds (-0.024 SDs; p-value=0.850), and insignificant and slightly positive for three year olds (.026 SDs; p-value=0.736). In contrast, the five and seven year olds of female scholarship recipients score substantially higher on the cognitive development tests, 0.238 and 0.252 SDs respectively (p-values=0.005 and 0.035). These results are robust to multiple hypotheses testing adjustment (Table A8) and to excluding measurement round fixed effects (Table A9). These effects fall between the 75th and 80th percentile of effect sizes for the 96 RCTs measuring the impacts of educational interventions on learning in low-and-middle-income countries considered in a recent meta-analysis (Evans and Yuan, 2022). In terms of early childhood education interventions, these effects are close to those of the most effective rigorously evaluated interventions, such as hiring an additional teacher focused on preschool instruction (0.29 and 0.46 SD increases in math and language scores; Ganimian et al. (2021)), offering scholarships for high-quality kindergartens (0.40 SDs; Dean and Jayachandran (2019)), and improving preschool curricula (0.11-0.26 SDs) (Dillon et al., 2017; Gallego et al., 2021; Oreopoulos et al., 2020). Breaking the results for female respondents down by cognitive domain (tables C2 to C6), we find strong effects on language skills (0.15 SDs for five year olds; 0.27 SDs for seven year olds), math and numeracy (0.15 SDs; 0.26 SDs), spatial reasoning (0.20 SDs; 0.12 SDs), and executive function (0.25 SDs; 0.20 SDs) but no effect on socio-cognitive development.[17,18]

It is noteworthy that a treatment effect emerges only once children reach age 5, increases from age 5 to 7, and focuses primarily on cognitive skills that underlie, and are enhanced by, learning to read and calculate in school. These findings suggest that having a more educated mother leads

[17] Note that Coffey and Spelke (2023) documents correlations across children and time that suggest that the socio-cognitive development tests may not have measured the underlying trait as intended. We keep these in the overall index shown in the main tables since this is what we had pre-specified.

[18] We break down the impacts on child cognitive development by child gender in Table A10. The sample sizes become small with this breakdown (especially for children of male respondents) and the results are quite imprecise. We cannot reject equality of the treatment effect by child gender among children of female scholarship-lottery participants.

to gains in children's readiness for learning in school, perhaps because an educated mother is more likely to bridge the gap between school and home. Another, more mechanical, interpretation could be that the cognitive tests are more robust at older ages, and that the tests of language, math, and executive function are more robust than the tests of socio-emotional development. Coffey and Spelke (2023) tested construct validity of our tests, by measuring overall correlations between game scores within the same domain cross-sectionally and longitudinally, and indeed the five and seven year old games appear to be significantly more reliable measures of the targeted cognitive abilities than the 1.5 years, 2.5 years and 3.5 years old games.[19] It is thus possible that the null effects on younger age groups are driven by noisy measurement. In other words, we cannot exclude the possibility that the mother's receipt of the scholarship mattered for the younger children on dimensions that we either didn't measure or measured less effectively, due to limitations on the number of questions young children would sit through and on the number of response options they were able to consider for each question. On the other hand, we see null effects on caregiver-reported language outcomes (Table A11).[20] The most plausible interpretation of the results may be that impacts on cognitive development only emerge after a few years. Given the treatment effect on mortality, it could also be that marginal children in the treatment group start with a cognitive deficit (say, because they survived a cerebral malaria episode, so they are alive but weakened), and it takes time for the impact of maternal education on child cognitive development to overcome this initial deficit.

Turning to children of male scholarship-lottery participants, we find no significant effects on cognitive development at any age, and the point estimates are negative at all ages except for 1.5 years old, and even marginally significantly negative for 5 year olds (Table 3, point estimate -0.22, p-value=0.069). The difference in effect sizes for children of male scholarship-lottery participants compared to children of female scholarship-lottery participants is significant for five year olds (Table 3 column 4, p-value=0.005).

Recall that for male scholarship-lottery participants, selection into parenthood is large since only about half report that they ever had a child (Table A2), and we see some imbalance in 2008 baseline characteristics between treatment and control groups within that subsample (Table A3).

[19] If the tests are measuring a cognitive domain accurately, performance on one of the tests should predict performance on a subsequent test in the same domain. We find that, for children who took both sets of tests, the five year old game scores are highly correlated with seven year old game scores (.53), while 1.5 year old game scores have little correlation with 2.5 or 3.5 year old game scores (0.15 and 0.07 respectively—see Figure A.3.)

[20] Dupas et al. (2024) found significant positive effects of a light-touch infant-directed speech information treatment among infants aged 0-24 months in Northern Ghana using our caregiver-reported language measure, suggesting this measure has some signal.

In Table A6, we use entropy balancing to reweigh observations so as to obtain balance on 2008 characteristics. The results are qualitatively unchanged.

Overall, these results suggest that investing in universal female secondary school education improves the cognitive abilities of the next generation, especially those that are most directly tied to learning in school, while additional investments in males' education alone does not appear to have the same magnitude of effects.

4.4 Channels

What are the likely explanations of the positive impact for children of female scholarship recipients, and the lack of effects for male recipients?

4.4.1 Parental education

Table A12 shows the scholarship treatment effect on the subsample of scholarship-lottery participants whose children could be surveyed and hence form the sample for the results on cognitive development.[21] The results are nearly identical to those reported for the full sample (Table A1), confirming large differences in parental education.

Notably, conditional on having a partner, female scholarship recipients are significantly more likely to have partners with tertiary education (+8.5 percentage points on a basis of 16.6%, column 5 of Table A12). However the opposite holds for men: while only 4.3% have a partner who has tertiary education in the control group, this reduces further by a significant 3.6 percentage points in the treatment group.

Maternal and parternal education do not have the same effect on the education of a child's primary caregiver. For 84% of the children, the primary caregiver is the child's mother (column 1 of Table 4). As a result, the primary caregiver for the children of female recipients is 25 percentage points more likely to have completed secondary school and 5 percentage points more likely to have completed tertiary (Table 4, columns 3 and 4). In contrast, children of male scholarship recipients have caregivers with the same level of education as children of males in the control group. In Table A7, we show that the treatment effects for female scholarship recipients are not altered when we add a control for father's education—suggesting that maternal education is the main driver.

Another reason why paternal and maternal education may not have symmetric impacts on a

[21] The results are identical if we include those who had a child who could not be surveyed.

child's environment is that children of male scholarship recipients appear significantly less likely to live with their father (-8.7 percentage points, p-value=0.024, column 4 Table A13). We see no such effect on the probability that children of female scholarship recipients live with their mother. Overall, the probability a child lives with their scholarship-recipient parent is much lower for children of male compared to female scholarship recipients (62% vs. 92%).

4.4.2 How does maternal education affect children's outcome?

Maternal education could affect child survival and cognitive development through lower or delayed fertility, higher, less volatile parental income, greater health knowledge, better parenting skills, higher valuation of a child's education or health, and/or higher bargaining power for women/improved marriage market prospects.

Maternal age The striking reduction in child mortality among female scholarship recipients could be a direct result of an increase in maternal age. Using DHS data from 55 low- and middle-income countries, Finlay et al. (2011) document a strong correlation between maternal age at birth and child health outcomes, including survival, even after controlling for socioeconomic status. They conclude that there is likely a causal effect of the biological maturity of the mother. Child marriage bans improve child health (Le et al., 2024) and educational outcomes (Chari et al., 2017), with higher maternal age at marriage and first birth being a likely mechanism. In our setting, maternal age at birth increases by 0.349 years for children born to female scholarship recipients (p-value=0.142) compared to the control group (column 7 of Table A5). For first-born children, the gap in maternal age at birth between treatment and control group rises to 0.64 years (p-value=0.040, Table A14, col 3). We are under-powered to detect mortality effects on first-born children (which make up 44% of the sample), but the point estimates in Table A14 are somewhat *smaller* compared to the overall mortality effects in Table 2, suggesting that maternal age at birth may not be the primary driver of the mortality results in our sample. Another piece of evidence suggesting a minor role for maternal age at birth in our context is found in the results for children of male scholarship recipients: maternal age at birth falls for them (Table A5, -0.46 years; p-value=0.22), yet the point estimates on survival are all positive.

Resources Duflo et al. (2024) report no treatment effect on income until 2019, but, by 2020, women who received a scholarship start showing 24% higher earnings than women who did not receive a scholarship. This means that the children of scholarship recipients who are old enough to have been tested at the ages of 5 or 7 by 2023 likely did not have improved economic resources in their early years of life, though their mother may have achieved better material circumstances by the time we conducted the tests (recall that we conducted the children tests from 2017 to 2023). Consistent with this, we see only a modest difference in socio-economic status on average (0.107 SDs gain in the SES index, p-value=0.103, not at all distinguishable from the effect for children of male scholarship recipients, column 5 of Table 4).

Resources *per child* could be greater since female scholarship recipients started bearing children later. We do not find any evidence for a quality-quantity trade-off, however. In fact, surprisingly, children of scholarship recipients do *not* have caregivers with fewer children to care for (column 6 of Table 4). This is surprising since we know female scholarship recipients had fewer children when we began the cognitive assessments. Possible explanations for this include: (i) the survival impact documented earlier; (ii) the fact that the sample of scholarship recipients who do have at least one child old enough to be tested is positively selected in terms of fertility; and (iii) children born from unwanted teen pregnancies may be more likely to be fostered by grandparents, and therefore the older children of non-recipients may not be living with their mothers. We do not have data on the location of children older than 7 so we cannot test whether this is a factor. We can however rule out yet a fourth explanation, namely, that scholarship recipients are more likely to foster children who are not their biological children, e.g., nephews and nieces: the (lack of) effect in column 6 of Table 4 is found even among biological children.

Aspirations and knowledge about parenting We see no significant differences in caregiver's aspirations for their children, or in their knowledge about the role of parental stimulation in children's cognitive development (columns 7 and 8 of Table 4, and tables B1 and B2).

Caregiver practices and behaviors Table 5 turns to parental/caregiver behavior, and a number of meaningful differences emerge. Turning first to health investments, we find that female scholarship recipients are significantly more likely to receive prenatal care during pregnancy (column 1). We see a positive and significant (p-value=0.068) effect on preventive health behaviors (column 2).[22] Moreover, we observe an improvement in caregiver-reported child health (Table A15 and Table B4). Finally, for the subset of children for which anthropometric outcomes

[22]The preventive health behaviors in the index are shown in Table B3. We see a significant impact of bednet usage and usage of a private toilet, but no change in water treatment. The most common source of drinking water for the children in our sample is sachet/bottled water (main source for 55.3% of children).

could be measured (see breakdown in Table B5), we find that stunting and wasting are not differential across treatment and comparison groups—if anything, they are worse for children of female scholarship recipients, which may be due to the fact that the frailest children of non-recipients were more likely to pass away prematurely.

Turning to potential channels for the effects on cognitive development, we see no significant impact on a child investment index—monetary investments in food or education-related supplies (see column 4 of Table 5, with breakdown shown in Table B7) nor on a schooling index (column 5 of Table 5, with breakdown shown in Table B8). However, among children of female scholarship recipients, the caregiver reports more interacting with the child in stimulating ways (column 3 of Table 5, with breakdown in Table B6). Drawing upon our objective measurements of adult-child interactions through day-long recordings of the child's auditory environment via the LENA (Table 6), we confirm the presence of increased adult-child engagement for children of female scholarship recipients. In particular, we see increases in conversational turns per minute (0.068 effect size, a 20% increase; p-value=0.005) and child vocalization per minute (0.32 effect size, a 17% increase, p-value 0.014).[23]

Summing up, female scholarship recipients do not invest greater financial resources in their children; but they seek more preventive care, and engage in more of the parent-child interactions that help children develop— not because they are more informed that such interactions are particularly important, but likely because their greater level of education enables them to interact with others (including their children) in a different way.

Consistent with the fact that caregivers of children of male scholarship recipients do not have more education than their counterparts in the control group, the impact on caregiver behavior for children of male respondents are never positive—all the coefficients in Table 5 are insignificant. Turning to LENA measures, we find *negative* treatment effects for children of male scholarship recipients (Table 6), consistent with the finding that they are less likely to live with their father (Table A13). Absence of the father mechanically reduces exposure to male adult words, and may also cause reduced exposure to female adult words if single mothers have less time on hand to verbally engage with each child individually. Table A17 shows that the negative LENA results for children of male scholarship recipients still hold with entropy balancing.

[23] We have non-trivial attrition in the LENA measurements, as shown in Table A2, but we find no imbalance in child characteristics (Table A4) nor in 2008 baseline characteristics or current household environment (Table A16) for female respondents. As expected given the imbalances in Table A3, there is imbalance on 2008 baseline characteristics for male respondents. We show results with entropy balancing in Table A17.

5 Cost-effectiveness and Cost-benefit analysis

Intergenerational impacts rarely factor into policy debates around subsidizing secondary education. Yet, our findings of substantial intergenerational impacts for female students suggest that such impacts may be the strongest basis for public subsidies of education. We estimate the cost per child death averted, cost per standard deviation improvement in child cognitive development, the benefit-cost ratio, and the internal rate of return (IRR) for a program providing means-tested secondary education scholarships. While such an exercise is, as always, sensitive to a host of assumptions, it provides a helpful sense of the relative magnitude of the benefits.

In Table 7, we present these estimates for a means-tested scholarship program open to both genders or open only to female students. We assume that the at-scale scholarship program would be means-tested in a manner similar to the program we evaluated (i.e., only those facing some difficulty paying for SHS on their own are eligible). To account for the fact that means-testing procedures may be less effective at-scale, we also present the benefit-cost ratio and internal rate of return of the program if targeting accuracy fell by 25% relative to the procedures used in this study, meaning that there would be 25% fewer students who complete secondary high school due to the program and thus, the benefits are 25% lower.[24]

5.1 Measuring costs

To measure costs from a social perspective, we add together school fees, expenses on school materials, and wages foregone due to receiving the scholarship. Duflo et al. (2024) directly estimate the effect of the scholarship for expenses on school materials and foregone wages. To estimate the effect on school fees paid, we combine Duflo et al. (2024)'s estimate of the scholarship's effect on years in SHS with administrative data on school fees for scholarship recipients. Since most of our benefits are per *second-generation child*, we divide these costs by the number of children per scholarship recipient to get the cost per second-generation child. Note that this is conservative since we don't observe complete fertility yet and thus, underestimate the fertility rate. We estimate a cost of $585 per scholarship recipient if the program covers males and females and a cost of $505 per scholarship recipient if the program only covers females.

[24] To select scholarship eligible students, Duflo et al. (2024) identified students who were admitted to secondary school but had not enrolled by the end of the first semester of the school year. 95% of these students cited financial difficulties as the reason they did not enroll. Such a targeting strategy would not be feasible at-scale so an alternative (and potentially less accurate) method would need to be used.

5.2 Cost-effectiveness

In column 1 Panel A of Table 7, we calculate the cost-effectiveness of secondary school scholarships in terms of averting child deaths. Combining the estimate of cost per scholarship recipient with the child mortality reduction (Table 2) gives us a cost per under-3 death averted of $23,582 for scholarships given to both genders, and of $15,184 for scholarships given to females only. We use the ITT estimate on mortality, which means that the cost of subsidizing inframarginal students who would go to school anyways is taken into account. To put these estimates in perspective, we convert Stenberg et al. (2021)'s estimates of the cost per healthy life year delivered by 36 WHO-recommended neo-natal and/or child health interventions into cost per under-3 death averted, using the WHO's 2019 recommendations. Stenberg et al. (2021)'s converted estimates imply that the median WHO-recommended intervention costs $2,300 to $8,200 per under-3 death averted, the 25th percentile intervention costs $3,600 to $16,100, and the 10th percentile intervention costs $11,300 to $33,200.[25] Thus, our estimates imply that women-only scholarship could be in the recommended set of interventions to reduce child mortality, even if they had no other impacts.

In Column 1 Panel B of Table 7, we evaluate the cost-effectiveness of secondary school scholarships in terms of improving children's cognitive development. As our estimate of the impact on cognitive development, we take the average of the intergenerational impact of the scholarships on cognitive scores for five and seven year-olds (Table 3). Using this estimate yields a cost per SD increase in early childhood cognitive test scores of $10,986 if the program covers both genders and $3,521 if it covers females only (Table 7).

5.3 Measuring benefits

We measure the benefits of the scholarship program from averted child deaths, improved early childhood cognitive development, and greater earnings for the scholarship recipients. We ignore the other effects of the scholarship (e.g., effects on caregiver-reported child health or first-generation life satisfaction) because we expect these benefits to be smaller in magnitude than the three benefits we consider, and we want to avoid double counting benefits (for example improved child health may result into lower child mortality).

We calculate the total benefits from averted child deaths using estimates of the value of statistical

[25] The ranges reflect differing estimates across region (Stenberg et al. (2021) evaluates interventions in East Africa or South Asia) and assumed coverage rate (50%, 80%, or 95%). The lower bound assumes a 95% coverage rate in East Africa. The upper bound assumes a 50% coverage rate in South Asia.

life (VSL) for our sample. We follow Robinson et al. (2019)'s recommendation for conducting VSL sensitivity analyses.[26] The three methods proposed by Robinson et al. (2019) suggest that we multiply Ghana's GNI per capita by 100, 33, or 160, yielding VSLs of $235,000 (refer to 'Medium B-C ratio' and 'Medium IRR' columns in Table 7), $76,862 (refer to 'Low B-C ratio' and 'Low IRR' columns), and $376,000 (refer to 'High B-C ratio' and 'High IRR' columns), respectively. Note that all these VSLs (even the "high" one) are very low compared to the numbers used in the United States.[27]

We quantify the value of early childhood cognitive gains by projecting their effect on adult earnings and taking the present value of this income stream. To model the relationship between early childhood cognitive development and adult earnings, we rely upon Gertler et al. (2014). They exploit a randomized-controlled trial of an early childhood stimulation program to estimate the effect of a s.d. increase in early childhood cognitive scores on adult earnings. They estimate that a 1 s.d. increase in early childhood cognitive scores translates to a ≈33% increase in annual adult income. While this study is in Jamaica rather than Ghana, we use their estimate as our 'best guess' of the relationship between early childhood cognitive gains and adult earnings. Assuming that this relationship is linear, we multiply the treatment effect on five and seven year-old cognitive test scores for our sample (Table 3) by 0.33 to project the percentage increase in adult earnings for the second-generation. This method generates a 3.5% increase in adult earnings for children of male or female scholarship recipients and a 7.5% for children of female scholarship recipients only.

To calculate the present value of these additional earnings, we assume that the benefits begin when a child turns 20 and persist for their working life (40 years). Given our uncertainty about the second-generation's adult income absent the intervention, we use three different estimates: GNI per capita (refer to 'Medium' columns of Table 7), first-generation control group mean earnings in 2023 (refer to 'Low' columns), and projected GNI per capita in each year if GNI per capita grows at 3% annually (refer to 'High' columns)[28]. The 'low' estimate assumes that the second-generation's income absent the intervention is the same as the first-generation's income in 2023 (when the first-generation was 31 years old on average) among those who did not receive the scholarship. The

[26]Refer to Appendix F for more details on the three methods recommended by Robinson et al. (2019).

[27]For example, the Environmental Protection Agency (EPA)'s website recommends using a VSL of $7.4 million USD2006, updated to the year of the analysis (see https://www.epa.gov/environmental-economics/mortality-risk-valuation#means, last accessed July 8, 2024.)

[28]This projection is consistent with Ghana's average GNI per capita growth from 2014-2022. It is important to note that this projection is highly speculative and slightly changing the projected growth rate significantly affects the estimated 'high' benefit-cost ratios.

'medium' estimate accounts for the fact that earnings increase over one's career, generally and in Ghana specifically.[29] Given this fact, the average earnings over the entire population is a more appropriate counterfactual income than first-generation control group earnings at 31 years old, even though our sample likely differs in important ways from the average Ghanaian.[30] The 'high' estimate accounts for the likelihood that the average income in Ghana will increase over time. This consideration is particularly important for projecting second-generation income since these benefits only begin to accrue in the 2030s and beyond.

We take a similar approach to quantifying the value of first-generation labor market benefits. Since Duflo et al. (2024) estimates no effect on earnings until 2020, we model the first-generation labor market benefits as a stream of income starting at 31 years old and lasting for the remainder of their working life (34 years). We assume that Duflo et al. (2024)'s estimated effect of the scholarship on 2023 earnings (13% for all recipients; 30% increase for female recipients only) persists throughout this period. For income absent the intervention, we, once again, have three difference estimates: GNI per capita ('Medium'), first-generation control group mean earnings ('Low'),[31] and projected GNI per capita in each year if GNI per capita grows at 3% annually ('High').

5.3.1 Benefit-cost ratios and internal rates of return

In Columns 3-5 of Table 7, we estimate a range of benefit-cost ratios and internal rates of return for secondary school scholarships. In Panels A-C, we consider the intergenerational benefits (averting child deaths and/or improving child cognitive development). In Panel D, we consider the first-generation labor market benefits estimated in Duflo et al. (2024). Panel E combines the first-generation and intergenerational benefits, providing an 'all-things-considered' benefit-cost ratio for comparison with other potential social investments. To account for potential slippage in targeting of the scholarship at-scale, we present the intergenerational benefits and 'all-things-considered' benefits if targeting accuracy fell by 25% in Panels F and G.

Our 'all-things-considered' benefit-cost ratios indicate that secondary school scholarships produce substantial social benefits, and even more so when targeted to female students. The

[29] In Round 6 of the Ghana Living Standards Survey, reported earnings increased up to one's late-fifties (Ghana Statistical Service, 2016)

[30] In general, we would expect our sample to earn more over their lifetime than the average Ghanaian since the first-generation respondents qualified for Senior High School and only ≈ 30% of Ghanaians qualified for Senior High School in this time period (Duflo et al., 2024).

[31] For first-generation earnings, we use the control group females yearly earnings ($417) rather than the overall control group yearly earnings ($687) when estimating benefits of the females-only scholarship. We do not do this for second-generation earnings because female recipients will have male and female offspring.

benefit-cost ratio estimates for female-only scholarships are 44 ('Medium B-C ratio'), 10 ('Low B-C ratio'), and 111 ('High B-C ratio') (Panel E; Table 7). The internal rates of return for female-only scholarships are estimated as 63% ('Medium'), 27% ('Low'), and 76% ('High'). The magnitude of the social benefits means that the social efficiency of the policy is robust to targeting accuracy falling by 25% when the policy is scaled up (Panel G; Table 7). An at-scale scholarship program for both genders is still socially-beneficial, but, due to the low impact of subsidizing boys' scholarships, the benefit-cost ratio falls by ≈50% and the IRR falls by ≈30% relative to a scholarship program restricted to female students (Panel E; Table 7).

Importantly, the benefits of secondary education for the next generation outstrip the benefits to the first-generation. As one can see by comparing Panel C and E of Table 7, the benefits to the second-generation account for 55-75% of the total benefits. Since the intergenerational survival benefits occur earlier than the first-generation labor market benefits, the internal rate of return for the total benefits is largely determined by the internal rate of return on the intergenerational survival benefits. Moreover, the inter-generational impacts are likely to be much less sensitive to possible general equilibrium impacts than the first-generation impacts. Duflo et al. (2024) show that the first-generation labor market impacts are primarily due to access to scarce public sectors job, and therefore may not obtain if the program was generalized. In contrast, the second generation estimates probably reflects the impacts of improved child-rearing human capital, and we would not expect negative externalities on other parents. This result underscores the importance of accounting for intergenerational effects when considering investments in secondary education.

Averting child deaths accounts for a large share of the benefits of secondary school scholarships. Depending on the scenario, averting child deaths accounts for 36-66% of total benefits (Panel A and E; Table 7). In fact, the scholarship program would still be socially-beneficial even if averting child deaths were the only benefit (benefit-cost ratios range from 3-57 and IRRs range from 13-52%).

Given the potential sensitivity of the mortality estimate, it is worth noting that the benefit-cost ratios of an at-scale scholarship program considering only intergenerational cognitive development benefits are 3 for both genders or 9 for females-only ('medium' benefits scenario), 0.8 for both genders or 1.6 for females-only ('low' benefits scenario), and 7 for both genders or 21 for females-only ('high' benefits scenario) (Panel B of Table 7). This means that these benefits alone would essentially cover the costs of female-only scholarships, even under pessimistic assumptions about future benefits. Note that these estimates rely on our assumption that the relationship between early childhood cognitive gains and adult income estimated in Gertler et al. (2014) holds in our context. First-generation labor market benefits-costs ratios are larger in magnitude (7, 2, and 12 for a program open to both genders; 19, 3, and 33 for a females-only program; Panel D of Table 7), but these estimates do not account for the possible dilution from general equilibrium impacts.

6 Conclusion

The UN *Call to Action on Education Investment* states that governments in low and lower-middle income countries "shall allocate at least 4-6% of GDP and at least 15-20% of total public expenditure to education, protecting public education budgets from the constrained fiscal environment resulting from the COVID 19 pandemic and the global economic crisis." Yet, the IMF estimates that the median education budget in sub-Saharan Africa was equal to about 3.5 percent of gross domestic product in 2020.[32] As sub-Saharan governments struggle to find financing to maintain, let alone expand, free education, understanding the long-term gains from such investments is key. While mothers' education has long been assumed to have a range of positive impacts on children, rigorous evidence to back this claim had been lacking.

By studying the long-term impacts of a randomized scholarship program in Ghana, this paper provides evidence that secondary school education for women does indeed have strong positive impacts on the next generation. Given the size of the child mortality and cognitive development gains, this externality should be considered when governments or international donors consider whether to fund the expansion of free secondary education, particularly in environments where women are disproportionately likely to drop out absent this policy. Our results indicate the primary mechanism through which women's education benefits the next generation is by raising non-monetary investments (child care and cognitive stimulation) by the primary caregiver of the child. Our rich data shows that access to secondary education causes these caregivers to gain the skills to safeguard their children's health and stimulate their children's cognitive development. One interesting question for future research is whether these parenting aptitudes were improved directly by secondary school instruction, indirectly through secondary education enabling students to learn how to learn or improving students' cognitive abilities. Alternatively, they could be learned from secondary school peers who were generally of higher SES than the marginal students in our study.

[32] see https://www.imf.org/en/Blogs/Articles/2024/04/25/sub-saharan-africas-growth-requires-quality-education-for-growing-population, last accessed April 29 2024.

In contrast, we find no evidence of positive impacts on the children of male scholarship recipients. Duflo et al. (2024) models the labor market outcomes for scholarship recipients and demonstrates how the expectation that women have a lower labor force participation rate might mean that the marginal male induced to attend secondary school by the scholarship would have lower returns to secondary school than the marginal female. Our work supports this interpretation by indicating that the non-labor market returns for the marginal male were also lower than for the marginal female.

References

Aiyar, Anaka and Joseph R. Cummins, "An age profile perspective on two puzzles in global child health: The Indian Enigma & economic growth," *Journal of Development Economics*, 1 2021, *148*, 102569.

Akresh, Richard, Daniel Halim, and Marieke Kleemans, "Long-Term and Intergenerational Effects of Education: Evidence from School Construction in Indonesia," *Econ. J.*, jan 2023, *133* (650), 582–612.

Anderson, Peter J and Natalie Reidy, "Assessing executive function in preschoolers," 2012.

Attanasio, Orazio, Helen Baker-Henningham, Raquel Bernal, Costas Meghir, Diana Pineda, and Marta Rubio-Codina, "Early Stimulation and Nutrition: The Impacts of a Scalable Intervention," *Journal of the European Economic Association*, 8 2022, *20* (4), 1395–1432.

_ , Sarah Cattan, Emla Fitzsimons, Costas Meghir, and Marta Rubio-Codina, "Estimating the Production Function for Human Capital: Results from a Randomized Control Trial in Colombia," *American Economic Review*, 2020.

Baird, Sarah, Craig McIntosh, and Berk Özler, "When the money runs out: Do cash transfers have sustained effects on human capital accumulation?," *J. Dev. Econ.*, 2019, *140* (May), 169–185.

Becker, Gary, *A treatise on the family*, Harvard university press, 1991.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital," *Am. Econ. Rev.*, mar 2005, *95* (1), 437–449.

_ , _ , and _ , "Staying in the Classroom and out of the maternity ward? The effect of compulsory schooling laws on teenage births*," *The Economic Journal*, 2008, *118* (530), 1025–1054.

Chari, A V, Rachel Heath, Annemie Maertens, and Freeha Fatima, "The causal effect of maternal age at marriage on child wellbeing: Evidence from India," *J. Dev. Econ.*, 2017, *127*, 42–55.

Chevalier, Arnaud, "Parental Education and Child's Education: A Natural Experiment," 2004.

Coffey, Joseph R. and Elizabeth S. Spelke, "Evaluating a battery of direct assessments of the cognitive skills of preschool children in Ghana," *Working Paper*, 2023.

Currie, Janet and Enrico Moretti, "Mother's education and the intergenerational transmission of human capital: Evidence from college openings," 2003.

Dean, Joshua T and Seema Jayachandran, "Attending kindergarten improves cognitive but not socioemotional development in India," 2019.

Dillon, Moira R., Harini Kannan, Joshua T. Dean, Elizabeth S. Spelke, and Esther Duflo, "Cognitive science in the field: A preschool intervention durably enhances intuitive but not formal mathematics," *Science*, 2017, *357* (6346), 47–55.

Duflo, Esther, Pascaline Dupas, and Michael Kremer, "Education, HIV, and early fertility: Experimental evidence from Kenya," *American Economic Review*, 2015, *105* (9), 2757–2797.

_ , _ , and _ , "The Impact of Secondary School Subsidies on Career Trajectories in a Dual Labor Market: Experimental Evidence from Ghana," *Working Paper*, 2024.

Duhon, Madeline, Lia Fernald, Joan Hamory, Edward Miguel, Eric Ochieng, and Michael W Walker, "Intergenerational Human Capital Impacts and Complementarities in Kenya," Technical Report, National Bureau of Economic Research 2024.

Dupas, Pascaline, Camille Falezan, Seema Jayachandran, and Mark Walsh, "Informing Mothers about the Benefits of Conversing with Infants: Experimental Evidence from Ghana," *Working Paper*, 2024.

Evans, David K. and Fei Yuan, "How Big Are Effect Sizes in International Education Studies?," *Educational Evaluation and Policy Analysis*, 2022, *44* (3), 532–540.

Fernandes, Michelle, "The Oxford Neurodevelopment Assessment (OX-NDA) for Infants aged 10 to 14 months. THE Ox-NDA MANUAL – V2.1," 2021.

Finlay, Jocelyn E, Emre Özaltin, and David Canning, "The association of maternal age with infant mortality, child anthropometric failure, diarrhoea and anaemia for first births: Evidence from 55 low- and middle-income countries," *BMJ Open*, 2011, *1* (2).

Gallego, Francisco A, Emma Näslund-Hadley, and Mariana Alfonso, "Changing Pedagogy to Improve Skills in Preschools: Experimental Evidence from Peru," *World Bank Economic Review*, 2021, *35* (1), 261–286.

Ganimian, Alejandro J, Karthik Muralidharan, and Christopher R Walters, "Augmenting State Capacity for Child Development: Experimental Evidence from India," Working Paper 28780, National Bureau of Economic Research May 2021.

Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M Chang, and Sally Grantham-McGregor, "Labor market returns to an early childhood stimulation intervention in Jamaica," *Science*, 2014, *344* (6187), 998–1001.

Ghana Statistical Service, "2010 population & housing census: National analytical report," 2013.

_ , "Ghana Living Standards Survey, Round 6," 2016.

Gilkerson, Jill and Jeffrey A Richards, "A Guide to Understanding the Design and Purpose of the LENA® System," *LENA Foundation Technical Report*, 2020.

Hahn, Youjin, Asadul Islam, Kanti Nuzhat, Russell Smyth, and Hee-Seung Yang, "Education, marriage, and fertility: Long-term evidence from a female stipend program in Bangladesh," *Economic Development and Cultural Change*, 2018, *66* (2), 383–415.

Hainmueller, Jens, "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies," *Political Analysis*, 2012, *20* (1), 25–46.

Keats, Anthony, "Women's schooling, fertility, and child health outcomes: Evidence from Uganda's free primary education program," *Journal of Development Economics*, 2018, *135* (May 2016), 142–159.

Le, Dung D, Teresa Molina, Yoko Ibuka, and Rei Goto, "The Intergenerational Health Effects of Child Marriage Bans," 2024.

Leung, Christy YY, Marc W Hernandez, and Dana L Suskind, "Enriching home language environment among families from low-SES backgrounds: A randomized controlled trial of a home visiting curriculum," *Early Childhood Research Quarterly*, 2020, *50*, 24–35.

Libertus, Melissa E., Lisa Feigenson, and Justin Halberda, "Preschool acuity of the approximate number system correlates with school math ability," *Dev. Sci.*, nov 2011, *14* (6), 1292–1300.

Mohanty, Sanjay K, Günther Fink, and Rajesh K Chauhan, "Distal determinants of fertility decline: Evidence from 640 Indian districts," 2016.

Oreopoulos, Philip, Uros Petronijevic, Christine Logel, and Graham Beattie, "Improving non-academic student outcomes using online and text-message coaching," *Journal of Economic Behavior and Organization*, 2020, *171*, 342–360.

Osili, Una Okonkwo and Bridget Terry Long, "Does female schooling reduce fertility? Evidence from Nigeria," *J. Dev. Econ.*, 2008, *87* (1), 57–75.

Ozier, Owen, "The impact of secondary schooling in Kenya: A regression discontinuity analysis," *Journal of Human Resources*, 2018, *53* (1), 157–188.

Robinson, Lisa A, James K Hammitt, and Lucy O Keeffe, "Valuing Mortality Risk Reductions in Global Benefit-Cost Analysis," *J. Benefit Cost Anal.*, 2019, *10*, 15–50.

Romano, Joseph P and Michael Wolf, "Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing," *Journal of the American Statistical Association*, 2005, *100* (469), 94–108.

Rosenzweig, Mark R. and T. Paul Schultz, "Schooling, Information and Nonmarket Productivity: Contraceptive Use and Its Effectiveness," *International Economic Review*, 1989, *30* (2), 457–477.

Simon, J. and James D Wolf, "Choice Reaction Time As A Function Of Angular Stimulus-Response Correspondence And Age," *Ergonomics*, 1963, *6* (1), 99–105.

Stenberg, Karin, Rory Watts, Melanie Y Bertram, Kaia Engesveen, Blerta Maliqi, Lale Say, and Raymond Hutubessy, "Cost-Effectiveness of Interventions to Improve Maternal, Newborn and Child Health Outcomes: A WHO-CHOICE Analysis for Eastern Sub-Saharan Africa and South-East Asia," *Int. J. Heal. policy Manag.*, 2021, *10* (11), 706–723.

Suhonen, Tuomo and Hannu Karhunen, "The intergenerational effects of parental higher education: Evidence from changes in university accessibility," *J. Public Econ.*, 2019, *176*, 195–217.

Suskind, Dana, Kristin R Leffel, Marc W Hernandez, Shannon G Sapolich, Elizabeth Suskind, Erin Kirkham, and Patrick Meehan, "An exploratory study of "Quantitative Linguistic Feedback" effect of LENA feedback on adult language production," *Communication Disorders Quarterly*, 2013, *34* (4), 199–209.

Thomas, Duncan, John Strauss, and Maria-Helena Henriques, "How Does Mother's Education Affect Child Height?," *The Journal of Human Resources*, 1991, *26* (2), 183.

United Nations, "United Nations World Population Prospects: 2022 Revision," Technical Report 2022.

Walker, Susan P., Theodore D. Wachs, Julie Meeks Gardner, Betsy Lozoff, Gail A. Wasserman, Ernesto Pollitt, and Julie A. Carter, "Child development: risk factors for adverse outcomes in developing countries," *Lancet*, 1 2007, *369* (9556), 145–157.

Walker, Susan, Theodore Wachs, Sally Grantham-Mcgregor, Maureen Black, Charles Nelson, Sandra Huffman, Helen Baker-Henningham, Susan Chang, Jena Hamadani, Betsy Lozoff, Julie Gardner, Christine Powell, Atif Rahman, and Linda Richter, "Inequality in early childhood: Risk and protective factors for early child development," *Lancet*, 09 2011, *378*, 1325–38.

Wantchekon, Leonard, Natalija Novta, Marko Klasnja, and Natalija Novta, "Education and Human Capital Externalities : Evidence from Colonial Benin," *Q. J. Econ.*, 2013, *130* (2), 52.

Wietzke, Frank-Borge, "Poverty, Inequality, and Fertility: The Contribution of Demographic Change to Global Poverty Reduction," *Population and Development Review*, 2020, *46* (1), 65–99.

Wodon, Quentin, Claudio Montenegro, Hoa Nguyen, and Adenike Onagoruwa, "The High Cost of Not Educating Girls Missed Opportunities," 2018.

World Health Organization, "WHO methods and data sources for global burden of disease estimates 2000-2019," Technical Report December 2020.

Xu, Dongxin, Umit Yapanel, and Sharmi Gray, "Reliability of the LENA Language Environment Analysis System in young children's natural home environment," *Boulder, CO: Lena Foundation*, 2009, pp. 1–16.

Table 1: First Generation: Impact of Scholarship on Fertility and Marriage

	(1) Ever pregnant/ had a pregnant partner (2013)	(2) Had unwanted pregnancy (2013)	(3) Number of children ever had (2019)	(4) Number of children ever had (2022)	(5) Ever lived with partner (2016)	(6) Currently married or cohabitating (2019)	(7) Still living with parents (2019)	(8) Most recent partner completed tertiary education (2019)
Panel A: Female Scholarship-lottery participants								
Treatment	-0.069**	-0.067**	-0.152*	-0.135	-0.121***	-0.062*	0.003	0.071*
	(0.033)	(0.032)	(0.082)	(0.100)	(0.033)	(0.034)	(0.033)	(0.039)
P-value	0.039	0.038	0.065	0.177	0.000	0.067	0.933	0.071
Comparison mean	0.483	0.390	1.332	1.771	0.498	0.475	0.355	0.195
N	1,009	985	986	877	1,007	986	986	575
Panel B: Male Scholarship-lottery participants								
Treatment	-0.018	-0.012	-0.026	0.022	-0.058**	-0.047	0.078**	-0.051**
	(0.020)	(0.017)	(0.060)	(0.083)	(0.026)	(0.030)	(0.031)	(0.022)
P-value	0.368	0.475	0.671	0.790	0.027	0.117	0.011	0.021
Comparison mean	0.112	0.075	0.568	0.927	0.229	0.291	0.242	0.072
N	982	980	965	862	988	965	966	371
P-val male=fem	0.210	0.136	0.246	0.289	0.138	0.703	0.097	0.008

*** p<0.01, ** p<0.05, * p<0.1

An observation is someone enrolled in the 2008 lottery for secondary school scholarships. Panel A shows results for female lottery participants; Panel B shows results for male lottery participants. "Treatment" means having won the scholarship lottery for Senior High School (SHS). Data Sources: surveys conducted in 2013, 2016, 2019 and 2022. Year of survey in parentheses. "Last pregnancy prenatal care index" is an index over dummies for reporting having gotten prenatal care at last pregnancy in survey rounds 2017, 2019 and 2022. The last row shows the p-values for tests that the effects are identical between males and females. The estimated treatment effects are in each panel's first row; standard errors are in each panel's second row in parentheses clustered at scholarship-eligible respondent-level; p-values from the test that a respective treatment effect is non-zero are reported in the third row; control group means are in each panel's fourth row; sample size for the estimation is in each panel's fifth row. Controls include JHS finishing exam score and baseline region fixed effects.

Table 2: Second Generation Impact: Child Survival

	(1) Survived to one yr (2023)	(2) Survived to three yrs (2023)	(3) Survived to one yr (2023)	(4) Survived to three yrs (2023)
Panel A: Children of Female Scholarship-lottery participants				
Treatment	0.018**	0.018*	0.017**	0.018*
	(0.008)	(0.010)	(0.008)	(0.010)
P-value	0.028	0.065	0.032	0.065
Comparison mean	0.965	0.960	0.965	0.960
N	1,707	1,395	1,707	1,395
Panel B: Children of Male Scholarship-lottery participants				
Treatment	0.014	0.007	0.014	0.009
	(0.010)	(0.012)	(0.010)	(0.012)
P-value	0.161	0.549	0.147	0.477
Comparison mean	0.970	0.971	0.970	0.971
N	1,016	772	1,016	772
P-val male=fem	0.706	0.532	0.612	0.473
Linear Year of birth Control	✓	✓		
Year of birth Fixed Effects			✓	✓

*** p<0.01, ** p<0.05, * p<0.1

An observation is a child of a participant in the lottery for secondary school scholarships. Cols 1 and 3 (respectively, cols 2 and 4) include all children who had reached 12 (respectively, 36) months as of the 2023 survey. Panel A shows results for children of female scholarship lottery participants; Panel B shows results for children of male scholarship-lottery participants. "Treatment" means the child's parent won the scholarship lottery for Senior High School (SHS). The estimated treatment effects are in the first row; standard errors clustered at the scholarship-lottery participant level are in the second row in parentheses; the third row reports the p-value; comparison group means are in the fourth row; the fifth row reports the number of observations. For the regressions in the first two columns, controls include birth order, year of birth, scholarship-lottery participant baseline region fixed effects and the JHS finishing exam score of the scholarship-lottery participant. Regressions in the last two columns include the same controls but instead of controlling for year of birth in a linear way, controls for year of birth fixed effects. Standard errors are clustered at the scholarship-lottery participant-level.

Table 3: Second Generation: Cognitive Development by Milestone Age Window

	(1)	(2)	(3)	(4)	(5)
	Cognitive ability index score				
Age:	1.5 years	2.5 years	3 years	5 years	7 years
Panel A: Children of female scholarship-lottery participants					
Treatment	-0.066	-0.024	0.026	0.238***	0.252**
	(0.095)	(0.127)	(0.079)	(0.084)	(0.119)
P-value	0.489	0.850	0.736	0.005	0.035
Comparison mean	0.005	0.025	-0.016	0.023	0.064
N	560	275	630	667	358
Panel B: Children of male scholarship-lottery participants					
Treatment	0.141	-0.217	-0.004	-0.222*	-0.100
	(0.118)	(0.153)	(0.095)	(0.121)	(0.194)
P-value	0.233	0.156	0.970	0.069	0.607
Comparison mean	-0.009	-0.034	0.041	-0.043	-0.114
N	342	207	342	298	174
P-val male=fem	0.310	0.263	0.938	0.005	0.115

*** p<0.01, ** p<0.05, * p<0.1

An observation is a child of a participant in the lottery for secondary school scholarships at a given age window (there can be multiple observations per child if the child was surveyed at multiple age windows). Panel A shows results for children of female scholarship-lottery participants; Panel B shows results for children of male scholarship-lottery participants. The estimated treatment effects are in the first row; standard errors clustered at the scholarship-lottery participant level are in the second row in parentheses; the third row reports the p-value; comparison group means are in the fourth row; the fifth row reports the number of observations. All regressions control for child age in months, child gender, child birth order, measurement round fixed effects, scholarship-lottery participant baseline region fixed effects, and the JHS finishing exam score of the scholarship-lottery participant. The latent abilities of the child are estimated using a one parameter logistic item response theory model. The results when we score unattempted questions as zeroes instead of missing are shown in Table C1. We have fewer observations at 2.5 years (column 2) and 7 years (column 5) because these tests were introduced later (July 2021 and May 2019, respectively).

Table 4: Caregiver Characteristics, Aspirations and Beliefs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Caregiver is Mother	Completed Secondary Education	Completed Tertiary Education	Earns income	SES index	Num. other children in household	Aspiration: child's years of education	Beliefs: Role of parental stimulation index
Panel A: Children of female scholarship-lottery participants								
Treatment	-0.003	0.248***	0.053**	0.021	0.107	0.059	0.025	0.046
	(0.017)	(0.040)	(0.021)	(0.030)	(0.066)	(0.117)	(0.039)	(0.075)
P-value	0.841	0.000	0.010	0.491	0.103	0.615	0.527	0.546
Comparison mean	0.906	0.218	0.035	0.748	-0.106	1.642	16.753	0.060
N	3,087	2,756	2,756	2,756	2,686	2,757	2,726	2,741
Panel B: Children of male scholarship-lottery participants								
Treatment	0.036	0.010	-0.002	-0.049	0.107	0.160	0.098	0.001
	(0.024)	(0.036)	(0.013)	(0.034)	(0.093)	(0.133)	(0.070)	(0.072)
P-value	0.136	0.785	0.858	0.153	0.252	0.227	0.165	0.987
Comparison mean	0.740	0.196	0.026	0.818	0.079	1.323	16.562	-0.105
N	1,767	1,536	1,536	1,536	1,511	1,536	1,522	1,524
P-val male=fem	0.138	0.000	0.026	0.140	0.908	0.604	0.363	0.610

*** p<0.01, ** p<0.05, * p<0.1

An observation is a child of a participant in the lottery for secondary school scholarships at a given age window (there can be multiple observations per child if the child was surveyed at multiple age windows). Panel A shows results for caregivers of children of female scholarship-lottery participants; Panel B shows results for children of male scholarship-lottery participants. The estimated treatment effects are in the first row; standard errors clustered at the scholarship-lottery participant level are in the second row in parentheses; the third row reports the p-value; comparison group means are in the fourth row; the fifth row reports the number of observations. All regressions control for child age in months, child gender, scholarship-lottery participant baseline region fixed effects, and the junior high school finishing exam score of the scholarship-lottery participant. All columns except (6) also control for child birth order. Components of the SES index are shown in Table B1. "Aspiration": shows the answer to the question "What is the highest level of education that you would like [child name] to complete?". Beliefs on the role of parental stimulation index: A higher value means that the caregiver is more aware of the positive impact of parental stimulation on infant brain development (see components in Table B2).

Table 5: Parental / Caregiver Behavior (Survey Measures)

	(1) Last pregnancy prenatal care index	(2) Preventive health behaviors index	(3) Child stimulation index	(4) Child investment index	(5) Schooling Index
Panel A: Children of female scholarship-lottery participants					
Treatment	0.129**	0.120*	0.147**	0.010	0.020
	(0.057)	(0.066)	(0.059)	(0.053)	(0.069)
P-value	0.023	0.068	0.013	0.850	0.770
Comparison mean	0.022	0.014	-0.006	0.047	0.065
N	795	2,743	2,739	2,742	1,833
Panel B: Children of male scholarship-lottery participants					
Treatment	0.041	0.006	-0.097	-0.047	0.062
	(0.093)	(0.079)	(0.083)	(0.074)	(0.098)
P-value	0.660	0.944	0.241	0.525	0.527
Comparison mean	-0.036	-0.018	0.020	-0.079	-0.130
N	504	1,528	1,528	1,528	941
P-val male=fem	0.328	0.259	0.019	0.576	0.678

*** p<0.01, ** p<0.05, * p<0.1
The unit of observation for column 1 is the scholarship-lottery participant (Panel A) or partner of the scholarship-lottery participant (Panel B). The unit of observations for columns 2 to 5 is a caregiver-child pair, with standard errors clustered at the scholarship-lottery participant level. For the 'Child education index', the sample is restricted to caregiver-child pairs with children over 36 months old. Refer to Appendix B for components of the 'Preventive health behaviors index' (Table B3), 'Child stimulation index' (Table B6), 'Child investment index' (Table B7), 'Child schooling index' (Table B8), and 'Last pregnancy prenatal care index'. All regressions control for scholarship-lottery participant baseline region fixed effects, and the junior high school finishing exam score of the scholarship-lottery participant. Columns 2-5 also control for child age in months, child gender and child birth order.

Table 6: Objectively Measured Child Language and Stimulation: LENA Measurements

	(1) Child vocalizations per min	(2) Conversational turns per min	(3) Meaningful speech	(4) Adult word count per min	(5) Female adult word count per min	(6) Male adult word count per min
Panel A: Children of female scholarship-lottery participants						
Treatment	0.324**	0.068***	0.009	0.423	0.371	0.052
	(0.131)	(0.024)	(0.008)	(0.725)	(0.619)	(0.291)
P-value	0.014	0.005	0.265	0.560	0.549	0.859
Comparison mean	1.956	0.336	0.156	12.952	9.814	3.138
N	560	560	560	560	560	560
Panel B: Children of male scholarship-lottery participants						
Treatment	-0.198	-0.053*	-0.017**	-2.590***	-1.927**	-0.662*
	(0.154)	(0.030)	(0.008)	(0.904)	(0.748)	(0.339)
P-value	0.201	0.072	0.047	0.005	0.011	0.052
Comparison mean	2.210	0.380	0.171	14.260	10.677	3.584
N	391	391	391	391	391	391
P-val male=fem	0.009	0.002	0.020	0.012	0.023	0.118

*** p<0.01, ** p<0.05, * p<0.1

Notes: An observation is a child of a participant in the lottery for secondary school scholarships. Only children aged 14-22 months old between January 2020 and February 2022 were eligible to be recorded using the LENA device. Panel A shows results for children of female scholarship-lottery participants; Panel B shows results for children of male scholarship-lottery participants. “Treatment” means the child’s parent won the scholarship lottery for Senior High School (SHS). The estimated treatment effects are in the first row; standard errors clustered at the scholarship-lottery participant level are in the second row in parentheses; the third row reports the p-value; comparison group means are in the fourth row; the fifth row reports the number of observations. The analysis is restricted to recording times between 7am to 6pm included. Only files with at least 5 hours of recording between 7am and 6pm are kept in the analysis. All regressions control for child age in months, child gender, scholarship-lottery participant baseline region fixed effects, and the junior high school finishing exam score of the scholarship-lottery participant. Conversational turns per minute are measured by the number of times there is one utterance by an adult/child and then one by child/adult in response (within five seconds). “% meaningful speech” is the share of the audio categorized as vocalizations from the target child or speech/vocalizations from adults or other children near the target child.

Table 7: Cost-Effectiveness and Cost-Benefit Analysis of Secondary School Scholarships

	Cost (in USD) per desirable outcome	Medium B-C ratio	Low B-C ratio	High B-C ratio	Medium IRR (in %)	Low IRR (in %)	High IRR (in %)
Panel A: Intergenerational survival benefits only							
Scholarship program for males and females	23,582 per death averted	11	3	37	28	13	41
Scholarship program for females only	15,184 per death averted	16	5	57	37	18	52
Panel B: Intergenerational cognitive development benefits only							
Scholarship program for males and females	10,986 per s.d. increase	3	.8	7	10	4	13
Scholarship program for females only	3,521 per s.d. increase	9	1.6	21	16	7	19
Panel C: Intergenerational benefits							
Scholarship program for males and females	-	13	4	43	29	15	41
Scholarship program for females only	-	25	7	78	38	20	53
Panel D: First-generation labor market benefits only							
Scholarship program for males and females	-	7	2	12	24	10	27
Scholarship program for females only	-	19	3	33	45	15	48
Panel E: Total benefits							
Scholarship program for males and females	-	20	6	56	41	20	51
Scholarship program for females only	-	44	10	111	63	27	76
Panel F: Intergenerational benefits if targeting accuracy falls by 25%							
Scholarship program for males and females	-	10	3	33	24	12	35
Scholarship program for females only	-	19	5	59	32	16	45
Panel G: Total benefits if targeting accuracy falls by 25%							
Scholarship program for males and females	-	15	5	42	34	17	44
Scholarship program for females only	-	33	8	83	53	22	64

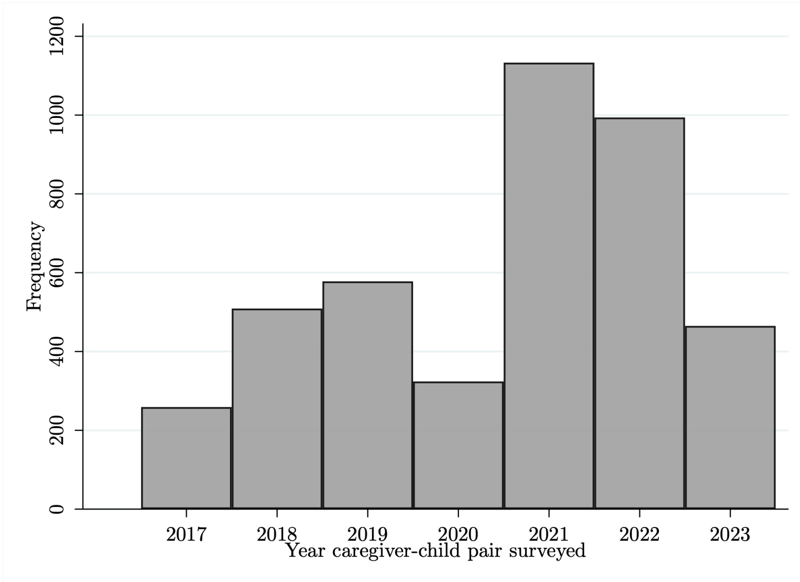
Notes: Panel A only considers the benefits from averting child deaths (Table 2). Panel B only considers the benefits from improving children's cognitive development (Table 3). Panel C jointly considers the benefits from averting child deaths and improving their cognitive development. Panel D only considers the benefits from first-generation labor market gains, in terms of total earnings in the last 6 months for scholarship recipient (estimated for 2023 earnings) Duflo et al. (2024). Panel E considers averting child deaths, improving their cognitive development, and first-generation labor market gains. Panel F and G consider the same benefits as Panel C and E, respectively, but assume that there are 25% fewer students who complete secondary school because of the scholarship program. The first row of each panel presents estimates of the cost effectiveness and benefit-cost ratio if the scholarship were given to male and female students. The second row presents these estimates if the scholarships were only given to female students. The first column calculates how cost-effective scholarships were in averting child deaths (Panel A) or improving cognitive development (Panel B; 's.d. increase' means standard deviation increase in the child's cognitive test scores). The second-fourth columns calculate the benefit-cost ratio using a discount rate of 5%. For the second column (Medium B-C ratio), we use 100*GNI per capita in Ghana for value of a statistical life (VSL) and assume GNI per capita in Ghana is expected adult income absent the intervention. In the third column, we present a less-optimistic scenario, using 33*GNI per capita and first-generation control group income by age ≈ 31 instead. In the fourth column, we present a more-optimistic scenario, using 160*GNI per capita and projected GNI per capita assuming a 3% growth rate. The fifth-seventh columns calculate the internal rate of return under the medium, low, and high scenarios for VSL and expected adult income absent the intervention described above. Refer to Appendix F for more details on these calculations.

Online Appendix

Table of Contents

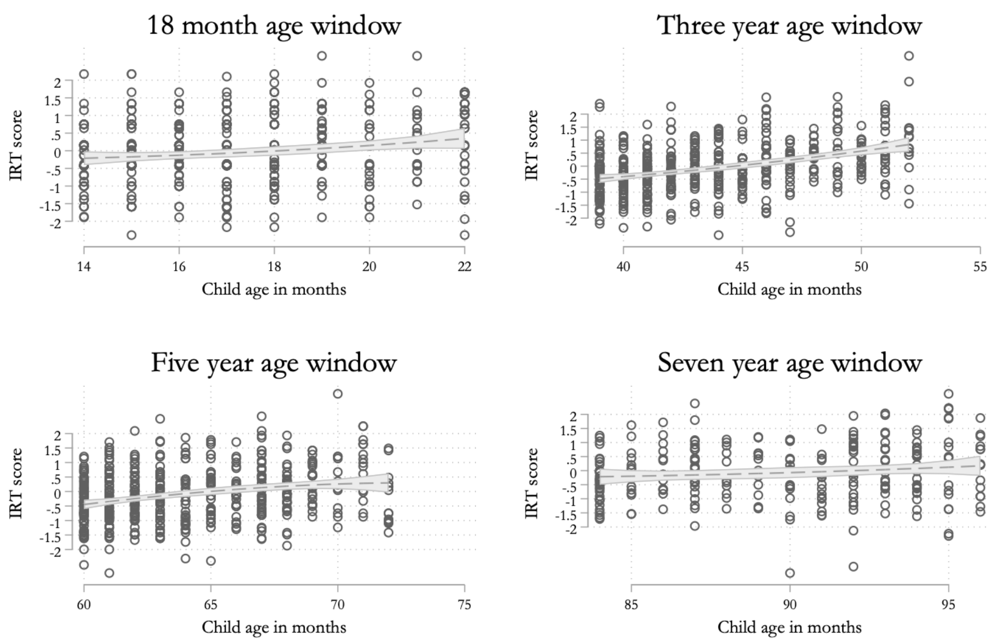
A Appendix Figures and Tables

Figure A.1: Distribution of caregiver-child in-person surveys by year



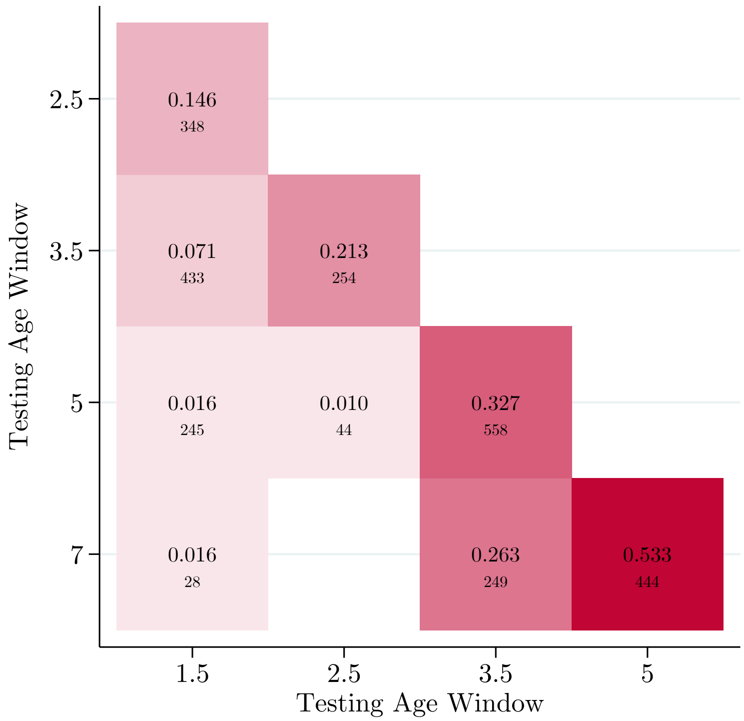
Notes: Caregiver-child in-person surveys refer to the surveys where the caregiver answered a series of questions and the child attempted the child cognitive games.

Figure A.2: Cognitive Development Scores by Age in Months



Notes: Plots only include the control group's children. Data are from 2017-2022 in-person measures administered by the surveyors with the target children at the following target ages: 14-22 months old ("18 month"), 39-52 months old ("Three"), 60-72 months old ("Five"), 84-96 months old ("Seven"). The latent abilities ("IRT score") of the child is estimated using a one parameter logistic item response theory model.

Figure A.3: Correlations Between Cognitive Development Scores Across Age Windows



Notes: This figure shows the pairwise correlations between cognitive test scores for children tested at least twice at different milestone age windows.

Table A1: First Generation: Impact of scholarship offer on education outcomes

	(1)	(2)	(3)	(4)	(5)
	Total years of SHS education (2017)	Total years of tertiary education (2017)	Total years of tertiary education (2022)	Completed SHS (2017)	Completed tertiary (2023)
Panel A: Female Scholarship-lottery participants					
Treatment	1.198***	0.152***	0.234***	0.274***	0.108***
	(0.119)	(0.055)	(0.078)	(0.032)	(0.028)
P-value	0.000	0.006	0.003	0.000	0.000
Comparison mean	1.651	0.210	0.323	0.398	0.118
N	983	996	880	997	836
Panel B: Male Scholarship-lottery participants					
Treatment	1.310***	0.047	0.077	0.282***	0.025
	(0.103)	(0.060)	(0.087)	(0.030)	(0.027)
P-value	0.000	0.437	0.377	0.000	0.356
Comparison mean	2.066	0.316	0.444	0.497	0.153
N	961	971	867	973	826
P-val male=fem	0.375	0.204	0.202	0.745	0.036

*** p<0.01, ** p<0.05, * p<0.1

Notes: This table replicates Duflo et al. (2023). An observation is someone enrolled in the 2008 lottery for secondary school scholarships. Panel A shows results for female scholarship-lottery participants; Panel B shows results for male scholarship-lottery participants. "Treatment" means having won the scholarship lottery for Senior High School (SHS). Data Sources: surveys conducted in 2013, 2016, 2019 and 2022. Year of survey in parentheses. The last row shows the p-values for tests that the effects are identical between males and females. The estimated treatment effects are in each panel's first row; standard errors are in each panel's second row in parentheses clustered at scholarship-eligible respondent-level; p-values from the test that a respective treatment effect is non-zero are reported in the third row; control group means are in each panel's fourth row; sample size for the estimation is in each panel's fifth row. Controls include JHS finishing exam score and baseline region fixed effects.

Table A2: Survey Rates

	All	Control		Treatment		T=C
	Mean	Mean	N	Mean	N	P-value
Panel A: Female Scholarship-lottery participant						
Scholarship-lottery Sample (Parent-level)						
Surveyed (2019)	0.95	0.95	703	0.96	335	0.331
Surveyed (2023)	0.80	0.79	703	0.83	335	0.154
Ever had a child	0.75	0.75	703	0.74	335	0.670
Any child ever elig. during tracking	0.69	0.70	703	0.68	335	0.625
All children too old when tracking began	0.05	0.05	703	0.05	335	0.711
At least one eligible child measured	0.64	0.65	703	0.64	335	0.791
At least one eligible child measured (if had any)	0.86	0.86	527	0.86	247	0.862
Mortality Status available (Child-level)						
1 yr old sample	0.93	0.93	1,268	0.93	559	0.880
3 yrs old sample	0.92	0.92	1,054	0.92	452	0.948
Cognitive Games Sample (Child-level)						
Administered infant assessment if sampled	0.94	0.94	449	0.93	229	0.410
Administered two-yr old assessment if sampled	0.93	0.93	207	0.94	125	0.769
Administered three-yr old assessment if sampled	0.93	0.92	502	0.96	240	0.053
Administered five-yr old assessment if sampled	0.92	0.93	523	0.91	237	0.391
Administered seven-yr old assessment if sampled	0.79	0.78	348	0.81	165	0.481
LENA Sample (Child-level)						
LENA recording available if sampled	0.70	0.69	522	0.72	280	0.296
Panel B: Male Scholarship-lottery participant						
Scholarship-lottery Sample (Parent-level)						
Surveyed (2019)	0.94	0.93	679	0.96	347	0.099
Surveyed (2023)	0.80	0.78	679	0.85	347	0.005
Ever had a child	0.52	0.51	679	0.53	347	0.622
Any child ever elig. during tracking	0.48	0.49	679	0.48	347	0.750
All children too old when tracking began	0.02	0.02	679	0.03	347	0.277
At least one eligible child measured	0.41	0.41	679	0.41	347	0.957
At least one eligible child measured (if had any)	0.79	0.80	349	0.78	184	0.602
Mortality Status available (Child-level)						
1 yr old sample	0.98	0.98	672	0.98	350	0.676
3 yrs old sample	0.98	0.98	511	0.97	265	0.537
Cognitive Games Sample (Child-level)						
Administered infant assessment if sampled	0.91	0.92	266	0.90	155	0.529
Administered two-yr old assessment if sampled	0.89	0.90	173	0.86	86	0.323
Administered three-yr old assessment if sampled	0.91	0.93	267	0.90	153	0.297
Administered five-yr old assessment if sampled	0.88	0.86	250	0.92	123	0.084
Administered seven-yr old assessment if sampled	0.81	0.79	156	0.85	74	0.307
LENA Sample (Child-level)						
LENA recording available if sampled	0.65	0.64	400	0.67	199	0.387

Notes: For the mortality data sample, an observation is a child of a participant in the lottery for secondary school scholarships. It includes all children born as of the 2023 survey. The Cognitive Games sample includes children that were selected to be assessed through the cognitive games. The LENA Sample includes children whose recordings were kept in analysis. Panel A shows results for female scholarship-lottery participants; Panel B shows results for male scholarship-lottery participants.

Table A3: Baseline (2008) Characteristics and Balance—Scholarship-Eligible Students (subsamples with at least one child born by 2023 and at least one child surveyed by 2023)

	(1)	(2)	(3)	(4)	(5)	(6)
	Age in 2008	BECE exam performance	No male head in the household	Number of HH members	Highest education of HH head: primary or less	Highest education of HH head: SHS or more
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Sample with at least one child born by 2023						
Female Scholarship-lottery participants:						
Treatment	-0.174	0.005	-0.022	-0.403**	-0.005	0.020
	(0.133)	(0.007)	(0.043)	(0.181)	(0.042)	(0.031)
P-value	0.191	0.523	0.615	0.026	0.913	0.506
Comparison mean	17.470	0.617	0.468	5.565	0.503	0.151
N	818	759	813	818	814	814
Panel B:						
Male Scholarship-lottery participants:						
Treatment	-0.249	0.007	0.099*	-0.263	0.035	-0.093***
	(0.185)	(0.008)	(0.050)	(0.242)	(0.050)	(0.032)
P-value	0.179	0.382	0.050	0.278	0.483	0.004
Comparison mean	17.600	0.624	0.381	5.803	0.503	0.194
N	568	534	566	568	563	563
P-val male=fem	0.738	0.843	0.070	0.627	0.555	0.012
Panel C: Sample with at least one child surveyed by 2023						
Female Scholarship-lottery participants:						
Treatment	-0.225	0.005	-0.022	-0.424**	-0.032	0.045
	(0.154)	(0.008)	(0.048)	(0.214)	(0.048)	(0.036)
P-value	0.145	0.525	0.643	0.048	0.502	0.206
Comparison mean	17.534	0.619	0.461	5.670	0.519	0.149
N	670	620	666	670	666	666
Panel D:						
Male Scholarship-lottery participants:						
Treatment	-0.358	0.014	0.145**	-0.309	0.001	-0.079**
	(0.236)	(0.010)	(0.060)	(0.281)	(0.060)	(0.037)
P-value	0.130	0.171	0.016	0.273	0.986	0.035
Comparison mean	17.648	0.620	0.390	5.799	0.537	0.181
N	429	403	427	429	425	425
P-val male=fem	0.766	0.521	0.044	0.754	0.632	0.015

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: The unit of observation is a scholarship-lottery participant. Regressions use the number of children of scholarship-lottery participants as weights. Panel A shows results for sample limited to scholarship-lottery participants who ever had a child. Panel B shows results for sample limited to scholarship-lottery participants who ever had a child surveyed. Data Source: Baseline survey conducted in 2008 with scholarship-lottery participants and their guardians. Controls include region fixed effects. Refer to Table 1 for other notes.

Table A4: Characteristics and Balance, per survey sample

	All		Control			Treatment			T=C
	Mean	SD	Mean	SD	N	Mean	SD	N	P-value
Panel A: Children of Female Scholarship-lottery participants									
Mortality Sample									
Child age in months (2023)	84.39	45.31	85.87	45.84	983	81.16	44.02	453	0.067
Child is first-born	0.40	0.49	0.40	0.49	1,186	0.42	0.49	521	0.330
Mom's age at birth	25.32	3.88	25.26	3.90	1,186	25.47	3.85	521	0.307
Cognitive Games Sample									
Child age in months	57.52	91.84	57.05	86.85	1,671	58.46	101.13	834	0.718
Child is female	0.50	0.50	0.49	0.50	1,648	0.52	0.50	817	0.231
Child is first-born	0.37	0.48	0.36	0.48	1,623	0.41	0.49	785	0.031
LENA Sample									
Child age in months	26.09	8.38	26.25	8.69	359	25.80	7.82	202	0.543
Child is female	0.52	0.50	0.50	0.50	358	0.55	0.50	201	0.284
Child is first-born	0.25	0.44	0.25	0.43	346	0.26	0.44	186	0.707
Panel B: Children of Male Scholarship-lottery participants									
Mortality Sample									
Child age in months (2023)	69.69	39.99	69.66	39.81	529	69.74	40.37	299	0.978
Child is first-born	0.51	0.50	0.51	0.50	665	0.52	0.50	344	0.785
Mom's age at birth	22.96	4.28	23.11	4.43	665	22.67	3.97	344	0.127
Cognitive Games Sample									
Child age in months	52.53	80.72	52.75	81.80	884	48.15	49.54	478	0.261
Child is female	0.51	0.50	0.50	0.50	873	0.54	0.50	476	0.176
Child is first-born	0.54	0.50	0.56	0.50	841	0.52	0.50	456	0.168
LENA Sample									
Child age in months	25.83	8.21	25.82	8.20	254	25.84	8.28	134	0.981
Child is female	0.50	0.50	0.48	0.50	250	0.54	0.50	134	0.253
Child is first-born	0.42	0.49	0.43	0.50	233	0.39	0.49	126	0.461

*** p<0.01, ** p<0.05, * p<0.1

Notes: For the mortality data sample, an observation is a child of a participant in the lottery for secondary school scholarships. It includes all children born as of the 2022 survey. The Cognitive Games sample includes children that were selected to undergo the cognitive games. The LENA Sample includes children whose recordings were kept in analysis. Panel A shows results for children of female scholarship-lottery participants; Panel B shows results for children of male scholarship-lottery participants.

Table A5: Child Mortality Results: Robustness to Controls Choices

	(1) Survived to one yr (2023)	(2) Survived to three yrs (2023)	(3) Survived to one yr (2023)	(4) Survived to three yrs (2023)	(5) Survived to one yr (2023)	(6) Survived to three yrs (2023)	(7) Mom's age at birth
Panel A: Children of female scholarship-lottery participants							
Treatment	0.017**	0.018*	0.018**	0.018*	0.019**	0.018*	0.349
	(0.008)	(0.010)	(0.008)	(0.009)	(0.008)	(0.010)	(0.238)
P-value	0.034	0.065	0.026	0.062	0.020	0.053	0.142
Comparison mean	0.965	0.960	0.965	0.960	0.965	0.960	24.797
N	1,707	1,395	1,707	1,395	1,707	1,395	1,825
Panel B: Children of male scholarship-lottery participants							
Treatment	0.013	0.007	0.013	0.006	0.013	0.007	-0.460
	(0.010)	(0.012)	(0.010)	(0.012)	(0.010)	(0.012)	(0.376)
P-value	0.164	0.578	0.174	0.621	0.160	0.549	0.222
Comparison mean	0.970	0.971	0.970	0.971	0.970	0.971	23.022
N	1,016	772	1,016	772	1,016	772	1,033
P-val male=fem	0.748	0.509	0.718	0.520	0.697	0.529	0.111
Birth order fixed effects	No	No	Yes	Yes	Yes	Yes	Yes
Mother's age at birth control	No	No	Yes	Yes	No	No	No
Birth year control	Yes	Yes	Yes	Yes	No	No	No

*** p<0.01, ** p<0.05, * p<0.1

Notes: See Table 2 notes. Columns (1) and (2) replicate columns (1) and (2) of Table 2 but exclude birth order controls. Columns (3) and (4) replicate columns (1) and (2) of Table 2 but controls for mother's age at birth. Columns (5) and (6) replicate columns (1) and (2) of Table 2 but exclude controls for birth year. For column (7), the controls include birth order, scholarship-lottery participant baseline region fixed effects and the JHS finishing exam score of the scholarship-lottery participant.

Table A6: Main Results Using Entropy Balancing

	(1) Survived to one yr (2023)	(2) Survived to three yrs (2023)	(3) 5 years	(4) 7 years
Panel A: Children of Female Scholarship-lottery participants				
Treatment	0.018**	0.017*	0.250***	0.247**
	(0.008)	(0.009)	(0.075)	(0.106)
P-value	0.027	0.073	0.001	0.020
Comparison mean	0.966	0.960	0.026	0.072
N	1707	1395	669	363
Panel B: Children of Male Scholarship-lottery participants				
Treatment	0.012	0.009	-0.259**	-0.048
	(0.009)	(0.012)	(0.121)	(0.169)
P-value	0.212	0.469	0.033	0.775
Comparison mean	0.970	0.971	-0.028	-0.093
N	1016	772	298	175
P-val male=fem	0.712	0.539	0.005	0.109

*** p<0.01, ** p<0.05, * p<0.1

Notes: See Table 2 notes for columns (1) and (2) and notes Table 3 for columns (3) and (4). Entropy balancing is used as a weighting method to match the characteristics of the original scholarship-lottery sample. The characteristics used for balancing are: age in 2008, BECE exam performance, no male head in the household, highest education of household head: SHS or more, highest education of household head: primary or less, and perceived returns to SHS.

Table A7: Main results: Robustness to including partner's education

	(1) Survived to one yr (2023)	(2) Survived to three yrs (2023)	(3) 5 years	(4) 7 years
Panel A: Children of Female Scholarship-lottery participants				
Treatment	0.018**	0.018*	0.217***	0.247**
	(0.008)	(0.009)	(0.084)	(0.119)
Most recent partner completed SHS	-0.000	0.003	0.128	0.055
	(0.011)	(0.012)	(0.097)	(0.125)
Most recent partner completed tert.	-0.002	-0.013	0.258**	0.042
	(0.015)	(0.018)	(0.122)	(0.208)
P-value	0.023	0.054	0.010	0.038
Comparison mean	0.965	0.960	0.023	0.064
N	1707	1395	667	358
Panel B: Children of Male Scholarship-lottery participant				
Treatment	0.014	0.006	-0.203	-0.121
	(0.010)	(0.012)	(0.123)	(0.201)
Most recent partner completed SHS	-0.012	-0.004	0.066	-0.253
	(0.014)	(0.016)	(0.147)	(0.307)
P-value	0.161	0.590	0.100	0.548
Comparison mean	0.970	0.971	-0.043	-0.114
N	1016	772	298	174
P-val male=fem	0.686	0.493	0.009	0.121

*** p<0.01, ** p<0.05, * p<0.1

Notes: Columns (1) and (2): an observation is a child of a participant in the lottery for secondary school scholarships. Controls include birth order, year of birth, scholarship-lottery participant baseline region fixed effects and the JHS finishing exam score of the scholarship-lottery participant. Standard errors are clustered at the scholarship-lottery participant-level.

Columns (3) and (4): an observation is a child of a participant in the lottery for secondary school scholarships at a given age window (there can be multiple observations per child if the child was surveyed at multiple age windows). Regressions control for child age in months, child gender, child birth order, measurement round fixed effects, scholarship-lottery participant baseline region fixed effects, and the JHS finishing exam score of the scholarship-lottery participant. The latent abilities of the child is estimated using a one parameter logistic item response theory model.

Panel A shows results for children of female scholarship-lottery participants; Panel B shows results for children of male scholarship-lottery participants. There are less than 10 partners of male scholarship-lottery participants who completed tertiary.

Table A8: Multiple Hypothesis Testing Adjustments

	(1)	(2)	(3)	(4)	(5)	(6)
	Survival		Cognitive ability index score			
Age:	until 1 year	until 3 year	1.5 years	3 years	5 years	7 years
Panel A: Children of female scholarship-lottery participants						
Treatment	0.018**	0.018*	-0.066	0.026	0.238***	0.252**
	(0.008)	(0.010)	(0.095)	(0.079)	(0.084)	(0.119)
P-value	0.028	0.065	0.489	0.736	0.005	0.035
Step-down p-val	0.119	0.185	0.734	0.736	0.025	0.135
Comparison mean	0.965	0.960	0.005	-0.016	0.023	0.064
N	1707	1395	560	630	667	358
Panel B: Children of male scholarship-lottery participants						
Treatment	0.014	0.007	0.141	-0.004	-0.222*	-0.100
	(0.010)	(0.012)	(0.118)	(0.095)	(0.121)	(0.194)
P-value	0.161	0.549	0.233	0.970	0.069	0.607
Step-down p-val	0.562	0.922	0.669	0.967	0.335	0.922
Comparison mean	0.970	0.971	-0.009	0.041	-0.043	-0.114
N	1016	772	342	342	298	174
P-val male=fem	0.706	0.532	0.310	0.938	0.005	0.115

Notes: Step-down p-values computed using the Romano-Wolf stepdown adjusted p-values procedure. Multiple hypothesis testing procedures are carried out separately for the male and female scholarship-lottery participant sample. Standard errors clustered at the scholarship-lottery participant level shown in parentheses.
In column (1)-(2), an observation is a child of a participant in the lottery for secondary school scholarships. See Table 2 notes.
In columns (3)-(6), an observation is a child of a participant in the lottery for secondary school scholarships at a given age window. See Table 3 notes.
*** p<0.01, ** p<0.05, * p<0.1

Table A9: Second Generation Cognitive Development:
Robustness to excluding Survey-Round Fixed Effects

	(1)	(2)	(3)	(4)	(5)
	Cognitive ability index score				
Age:	1.5 years	2.5 years	3 years	5 years	7 years
Panel A: Children of female scholarship-lottery participants					
Treatment	-0.047	-0.037	0.023	0.232***	0.250**
	(0.095)	(0.129)	(0.083)	(0.085)	(0.118)
P-value	0.620	0.776	0.782	0.006	0.034
Comparison mean	0.005	0.025	-0.016	0.023	0.064
N	560	275	630	667	358
Panel B: Children of male scholarship-lottery participants					
Treatment	0.120	-0.283*	-0.037	-0.203*	0.020
	(0.116)	(0.152)	(0.100)	(0.113)	(0.186)
P-value	0.302	0.065	0.711	0.074	0.915
Comparison mean	-0.009	-0.034	0.041	-0.043	-0.114
N	342	207	342	298	174
P-val male=fem	0.394	0.273	0.656	0.003	0.251

This table reproduces Table 3, excluding survey-round fixed effects. See Table 3 notes.
** p<0.01, ** p<0.05, * p<0.1

Table A10: Second Generation Cognitive Development: Breakdown by Child Gender

	(1)	(2)	(3)	(4)	(5)
	Cognitive ability index score				
Age:	1.5 years	2.5 years	3 years	5 years	7 years
Panel A: Children of Female Scholarship lottery participants					
Treatment	0.082	-0.064	0.032	0.133	0.333**
	(0.138)	(0.200)	(0.100)	(0.117)	(0.152)
Treatment x Female Child	-0.282	0.077	-0.011	0.219	-0.159
	(0.187)	(0.263)	(0.139)	(0.162)	(0.222)
Female child	0.153	-0.095	0.167**	-0.084	0.247**
	(0.108)	(0.153)	(0.083)	(0.085)	(0.124)
P-value	0.553	0.749	0.750	0.255	0.030
Comparison mean	0.005	0.025	-0.016	0.023	0.064
N	560	275	630	667	358
Panel B: Children of Male Scholarship-lottery participants					
Treatment	0.241*	-0.208	0.136	-0.387***	-0.405*
	(0.145)	(0.244)	(0.122)	(0.146)	(0.222)
Treatment x Female Child	-0.192	-0.017	-0.272	0.318	0.612*
	(0.215)	(0.307)	(0.180)	(0.212)	(0.368)
Female child	0.216	-0.099	0.124	0.043	-0.048
	(0.132)	(0.165)	(0.116)	(0.144)	(0.206)
P-value	0.099	0.395	0.265	0.009	0.070
Comparison mean	-0.009	-0.034	0.041	-0.043	-0.114
N	342	207	342	298	174
P-val male=fem	0.310	0.263	0.938	0.005	0.115

*** p<0.01, ** p<0.05, * p<0.1

See Table 3 notes.

Table A11: Caregiver-reported Language Outcomes

	(1) IRT: 18 mo language	(2) IRT: 18 mo says	(3) IRT: 18 mo understands	(4) IRT: 18 mo gestures sometimes
Panel A: Children of female scholarship-lottery participants				
Treatment	-0.022	-0.021	-0.041	-0.016
	(0.075)	(0.066)	(0.079)	(0.072)
P-value	0.772	0.752	0.600	0.821
Comparison mean	-0.025	-0.026	-0.018	0.008
N	645	645	645	645
Panel B: Children of male scholarship-lottery participants				
Treatment	-0.110	-0.092	-0.116	-0.031
	(0.082)	(0.081)	(0.086)	(0.083)
P-value	0.182	0.261	0.177	0.713
Comparison mean	0.083	0.073	0.093	-0.001
N	393	393	393	393
P-val male=fem	0.682	0.774	0.764	0.924

*** p<0.01, ** p<0.05, * p<0.1

Index based on answers to questions adapted from the MacArthur-Bates Communicative Development Inventories (MB-CDI) to our context. All regressions control for child age in months, child gender, child birth order, scholarship-lottery participant baseline region fixed effects, and the junior high school finishing exam score of the scholarship-lottery participant.

Table A12: First Generation Impact of Scholarship on Education Outcomes—scholarship-lottery participants with at least one child surveyed

	(1) Total standardized score (2013)	(2) Total years of education to date (2019)	(3) Completed SHS (2019)	(4) Completed tertiary (2019)	(5) Most recent partner completed tertiary education (2019)
Panel A: Female scholarship-lottery participants					
Treatment	0.193**	1.509***	0.301***	0.054**	0.085**
	(0.088)	(0.188)	(0.040)	(0.022)	(0.041)
P-value	0.028	0.000	0.000	0.016	0.038
Comparison mean	-0.323	10.491	0.288	0.045	0.166
N	648	644	651	651	482
Panel B: Male scholarship-lottery participants					
Treatment	0.055	1.317***	0.274***	0.042	-0.036*
	(0.094)	(0.198)	(0.049)	(0.029)	(0.019)
P-value	0.559	0.000	0.000	0.141	0.063
Comparison mean	0.037	11.184	0.401	0.061	0.043
N	411	421	423	423	284
P-val male=fem	0.259	0.509	0.705	0.667	0.008

*** p<0.01, ** p<0.05, * p<0.1

Notes: Observations are scholarship-lottery participants who ever had a child surveyed. Refer to Table 1 for notes.

Table A13: Household Composition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Caregiver is Father	Caregiver is Grandmother	Lives with Mother	Lives with Father	Lives with both parents	Number of siblings	Number of adults in household
Panel A: Children of female scholarship-lottery participants							
Treatment	-0.001	0.006	-0.009	0.004	0.007	0.004	0.043
	(0.006)	(0.014)	(0.014)	(0.034)	(0.035)	(0.070)	(0.088)
P-value	0.891	0.691	0.515	0.906	0.851	0.952	0.623
Comparison mean	0.014	0.061	0.933	0.646	0.632	1.926	2.380
N	3,087	3,087	3,042	2,695	2,669	2,996	2,733
Panel B: Children of male scholarship-lottery participants							
Treatment	-0.021	-0.022	0.024	-0.087**	-0.092**	0.137	0.064
	(0.018)	(0.015)	(0.019)	(0.039)	(0.043)	(0.093)	(0.099)
P-value	0.238	0.138	0.215	0.024	0.032	0.138	0.521
Comparison mean	0.158	0.077	0.888	0.711	0.656	1.416	2.369
N	1,767	1,767	1,603	1,647	1,491	1,686	1,523
P-val male=fem	0.280	0.133	0.114	0.115	0.120	0.251	0.953

*** p<0.01, ** p<0.05, * p<0.1

Notes: An observation is a child of a participant in the lottery for secondary school scholarships, at a given age window (there can be multiple observations per child if the child was surveyed at multiple age windows). Panel A shows results for children of female Scholarship Study participants; Panel B shows results for children of male Scholarship Study participants. The estimated treatment effects are in the first row; standard errors clustered at the scholarship-lottery participant level are in the second row in parentheses; the third cell row reports p-values of tests of hypotheses of equality of treatment effects; comparison group means are in the fourth cell row; the fifth cell row reports no. of observations. Regression controls include child age, child gender, scholarship-lottery participant baseline region fixed effects and the JHS finishing exam score of the scholarship-lottery participant. Columns (1) - (7) also control for the child birth order. Standard errors are clustered at the scholarship-lottery participant-level.

Table A14: Child Mortality, restricted to firstborns

	(1) Survived to one yr (2023)	(2) Survived to three yrs (2023)	(3) Mom's age at birth
Panel A: Children of female scholarship-lottery participants			
Treatment	0.013	0.013	0.637**
	(0.015)	(0.016)	(0.309)
P-value	0.390	0.415	0.040
Comparison mean	0.955	0.952	22.955
N	691	635	691
Panel B: Children of male scholarship-lottery participants			
Treatment	0.009	0.005	0.149
	(0.014)	(0.016)	(0.366)
P-value	0.512	0.740	0.685
Comparison mean	0.971	0.969	21.643
N	522	440	523
P-val male=fem	0.826	0.754	0.301

*** p<0.01, ** p<0.05, * p<0.1

Notes: An observation is a child of a participant in the lottery for secondary school scholarships. Panel A shows results for children of female Scholarship Study participants; Panel B shows results for children of male Scholarship Study participants. The estimated treatment effects are in the first row; standard errors clustered at the scholarship-lottery participant level are in the second row in parentheses; the third cell row reports p-values of tests of hypotheses of equality of treatment effects; comparison group means are in the fourth cell row; the fifth cell row reports no. of observations. Regression controls in columns (3) include scholarship-lottery participant baseline region fixed effects and the JHS finishing exam score of the scholarship-lottery participant. Controls in column (1)-(2) also include child year of birth fixed effects. Standard errors are clustered at the scholarship-lottery participant-level. All the variables have been restricted to the sample of firstborns.

Table A15: Child Health and Location

	(1) Caregiver reported child health index	(2) Physical development index	(3) Child lives in urban area	(4) Under 3 yrs when began creche/daycare/nursery
Panel A: Children of female scholarship-lottery participants				
Treatment	0.092*	-0.050	0.004	0.019
	(0.047)	(0.037)	(0.037)	(0.029)
P-value	0.050	0.171	0.923	0.518
Comparison mean	0.056	-0.005	0.442	0.757
N	2,742	2,614	2,891	1,833
Panel B: Children of male scholarship-lottery participants				
Treatment	0.008	-0.090	-0.022	0.037
	(0.077)	(0.063)	(0.043)	(0.040)
P-value	0.914	0.155	0.604	0.353
Comparison mean	-0.099	0.009	0.404	0.682
N	1,528	1,478	1,645	941
P-val male=fem	0.372	0.526	0.660	0.597

*** p<0.01, ** p<0.05, * p<0.1

Notes: The details of the indices are shown in Table B4 and Table B5. Refer to Table 4 for other details on specifications and outcomes.

Table A16: LENA Sample, Attrition, Characteristics and Balance

	All		Control			Treatment			T=C
	Mean	SD	Mean	SD	N	Mean	SD	N	P-value
Panel A: Children of Female Scholarship-lottery participant									
LENA Attrition									
Refusal	0.18	0.39	0.20	0.40	474	0.15	0.36	251	0.105
If consent: LENA complete	0.85	0.36	0.85	0.35	428	0.84	0.37	245	0.677
If complete: kept in analysis	0.98	0.14	0.98	0.14	365	0.98	0.14	206	0.984
Baseline Characteristics									
Age in 2008	17.46	1.46	17.54	1.45	358	17.32	1.46	202	0.095
BECE exam performance	0.61	0.09	0.61	0.10	338	0.62	0.08	189	0.288
No male head in the household	0.45	0.50	0.46	0.50	353	0.41	0.49	202	0.221
Highest education of HH head: primary or less	0.55	0.50	0.55	0.50	351	0.54	0.50	202	0.904
Highest education of HH head: SHS or more	0.16	0.37	0.14	0.35	351	0.20	0.40	202	0.072
Current Household Environment									
Nb adults in HH	2.31	1.22	2.29	1.27	357	2.35	1.12	200	0.582
Nb siblings	1.76	1.25	1.76	1.27	345	1.77	1.23	186	0.934
Children 0 to 5 yrs in HH	0.78	0.85	0.74	0.83	357	0.84	0.87	200	0.169
Children 6 to 18 yrs in HH	1.15	1.14	1.14	1.12	357	1.18	1.20	200	0.693
Number of Bedrooms	1.48	0.98	1.44	0.92	357	1.57	1.09	201	0.134
Child is female	0.51	0.50	0.50	0.50	358	0.54	0.50	202	0.283
LENA Measurement									
Nb times device removed	0.22	0.49	0.19	0.49	342	0.26	0.49	197	0.133
Total minutes device removed	10.63	35.26	9.72	36.60	342	12.22	32.83	197	0.428
Nb times held on back	0.27	0.67	0.27	0.67	342	0.27	0.67	197	0.990
Total minutes held on back	10.75	34.03	10.52	36.20	339	11.14	30.02	197	0.840
LENA Recording Duration	906.15	176.42	912.47	164.56	358	894.94	195.60	202	0.259
LENA day reported as unusual	0.06	0.23	0.05	0.21	342	0.08	0.27	197	0.159
Panel B: Children of Male Scholarship-lottery participant									
LENA Attrition									
Refusal	0.22	0.42	0.23	0.42	376	0.20	0.40	177	0.421
If consent: LENA complete	0.84	0.37	0.83	0.37	315	0.85	0.36	163	0.677
If complete: kept in analysis	0.97	0.16	0.97	0.16	262	0.97	0.17	138	0.895
Baseline Characteristics									
Age in 2008	17.60	1.84	17.70	1.81	255	17.42	1.89	134	0.148
BECE exam performance	0.62	0.07	0.61	0.06	241	0.63	0.08	121	0.009
No male head in the household	0.40	0.49	0.36	0.48	254	0.48	0.50	134	0.028
Highest education of HH head: primary or less	0.53	0.50	0.52	0.50	254	0.55	0.50	132	0.580
Highest education of HH head: SHS or more	0.19	0.39	0.22	0.41	254	0.14	0.34	132	0.057
Current Household Environment									
Nb adults in HH	2.40	1.21	2.36	1.15	248	2.47	1.31	134	0.409
Nb siblings	1.21	1.07	1.16	1.03	234	1.32	1.15	126	0.179
Children 0 to 5 yrs in HH	0.81	0.94	0.75	0.86	248	0.94	1.07	134	0.053
Children 6 to 18 yrs in HH	0.84	1.23	0.81	1.25	248	0.90	1.19	134	0.540
Number of Bedrooms	1.40	0.82	1.39	0.83	249	1.43	0.79	134	0.589
Child is female	0.49	0.50	0.47	0.50	255	0.54	0.50	134	0.212
LENA Measurement									
Nb times device removed	0.22	0.49	0.24	0.53	240	0.18	0.41	131	0.270
Total minutes device removed	10.06	34.87	10.93	31.69	238	8.47	40.09	131	0.517
Nb times held on back	0.28	0.63	0.28	0.63	240	0.29	0.64	131	0.874
Total minutes held on back	16.37	63.50	15.07	53.14	238	18.75	79.11	131	0.595
LENA Recording Duration	899.56	192.84	889.81	201.94	255	918.10	173.43	134	0.169
LENA day reported as unusual	0.04	0.20	0.05	0.21	240	0.04	0.19	131	0.729

*** p<0.01, ** p<0.05, * p<0.1

Notes: LENA data. Unit: recording. Sample: LENA surveyed individuals. The analysis is restricted to recording times between 7am to 6pm included. Only files with at least 5 hours of recording between 7am and 6pm are kept in the analysis.

Table A17: LENA Results Using Entropy Balancing

	(1) Child vocalizations per min	(2) Conversational turns per min	(3) Meaningful speech	(4) Adult word count per min	(5) Female adult word count per min	(6) Male adult word count per min
Panel A: Children of female scholarship-lottery participants						
Treatment	0.366***	0.076***	0.010	0.530	0.525	0.005
	(0.118)	(0.022)	(0.006)	(0.668)	(0.529)	(0.301)
P-value	0.002	0.000	0.111	0.428	0.321	0.988
Comparison mean	1.956	0.336	0.156	12.971	9.834	3.137
N	560	560	560	560	560	560
Panel B: Children of male scholarship-lottery participants						
Treatment	-0.183	-0.059**	-0.015*	-2.664***	-2.060***	-0.604*
	(0.138)	(0.026)	(0.008)	(0.876)	(0.735)	(0.313)
P-value	0.186	0.024	0.060	0.003	0.005	0.054
Comparison mean	2.208	0.380	0.171	14.276	10.699	3.578
N	391	391	391	391	391	391
P-val male=fem	0.009	0.002	0.020	0.012	0.024	0.117

*** p<0.01, ** p<0.05, * p<0.1

Notes: See notes of Table 6. Entropy balancing is used as a weighting method to match the characteristics of the original scholarship-lottery sample. The characteristics used for balancing are: age in 2008, BECE exam performance, no male head in the household, highest education of household head: SHS or more, highest education of household head: primary or less, and perceived returns to SHS.

B Indices Components

Table B1: SES—Index Components

	(1) Num. bedrooms per a.e.	(2) Food consumption per a.e.	(3) Metal sheet roof	(4) Mud walls (reversed in index)	(5) Wage-employed
Panel A: Children of female scholarship-lottery participants					
Treatment	0.014	-1.135	0.010	-0.027	0.033
	(0.014)	(3.254)	(0.012)	(0.025)	(0.034)
P-value	0.319	0.727	0.420	0.284	0.341
Comparison mean	0.403	74.497	0.960	0.151	0.158
N	2,729	2,734	2,741	2,741	3,028
Panel B: Children of male scholarship-lottery participants					
Treatment	0.021	-2.539	-0.007	-0.041	0.053
	(0.019)	(4.365)	(0.013)	(0.037)	(0.058)
P-value	0.257	0.561	0.616	0.267	0.365
Comparison mean	0.409	86.255	0.966	0.241	0.406
N	1,518	1,523	1,525	1,525	1,760
P-val male=fem	0.758	0.743	0.368	0.810	0.844

*** p<0.01, ** p<0.05, * p<0.1

Notes: An observation is a child of a scholarship-lottery participant at a given age window (meaning there can be multiple observations per child). Panel A shows results for children of female scholarship study participants; Panel B shows results for children of male scholarship study participants. The estimated treatment effects are in the first row; standard errors clustered at the Scholarship Study participant level are in the second row in parentheses; the third cell row reports p-values of tests of hypotheses of equality of treatment effects; comparison group means are in the fourth cell row; the fifth cell row reports no. of observations. All regressions control for child age, child gender, child birth order, scholarship-lottery participant baseline region fixed effects, and the junior high school finishing exam score of the scholarship-lottery participant. In column names,"reversed" means this component was reverse-scored when we created the relevant index and a.e. stands for number of adult equivalents in the household. The variable wage-employed is based on the 2019 follow-up survey of the scholarship-lottery participant.

Table B2: Caregiver Beliefs on parental stimulation —Index Components

	(1) Believes parents should sing songs to child before turns 7 mos	(2) Believes parents should read stories to child before turns 1	(3) Believes should talk to child in full sentences at birth	(4) Believes should talk to child in full sentences before turns 7 mos
Panel A: Children of female scholarship-lottery participants				
Treatment	0.034	0.033	-0.008	-0.030
	(0.024)	(0.028)	(0.017)	(0.026)
P-value	0.157	0.231	0.651	0.250
Comparison mean	0.818	0.182	0.072	0.208
N	2,741	2,741	2,740	2,740
Panel B: Children of male scholarship-lottery participants				
Treatment	0.028	0.009	-0.014	-0.032
	(0.034)	(0.029)	(0.013)	(0.028)
P-value	0.416	0.752	0.308	0.253
Comparison mean	0.776	0.156	0.041	0.162
N	1,524	1,521	1,524	1,524
P-val male=fem	0.785	0.523	0.751	0.980

*** p<0.01, ** p<0.05, * p<0.1
Refer to Table B1.

Table B3: Preventive Health Behaviors—Index Components

	(1)	(2)	(3)	(4)	(5)	(6)
	Took child for check-up in past 12 mo	Child sleeps under mosquito net	Buys packaged drinking water	Shows card and has all vaccines	Improved Toilet	HH has priv. toilet
Panel A: Children of female scholarship-lottery participants						
Treatment	0.011	0.057*	-0.050	0.013	-0.004	0.076**
	(0.020)	(0.034)	(0.036)	(0.028)	(0.033)	(0.034)
P-value	0.598	0.093	0.172	0.630	0.900	0.027
Comparison mean	0.368	0.610	0.558	0.481	0.302	0.234
N	2,742	2,742	2,742	2,734	2,741	2,741
Panel B: Children of male scholarship-lottery participants						
Treatment	-0.026	0.066	-0.002	-0.024	-0.004	-0.012
	(0.025)	(0.042)	(0.044)	(0.036)	(0.039)	(0.039)
P-value	0.296	0.115	0.968	0.497	0.912	0.753
Comparison mean	0.400	0.588	0.573	0.486	0.270	0.217
N	1,530	1,530	1,530	1,523	1,525	1,525
P-val male=fem	0.314	0.934	0.319	0.327	0.859	0.070

*** p<0.01, ** p<0.05, * p<0.1
Notes: Refer to Table B1.

Table B4: Caregiver-reported Child Health—Index Components

	(1) Cg-report of health	(2) Fevers over 3 mos	(3) Diarrhea over 3 mos	(4) Burned badly ever	(5) Broke bone ever	(6) Concussed ever
Panel A: Children of female scholarship-lottery participants						
Treatment	0.033	-0.067*	-0.012	-0.024	-0.003	-0.005
	(0.041)	(0.037)	(0.017)	(0.018)	(0.008)	(0.008)
P-value	0.421	0.070	0.470	0.186	0.734	0.511
Comparison mean	4.215	0.505	0.147	0.138	0.034	0.041
N	2,742	2,741	2,740	2,742	2,742	2,742
Panel B: Children of male scholarship-lottery participants						
Treatment	-0.047	0.011	-0.012	0.010	0.001	-0.018*
	(0.053)	(0.051)	(0.025)	(0.027)	(0.013)	(0.010)
P-value	0.374	0.833	0.637	0.695	0.933	0.075
Comparison mean	4.150	0.523	0.204	0.175	0.043	0.047
N	1,530	1,528	1,529	1,530	1,530	1,530
P-val male=fem	0.283	0.225	0.991	0.311	0.848	0.416

*** p<0.01, ** p<0.05, * p<0.1

Notes: Refer to Table B1. Caregiver-report of child health was on a Likert scale from 1 (very bad) to 5 (very good). When constructing the index, the injury-related variables shown here are reversed so that a higher Child Health Index means the caregiver reported that the child is healthier.

Table B5: Physical Development—Index Components (Anthropometrics)

	(1) Weight-for-age Z-score	(2) Body mass index-for-age	(3) Length/height for-age	(4) Weight-for length/ height
Panel A: Children of female scholarship-lottery participants				
Treatment	-0.051	-0.039	-0.005	-0.027
	(0.075)	(0.083)	(0.087)	(0.082)
P-value	0.499	0.636	0.951	0.743
Comparison mean	-0.666	-0.489	-0.591	-0.539
N	2,609	1,990	1,999	1,745
Panel B: Children of male scholarship-lottery participants				
Treatment	-0.244**	-0.195*	-0.201*	-0.177
	(0.097)	(0.118)	(0.120)	(0.108)
P-value	0.012	0.099	0.096	0.101
Comparison mean	-0.695	-0.424	-0.657	-0.463
N	1,473	1,090	1,100	974
P-val male=fem	0.138	0.438	0.171	0.432

*** p<0.01, ** p<0.05, * p<0.1

Notes: We calculate z-scores for height-for-age, weight-for-age, weight-for-height, and body mass index using the WHO growth standards. All regressions control for child age in months, child gender, measurement round fixed effects, scholarship-lottery participant baseline region fixed effects, and the junior high school finishing exam score of the scholarship-lottery participant. Column (1) includes all age groups. Columns (2) to (4) exclude the infant/1.5 year-old group.

Table B6: Child Stimulation—Index Components

	(1) Sang to child in past month	(2) Read to child in past month	(3) Told stories to child in past month	(4) Played with child in past month	(5) Named/counted/drew with child in past month
Panel A: Children of female scholarship-lottery participants					
Treatment	0.059**	0.018	0.041	0.028**	0.053**
	(0.024)	(0.025)	(0.029)	(0.014)	(0.023)
P-value	0.014	0.469	0.149	0.043	0.022
Comparison mean	0.661	0.622	0.388	0.882	0.691
N	2,736	2,733	2,730	2,735	2,735
Panel B: Children of male scholarship-lottery participants					
Treatment	-0.029	0.020	-0.037	-0.053**	-0.010
	(0.036)	(0.036)	(0.035)	(0.021)	(0.035)
P-value	0.414	0.582	0.292	0.012	0.775
Comparison mean	0.691	0.546	0.396	0.922	0.685
N	1,524	1,526	1,526	1,528	1,526
P-val male=fem	0.034	0.923	0.098	0.002	0.173

*** p<0.01, ** p<0.05, * p<0.1

Notes: Refer to Table B1.

Table B7: Child Investment- Index Components

	(1) Child ate protein in the morning	(2) Child ate protein in the evening	(3) Number of books	(4) HH has writing materials
Panel A: Children of female scholarship-lottery participants				
Treatment	-0.032	0.020	0.025	0.011
	(0.026)	(0.015)	(0.127)	(0.018)
P-value	0.215	0.191	0.847	0.546
Comparison mean	0.674	0.887	1.571	0.799
N	2,593	2,678	2,722	2,734
Panel B: Children of male scholarship-lottery participants				
Treatment	-0.015	0.004	-0.093	-0.028
	(0.031)	(0.022)	(0.145)	(0.031)
P-value	0.628	0.843	0.521	0.368
Comparison mean	0.682	0.883	1.252	0.742
N	1,474	1,500	1,521	1,523
P-val male=fem	0.617	0.442	0.690	0.284

*** p<0.01, ** p<0.05, * p<0.1

Notes: Refer to Table B1.

Table B8: Schooling—Index Components

	(1) Currently attends school	(2) Currently private school	(3) Mins. in school per day	(4) Under 3 yrs when began creche/daycare/nursery
Panel A: Children of female scholarship-lottery participants				
Treatment	-0.007	0.015	-4.676	0.019
	(0.015)	(0.037)	(9.618)	(0.029)
P-value	0.629	0.691	0.627	0.518
Comparison mean	0.939	0.629	453.659	0.757
N	1,833	1,833	1,833	1,833
Panel B: Children of male scholarship-lottery participants				
Treatment	0.022	-0.002	6.802	0.037
	(0.025)	(0.045)	(14.621)	(0.040)
P-value	0.375	0.961	0.642	0.353
Comparison mean	0.894	0.576	424.507	0.682
N	941	941	941	941
P-val male=fem	0.364	0.816	0.495	0.597

*** p<0.01, ** p<0.05, * p<0.1

Notes: The sample is restricted to children over 36 months old. Refer to Table B1 for other table notes.

C Cognitive Development Results: Robustness to Scoring Decisions, and Results by Skill Type

Table C1: Robustness to Scoring Decisions—unattempted questions scored as incorrect

	(1)	(2)	(3)	(4)	(5)
	Cognitive ability index score				
Age:	1.5 years	2.5 years	3 years	5 years	7 years
Panel A: Children of female scholarship-lottery participants					
Treatment	-0.103	-0.020	0.040	0.237***	0.244**
	(0.101)	(0.128)	(0.080)	(0.083)	(0.119)
P-value	0.306	0.878	0.616	0.005	0.041
Comparison mean	0.007	0.013	-0.010	0.022	0.065
N	560	275	630	667	358
Panel B: Children of male scholarship-lottery participants					
Treatment	0.133	-0.213	-0.005	-0.230*	-0.102
	(0.118)	(0.151)	(0.096)	(0.121)	(0.194)
P-value	0.263	0.160	0.958	0.060	0.600
Comparison mean	-0.010	-0.020	0.030	-0.041	-0.115
N	342	207	342	298	174
P-val male=fem	0.232	0.244	0.822	0.004	0.121

*** p<0.01, ** p<0.05, * p<0.1

Notes: Refer to Table 3. The only difference with Table 3 is that in this case we consider unattempted questions as failed/incorrect (i.e., scored 0) instead of missing.

Table C2: Language Skills Development

	(1) 3.5 years	(2) 5 years	(3) 7 years
Panel A: Children of female scholarship-lottery participants			
Treatment	-0.017	0.156*	0.271**
	(0.083)	(0.088)	(0.113)
P-value	0.841	0.075	0.018
Comparison mean	-0.018	-0.020	0.025
N	630	667	358
Panel B: Children of male scholarship-lottery participants			
Treatment	-0.069	-0.419***	-0.110
	(0.101)	(0.118)	(0.180)
P-value	0.494	0.000	0.544
Comparison mean	0.049	0.046	-0.033
N	342	298	174
P-val male=fem	0.861	0.001	0.049

*** p<0.01, ** p<0.05, * p<0.1

Notes: Refer to Table 3. The 18 months language outcome is based on caregiver reports so it is not included in the overall 18 months cognitive score.

Table C3: Math and Numeracy Development

	(1) 2.5 years	(2) 3.5 years	(3) 5 years	(4) 7 years
Panel A: Children of female scholarship-lottery participants				
Treatment	0.007	0.085	0.155*	0.261**
	(0.132)	(0.079)	(0.085)	(0.118)
P-value	0.957	0.282	0.068	0.027
Comparison mean	-0.005	-0.005	0.043	0.070
N	275	630	667	358
Panel B: Children of male scholarship-lottery participants				
Treatment	-0.173	0.089	-0.075	-0.009
	(0.147)	(0.102)	(0.131)	(0.192)
P-value	0.240	0.382	0.568	0.961
Comparison mean	0.014	0.021	-0.081	-0.126
N	207	342	298	174
P-val male=fem	0.336	0.848	0.234	0.288

*** p<0.01, ** p<0.05, * p<0.1

Notes: Refer to Table 3.

Table C4: Spatial Reasoning

	(1) 2.5 years	(2) 3.5 years	(3) 5 years	(4) 7 years
Panel A: Children of female scholarship-lottery participants				
Treatment	-0.047	0.086	0.195**	0.203*
	(0.123)	(0.085)	(0.087)	(0.117)
P-value	0.699	0.310	0.025	0.084
Comparison mean	0.058	-0.028	0.007	0.067
N	275	630	667	358
Panel B: Children of male scholarship-lottery participants				
Treatment	-0.148	-0.102	-0.205*	0.035
	(0.155)	(0.112)	(0.121)	(0.210)
P-value	0.340	0.367	0.092	0.868
Comparison mean	-0.076	0.044	-0.007	-0.126
N	207	342	298	174
P-val male=fem	0.513	0.134	0.014	0.350

*** p<0.01, ** p<0.05, * p<0.1
Notes: Refer to Table 3.

Table C5: Executive Function

	(1) 1.5 years	(2) 2.5 years	(3) 3.5 years	(4) 5 years	(5) 7 years
Panel A: Children of female scholarship-lottery participants					
Treatment	-0.066	0.005	-0.030	0.242***	0.193
	(0.095)	(0.149)	(0.083)	(0.079)	(0.128)
P-value	0.489	0.972	0.715	0.002	0.133
Comparison mean	0.005	-0.043	0.039	0.033	0.040
N	560	275	630	667	358
Panel B: Children of male scholarship-lottery participants					
Treatment	0.141	-0.111	0.133	-0.015	-0.296
	(0.118)	(0.156)	(0.111)	(0.135)	(0.184)
P-value	0.233	0.480	0.232	0.909	0.110
Comparison mean	-0.009	0.054	-0.070	-0.077	-0.072
N	342	207	342	298	174
P-val male=fem	0.310	0.674	0.226	0.110	0.044

*** p<0.01, ** p<0.05, * p<0.1
Notes: Refer to Table 3.

Table C6: Social Cognitive Development

	(1) 2.5 years	(2) 3.5 years	(3) 5 years	(4) 7 years
Panel A: Children of female scholarship-lottery participants				
Treatment	0.054	-0.034	0.008	0.021
	(0.132)	(0.085)	(0.094)	(0.118)
P-value	0.683	0.692	0.935	0.857
Comparison mean	-0.047	-0.009	0.008	0.083
N	275	630	667	358
Panel B: Children of male scholarship-lottery participants				
Treatment	-0.115	0.014	-0.154	0.221
	(0.170)	(0.130)	(0.148)	(0.184)
P-value	0.500	0.912	0.298	0.233
Comparison mean	0.054	0.029	-0.014	-0.156
N	207	342	298	174
P-val male=fem	0.442	0.795	0.418	0.382

*** p<0.01, ** p<0.05, * p<0.1

Notes: Refer to Table 3.

D Description of the Cognitive Development Measures

In this section, we briefly describe the novel assessments used to measure child cognitive development across a number of domains. Coffey and Spelke (2023) provides further details on the assessments and presents evidence for the validity of each test.

D.1 Executive function assessments

Executive function captures the set of "cognitive skills responsible for formulating goals, planning how to achieve them, and carrying out these plans effectively." (Anderson and Reidy, 2012) For the younger age groups (14-22 months old, 30-36 months old and 3-4 years old), the surveyors administered the executive function games using toys and cups. In these games, the surveyors tested the child's working memory, object permanence, attention switching, and mental simulation/rotation by hiding toys under the cups, moving the cups, and asking the child to identify the location of the toy. Older age groups (5-6 years old and 7-8 years old) completed a Simon task on a computer (Simon and Wolf, 1963).

D.2 Language skill assessments

For 14-22 months old and 30-36 months old, language skills were assessed through caregiver reports on a version of the MacArthur-Bates Communicative Development Inventories Words & Gestures (MB-CDI-WG) adapted to our context through piloting. Surveyors assessed the vocabulary of 3-4 years old, 5-6 years old or 7-8 years old by asking them to point ot the image corresponding to a given word on a laminated card for 3-4 year olds (Figure D.1) and on a computer screen for children above 4 years old. For 7-8 year olds, surveyors also tested their ability to identify letters or words, and their ability to read simple sentences. These assessments were carried in the primary language spoken to the child in their home according to the caregiver.

Figure D.1: 3-year old playing Vocabulary game

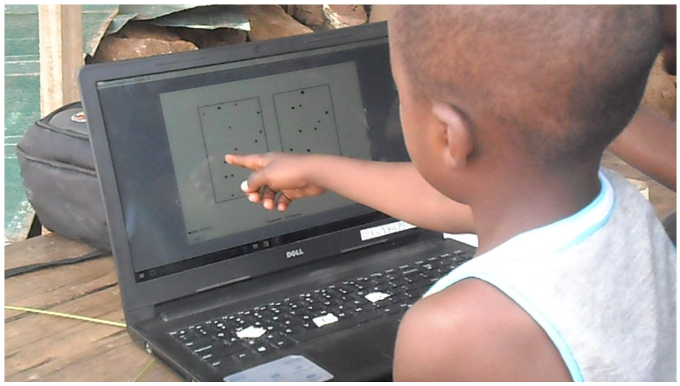
Note: A surveyor testing a 3-year old's vocabulary during piloting using laminated cards. The surveyor says a word in the child's primary language and asks the child to point to the picture of this word.

D.3 Math & numeracy assessments

For 14-22 months, 30-36 months old, 3-4 year olds and 5-6 years old, we tested their approximate number sense (ANS) aptitude through the Panamath game (Figure D.2). Research has shown that

ANS aptitude at early ages is predictive of mathematical ability later on in life (Libertus et al., 2011) Surveyors used laminated cards for children under 4 years old and computers for children above 4 years old. For 30-36 months old, 3-4 year olds and 5-6 years old, surveyors also tested the child's ability to identify Arabic numbers and correctly count the numbers of objects in a given picture. For children over 3 years old, we tested addition and subtraction skills as well.

Figure D.2: 5-year old playing Panamath game



Notes: A surveyor playing the Panamath game with a 5-year old during piloting. This game measures the child's approximate number system aptitude. The child is asked which box has more dots or which dots has more dots of a particular color.

D.4 Social cognition assessments

To measure social cognition, we tested the child's ability to interpret the gaze and emotion of faces. In these tests, the child had to correctly interpret the gaze or emotion of the faces shown to them on the computer or laminated cards. The gaze-related tests were conducted with children over 30 months old, while the emotion-related tests were conducted with children over 3 years old.

D.5 Spatial reasoning assessments

The spatial reasoning tests measured the child's ability to concerns skills think about and manipulate objects in space. The spatial reasoning assessments asked 30-36 month-olds to identify the block that matched the shape of a given hole. Children above 3 years old were tesed on their ability to read maps. They were shown a laminated map that charted the shapes on a large mat places on the floor (Figure D.3). The surveyor pointed to the map and asked the child to place an object on the corresponding spot on the mat. All children above 3 years old were tested on their understanding of geometry as well.

Figure D.3: 5-year old playing Reading Maps game

Notes: A five-year old placing an object on the indicated spot during the Reading Maps game in piloting.

References

Anderson, Peter J. and Natalie Reidy, "Assessing Executive Function in Preschoolers," *Neuropsychology Review*, 2012, *22 (4)*, 345-360.

Coffey, Joseph R. and Elizabeth S. Spelke, "Evaluating a battery of direct assessments of the cognitive skills of preschool children in Ghana," *Working Paper*, 2023.

Simon, J. Richard and James D. Wolf, "Choice Reaction Time As A Function Of Angular Stimulus-Response Correspondence And Age," *Ergonomics*, 1963, *6 (1)*, 99-105.

E Description of the Language and ENvironmental Analysis (LENA) System for Measuring the Child's Auditory Environment

E.1 Description of technology and survey protocols

The Language Environment Analysis (LENA) System is a recording technology that is used to record the naturalistic speech that a child hears and produces throughout the day. In this study, we used the LENA to measure the auditory environment of children aged 14-22 months and 30-36 months.

The LENA device was given to the family by the surveyor team that administered the caregiver survey and cognitive assessments for children 14-22 months and 30-36 months. The device was held in a T-shirt so that it could easily be worn by the target child the next day. The family was instructed that the child should wear the shirt and device the entire day if possible (refer to Figure E.1 for example of a child wearing the LENA shirt). The device has a battery life of 16 hours.

A separate team of surveyors was tasked with collecting the device from the family and administering a short debrief survey. The short debrief survey included questions about factors that may have influenced the quality of the data (e.g., carrying the child on your back, taking off the LENA device, or having an unusual day).

Figure E.1: Child wearing LENA shirt

Notes: A child wearing the LENA shirt during piloting. The LENA device is stored in the patterned pocket on the front of the shirt.

E.2 Description of analysis of LENA data

The audio from the LENA device was processed using LENA Pro software. LENA Pro software automatically generates a number of variables relating to the child's auditory environment. Below we define the variables used in this paper:[33]

- **Adult words count:** Number of words spoken by adults near the child during the recording.
- **Child vocalization count:** Number of speech-like utterances by the child wearing the LENA device. Speech-like utterances include words or pre-speech communicative sounds (e.g., babbles or squeals), but excludes non-speech sounds (e.g., breathing or burping).
- **Conversational turn count:** Number of times an adult and the child wearing the LENA device have a back-and-forth vocal interaction (i.e. the adult speaks and child vocalizes or vice verse.)
- **% audio of meaningful speech:** Percentage of the total audio recorded by the device containing speech where the LENA could identify the primary speaker as a female adult, male adult, child wearing the LENA device, or other child. Excludes overlapping speech where the LENA could not identify a primary speaker.

References

Gilkerson, Jill and Jeffrey A Richards, "A Guide to Understanding the Design and Purpose of the LENA System," *LENA Foundation Technical Report*, 2020

Xu, Dongxin, Umit Yapanel, and Sharmi Gray, "Reliability of the LENA Language Environment Analysis System in Young Children's Natural Home Environment" *LENA Foundation Technical Report*, 2009

[33] These definitions draw upon Xu et al. (2009) and Gilkerson and Richards (2020).

F Details of Cost-Effectiveness and Cost-Benefit Analyses

This section describes in-detail how we calculate the cost-effectiveness, benefit-cost ratios, and internal rates of return (IRRs) presented in Table 7.

F.1 Cost Calculations

From a social perspective, the costs of the secondary school scholarship program are the additional expenditures on education and the foregone wages of the recipients. We calculate costs for two types of programs: a means-tested scholarship program for both genders and a means-tested scholarship program restricted to females-only. In the first case, we use the cost estimates for our entire sample. In the second, we use the cost estimates for the females in our sample.

In all scenarios, we use the following formula to calculate the cost of the program per child:

$$\text{Cost per scholarship recipient}_p = \beta_p^{\text{SHS years}} \times \text{Yearly SHS fees} + \beta_p^{\text{Foregone Wages}} + \beta_p^{\text{Non-fee Education Expenses}} \quad (2)$$

Where p denotes the program type. Note that each of the costs is characterized in terms of the spending induced by the treatment, regardless of whether the spending is by the scholarship program or the student's family.[34]

$\beta_p^{\text{SHS years}}$, $\beta_p^{\text{Foregone Wages}}$, and $\beta_p^{\text{School Materials Expenses}}$ are the treatment effect of the scholarship on these outcomes. In our sample, years of SHS increased by 1.27 (1.20 for females), non-fee education expenses by \$206 per recipient (\$186 for females),[35] and wages decreased by \$180 (\$132 for females) between Winter 2009 and Summer 2013 (the time when the first-generation would have gone to SHS).[36] These estimates are based on self-reported SHS attendance, expenditures on school materials, and wages. Since Duflo et al. (2024) were involved in the administration of the scholarship, they have detailed administrative data on the school fees covered. On average, the scholarships paid \$157 in SHS fees per year of SHS attended by a scholarship recipient. Taken together the costs per recipient are \$584 (\$505 for females only).

F.2 Benefits Calculations

We consider the benefits from the scholarship in terms of averted child deaths, improvements in child cognitive development, and increases in scholarship recipient wages. As inputs into the benefit-cost ratios and internal rates of return, we calculate the benefits for the average scholarship recipient for each year after completion of the intervention (i.e. completion of secondary school).

F.2.1 Benefit from child cognitive gains

We assume that the primary benefit from cognitive gains for the child are through labor market outcomes. Since we would need to wait decades to gather labor market information for the children of scholarship recipients, we use the estimates from Gertler et al. (2014) to proxy for these gains. Gertler et al. (2014) exploit the randomization of an early childhood stimulation program in Jamaica in the 1980s. They follow-up with the beneficiaries of this program as adults, finding significant impacts of the program on their adult

[34] In our sample, the school fees were covered by the scholarship but the school materials and forgone wages were not.

[35] For additional school expenses, we rely on reports of spending on transportation and other school materials that were not covered by the scholarship. Then multiply these additional expenses by the number of additional school terms attended by scholarship recipients (3*additional SHS years)

[36] We calculate foregone wages based on wages reported by respondents in 2013 (when respondents would have still been in SHS). The scholarship led to a decrease in monthly wages. Assuming this effect was constant over the ≈ 42 months of SHS, we get the total foregone wages.

wages. They estimate that that a 1 s.d. improvement in early childhood cognitive development results in a 33% increase in adult wages, assuming that the effect of the program on adult wages comes solely through improved early childhood cognitive development. We assume that this relationship holds in our sample. We multiply the average of the effect of the scholarship on cognitive scores of the five and seven year olds by 33% to get the increase in adult wages due to the early childhood cognitive development caused by the scholarship program, Effect on earnings_p.

We assume that the second-generation will start earning wages at 18 years old and that their careers will last 40 years (i.e. $t = 40$). Since the median child was born 5 years after the completion of the intervention, this means that the benefits for child cognitive gains begin 23 years after the completion of the intervention and end 63 years after the completion of the intervention. Using these assumptions, the benefits from child cognitive gains for the average scholarship recipient t years after the completion of the intervention is:

$$\text{Child cog. benefits}_{p,s,t} = \begin{cases} \text{Effect on child's earnings}_p \times \text{Child's earnings absent the intervention}_s & \text{if } 23 < t \leq 63, \\ \times \text{ Num. children}_p & \\ 0 & \text{otherwise.} \end{cases}$$

Where s denotes whether this is the medium, low, or high benefits scenario. We estimate 'Num. children_p' based on the average number of children for scholarship recipients in our sample for both genders (1.5) and for females only (1.9) as of 2023 (9 years after the scholarship when respondents are 31 years old on average) which is conservative since recipients are likely to have more children in their 30s. Given the uncertainty around the child's future earnings absent the intervention, we vary our estimate of 'Child's earnings absent the intervention_s' across the medium, low, and high benefits scenarios. In the medium scenario, we use Ghana's current GNI per capita. We use the first-generation control group's earnings in 2023 (when they were 31 years old; measured by Duflo et al. (2024)) in the low scenario.[37] In the high scenario, we project yearly earnings by assuming a 3% growth rate of GNI per capita. Essentially, this scenario means that the discount rate is $r - g$ rather than just r. Since we use a discount rate of 5% ($r = .05$), we can calculate benefits in the high scenario by using a 2% discount rate and Earnings absent the intervention equal to current GNI per capita.

F.2.2 Benefit from first-generation labor market gains

We apply a similar method to calculate the benefits from the labor market gains for the scholarship recipients. Based on Duflo et al. (2024), we model the gains as beginning 8 years after the initial scholarship ends ($t = 8$) and lasting for the remainder of the first-generation respondent's working life (29 years; until $t = 37$). Duflo et al. (2024) estimate a 30% increase in earnings for females by 2023 and a 13% increase in earnings overall. With this method, the benefits from first-generation labor market gains for the average scholarship recipient t years after the completion of the intervention is:

$$\text{1st gen. labor market benefits}_{p,s,t} = \begin{cases} \text{Effect on 1st gen. earnings}_p & \\ \times \text{ 1st gen. earnings absent the intervention}_s & \text{if } 8 < t \leq 37, \\ \times \text{ Num. children}_p & \\ 0 & \text{otherwise.} \end{cases}$$

Once again, we use three estimates of counterfactual yearly earnings: GNI per capita (medium scenario), first-generation control group mean earnings in 2023 (low scenario), and projected GNI per capita in each year using a 3% growth rate (high scenario). The counterfactual yearly earnings depend upon the program type in this case, because in the low scenario, first-generation control group mean earnings in 2023 are lower

[37]Since children are male and female regardless of the scholarship recipient's gender, we use the same Yearly earnings absent the intervention estimate for each program type.

for females than overall ($417 vs. $687).

F.2.3 Benefits from averted child deaths

We quantify the benefits averted child deaths using the Value of a Statistical Life (VSL) benchmarks. These benchmarks come from the recommendations of Robinson et al. (2019):

1. Low scenario: "VSL extrapolated from a U.S. estimate to the target country using an income elasticity of 1.5. The starting point for this calculation should be a U.S. VSL to GNI per capita ratio of 160 , based on a U.S. VSL of $9.4 million and U.S. GNI per capita of $57,900."

2. High scenario: "VSL = 160* GNI per capita in the target country. This calculation applies the U.S. ratio to all countries, which is equivalent to using that ratio as the starting point and assuming income elasticity is 1.0."

3. Medium scenario: "VSL = 100* GNI per capita in the target country. This calculation applies the OECD ratio to all countries, which is equivalent to using that ratio as the starting point and assuming income elasticity is 1.0."

With the VSL estimates in hand, we can calculate the survival benefits per child using the following formula:

$$\text{Survival benefits per child}_{p,s} = \text{VSL}_s * \text{Age weight} * \text{Effect on mortality}_p$$

We multiply the VSL by an age weight of 1.26 to reflect the fact that more years of life are lost due to under-3 deaths relative to deaths of the average Ghanaian.[38]

Effect on mortality_p differs by the program type. Given our assumptions, Effect on mortality_p is equal to the effect on the children of female scholarship recipients (1.8 pp decrease; Table 2) for the scholarships available to females-only, while it is equal to the effect on scholarship recipients of either gender (1.6 pp decrease) when we evaluate the program open to both genders.

To ensure we apply the same assumptions for discount rates and GNI per capita growth to all benefits, we construct benefits for each year after completion of the scholarship by assuming that these VSL estimates are the net present value of the value of life year over 70 years with a discount rate of 5%. This allows us to use projected GNI per capita in each year using a 3% growth rate in the high scenario.[39]. For the IRR, it allows us to identify the discount rate where the investment would still be socially-efficient accounting for the fact that the discount rate will affect the value of an additional year of life just as it will affect the value of an additional dollar in the future. Importantly, we assume an equal number of child deaths are averted in the first 10 years after the end of the intervention and (conservatively) assume that no child deaths are averted after this point. Therefore, the effect in terms of survival benefits phases-in over 10 years and phases out over 10 years. Together, these assumptions give us the following formula for the benefits from averted child deaths for the average scholarship recipient t years after the completion of the intervention:

$$\text{Child survival benefits}_{p,s,t} = \begin{cases} \dfrac{\text{Survival benefits per child}_{p,s} \times 0.05}{1-(1+0.05)^{-70}} * \dfrac{t-1}{10} & \text{if } 1 < t < 11, \\ \dfrac{\text{Survival benefits per child}_{p,s} \times 0.05}{1-(1+0.05)^{-70}} & \text{if } 11 \leq t \leq 71, \\ \dfrac{\text{Survival benefits per child}_{p,s} \times 0.05}{1-(1+0.05)^{-70}} * \dfrac{81-t}{10} & \text{if } 71 < t \leq 81. \end{cases}$$

[38] The average Ghanaian is 20 years old, so we compute $\frac{\text{Years of Life Lost due to 1-5 year old death}}{\text{Years of Life Lost due to 20-25 year old death}}$ using the WHO's 2019 guidelines (World Health Organization, 2020).

[39] In the low and medium scenario, we use current GNI per capita

F.3 Benefit-Cost Ratio and Internal Rate of Return Calculations

Unsurprisingly, the formula for the benefit cost ratio is:

$$\text{Benefit-Cost ratio}_{p,s} = \frac{\text{Benefits per scholarship recipient}_{p,s}}{\text{Costs per scholarship recipient}_{p,s}}$$

. To calculate the benefits, we take the net present value of the benefits in each year after completion of the intervention:

$$\text{Benefits per scholarship recipient}_{p,s} = \sum_{t=0}^{T} \frac{\text{Benefits}_{p,s,t}}{(1+r)^t}$$

using $r = .05$.

We calculate the internal rate of return as the discount rate that makes the benefits of the program equal to the costs by solving the following formula for IRR:

$$0 = \sum_{t=0}^{T} \frac{\text{Benefits}_{p,s,t}}{(1+\text{IRR})^t}) - \text{Costs per scholarship recipient}_{p,s}$$

Across panels, we vary what benefits are included in $\text{Benefits}_{p,s,t}$:

- Panel A: $\text{Benefits}_{p,s,t} = \text{Child survival benefits}_{p,s,t}$
- Panel B: $\text{Benefits}_{p,s,t} = \text{Child cog. benefits}_{p,s,t}$
- Panel C: $\text{Benefits}_{p,s,t} = \text{1st-gen. labor market benefits}_{p,s,t}$
- Panel D:

$$\text{Benefits}_{p,s,t} = \text{Child survival benefits}_{p,s,t} + \text{Child cog. benefits}_{p,s,t}$$

- Panel E:

$$\text{Benefits}_{p,s,t} = \text{Child survival benefits}_{p,s,t} + \text{Child cog. benefits}_{p,s,t} + \text{1st-gen. labor market benefits}_{p,s,t}$$

- Panel F:

$$\text{Benefits}_{p,s,t} = (\text{Child survival benefits}_{p,s,t} + \text{Child cog. benefits}_{p,s,t}) * .75$$

- Panel G:

$$\text{Benefits}_{p,s,t} = (\text{Child survival benefits}_{p,s,t} + \text{Child cog. benefits}_{p,s,t} + \text{1st-gen. labor market benefits}_{p,s,t}) * .75$$

The program type, p, varies across the rows of Table 7 and the scenario presented, s, varies across the columns.